



# Probabilidade e Estatística para Engenharia e Ciências

**Jay L. Devore**

# **Probabilidade e Estatística: para Engenharia e Ciências**

Tradução da SEXTA EDIÇÃO Norte-Americana

**Dados Internacionais de Catalogação na Publicação (CIP)**  
**(Câmara Brasileira do Livro, SP, Brasil)**

Devore, Jay L.

Probabilidade e estatística : para engenharia e ciências / Jay L. Devore ; [tradução Joaquim Pinheiro Nunes da Silva]. -- São Paulo: Cengage Learning, 2006.

Título original: Probability and statistics for engineering and the sciences  
bibliografia.  
ISBN 978-85-221-0924-1

1. Engenharia 2. Estatística matemática I. Título.

05-1183

CDD-519.5

**Índice para catálogo sistemático:**

1. Estatística matemática 519.5

**Probabilidade e estatística: para engenharia e ciências****Jay L. Devore**

Gerente Editorial: Patrícia La Rosa

Editora de Desenvolvimento: Tatiana Pavanelli Valsi

Produtora Editorial: Ligia Cosmo Cantarelli

Produtora Gráfica: Fabiana Alencar Albuquerque

Título original: Probability and  
Statistics for Engineering and the Sciences  
ISBN: 0-534-39933-9Tradução: MGS Language Services  
Joaquim Pinheiro Nunes da Silva

Copidesque: Roberta Costa Gonçalves de Almeida

Revisão Técnica: Armando Zeferino Milioni

Revisão: Iná de Carvalho  
Cristina Paixão Lopes  
Silvana GouveiaDiagramação: ERJ – Composição Editorial e Artes  
Gráficas Ltda.

Capa: FZ. Dáblio

© 2004 Brook/Cole

© 2005 Cengage Learning Edições Ltda.

Todos os direitos reservados. Nenhuma parte deste livro poderá ser reproduzida, sejam quais forem os meios empregados, sem a permissão, por escrito, da Editora.

Aos infratores aplicam-se as sanções previstas nos artigos 102, 104, 106 e 107 da Lei nº 9610, de 19 de fevereiro de 1998.

Para informações sobre nossos produtos, entre em contato pelo telefone **0800 11 19 39**Para permissão de uso de material desta obra, envie seu pedido para **[direitosautorais@cengage.com](mailto:direitosautorais@cengage.com)**

© 2005 Cengage Learning. Todos os direitos reservados.

ISBN-13: 978-85-221-0924-1

**Cengage Learning**

Condomínio E-Business Park

Rua Werner Siemens, 111 – Prédio 20 – Espaço 04

Lapa de Baixo – CEP 05069-900 – São Paulo – SP

Tel.: (11) 3665-9900 – Fax: (11) 3665-9901

SAC: 0800 11 19 39

Para suas soluções de curso e aprendizado, visite  
**[www.cengage.com.br](http://www.cengage.com.br)**

Impresso no Brasil.

*Printed in Brazil.*

1 2 3 4 10 09 08 07 06 05

# Probabilidade e Estatística: para Engenharia e Ciências

Tradução da SEXTA EDIÇÃO Norte-Americana

**Jay L. Devore**

California Polytechnic State University  
San Luis Obispo

**Tradução:**

MGS Language Services e  
Joaquim Pinheiro Nunes da Silva

**Revisão Técnica:**

Professor Armando Zeferino Milioni  
Departamento de Organização  
Divisão de Engenharia Mecânica-Aeronáutica  
Instituto Tecnológico de Aeronáutica



*Para minha esposa, Carol:*

*Sua dedicação ao ensino  
é uma contínua inspiração para mim.*

*Para minhas filhas, Allison e Teresa:*

*O enorme orgulho que tenho de  
suas realizações não conhece fronteiras.*



# Sumário

<b>1</b>	<b>Visão geral e estatística descritiva</b>	1
	Introdução	1
	1.1 Populações, amostras e processos	2
	1.2 Métodos tabular e gráfico em estatística descritiva	9
	1.3 Medidas de localização	24
	1.4 Medidas de dispersão	31
	Exercícios suplementares	41
	Bibliografia	44
<b>2</b>	<b>Probabilidade</b>	45
	Introdução	45
	2.1 Espaços amostrais e eventos	45
	2.2 Axiomas, interpretações e propriedades da probabilidade	50
	2.3 Probabilidade condicional	66
	2.4 Técnicas de contagem	58
	2.5 Independência	75
	Exercícios suplementares	81
	Bibliografia	84
<b>3</b>	<b>Variáveis aleatórias discretas e distribuições de probabilidades</b>	85
	Introdução	85
	3.1 Variáveis aleatórias	85
	3.2 Distribuições de probabilidade para variáveis aleatórias discretas	89
	3.3 Valores Esperados de Variáveis Aleatórias Discretas	98
	3.4 Distribuição de probabilidade binomial	105
	3.5 Distribuições binomiais hipergeométrica e negativa	113
	3.6 Distribuição de probabilidade de Poisson	118
	Exercícios suplementares	
	Bibliografia	126
<b>4</b>	<b>Variáveis aleatórias contínuas e distribuições de probabilidade</b>	127
	Introdução	127
	4.1 Variáveis aleatórias contínuas e funções de densidade de probabilidade	127
	4.2 Funções de distribuição acumulada e valores esperados	133
	4.3 A distribuição normal	141
	4.4 A distribuição gama e seus parentes	152



- 4.5 Outras distribuições contínuas 158
- 4.6 Gráficos de probabilidade 165
  - Exercícios suplementares 173
  - Bibliografia 178

## **5 Distribuições de probabilidade conjunta e amostras aleatórias** 179

- Introdução 179
- 5.1 Variáveis aleatórias de distribuição conjunta 179
- 5.2 Valores esperados, covariância e correlação 191
- 5.3 Estatísticas e suas distribuições 197
- 5.4 A distribuição da média amostral 207
- 5.5 Distribuição de uma combinação linear 213
  - Exercícios suplementares 218
  - Bibliografia 220

## **6 Estimativa pontual** 221

- Introdução 221
- 6.1 Alguns conceitos gerais sobre estimativa pontual 221
- 6.2 Métodos de estimativa pontual 235
  - Exercícios suplementares 243
  - Bibliografia 244

## **7 Intervalos estatísticos baseados em uma única amostra** 245

- Intervalo 245
- 7.1 Propriedades básicas de intervalos de confiança 246
- 7.2 Intervalos de confiança para amostras grandes para uma média e proporção da população 254
- 7.3 Intervalos baseados em uma distribuição normal da população 261
- 7.4 Intervalos de confiança para variância e desvio padrão de uma população normal 269
  - Exercícios suplementares 272
  - Bibliografia 274

## **8 Testes de hipóteses com base em uma única amostra** 275

- Introdução 275
- 8.1 Hipóteses e procedimentos de teste 275
- 8.2 Testes quanto à média de uma população 285
- 8.3 Testes com relação a uma proporção da população 295
- 8.4 Valores  $P$  300
- 8.5 Alguns comentários sobre como selecionar um procedimento de teste 307
  - Exercícios suplementares 310
  - Bibliografia 313

## **9 Inferências baseadas em duas amostras** 314

- Introdução 314
- 9.1 Testes  $z$  e intervalos de confiança para diferença as médias de duas populações 314
- 9.2 Teste  $t$  e intervalo de confiança de confiança de duas amostras 324

- 9.3 Análise de dados pareados 332
- 9.4 Inferências quanto à diferença entre proporções de populações 341
- 9.5 Inferências sobre variâncias de duas populações 347
  - Exercícios suplementares 356
  - Bibliografia 356

## **10 A análise de variância** 357

- Introdução 357
- 10.1 ANOVA de fator único 358
- 10.2 Comparações múltiplas em ANOVA 367
- 10.3 Mais sobre ANOVA de fator único 373
  - Exercícios suplementares 382
  - Bibliografia 384

## **11 ANOVA Multifatores** 385

- Introdução 385
- 11.1 ANOVA de dois fatores com  $K_{ij} = 1$  386
- 11.2 ANOVA de dois fatores com  $K_{ij} > 1$  398
- 11.3 ANOVA de três fatores 406
- 11.4 Experimentos fatoriais  $2^p$  415
  - Exercícios suplementares 428
  - Bibliografia 431

## **12 Regressão linear simples e correlação** 432

- Introdução 432
- 12.1 O modelo de regressão linear simples 433
- 12.2 Estimando parâmetros de modelo 440
- 12.3 Inferências sobre o parâmetro do coeficiente angular  $\beta_1$  452
- 12.4 Inferências sobre  $\mu_{Y \cdot x^*}$  e a previsão de valores de  $Y$  461
- 12.5 Correlação 468
  - Exercícios suplementares 477
  - Bibliografia 482

## **13 Regressão não-linear e múltipla** 483

- Introdução 483
- 13.1 Aplicabilidade do modelo e verificação do modelo 484
- 13.2 Regressão com variáveis transformadas 491
- 13.3 Regressão polinomial 501
- 13.4 Análise de regressão múltipla 511
- 13.5 Outras questões em regressão múltipla 532
  - Exercícios suplementares 544
  - Bibliografia 549

## **14 Testes de aderência e análise de dados categorizados** 550

- Introdução 550
- 14.1 Testes de aderência quando as probabilidades das categorias são completamente especificadas 551
- 14.2 Testes de aderências para hipóteses compostas 558
- 14.3 Tabelas de contingência de entrada dupla 569
  - Exercícios suplementares 576
  - Bibliografia 579

## 15 Procedimentos livres de distribuição 580

- Introdução 580
- 15.1 O teste de postos com sinais de Wilcoxon 581
- 15.2 O teste de soma de classificação de Wilcoxon 588
- 15.3 Intervalos de confiança livres de distribuição 594
- 15.4 ANOVA livre de distribuição 598
  - Exercícios suplementares 602
  - Bibliografia 604

## 16 Métodos de controle de qualidade 605

- Introdução 605
- 16.1 Comentários gerais sobre gráficos de controle 606
- 16.2 Gráficos de controle para localização do processo 607
- 16.3 Gráficos de controle de variações de processo 616
- 16.4 Gráficos de controle de atributos 619
- 16.5 Procedimentos de CUSUM 624
- 16.6 Amostragem de aceitação 632
  - Exercícios suplementares 638
  - Bibliografia 639

### Tabelas do apêndice 639

- A.1 Probabilidades binomiais acumuladas 640
  - A.2 Probabilidades de Poisson acumuladas 642
  - A.3 Área sob a curva normal padronizada 643
  - A.4 Função Gama incompleta 644
  - A.5 Valores críticos para as distribuições  $t$  645
  - A.6 Valores críticos de tolerância para populações com distribuições 646
  - A.7 Valores críticos para distribuições Qui-quadrado 647
  - A.8 Áreas de cauda da curva  $t$  648
  - A.9 Valores críticos para distribuições  $F$  650
  - A.10 Valores críticos para distribuições de amplitude Studentizada 656
  - A.11 Áreas de cauda da curva Qui-Quadrada 658
  - A.12 Valores críticos para o teste de Ryan-Joiner de normalidade 659
  - A.13 Valores críticos para o Teste de Wilcoxon com pares de observações 660
  - A.14 Valores críticos do teste de soma de postos de Wilcoxon 661
  - A.15 Valores críticos para intervalo de pares de observações de Wilcoxon 662
  - A.16 Valores críticos para o intervalo de soma de postos de Wilcoxon 663
  - A.17 Curvas  $\beta$  para testes  $t$  664
- Respostas dos exercícios de numeração ímpar 665
- Índice 685

# Prefácio

## Objetivo

A utilização de modelos probabilísticos e de métodos estatísticos para análise de dados se tornou uma prática comum em quase todas as disciplinas científicas. Este livro visa fornecer uma introdução abrangente dos modelos e métodos mais passíveis de serem encontrados e usados por estudantes em cursos de engenharia e ciências naturais. Apesar de os exemplos e exercícios terem sido desenvolvidos para cientistas e engenheiros, a maior parte dos métodos abrangidos são fundamentais para análises estatísticas em várias outras disciplinas, de forma que os estudantes de administração e ciências sociais também se beneficiarão desta leitura.

## Abordagem

Estudantes de cursos de estatística, desenvolvidos como base para outras áreas, podem ser inicialmente céticos em relação ao valor e à relevância da disciplina, mas minha experiência mostra que eles podem se interessar por estatística com a utilização de bons exemplos e exercícios que combinem experiências diárias e interesses científicos. Conseqüentemente, esforcei-me para fornecer exemplos de dados reais em detrimento dos artificiais, dados cuja coleta e análise realmente valem a pena. Muitos dos métodos apresentados, especialmente nos últimos capítulos de inferência estatística, são ilustrados pela análise de dados de fontes publicadas, sendo que muitos exercícios também os envolvem. Algumas vezes, o leitor pode não estar familiarizado com o contexto de um determinado problema (como já aconteceu comigo), mas descobri que os alunos são mais atraídos por problemas reais com um contexto não-familiar do que por problemas claramente artificiais apresentados de forma familiar.

## Nível matemático

A exposição é relativamente modesta em termos de desenvolvimento matemático. Apenas no Capítulo 4 e em partes dos capítulos 5 e 6 é feito uso substancial de cálculo. Em particular, com exceção de uma observação ocasional ou de um aparte, o cálculo aparece apenas na parte de inferência do livro, na segunda seção do Capítulo 6. A álgebra matricial não é utilizada. Dessa forma, quase toda a exposição deve ser acessível àqueles cuja formação matemática inclua um semestre ou dois trimestres de cálculo integral e diferencial.

## Conteúdo

O Capítulo 1 começa com alguns conceitos e terminologias básicos: população, amostra, estatística descritiva, inferência estatística e estudos analíticos *versus* numéricos, e assim por diante, continuando com uma análise de métodos descritivos numéricos e gráficos importantes.

Um desenvolvimento particularmente tradicional de probabilidade é fornecido no Capítulo 2, seguido por distribuições probabilísticas de variáveis aleatórias discretas e contínuas nos capítulos 3 e 4, respectivamente. As distribuições conjuntas e suas propriedades são discutidas na primeira parte do Capítulo 5. A parte final desse capítulo apresenta a estatística e suas distribuições amostrais, que formam a ponte entre a probabilidade e a inferência. Os próximos três capítulos cobrem a estimativa pontual, os intervalos estatísticos e os testes de hipóteses com base em uma única amostra. Os métodos de inferência envolvendo duas amostras independentes e pares de dados são apresentados no Capítulo 9. A análise da variância é tratada nos capítulos 10 e 11 (fator único e vários fatores, respectivamente). A regressão começa a ser estudada no Capítulo 12 (o modelo de regressão linear simples e correlação) e depois retorna extensivamente no Capítulo 13. Os três últimos capítulos versam sobre métodos qui-quadrado, procedimentos de distribuição livre (não paramétrico) e técnicas de controle estatístico de qualidade.

### Ajuda na aprendizagem dos estudantes

Embora o nível matemático do livro ofereça pouca dificuldade aos estudantes de ciências e de engenharia, a compreensão dos conceitos e a valorização da compreensão do desenvolvimento lógico da metodologia podem, algumas vezes, exigir um esforço substancial. Para ajudar os estudantes a obter compreensão e interesse, forneço inúmeros exercícios com diferentes níveis de dificuldade, desde os que envolvem aplicação rotineira do material do texto até aqueles que exigem que o leitor extrapole os conceitos discutidos no livro para novas situações. Há muito mais exercícios do que a maioria dos instrutores normalmente atribui durante um curso específico, mas recomendo que os estudantes sejam estimulados a trabalhar com uma quantidade significativa deles. Em uma disciplina voltada para a resolução de problemas, o envolvimento ativo é a forma mais segura de identificar e solucionar as dúvidas que inevitavelmente surgem. As respostas dos exercícios ímpares são exibidas na seção de respostas no final do livro. Além disso, está disponível um Manual de Soluções do Aluno, que consiste em soluções desenvolvidas de praticamente todos os exercícios ímpares.

### Este livro contém

- Muitos exemplos e exercícios novos, em sua maioria envolvendo dados reais. Alguns deles requerem menos estrutura do que os exercícios característicos de livros-texto. Por exemplo: um exercício do Capítulo 13 fornece um conjunto de dados de regressão múltipla (que, à primeira vista, parece um conjunto de dados de regressão linear simples, exceto pelo fato de também haver informações sobre um previsor categórico); o estudante deve apenas adequar diversos modelos e tirar as próprias conclusões. Um exercício no final do Capítulo 14 inclui partes sobre correlação, ANOVA, uma análise da distribuição de duas amostras e uma análise de qui-quadrado, sem necessariamente mencioná-las. A ênfase normalmente está também na interpretação e não apenas na mecânica. Exemplificando: no Capítulo 1 o estudante deve não somente construir um *boxplot*, como também dizer o que ele sugere. Alguns exercícios apresentam alguns conceitos ou técnicas, como, por exemplo, uma plotagem de metade da curva normal, no Capítulo 4, e a autocorrelação para uma série temporal, no Capítulo 12.
- A definição ampliada de uma hipótese, no início do Capítulo 8, inclui uma afirmação de que a distribuição da população tem forma determinada (em vez de as hipóteses serem apenas solicitações de parâmetros).
- Descrição detalhada de procedimentos da distribuição  $t$  combinada para fornecer melhor equilíbrio entre as análises com e sem combinação.
- São fornecidas mais informações sobre alternativas para fórmulas tradicionais de obtenção de estimativas de intervalos de proporção única e de diferença entre proporções.

- A seção de comparações múltiplas em ANOVA foi completamente reescrita para enfatizar melhor a natureza do nível de confiança simultâneo e o procedimento de Tukey.
- Uma discussão sobre a capacidade de cálculo do MINITAB® em ANOVA de fator único.
- Ênfase na natureza da variação da estimativa do coeficiente angular na regressão linear simples.
- O método LOWESS, para ajustar uma função de regressão não-linear.
- Uma interpretação alternativa dos coeficientes de regressão em regressão múltipla via plotagens de resíduos parciais.

## Cobertura recomendada

Há material suficiente neste livro para um curso com duração de um ano. Se o curso tiver duração inferior, terá de ser seletivo na escolha dos tópicos a serem incluídos. Na Cal Poly, lecionamos em uma seqüência de dois trimestres, com quatro horas de aulas por semana. Durante as dez primeiras semanas, cobrimos a maior parte do material dos capítulos 1-7 (chegando em distribuições conjuntas, nos detalhes de estimativas de máxima verossimilhança e o método dos momentos). O segundo trimestre começa com os testes de hipóteses e segue para inferências com base em duas amostras, ANOVA, regressão e seleções dos capítulos sobre qui-quadrado, de distribuição livre e de controle de qualidade. A cobertura do material em um curso de um semestre obviamente deve ser mais restritiva. É claro que nunca há tempo suficiente para ensinarmos aos estudantes tudo o que desejamos que eles saibam.

## Agradecimentos

Meus colegas na Cal Poly me ofereceram apoio e estímulo inestimáveis ao longo dos anos. Também sou grato aos muitos usuários que fizeram sugestões para aprimoramentos (e apontaram erros ocasionais). Uma nota de agradecimento a Julie Seely, por seu trabalho no Manual de Soluções do Aluno.

Agradeço ao importante retorno dado pelos seguintes revisores deste livro: Robert L. Armacost, University of Central Florida; Douglas M. Bates, University of Wisconsin–Madison; David Clark, California State Polytechnic University em Pomona; David M. Cresap, University of Portland; Don E. Deal, University of Houston; Charles E. Donaghey, University of Houston; Patrick J. Driscoll, U. S. Military Academy; Mark Duva, University of Virginia; Nasser S. Fard, Northeastern University; Celso Grebogi, University of Maryland; James J. Halavin, Rochester Institute of Technology; Jennifer Hoeting, Colorado State University; Wei-Min Huang, Lehigh University; Stephen Kokoska, Colgate University; Arnold R. Miller, University of Denver; John J. Millson, University of Maryland; Don Ridgeway, North Carolina State University; Larry J. Ringer, Texas A&M University; Richard M. Soland, The George Washington University; Clifford Spiegelman, Texas A&M University; Jerry Stedinger, Cornell University; David Steinberg, Tel Aviv University; G. Geoffrey Vining, University of Florida; Bhutan Wadhwa, Cleveland State University; e Michael G. Zabetakis, University of Pittsburgh.

Agradeço muito aos serviços de produção fornecidos por Martha Emry (que sempre foi prestativa, atenciosa e providencial em suas conversas comigo). Obrigado também a Pamela Rockwell, editora de cópias. A equipe Duxbury and Brooks/Cole foi, como sempre, de grande valia, fornecendo grande apoio. Obrigado em especial a Carolyn Crockett, Curt Hinrichs, Ann Day, Julie Bliss, Joseph Rogove e Karen Haga. Também desejo elogiar os representantes de vendas da Thomson Learning por seu trabalho árduo, ao ouvirem opiniões sobre esta obra, bem como de outros livros escritos por mim.

Finalmente, as palavras não podem expressar adequadamente minha gratidão a minha esposa Carol, por seu apoio aos meus esforços literários no curso das três últimas décadas.



# Visão Geral e Estatística Descritiva

## Introdução

Os conceitos e métodos estatísticos não são apenas úteis, como também indispensáveis na compreensão do mundo ao nosso redor. Eles fornecem meios de obtenção de novas percepções no que diz respeito ao comportamento de diversos fenômenos que você encontrará em seu campo de especialização em engenharia ou ciência.

A disciplina estatística nos ensina a fazer julgamentos inteligentes e a tomar decisões na presença de incertezas e variações. Sem incertezas ou variações, haveria pouca necessidade de estatísticos ou métodos estatísticos. Se cada componente de um determinado tipo tivesse exatamente o mesmo tempo de vida, se todos os resistores produzidos por um determinado fabricante tivessem o mesmo valor de resistência, se as determinações de pH de espécimes de solo de um local determinado fornecessem resultados idênticos, e assim por diante, então uma única observação revelaria todas as informações desejadas.

Uma manifestação interessante das variações surge ao longo dos testes de desempenho de emissões em motores automotivos. Os requisitos de custos e tempo do FTP (*Federal Test Procedure*) impedem seu uso generalizado em programas de inspeção veiculares. Como resultado, muitas agências desenvolveram testes mais rápidos e baratos, que, espera-se, reproduzam os resultados do FTP. De acordo com o artigo de jornal "Motor Vehicle Emissions Variability" (*J. of the Air and Waste Mgmt. Assoc.*, 1996: 667-675), a aceitação do FTP como um padrão universal leva à crença de que medidas repetidas no mesmo veículo terão resultados idênticos (ou quase idênticos). Os autores do artigo aplicaram o FTP a sete veículos caracterizados como "altos emissores". Seguem os resultados de um dos veículos:

HC (gm/milha)	13,8	18,3	32,2	32,5
CO (gm/milha)	118	149	232	236



A variação substancial nas medidas de HC e CO apresenta dúvidas consideráveis sobre a sabedoria popular e dificulta a elaboração de avaliações precisas sobre níveis de emissão.

Como as técnicas estatísticas podem ser usadas para obter informações e tirar conclusões? Suponha, por exemplo, que um engenheiro de materiais tenha desenvolvido um revestimento para retardar a corrosão em tubulações de metal sob circunstâncias especificadas. Se esse revestimento for aplicado a diferentes segmentos do tubo, variações nas condições ambientais e nos próprios segmentos resultarão em uma corrosão maior em alguns segmentos do que em outros. Os métodos de análise estatística podem ser usados nos dados de um experimento como esse para decidir se a quantidade média de corrosão excede um limite superior especificado de algum tipo ou para prever a quantidade de corrosão que ocorrerá num único tubo.

Como alternativa, suponha que o engenheiro tenha desenvolvido tal revestimento acreditando que será superior àquele usado no momento. Um experimento comparativo pode ser efetuado para investigar essa questão, aplicando-se o revestimento atual a alguns segmentos do tubo e o novo a outros. Isso deve ser feito com cuidado, para que não surja uma conclusão errada. Por exemplo: talvez a quantidade média de corrosão seja idêntica para os dois revestimentos. Entretanto, o novo revestimento pode ter sido aplicado a segmentos que possuem uma capacidade superior de resistência à corrosão e sob condições ambientais menos severas, se comparados aos segmentos e condições do revestimento atual. O investigador provavelmente observaria então uma diferença causada não pelos próprios revestimentos, mas por variações externas. A estatística oferece métodos não somente para análise dos resultados de experimentos depois que foram executados, como também sugestões de como os experimentos devem ser executados de forma eficiente para diminuir os efeitos das variações e ter melhores chances de produzir conclusões corretas.

---

## 1.1 | Populações, amostras e processos

Os engenheiros e cientistas estão constantemente expostos a conjuntos de fatos ou **dados**, tanto em suas carreiras como em suas atividades diárias. A disciplina estatística fornece métodos para organizar e resumir os dados para tirar conclusões com base em informações contidas nos dados.

Uma investigação normalmente enfocará uma coleção bem definida de objetos que constituem uma **população** de interesse. Em um estudo, a população pode consistir em todas as cápsulas de gelatina de um determinado tipo produzidas durante um período especificado. Outra investigação pode envolver a população que consiste em todos os indivíduos que receberam um diploma de engenharia durante o ano acadêmico mais recente. Quando as informações desejadas estiverem disponíveis para todos os objetos da população, temos o que é denominado **censo**. Restrições de tempo, dinheiro e outros recursos escassos normalmente tornam um censo impraticável ou inviável. Em vez disso, um subconjunto da população – uma **amostra** – é selecionado de uma forma prescrita. Dessa maneira, podemos obter uma amostra de mancais de uma determinada produção como base de investigação da conformidade dos mancais com as especificações do fabricante; ou podemos selecionar uma amostra dos formandos em engenharia do ano anterior para obter um retorno sobre a qualidade dos currículos.

Normalmente, estamos interessados apenas em certas características dos objetos de uma população: o número de falhas na superfície de cada invólucro, a espessura de cada parede da cápsula, o sexo de um formando em engenharia, a idade com que um indivíduo se formou etc. Uma característica pode ser categorizada, como sexo ou tipo de defeito, ou pode ter natureza numérica. No primeiro caso, o valor da característica é uma categoria (por exemplo, feminino ou solda insuficiente), enquanto, no último caso, o valor é um número (por exemplo, idade = 23 anos ou diâmetro = 0,502 cm). Uma **variável** é qualquer característica cujo valor pode mudar de um objeto para outro na população. Inicialmente, devemos identificar as variáveis com letras minúsculas do final do nosso alfabeto. Os exemplos incluem:

- $x$  = marca da calculadora de um estudante
- $y$  = número de defeitos graves em um automóvel recentemente fabricado
- $z$  = distância de frenagem de um automóvel sob condições especificadas

Os dados resultam da observação de uma variável ou de duas ou mais variáveis simultaneamente. Um conjunto de dados **univariado** consiste em observações sobre uma única variável. Por exemplo: podemos determinar o tipo de transmissão, automática (A) ou manual (M), de cada um dentre 10 automóveis recentemente comprados em um determinado revendedor, resultando em um conjunto de dados categorizados.

M A A A M A A M A A

A amostra a seguir de vida útil (horas) de baterias da marca D colocadas em um determinado uso é um conjunto numérico de dados univariados:

5,6 5,1 6,2 6,0 5,8 6,5 5,8 5,5

Temos dados **bivariados** quando as observações são feitas em cada uma de duas variáveis. Nosso conjunto de dados pode consistir em um par (altura, peso) de cada jogador de basquete de um time, com a primeira observação como (72, 168), a segunda como (75, 212) e assim por diante. Se um engenheiro determinar o valor de  $x$  = vida útil do componente e  $y$  = motivo de falha do componente, o conjunto de dados resultante será bivariado com uma variável numérica e outra categorizada. Dados **multivariados** surgem quando são feitas observações sobre mais de duas variáveis. Por exemplo: um médico pesquisador pode determinar a pressão sanguínea sistólica, a pressão sanguínea diastólica e o nível de colesterol de cada paciente participante de um estudo. Cada observação seria um trio de números, como (120, 80, 146). Em muitos conjuntos de dados multivariados, algumas variáveis são numéricas e outras são categorizadas. Dessa forma, a edição automotiva anual de *Consumer Reports* fornece valores de tais variáveis como tipo do veículo (pequeno, esportivo, compacto, médio, grande), consumo de combustível na cidade (milhas/galão), consumo de combustível na estrada (milhas/galão), tipo de tração (traseira, dianteira, nas quatro rodas) e assim por diante.

## Ramos da estatística

Um investigador que tenha coletado dados pode simplesmente desejar resumir e descrever suas características importantes. Isso exige a utilização de métodos de **estatística descritiva**. Alguns desses métodos são gráficos por natureza: a construção de histogramas, *boxplots* e gráficos de dispersão são os exemplos principais. Outros métodos descritivos envolvem o cálculo de medidas numéricas, como médias, desvios padrão e coeficientes de correlação. A ampla disponibilidade de pacotes de softwares estatísticos facilitou bastante essas tarefas. Os computadores são muito mais eficientes que os seres humanos em cálculo e na criação de imagens (depois de terem recebido as instruções corretas do usuário). Isso significa que o investigador não tem de perder muito tempo com “trabalho braçal” e terá mais tempo para estudar os dados e extrair mensagens importantes. Neste livro serão apresentados resultados de vários pacotes, como MINITAB, SAS, e S-Plus.

### Exemplo 1.1

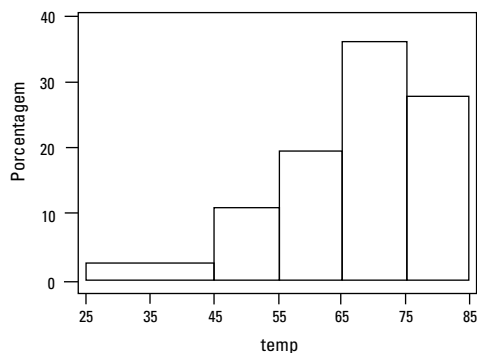
A tragédia que ocorreu com o ônibus espacial *Challenger* e seus astronautas, em 1986, levou a diversos estudos para investigar os motivos da falha da missão. A atenção rapidamente se voltou ao comportamento dos anéis de

vedação do motor do foguete. Aqui estão os dados resultantes de observações de  $x =$  temperatura do anel de vedação ( $^{\circ}\text{F}$ ) de cada teste de acionamento ou lançamento real do motor do foguete da nave (*Presidential Commission on the Space Shuttle Challenger Accident*, Vol. 1, 1986: 129-131).

84 49 61 40 83 67 45 66 70 69 80 58  
 68 60 67 72 73 70 57 63 70 78 52 67  
 53 67 75 61 70 81 76 79 75 76 58 31

Sem nenhuma organização, é difícil ter noção do que pode ser uma temperatura normal ou representativa, se os valores estão altamente concentrados em torno de um ponto ou se estão dispersos, se há lacunas nos dados, que porcentagem dos dados estão na faixa dos 60 e assim por diante. A Figura 1.1 mostra o que é chamado de *diagrama caule e folha* dos dados, assim como um *histograma*. Mais adiante serão discutidas a construção e a interpretação desses resumos ilustrativos; no momento, desejamos que você observe como eles começam a nos mostrar como os valores de temperatura são distribuídos ao longo da escala de medida. Alguns desses lançamentos/acionamentos tiveram sucesso e outros resultaram em falha. No Capítulo 13, consideraremos se a temperatura teve influência na probabilidade de um lançamento bem-sucedido.

```
Caule folha da temp N = 36
Unidade da folha = 1,0
1 3 1
1 3
2 4 0
4 4 59
6 5 23
9 5 788
13 6 0113
(7) 6 6777789
16 7 000023
10 7 556689
4 8 0134
```



**Figura 1.1** Um diagrama caule e folha MINITAB e histograma dos dados da temperatura do anel de vedação

Com uma amostra da população, um investigador freqüentemente usaria as informações da amostra para tirar algum tipo de conclusão (fazer uma inferência de algum tipo) sobre a população, ou seja, a amostra é um meio para chegar a um fim e não o fim em si. As técnicas de generalização de uma amostra para uma população são agrupadas no ramo da nossa disciplina denominado **inferência estatística**.

### Exemplo 1.2

As investigações de resistência de materiais fornecem um campo fértil para a aplicação de métodos estatísticos. O artigo “Effects of Aggregates and Microfillers on the Flexural Properties of Concrete” (*Magazine of Concrete Research*, 1997, p. 81-98) relatou um estudo de propriedades de resistência de concreto de alto desempenho obtidas pela utilização de superplásticos e determinados adesivos. A resistência à compressão desse concreto foi

investigada anteriormente, mas não se sabe muito sobre a resistência à flexão (uma medida da capacidade de resistência a falhas decorrentes de flexão). Os dados a seguir sobre resistência à flexão (em megapascal, MPa, onde  $1 \text{ Pa (pascal)} = 1,45 \times 10^{-4} \text{ psi}$ ) foram exibidos no artigo citado:

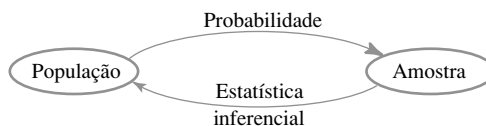
5,9	7,2	7,3	6,3	8,1	6,8	7,0	7,6	6,8	6,5	7,0	6,3	7,9	9,0
8,2	8,7	7,8	9,7	7,4	7,7	9,7	7,8	7,7	11,6	11,3	11,8	10,7	

Suponha que busquemos uma *estimativa* do valor médio da resistência à flexão de todas as vigas que podem ser feitas dessa forma (se considerarmos a população de todas as vigas, estaremos tentando estimar a média da população). Pode-se mostrar que, com alto nível de confiança, a resistência média da população está entre 7,48 MPa e 8,80 MPa. Isso é denominado *intervalo de confiança* ou *estimativa por intervalo*. De forma alternativa, esses dados podem ser usados para prever a resistência à flexão de uma *única* viga desse tipo. Com alto nível de confiança, a resistência de uma determinada viga excederá 7,35 MPa. O número 7,35 é denominado *limite inferior de previsão*. ■

Este livro enfoca principalmente os métodos de apresentação e ilustração de inferência estatística úteis ao trabalho científico. Os tipos mais importantes de procedimentos inferenciais (estimativa por pontos, teste de hipóteses e estimativa por intervalos de confiança) são apresentados nos capítulos 6-8 e usados em aplicações mais avançadas nos capítulos 9-16. O restante deste capítulo apresenta os métodos de estatística descritiva mais usados no desenvolvimento da inferência.

Os capítulos 2-5 apresentam material da disciplina de probabilidade. Esse material, em suma, faz uma ponte entre as técnicas descritiva e inferencial. A proficiência em probabilidade leva à melhor compreensão de como os procedimentos inferenciais são desenvolvidos e usados, como as conclusões estatísticas podem ser traduzidas para a linguagem do dia-a-dia e interpretadas, e quando e onde podem ocorrer ciladas na aplicação dos métodos. A probabilidade e a estatística lidam com questões que envolvem populações e amostras, mas o fazem de “maneira inversa” uma em relação a outra.

Em um problema de probabilidade, as propriedades da população, que são objeto de estudo, são assumidas como conhecidas (por exemplo: em uma população numérica, uma distribuição especificada dos valores da população pode ser assumida) e as questões relativas a uma amostra proveniente da população são propostas e respondidas. Em um problema de estatística, as características de uma amostra estão disponíveis ao investigador e essas informações permitem que ele tire conclusões sobre a população. A relação entre as duas disciplinas pode ser resumida da seguinte forma: a probabilidade faz suas considerações da população para a amostra (raciocínio dedutivo) e a inferência estatística faz considerações da amostra para a população (raciocínio indutivo). Isso é ilustrado na Figura 1.2.



**Figura 1.2** A relação entre probabilidade e inferência estatística

Antes de podermos entender o que uma determinada amostra pode nos dizer sobre a população, devemos entender a incerteza associada à tomada da amostra de uma dada população. É por isso que estudamos probabilidade antes de estatística.

Como exemplo do contraste entre os focos da probabilidade e da inferência estatística, considere o uso de cintos de segurança manuais de dois pontos em carros equipados com cintos automáticos de três pontos. (O artigo “Automobile Seat Belts: Usage Patterns in Automatic Belt Systems,” *Human Factors*, 1998, p. 126-135, resume os dados de utilização.) Em probabilidade, podemos assumir que 50% de todos os motoristas de carros equipados dessa forma em uma determinada área metropolitana usam regularmente o cinto de dois pontos (uma hipótese sobre a população), de forma que perguntamos: “Qual a probabilidade de que uma amostra de 100 motoristas inclua ao menos 70 que usam regularmente o cinto de dois pontos?” ou “Em uma amostra de tamanho

100, quantos motoristas podemos esperar que usem o cinto de dois pontos?” Por outro lado, em inferência estatística temos as informações da amostra disponíveis. Por exemplo: uma amostra de 100 motoristas de tais carros revelou que 65 usam o cinto de dois pontos regularmente. Podemos perguntar então: “Isso fornece evidência suficiente para a conclusão de que mais de 50% de todos os motoristas nessa área usam regularmente o cinto de dois pontos?” Nesse último cenário, tentamos usar as informações da amostra para responder a uma pergunta sobre a estrutura de toda a população a partir da qual a amostra foi selecionada.

No exemplo do cinto de dois pontos, a população está bem definida e concreta: todos os motoristas de carros equipados de uma forma em uma determinada área metropolitana. No Exemplo 1.1, entretanto, está disponível uma amostra de temperaturas de anéis de vedação de uma população que não existe realmente. Em vez disso, é conveniente pensarmos na população como consistindo de todas as medidas possíveis de temperatura que podem ser feitas em condições experimentais similares. Tal população é denominada **população conceitual** ou **hipotética**. Há diversas situações de problemas em que se encaixam questões na estrutura de inferência estatística pela conceitualização de uma população.

## Estudos enumerativos versus analíticos

W. E. Deming, um influente estatístico norte-americano, força motriz na revolução de qualidade do Japão nos anos 50 e 60, apresentou a distinção entre *estudos enumerativos* e *estudos analíticos*. No primeiro, o interesse se foca em uma coleção finita, identificável e imutável de indivíduos ou objetos que formam uma população. Uma *estrutura de amostragem*, isto é, uma listagem de todos os indivíduos ou objetos a fazerem parte da amostra, está disponível a um investigador ou pode ser construída. Por exemplo: a estrutura pode ser constituída por todas as assinaturas em uma petição para qualificação de uma determinada iniciativa de voto secreto em uma eleição iminente; uma amostra normalmente é selecionada para apurar se o número de assinaturas *válidas* excede um valor especificado. Como outro exemplo, a estrutura pode conter números de série de todos os fornos fabricados por uma determinada empresa durante certo período de tempo; uma amostra pode ser selecionada para inferir algo sobre a vida útil média dessas unidades. A utilização de métodos inferenciais a ser desenvolvida neste livro é razoavelmente não-controvertida nesse cenário (apesar de os estatísticos ainda poderem discutir sobre que métodos em particular devem ser usados).

Um estudo analítico é definido, de modo geral, como aquele que não é de natureza enumerativa. Esses estudos são normalmente executados com o objetivo de melhorar um produto futuro por meio da ação em um processo de algum tipo (por exemplo: recalibragem de equipamentos ou ajuste do nível de algum insumo, como a quantidade de um catalisador). Frequentemente, os dados podem ser obtidos apenas em um processo existente, que pode diferir do processo futuro em aspectos importantes. Dessa forma, não há uma estrutura de amostragem que relacione os indivíduos ou objetos de interesse. Por exemplo: uma amostra de cinco turbinas com um novo *design* pode ser fabricada e testada em caráter experimental para investigar sua eficiência. Essas cinco turbinas podem ser vistas como uma amostra da população conceitual de todos os protótipos que podem ser fabricados em condições similares, mas *não* necessariamente como representantes da população de unidades fabricadas depois que a produção for iniciada. Os métodos de utilização de informações de amostras para obtenção de conclusões sobre a produção futura podem ser problemáticos. Alguém com experiência na área de projetos e engenharia de turbinas (ou de qualquer outra área relevante para a disciplina) deve ser chamado para julgar se essa extrapolação é sensata. Uma boa exposição dessas questões está no artigo “Assumptions for Statistical Inference” de Gerald Hahn e William Meeker (*The American Statistician*, 1993, p. 1-11).

## Coletando dados

A estatística lida não somente com a organização e análise de dados depois de sua coleta, como também com o desenvolvimento de técnicas de coleta. Se os dados não são coletados de forma correta, um investigador pode não ter condições de responder às questões em consideração com um nível de confiança razoável. Um problema comum é que a população-alvo — a respeito da qual serão tiradas conclusões — pode ser diferente da população da qual se obteve a amostra. Por exemplo: publicitários podem desejar diversos tipos de informações sobre os hábitos televisivos de clientes potenciais. As informações mais sistemáticas desse tipo são provenientes de dispositivos de monitoramento locais em um pequeno número de lares nos Estados Unidos. Já se presumiu que a

colocação em si desses dispositivos afeta o comportamento dos telespectadores, de forma que as características da amostra podem ser diferentes daquelas da população-alvo.

Quando a coleta de dados exige a seleção de indivíduos ou objetos a partir de uma estrutura, o método mais simples de assegurar uma seleção representativa é tomar uma *amostra aleatória simples*. Trata-se de uma amostra em que qualquer subconjunto de tamanho especificado (como uma amostra de tamanho 100) tem a mesma chance de ser selecionado. Por exemplo: se a estrutura consistir em 1.000.000 de números de série, os números 1, 2, ... até 1.000.000 podem ser colocados em tiras idênticas de papel. Após colocá-las em uma caixa e mexer bastante, elas podem ser retiradas uma a uma, até que a amostra de tamanho requerido seja obtida. De forma alternativa (e normalmente preferida), uma tabela de números aleatórios ou um gerador de números aleatórios pode ser usado.

Algumas vezes, os métodos de amostragem alternativos podem ser usados para facilitar o processo de seleção, para obter informações extras ou para aumentar o nível de confiança das conclusões. Um desses métodos, a *amostragem estratificada*, exige a separação das unidades da população em grupos não-passíveis de sobreposição e a tomada de uma amostra de cada um. Por exemplo: um fabricante de videocassetes pode desejar informações sobre a satisfação dos clientes com as unidades produzidas no ano anterior. Se foram fabricados e vendidos três modelos diferentes, pode ser selecionada uma amostra de cada um dos três modelos correspondentes, o que resultaria em informações sobre todos os modelos e asseguraria que nenhum deles teve mais ou menos representatividade na amostra inteira.

Freqüentemente, uma amostra de “conveniência” é obtida pela seleção de indivíduos ou objetos sem aleatoriedade sistemática. Como exemplo, um grupo de tijolos pode ser empilhado de forma que seja extremamente difícil selecionar as peças centrais. Se os tijolos do topo e das laterais forem de alguma forma diferentes dos outros, os dados resultantes da amostra não serão representativos da população. Um investigador assumirá, com freqüência, que essa amostra de conveniência se aproxima de uma amostra aleatória. Nesse caso, o repertório de métodos inferenciais do estatístico pode ser usado, o que é, entretanto, de julgamento do profissional. A maioria dos métodos discutidos aqui em diante são baseadas em variações da amostragem aleatória simples, descritas no Capítulo 5.

Os engenheiros e cientistas freqüentemente coletam dados executando algum tipo de experimento, o que pode envolver a decisão de como alocar diferentes tratamentos (como fertilizantes ou revestimentos para proteção de corrosão) às diversas unidades experimentais (lotes de terra ou segmentos de tubo). Além disso, um investigador pode variar sistematicamente os níveis ou categorias de determinados fatores (como pressão ou tipo de material isolante) e observar o efeito em alguma variável de resposta (como o resultado de um processo de produção).

### Exemplo 1.3

Um artigo no *New York Times* (27 de janeiro de 1987) reportou que o risco de ataques cardíacos pode ser reduzido pela ingestão de aspirina. Essa conclusão foi baseada em um experimento planejado que envolveu um grupo de controle de indivíduos que tomaram um placebo com aparência de aspirina, mas reconhecidamente inócua, e outro que tomou aspirina de acordo com um regime especificado. Os comprimidos foram atribuídos aleatoriamente aos grupos para evitar desvios, tornando possível a utilização de métodos probabilísticos para análise dos dados. Dos 11.034 componentes do grupo de controle, 189 tiveram ataques cardíacos posteriormente, enquanto apenas 104 dos 11.037 indivíduos do grupo da aspirina tiveram o problema. A taxa de incidência de ataques cardíacos no grupo de tratamento foi cerca de metade da taxa do grupo de controle. Uma explicação possível para esse resultado é a variação devida ao acaso: que a aspirina não tem o efeito desejado e que a diferença observada é uma variação normal, da mesma forma que jogar duas moedas idênticas normalmente produziria números diferentes de caras. Entretanto, nesse caso, os métodos inferenciais sugerem que a variação devida ao acaso em si não pode explicar adequadamente a magnitude da diferença observada. ■

### Exemplo 1.4

Um engenheiro deseja investigar os efeitos de um tipo de adesivo e de um material condutor na resistência do contato ao montar um IC (Circuito Integrado) em uma determinada base. Dois tipos de adesivo e dois materiais condutores estão sendo considerados. Duas observações são feitas para cada combinação – tipo de adesivo/material condutor – resultando nos dados a seguir:

Tipo de adesivo	Material condutor	Resistência observada do contato	Média
1	1	82,77	79,5
1	2	75,87	81,0
2	1	84,80	82,0
2	2	78,90	84,0

As resistências de contato médias são ilustradas na Figura 1.3. Parece que o tipo de adesivo 2 melhora a resistência do contato se comparado ao tipo 1 mais ou menos no mesmo valor, não importando o material condutor usado, com a combinação 2, 2 sendo a melhor. Os métodos inferenciais podem novamente ser usados para julgar se esses efeitos são reais ou simplesmente consequência da variância devido ao acaso.

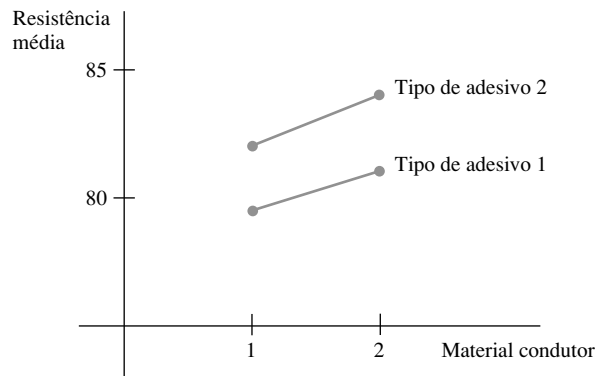


Figura 1.3 Resistências de contato médias no Exemplo 1.4

Suponha, adicionalmente, que haja dois períodos de cura em consideração e também dois tipos de circuito integrado após o revestimento. Há, portanto,  $2 \cdot 2 \cdot 2 \cdot 2 = 16$  combinações desses quatro fatores e nosso engenheiro não possui recursos suficientes nem mesmo para fazer uma única observação para cada uma dessas combinações. No Capítulo 11 veremos como a seleção cuidadosa de uma fração dessas possibilidades normalmente fornecerá as informações desejadas. ■

## Exercícios | Seção 1.1 (1–9)

- Forneça uma amostra possível, de tamanho 4, de cada uma das populações a seguir:
  - Todos os jornais diários publicados nos Estados Unidos.
  - Todas as empresas listadas na New York Stock Exchange.
  - Todos os alunos de sua universidade ou faculdade.
  - Todas as médias, em pontos, dos alunos de sua universidade ou faculdade.
- Para cada uma das populações hipotéticas a seguir, forneça uma amostra plausível de tamanho 4:
  - Todas as distâncias que podem resultar quando uma bola de futebol é arremessada.
  - O tamanho das páginas dos livros publicados nos últimos cinco anos.
  - Todas as medidas de intensidade de terremotos (escala Richter) que podem ser registradas na Califórnia durante o próximo ano.
  - Todos os possíveis produtos (em gramas) de uma determinada reação química feita em um laboratório.
- Considere a população que consiste em todos os videocassetes de uma determinada marca e modelo, e enfoque se um videocassete precisa de manutenção durante o tempo de garantia.
  - Formule diversas questões sobre probabilidade baseadas em uma amostra de 100 desses videocassetes.
  - Qual questão sobre inferência estatística pode ser respondida ao determinar o número de videocassetes que precisam de serviço de garantia em uma amostra de tamanho 100?
- Dê três exemplos diferentes de populações concretas e três exemplos diferentes de populações hipotéticas.
  - Para cada uma de suas populações hipotéticas e concretas, dê um exemplo de uma questão sobre probabilidade e um exemplo de uma questão sobre inferência estatística.
- Diversas universidades e faculdades instituíram programas de Instrução Suplementar (SI), em que um monitor se encontra regularmente com um grupo de estudantes matriculados em um curso para promover discussões

sobre o material desse curso e melhorar o domínio da disciplina. Suponha que os estudantes de um grande curso de estatística (o que mais poderia ser?) são aleatoriamente divididos em um grupo de controle que não participará do SI e um grupo de tratamento que o fará. No final do período, é determinada a pontuação total de cada estudante no curso.

- a. As pontuações do grupo de SI são uma amostra da população existente? Caso seja, qual é? Caso contrário, qual é a população conceitual relevante?
  - b. Qual você acha que é a vantagem de dividir aleatoriamente os estudantes em dois grupos em vez de deixar cada estudante escolher o grupo do qual participará?
  - c. Por que os investigadores não colocaram todos os estudantes no grupo de tratamento? *Nota:* O artigo “Supplemental Instruction: An Effective Component of Student Affairs Programming” (*J. of College Student Devel.*, 1997, p. 577-586) discute a análise de dados de diversos programas de SI.
6. O sistema da CSU (California State University) consiste em 23 *campi*, de San Diego State, no sul, até Humboldt State, perto da fronteira com Oregon. Um administrador da CSU deseja fazer uma inferência sobre a distância média entre as cidades natais de seus alunos e seus *campi*. Descreva e discuta diversos métodos de amostragem que podem ser empregados. Esse estudo seria enumerativo ou analítico? Explique seu raciocínio.
7. Certa cidade é dividida naturalmente em 10 bairros. Como um avaliador imobiliário deve selecionar uma amostra de casas de uma única família que pode ser usada como base para o desenvolvimento de uma equação para previsão do valor avaliado a partir de características como idade, tamanho, número de banheiros, distância até a escola mais próxima e assim por diante? Esse estudo é enumerativo ou analítico?
8. A quantidade de fluxo que passa através de uma válvula solenóide em um sistema de controle de poluição de um automóvel é uma característica importante. Foi executado um experimento para estudar como a taxa do fluxo depende de três fatores: comprimento do núcleo, carga da mola e largura da bobina. Foram escolhidos dois níveis diferentes (alto e baixo) de cada fator e foi feita uma única observação sobre o fluxo para cada combinação de níveis.
- a. O conjunto de dados resultante consistiu em quantas observações?
  - b. Esse estudo é enumerativo ou analítico? Explique seu raciocínio.
9. Em um famoso experimento executado em 1882, Michelson e Newcomb fizeram 66 observações do tempo levado pela luz para percorrer a distância entre dois locais em Washington, D.C. Algumas das medidas (codificadas de certa forma) foram 31, 23, 32, 36, -2, 26, 27 e 31.
- a. Por que essas medidas não são idênticas?
  - b. Esse estudo é enumerativo? Por quê?

---

## 1.2 Métodos tabular e gráfico em estatística descritiva

A estatística descritiva pode ser dividida em duas áreas gerais. Nesta seção, discutiremos a primeira dessas áreas, que representa um conjunto de dados usando técnicas visuais. Nas seções 1.3 e 1.4, desenvolveremos algumas medidas numéricas simples para conjuntos de dados. Muitas técnicas visuais podem ser conhecidas: tabelas de frequência, folhas de contagem, histogramas, gráficos de pizza, gráficos de barras, digramas de dispersão e afins. Aqui enfocamos algumas dessas técnicas, que são mais úteis e relevantes para probabilidade e inferência estatística.

### Notações

Algumas notações gerais facilitarão a aplicação de nossos métodos e fórmulas a uma ampla gama de problemas práticos. O número de observações em uma única amostra, isto é, o *tamanho da amostra*, normalmente será representado por  $n$ , de forma que  $n = 4$  para a amostra de universidades {Stanford, Iowa State, Wyoming, Rochester} e também para a amostra de medidas de pH {6,3; 6,2; 5,9; 6,5}. Se duas amostras estiverem sendo consideradas simultaneamente,  $m$  e  $n$  ou  $n_1$  e  $n_2$  podem ser usados para representar os números de observações. Portanto, se {29,7; 31,6; 30,9} e {28,7; 29,5; 29,4; 30,3} forem medidas de eficiência térmica para dois tipos de motor diesel, então  $m = 3$  e  $n = 4$ .

Dado um conjunto de dados que consiste de  $n$  observações de uma variável  $x$ , as observações individuais serão representadas por  $x_1, x_2, x_3, \dots, x_n$ . O índice não tem nenhuma relação com a magnitude de uma



determinada observação. Dessa forma,  $x_1$  em geral não será a menor observação do conjunto e  $x_n$  normalmente não será a maior. Em diversas aplicações,  $x_1$  será a primeira observação coletada pelo investigador,  $x_2$  será a segunda e assim por diante. A  $i$ -ésima observação do conjunto de dados será representada por  $x_i$ .

### Diagramas de caule e folha

Considere um conjunto de dados numéricos  $x_1, x_2, \dots, x_n$ , para o qual cada  $x_i$  consiste de, pelo menos, dois dígitos. Uma forma rápida de obter uma representação visual informativa do conjunto de dados é construir um *diagrama de caule e folha*.

**Etapas de construção de um diagrama de caule e folha**

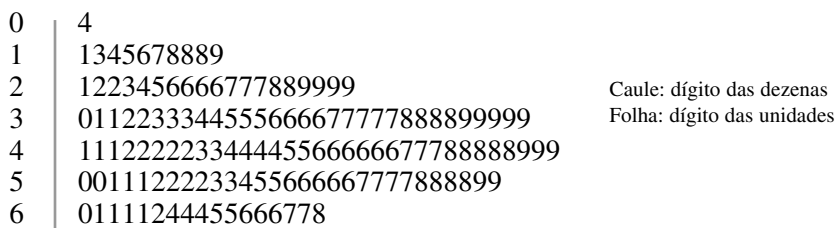
1. Selecione um ou mais dígitos de liderança para serem o caule. Os dígitos à direita desse dígito de liderança serão as folhas.
2. Relacione os valores de caule possíveis em uma coluna vertical.
3. Registre a folha de cada observação ao lado do caule correspondente.
4. Indique as unidades dos caules e das folhas em algum lugar do diagrama.

Se o conjunto de dados consistir em notas de provas, cada uma entre 0 e 100, a pontuação de 83 terá caule 8 e folha 3. Para um conjunto de dados de consumo de combustível de automóveis (milhas/galão), todos entre 8,1 e 47,8, podemos usar as dezenas como caule, de forma que 32,6 teria uma folha de 2,6. Em geral, recomenda-se que o diagrama tenha entre 5 e 20 caules.

#### Exemplo 1.5

O consumo de álcool por alunos de faculdades causa grande preocupação, não apenas para os membros da comunidade acadêmica, como também pelas conseqüências potenciais à saúde e à segurança da sociedade em geral. O artigo “Health and Behavioral Consequences of Binge Drinking in College” (*J. of the Amer Med. Assoc.*, 1994, p. 1672-1677) relatou em um abrangente estudo de consumo excessivo de álcool em diversos *campi* nos Estados Unidos. Um episódio de bebedeira foi definido como cinco ou mais bebidas em seqüência para os homens e quatro ou mais para as mulheres. A Figura 1.4 mostra um diagrama de caule e folha de 140 valores de  $x =$  ao percentual de estudantes de nível superior que se embriagam. (Esses valores não foram fornecidos no artigo citado, mas nosso diagrama apresenta-se de acordo com a ilustração exibida dos dados.)

A primeira folha na linha do caule 2 é 1, o que nos informa que 21% dos alunos de uma das faculdades da amostra se embriagavam. Sem a identificação dos dígitos do caule e das folhas no diagrama, não saberíamos se a observação do caule 2, folha 1, devia ser lida como 21%, 2,1% ou 0,21%.



**Figura 1.4** Diagrama de caule e folha de percentual de alunos que se embriagam em cada uma de 140 faculdades

Ao criar um diagrama a mão, a organização das folhas da menor para a maior em cada linha pode tomar muito tempo e essa organização normalmente contribui pouco ou quase nada para informações extras. Suponha que as observações tenham sido relacionadas em ordem alfabética por nome de escola, da seguinte forma:

16% 33% 64% 37% 31% ...

Colocar esses valores no diagrama nessa ordem resultaria no caule 1 tendo 6 como sua primeira folha e o começo da linha do caule 3 seria

3 | 371 . . .

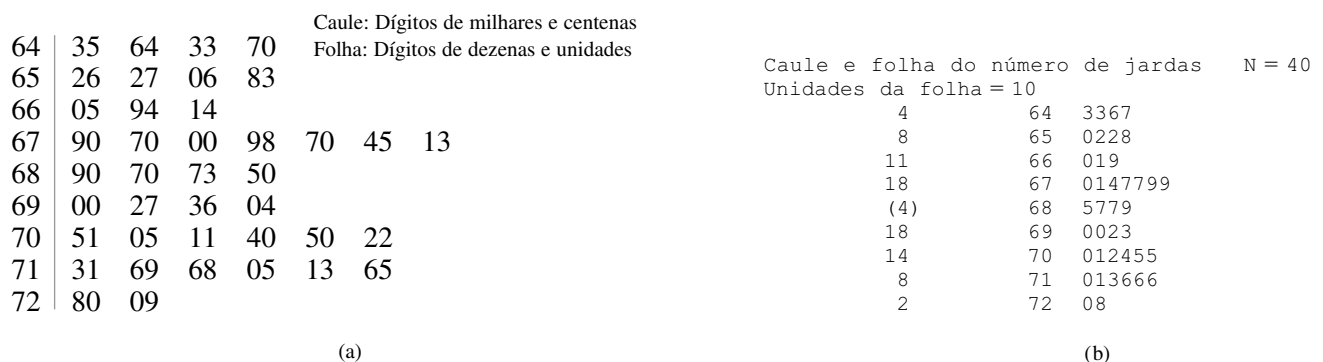
O diagrama sugere que um valor típico ou representativo está na linha do caule 4, talvez na metade da faixa de 40%. As observações não estão altamente concentradas em torno desse valor típico, como estariam se todos os valores estivessem entre 20% e 49%. O diagrama cresce para um único pico, quando nos movemos para baixo e então declina: não há lacunas no diagrama. Seu formato não é perfeitamente simétrico, parece aumentar mais na direção das folhas inferiores, do que na direção das folhas de cima. Por último, não há observações distantes da parte principal dos dados (sem *outliers*, ou seja, pontos fora da curva), como aconteceria se um dos valores fosse 86% em vez de 26%. A característica mais surpreendente dos dados é que, na maior parte das faculdades, pelo menos um quarto dos alunos se embriaga. O problema do consumo excessivo de álcool nos *campi* é muito mais difundido do que muitos haviam suspeitado. ■

Um diagrama de caule e folha transmite informações sobre os seguintes aspectos dos dados:

- identificação de um valor típico ou representativo;
- extensão da dispersão ao redor do valor típico;
- presença de lacunas nos dados;
- extensão da simetria na distribuição de valores;
- número e localização dos bicos;
- presença de valores fora da curva.

### Exemplo 1.6

A Figura 1.5 apresenta diagramas de caule e folha para uma amostra aleatória de comprimentos de campos de golfe (jardas) que foram designados pela *Golf Magazine* entre os mais desafiadores dos Estados Unidos. Entre a amostra de 40 campos, o mais curto tem 6.433 jardas de comprimento e o mais longo, 7.280. Os comprimentos parecem estar distribuídos de forma aproximadamente uniforme acima da faixa de valores da amostra. Observe que, aqui, uma escolha de caule de um único dígito (6 ou 7) ou de três dígitos (643, ... , 728) resultaria em diagramas não-informativos, no primeiro caso porque não haveria caules suficientes, no segundo, porque haveria caules demais.

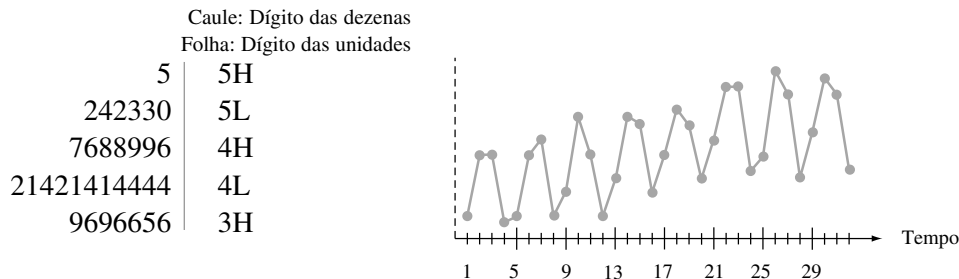


**Figura 1.5** Diagramas de caule e folha de comprimentos em jardas de campos de golfe: (a) folhas de dois dígitos; (b) diagrama de MINITAB com folhas truncadas de um dígito. ■

Um diagrama de caule e folha não mostra a ordem em que as observações foram obtidas, possivelmente ocultando importantes informações sobre o mecanismo gerador dos dados. Por exemplo: durante um período de tempo determinado, a largura de um calço de certa peça feito por uma fresa tende a aumentar em relação ao valor-alvo, indicando um processo “fora de controle”. Um *gráfico de linha com marcadores* combina o quadro das observações ao longo do tempo com um diagrama caule e folha.

**Exemplo 1.7**

Cada observação no diagrama de caule e folha da Figura 1.6 é o valor de produção de cerveja dos EUA (milhões de barris) para um trimestre diferente durante o período de 1975-1982. O diagrama usa *caules repetidos*; por exemplo: a linha 4L é para observações com uma folha “baixa” — 0, 1, 2, 3 ou 4 — e as observações com folhas mais altas são colocadas na linha 4H. O gráfico da *série temporal* à direita mostra tendência de aumento com o tempo e também maior produção nos segundo e terceiro trimestres de um determinado ano do que nos outros dois trimestres (um efeito *sazonal*).



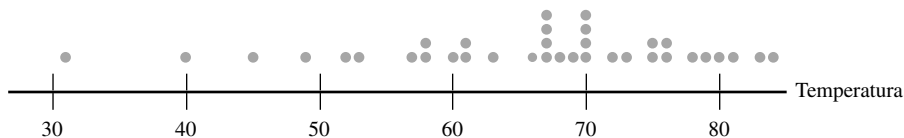
**Figura 1.6** Um gráfico de linha com marcadores de produção de cerveja dos Estados Unidos ■

**Gráfico de pontos**

Um gráfico de pontos é um interessante resumo dos dados numéricos quando esse conjunto é razoavelmente pequeno ou possui relativamente poucos valores distintos. Cada observação é representada por um ponto sobre o local correspondente em uma escala de medida horizontal. Quando um valor ocorre mais de uma vez, há um ponto para cada ocorrência e esses pontos são empilhados verticalmente. Como ocorre com o diagrama de caule e folha, um gráfico de pontos fornece informações sobre localização, dispersão, extremos e lacunas.

**Exemplo 1.8**

A Figura 1.7 mostra um gráfico de pontos para os dados de temperatura de Anel de Vedação apresentados no Exemplo 1.1, na seção anterior. Um valor de temperatura representativo é o que se apresenta na metade de 60 (°F) e há grande dispersão ao redor do centro. Os dados se estendem mais na extremidade inferior do que na extremidade superior e a menor observação, 31, pode ser claramente descrita como um *outlier*.



**Figura 1.7** Um gráfico de pontos dos dados de temperatura do Anel de Vedação (°F) ■

Se o conjunto de dados discutido no Exemplo 1.8 consistisse de 50 ou 100 observações de temperatura, cada uma registrada em décimos de graus, seria muito mais trabalhoso construir um gráfico de pontos. Nossa próxima técnica é melhor adaptável a tais situações.

**Histogramas**

Alguns dados numéricos são obtidos por contagem para determinar o valor de uma variável (o número de autuações de tráfego que uma pessoa recebeu no último ano, o número de pessoas que chegam para trabalhar durante um determinado período), enquanto outros dados são obtidos pela tomada de medidas (peso de um indivíduo,

tempo de reação a determinado estímulo). A recomendação para plotagem de um histograma geralmente é diferente para esses dois casos.

#### DEFINIÇÃO

Uma variável é **discreta** se o seu conjunto de valores possíveis é finito ou pode ser relacionado em uma seqüência infinita (em que haja um primeiro número, um segundo e assim por diante). Uma variável é **contínua** se os seus valores possíveis consistem de um intervalo completo na reta de numeração.

Uma variável discreta  $x$  quase sempre é resultante de contagem e, nesse caso, 0, 1, 2, 3... ou algum subconjunto desses inteiros são valores possíveis. Variáveis contínuas surgem da tomada de medidas. Por exemplo: se  $x$  é o pH de uma substância química, em teoria,  $x$  pode ser qualquer número entre 0 e 14: 7.0, 7.03, 7.032 e assim por diante. Claro que, na prática, há limitações no que diz respeito ao grau de precisão de qualquer instrumento de medida, de forma que podemos não ser capazes de determinar pH, tempo de reação, altura e concentração com um número arbitrariamente grande de casas decimais. Entretanto, do ponto de vista da criação de modelos matemáticos para a distribuição dos dados, é útil imaginar um intervalo contínuo de valores possíveis.

Considere os dados constituídos de observações de uma variável discreta  $x$ . A **freqüência** de qualquer valor particular de  $x$  é o número de vezes em que esse valor ocorre naquele conjunto. A **freqüência relativa** de um valor é a fração ou proporção de vezes em que o valor ocorre:

$$\text{freqüência relativa de um valor} = \frac{\text{número de vezes que o valor ocorre}}{\text{número de observações do conjunto de dados}}$$

Suponha, por exemplo, que o nosso conjunto de dados consista em 200 observações de  $x$  = o número de defeitos graves em um novo carro de certo tipo. Se 70 desses valores  $x$  forem 1, então

$$\text{freqüência do valor } x = 1: 70$$

$$\text{freqüência relativa do valor } x = 1: \frac{70}{200} = 0,35$$

A multiplicação da freqüência relativa por 100 fornece a porcentagem. No exemplo dos defeitos, 35% dos carros da amostra apresentaram apenas um defeito grave. As freqüências relativas, ou porcentagens, normalmente interessam mais do que as freqüências em si. Teoricamente, as freqüências relativas deveriam somar 1, mas, na prática, a soma pode ser ligeiramente diferente por causa do arredondamento. Uma **distribuição de freqüência** é uma tabulação das freqüências e/ou freqüências relativas.

#### Construção do histograma para dados discretos

Primeiro, determine a freqüência e a freqüência relativa de cada valor de  $x$ . Depois, marque os valores possíveis de  $x$  em uma escala horizontal. Acima de cada valor, desenhe um retângulo cuja altura seja a freqüência relativa (ou a freqüência, como alternativa) daquele valor.

Essa construção assegura que a *área* de cada retângulo seja proporcional à freqüência relativa do valor. Assim, se as freqüências relativas de  $x = 1$  e  $x = 5$  são 0,35 e 0,07, respectivamente, a área do retângulo acima de 1 será cinco vezes a área do retângulo acima de 5.

#### Exemplo 1.9

Quão incomum é um jogador que não atinge a bola ou a atinge uma única vez em um jogo de beisebol da liga principal e com que freqüência uma equipe consegue atingir a bola mais de 10, 15 ou mesmo 20 vezes? A Tabela 1.1 é uma distribuição de freqüência do número de acertos por equipe, por partida, para todos os jogos de nove séries entre 1989 e 1993.

**Tabela 1.1** Distribuição de frequência de acertos em jogos de nove séries

Acertos/Jogo	Número de Jogos	Frequência Relativa	Acertos/Jogo	Número de Jogos	Frequência Relativa
0	20	0,0010	14	569	0,0294
1	72	0,0037	15	393	0,0203
2	209	0,0108	16	253	0,0131
3	527	0,0272	17	171	0,0088
4	1048	0,0541	18	97	0,0050
5	1457	0,0752	19	53	0,0027
6	1988	0,1026	20	31	0,0016
7	2256	0,1164	21	19	0,0010
8	2403	0,1240	22	13	0,0007
9	2256	0,1164	23	5	0,0003
10	1967	0,1015	24	1	0,0001
11	1509	0,0779	25	0	0,0000
12	1230	0,0635	26	1	0,0001
13	834	0,0430	27	1	0,0001
				19,383	1,0005

O histograma correspondente da Figura 1.8 tem um leve aclive para um único pico e depois tem um declive. O histograma se estende um pouco mais do lado direito (em direção aos valores maiores) do que para o lado esquerdo – uma inclinação ligeiramente “positiva”.

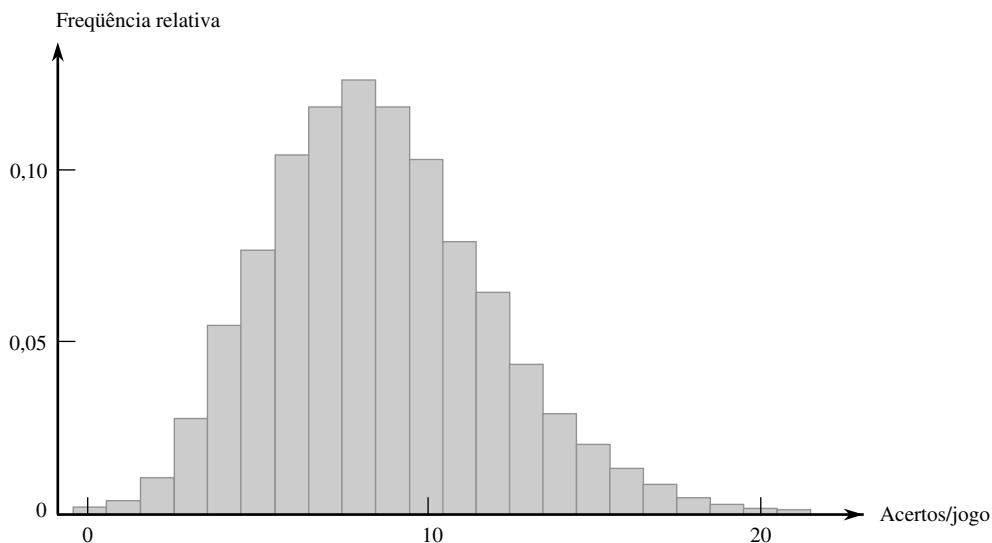
A partir das informações tabuladas ou do histograma em si, podemos determinar o seguinte:

$$\begin{aligned} \text{proporção de jogos com} & & \text{frequência} & & \text{frequência} & & \text{frequência} \\ \text{no máximo dois acertos} & = & \text{relativa para} & + & \text{relativa para} & + & \text{relativa para} \\ & & x = 0 & & x = 1 & & x = 2 \\ & & & & & & \\ & = & 0,0010 & + & 0,0037 & + & 0,0108 = 0,0155 \end{aligned}$$

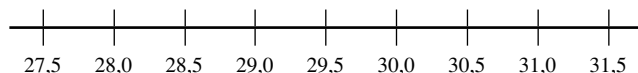
De forma similar,

$$\text{proporção de jogos com acertos entre 5 e 10 (inclusive)} = 0,0752 + 0,1026 + \dots + 0,1015 = 0,6361$$

Isto é, cerca de 64% de todos esses jogos tiveram entre 5 e 10 (inclusive) acertos.



A construção de um histograma de dados contínuos (medidas) exige que o eixo das medidas seja subdividido em um número aceitável de **intervalos de classe** ou **classes**, de forma que cada observação esteja contida completamente em uma classe. Suponha, por exemplo, que tenhamos 50 observações de  $x =$  consumo de combustível de um automóvel (milhas/galão), sendo o menor deles 27,8 e o maior, 31,4. Podemos, então, usar os limites de classe 27,5, 28,0, 28,5, ... , e 31,5, conforme mostrado abaixo:



Uma dificuldade potencial é que, ocasionalmente, uma observação fica exatamente sobre um dos limites de classe, não estando necessariamente em apenas um intervalo, por exemplo, 29,0. Uma forma de lidar com esse problema é usar limites como 27,55, 28,05, ... , 31,55. A adição do dígito de centésimos aos limites de classe evita que as observações estejam exatamente sobre os limites resultantes. Outra abordagem é usar as classes 27,5–<28,0, 28,0–<28,5, ... , 31,0–<31,5. Então, 29,0 estaria na classe 29,0–<29,5 em vez de na classe 28,5–<29,0. Em outras palavras, com essa convenção, uma observação sobre um limite é colocada no intervalo à direita do limite. Essa é a maneira pela qual o MINITAB constrói um histograma.

**Construção de histograma para dados contínuos: classes de larguras iguais**

Determine a frequência e a frequência relativa de cada classe. Marque os limites de classe em um eixo de medida horizontal. Acima de cada intervalo de classe, desenhe um retângulo cuja altura seja a frequência relativa correspondente (ou a frequência).

**Exemplo 1.10**

As empresas de energia necessitam de informações sobre o consumo de seus clientes para obterem previsões precisas da demanda. Investigadores da Wisconsin Power and Light determinaram que o consumo de energia (BTUs) dura um determinado período para uma amostra de 90 lares aquecidos a gás. O valor de consumo ajustado foi calculado conforme segue:

$$\text{consumo ajustado} = \frac{\text{consumo}}{(\text{clima, em grau dias})(\text{área da casa})}$$

O resultado apresenta-se nos dados anexos (parte do conjunto de dados armazenado FURNACE.MTW disponível no MINITAB) que ordenamos do menor para o maior.

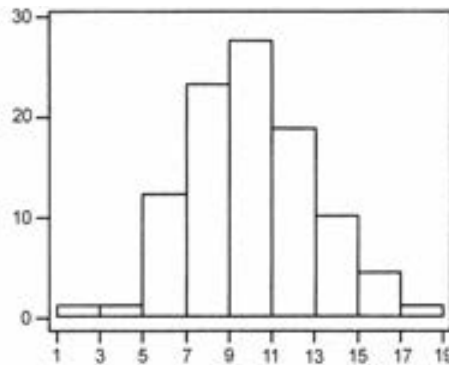
2,97	4,00	5,20	5,56	5,94	5,98	6,35	6,62	6,72	6,78
6,80	6,85	6,94	7,15	7,16	7,23	7,29	7,62	7,62	7,69
7,73	7,87	7,93	8,00	8,26	8,29	8,37	8,47	8,54	8,58
8,61	8,67	8,69	8,81	9,07	9,27	9,37	9,43	9,52	9,58
9,60	9,76	9,82	9,83	9,83	9,84	9,96	10,04	10,21	10,28
10,28	10,30	10,35	10,36	10,40	10,49	10,50	10,64	10,95	11,09
11,12	11,21	11,29	11,43	11,62	11,70	11,70	12,16	12,19	12,28
12,31	12,62	12,69	12,71	12,91	12,92	13,11	13,38	13,42	13,43
13,47	13,60	13,96	14,24	14,35	15,12	15,24	16,06	16,90	18,26

Deixamos o MINITAB selecionar os intervalos de classe. A característica mais impressionante do histograma da Figura 1.9 é sua semelhança com uma curva em forma de sino (e portanto simétrica), com o ponto de simetria próximo a 10.

Classe	1-<3	3-<5	5-<7	7-<9	9-<11	11-<13	13-<15	15-<17	17-<19
Frequência	1	1	11	21	25	17	9	4	1
Frequência relativa	0,011	0,011	0,122	0,233	0,278	0,189	0,100	0,044	0,011

Pelo histograma,

$$\begin{array}{l} \text{proporção de} \\ \text{observações} \\ \text{inferior a 9} \end{array} \approx 0,01 + 0,01 + 0,12 + 0,23 = 0,37 \quad (\text{valor exato} = \frac{34}{90} = 0,378)$$



**Figura 1.9** Histograma dos dados do consumo de energia do Exemplo 1.10

A frequência relativa da classe 9-<11 é cerca de 0,27, de forma que estimamos que metade desse valor, ou 0,135, esteja entre 9 e 10. Conseqüentemente,

$$\begin{array}{l} \text{proporção de observações} \\ \text{inferior a 10} \end{array} \approx 0,37 + 0,135 = 0,505 \quad (\text{pouco mais de } 50\%)$$

O valor exato dessa proporção é  $47/90 = 0,522$ . ■

Não há regras rápidas e absolutas sobre o número de classes ou a escolha das classes em si. Entre 5 e 20 classes serão satisfatórias para a maior parte dos conjuntos de dados. Normalmente, quanto maior o número de observações em um conjunto de dados, mais classes devem ser usadas. Uma regra prática razoável é

$$\text{número de classes} \approx \sqrt{\text{número de observações}}$$

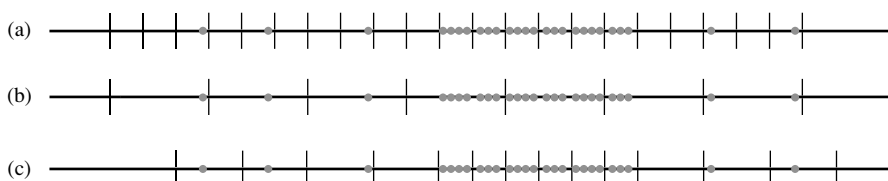
Classes de mesma largura podem não ser uma boa escolha se o conjunto de dados “se estender” para um lado ou para o outro. A Figura 1.10 exibe um gráfico de pontos desse conjunto de dados. Usar um número pequeno de classes de mesma largura resulta em quase todas as observações estarem em apenas uma ou duas classes. Se for usado um grande número de classes de mesma largura for usado, muitas classes terão frequência zero. Uma opção melhor é usar alguns intervalos mais amplos próximos às observações dos extremos e intervalos mais estreitos na região de alta concentração.

**Construindo um histograma para dados contínuos:  
classes de larguras diferentes**

Após determinar as frequências e as frequências relativas, calcule a altura de cada retângulo, usando a fórmula

$$\text{altura do retângulo} = \frac{\text{frequência relativa da classe}}{\text{largura da classe}}$$

As alturas resultantes dos retângulos normalmente são denominadas *densidades* e a escala vertical é a **escala de densidade**. Essa recomendação também funcionará quando as larguras das classes forem iguais.



**Figura 1.10** Seleção de intervalos de classe para pontos “estendidos”: (a) muitos intervalos estreitos de mesma largura; (b) poucos intervalos mais amplos; (c) intervalos de larguras diferentes

**Exemplo 1.11**

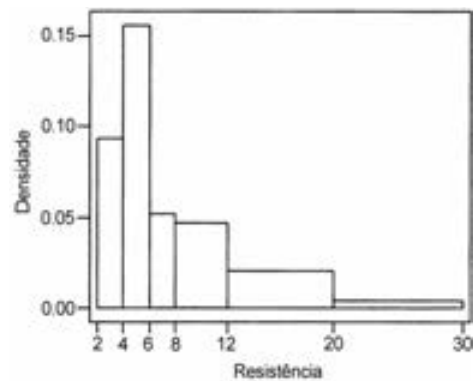
A corrosão das barras de aço da armação é um problema sério em estruturas de concreto localizadas em ambientes afetados por condições climáticas extremas. Por esse motivo, os pesquisadores têm investigado a utilização de barras de reforço feitas de material composto. Um estudo foi executado para desenvolver diretrizes sobre a aderência de barras plásticas reforçadas com fibra de vidro ao concreto (“Design Recommendations for Bond of GFRP Rebars to Concrete,” *J. of Structural Engr.*, 1996, p. 247-254). Considere as 48 observações da resistência da aderência medida:

11,5	12,1	9,9	9,3	7,8	6,2	6,6	7,0	13,4	17,1	9,3	5,6
5,7	5,4	5,2	5,1	4,9	10,7	15,2	8,5	4,2	4,0	3,9	3,8
3,6	3,4	20,6	25,5	13,8	12,6	13,1	8,9	8,2	10,7	14,2	7,6
5,2	5,5	5,1	5,0	5,2	4,8	4,1	3,8	3,7	3,6	3,6	3,6

<i>Classe</i>	2-<4	4-<6	6-<8	8-<12	12-<20	20-<30
<i>Frequência</i>	9	15	5	9	8	2
<i>Frequência relativa</i>	0,1875	0,3125	0,1042	0,1875	0,1667	0,0417
<i>Densidade</i>	0,094	0,156	0,052	0,047	0,021	0,004

O histograma resultante é exibido na Figura 1.11. A cauda direita ou superior se estende muito além da esquerda ou inferior – um desvio substancial da simetria.





**Figura 1.11** Um histograma de densidade em MINITAB dos dados de resistência à aderência do Exemplo 1.11

Quando as larguras de classe são diferentes, deixar de usar uma escala de densidade gera um gráfico com áreas distorcidas. Para classes de larguras iguais, o divisor é o mesmo em cada cálculo de densidade e o cálculo adicional é simplesmente uma nova escala do eixo vertical (ou seja, os histogramas que usam frequência relativa e os que usam densidade terão exatamente a mesma aparência). Um histograma de densidade possui uma propriedade interessante. Ao multiplicar os dois lados da fórmula pela largura da classe, teremos

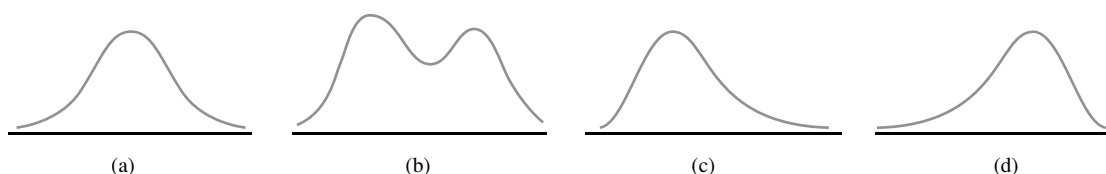
$$\begin{aligned} \text{frequência relativa} &= (\text{largura da classe})(\text{densidade}) = (\text{largura do retângulo})(\text{altura do retângulo}) \\ &= \text{área do retângulo} \end{aligned}$$

Isto é, *a área de cada retângulo é a frequência relativa da classe correspondente*. Além disso, como a soma das frequências relativas deve ser 1,0 (exceto por arredondamento), *a área total de todos os retângulos em um histograma de densidade é 1*. Sempre é possível desenhar um histograma de forma que a área seja igual à frequência relativa (isso também é verdade para um histograma de dados discretos). É só usar a escala de densidade. Essa propriedade terá um papel importante na criação de modelos de distribuições no Capítulo 4.

## Formatos de histogramas

Os histogramas podem ter diversos formatos. Um histograma **unimodal** é aquele que possui um aclave para um único pico e depois um declive. Um histograma **bimodal** possui dois picos diferentes. A bimodalidade pode ocorrer quando o conjunto de dados consistir em observações sobre dois tipos bastante diferentes de indivíduos ou objetos. Por exemplo: considere um grande conjunto de dados formado por tempos de viagem de automóveis entre San Luis Obispo, Califórnia, e Monterey, Califórnia (excluindo tempo de parada para apreciar a vista, comer etc.). Esse histograma mostraria dois picos: um para os carros que tomaram a rota do interior (cerca de 2,5 horas) e outro para os carros que foram pelo litoral (3,5–4 horas). A bimodalidade, entretanto, não acontece automaticamente nessas situações. A bimodalidade ocorrerá no histograma de dados combinados somente se os dois histogramas separados estiverem “distantes” em relação às suas dispersões. Assim, um grande conjunto de dados consistindo em alturas de alunos de faculdades não deve resultar em um histograma bimodal porque a altura típica dos homens, cerca de 69 polegadas, não está suficientemente distante da altura típica das mulheres, cerca de 64-65 polegadas. Um histograma com mais de dois picos é denominado **multimodal**. Claro que o número de picos pode depender da escolha dos intervalos de classe, particularmente com um pequeno número de observações. Quanto maior o número de classes, maior é a probabilidade de a bimodalidade ou de a multimodalidade se manifestar.

Um histograma é **simétrico** se a metade esquerda for uma imagem refletida da metade direita. Um histograma unimodal tem **inclinação positiva** se a cauda direita ou superior for estendida em comparação à cauda esquerda ou inferior e **inclinação negativa** desviar-se para a esquerda. A Figura 1.12 exhibe histogramas “ajustados”, obtidos pela sobreposição de uma curva ajustada sobre os retângulos, que ilustram as diversas possibilidades.



**Figura 1.12** Histogramas ajustados: (a) unimodal simétrico; (b) bimodal; (c) desvio positivo; (d) desvio negativo

### Dados qualitativos

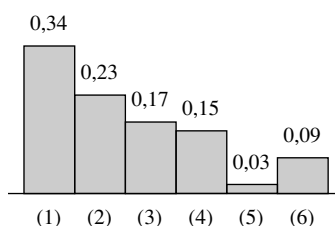
Tanto uma distribuição de freqüência como um histograma podem ser construídos quando o conjunto de dados for de natureza *qualitativa* (categorizada). Em alguns casos, haverá uma organização natural das classes, por exemplo: calouros, segundanistas, terceiranistas, formandos e graduados, enquanto em outros casos a organização será arbitrária, por exemplo: católicos, judeus, protestantes e assim por diante. Com esses dados categorizados, os intervalos sobre os quais os retângulos serão construídos devem ter a mesma largura.

#### Exemplo 1.12

Cada membro de uma amostra de 120 indivíduos proprietários de motocicletas foi indagado sobre a marca de sua moto. A distribuição de freqüência dos dados resultantes é fornecida na Tabela 1.2 e o histograma é exibido na Figura 1.13.

**Tabela 1.2** Distribuição de freqüência dos dados de motocicletas

Fabricante	Freqüência	Freqüência Relativa
1. Honda	41	0,34
2. Yamaha	27	0,23
3. Kawasaki	20	0,17
4. Harley-Davidson	18	0,15
5. BMW	3	0,03
6. Outro	11	0,09
	120	1,01



**Figura 1.13** Histograma dos dados de motocicletas

### Dados Multivariados

As técnicas apresentadas até agora referem-se exclusivamente a situações em que cada observação de um conjunto de dados é um único número ou uma única categoria. Os dados, entretanto, freqüentemente são de natureza multivariada. Isto é, se obtivermos uma amostra de indivíduos ou objetos e em cada um tivermos duas ou mais medidas, cada “observação” consistirá em diversas medidas de um indivíduo ou objeto. A amostra é bivariada se cada observação consistir em duas medidas ou respostas, de forma que o conjunto de dados possa ser representado como  $(x_1, y_1), \dots, (x_n, y_n)$ . Por exemplo:  $x$  pode se referir ao tamanho do motor e  $y$  ao seu deslocamento; ou  $x$  pode se referir à marca da calculadora de um formando e  $y$  à sua formação acadêmica. Nos capítulos 11-14, analisaremos conjuntos de dados multivariados desse tipo; assim, adiaremos uma discussão mais detalhada até lá.

## Exercícios | Seção 1.2 (10–32)

**10.** Considere os dados de resistência de vigas fornecidos no Exemplo 1.2.

- Construa um diagrama de caule e folha dos dados. O que parece ser um valor de resistência representativo? As observações parecem estar concentradas ao redor do valor representativo ou dispersas?
- O diagrama parece razoavelmente simétrico ao redor de um valor representativo ou seu formato pode ser descrito de outra forma?
- Parece haver valores aberrantes de resistência?
- Que proporção de observações de resistência dessa amostra excedem 10 MPa?

**11.** Cada pontuação do conjunto de notas de um exame a seguir está nas dezenas 60, 70, 80 ou 90. Um diagrama de caule e folha com apenas os quatro caules 6, 7, 8 e 9 não forneceria uma descrição muito detalhada da distribuição das pontuações. Nessas situações, é desejável usarmos caules repetidos. Aqui podemos repetir o caule 6 duas vezes, usando 6L para pontuações na parte inferior da dezena dos 60 (folhas 0, 1, 2, 3 e 4) e 6H para as pontuações na parte superior da dezena dos 60 (folhas 5, 6, 7, 8 e 9). De forma similar, os outros caules podem ser repetidos duas vezes para obtermos um diagrama consistindo em oito linhas. Construa esse diagrama para as pontuações fornecidas. Que característica dos dados é realçada por ele?

74 89 80 93 64 67 72 70 66 85 89 81 81  
71 74 82 85 63 72 81 81 95 84 81 80 70  
69 66 60 83 85 98 84 68 90 82 69 72 87  
88

**12.** Os dados anexos de densidade para diversos tipos de madeira usados em construção foram relatados no artigo “Bolted Connection Design Values Based on European Yield Model” (*J. of Structural Engr.*, 1993, p. 2169-2186):

0,31 0,35 0,36 0,36 0,37 0,38 0,40 0,40 0,40  
0,41 0,41 0,42 0,42 0,42 0,42 0,42 0,43 0,44  
0,45 0,46 0,46 0,47 0,48 0,48 0,48 0,51 0,54  
0,54 0,55 0,58 0,62 0,66 0,66 0,67 0,68 0,75

Construa um diagrama de caule e folha usando caules repetidos (veja o exercício anterior) e comente suas características interessantes.

**13.** As propriedades mecânicas permissíveis para projetos estruturais de veículos aeroespaciais metálicos exigem um método aprovado para análise estatística de dados de testes empíricos. O artigo “Establishing Mechanical Property Allowables for Metals” (*J. of Testing and Evaluation*, 1998, p. 293-299) usou os dados sobre resistência à tração (ksi) como base para definir as dificuldades de desenvolvimento do método.

122,2 124,2 124,3 125,6 126,3 126,5 126,5 127,2 127,3  
127,5 127,9 128,6 128,8 129,0 129,2 129,4 129,6 130,2  
130,4 130,8 131,3 131,4 131,4 131,5 131,6 131,6 131,8  
131,8 132,3 132,4 132,4 132,5 132,5 132,5 132,5 132,6  
132,7 132,9 133,0 133,1 133,1 133,1 133,1 133,2 133,2  
133,2 133,3 133,3 133,5 133,5 133,5 133,8 133,9 134,0  
134,0 134,0 134,0 134,1 134,2 134,3 134,4 134,4 134,6  
134,7 134,7 134,7 134,8 134,8 134,8 134,9 134,9 135,2  
135,2 135,2 135,3 135,3 135,4 135,5 135,5 135,6 135,6  
135,7 135,8 135,8 135,8 135,8 135,8 135,9 135,9 135,9  
135,9 136,0 136,0 136,1 136,2 136,2 136,3 136,4 136,4  
136,6 136,8 136,9 136,9 137,0 137,1 137,2 137,6 137,6  
137,8 137,8 137,8 137,9 137,9 138,2 138,2 138,3 138,3  
138,4 138,4 138,4 138,5 138,5 138,6 138,7 138,7 139,0  
139,1 139,5 139,6 139,8 139,8 140,0 140,0 140,7 140,7  
140,9 140,9 141,2 141,4 141,5 141,6 142,9 143,4 143,5  
143,6 143,8 143,8 143,9 144,1 144,5 144,5 147,7 147,7

- Construa um diagrama de caule e folha dos dados excluindo (truncando) inicialmente os dígitos decimais e depois repetindo cada caule cinco vezes (uma vez para as folhas 1 e 2, uma segunda vez para as folhas 3 e 4 etc.). Por que é relativamente fácil identificar um valor de resistência representativo?
- Construa um histograma usando classes de mesma largura em que a primeira classe possua um limite inferior a 122 e um limite superior a 124. Comente, então, características interessantes do histograma.

**14.** O conjunto de dados a seguir consiste de observações da vazão de chuviros (L/min) de uma amostra de  $n = 129$  lares em Perth, Austrália (“An Application of Bayes Methodology to the Analysis of Diary Records in a Water Use Study”, *J. Amer. Stat. Assoc.*, 1987, p. 705-711):

4,6 12,3 7,1 7,0 4,0 9,2 6,7 6,9 11,5 5,1  
11,2 10,5 14,3 8,0 8,8 6,4 5,1 5,6 9,6 7,5  
7,5 6,2 5,8 2,3 3,4 10,4 9,8 6,6 3,7 6,4  
8,3 6,5 7,6 9,3 9,2 7,3 5,0 6,3 13,8 6,2  
5,4 4,8 7,5 6,0 6,9 10,8 7,5 6,6 5,0 3,3  
7,6 3,9 11,9 2,2 15,0 7,2 6,1 15,3 18,9 7,2  
5,4 5,5 4,3 9,0 12,7 11,3 7,4 5,0 3,5 8,2  
8,4 7,3 10,3 11,9 6,0 5,6 9,5 9,3 10,4 9,7  
5,1 6,7 10,2 6,2 8,4 7,0 4,8 5,6 10,5 14,6  
10,8 15,5 7,5 6,4 3,4 5,5 6,6 5,9 15,0 9,6  
7,8 7,0 6,9 4,1 3,6 11,9 3,7 5,7 6,8 11,3  
9,3 9,6 10,4 9,3 6,9 9,8 9,1 10,6 4,5 6,2  
8,3 3,2 4,9 5,0 6,0 8,2 6,3 3,8 6,0

- Construa um diagrama de caule e folha dos dados.
- Qual taxa é considerada vazão típica ou representativa?
- O diagrama parece ser concentrado ou disperso?

- d. A distribuição dos valores parece razoavelmente simétrica? Em caso negativo, como você descreveria o desvio da simetria?
- e. Você descreveria alguma observação como distante do resto dos dados (um *outlier*)?

15. Um artigo da *Consumer Reports* sobre pasta de amendoim (setembro de 1990) relatou as seguintes pontuações para diversas marcas:

<i>Cremosa</i>	56	44	62	36	39	53	50	65	45	40
	56	68	41	30	40	50	56	30	22	
<i>Crocante</i>	62	53	75	42	47	40	34	62	52	
	50	34	42	36	75	80	47	56	62	

Construa um diagrama de caule e folha comparativo, relacionando caules na parte central da página e, então, exibindo as folhas de cremosa à direita e as de crocante à esquerda. Descreva semelhanças e diferenças para os dois tipos.

16. O artigo citado no Exemplo 1.2 também fornece as informações de resistência de corpos de prova a seguir:

6,1	5,8	7,8	7,1	7,2	9,2	6,6	8,3	7,0	8,3
7,8	8,1	7,4	8,5	8,9	9,8	9,7	14,1	12,6	11,2

- a. Construa um diagrama de caule e folha comparativo (veja o exercício anterior) dos dados de vigas e cilindros e, então, responda às questões nas partes (b)-(d) do Exercício 10 para as observações sobre os corpos de prova.
- b. De que formas os dois lados do diagrama são similares? Há diferenças óbvias entre as observações sobre vigas e corpos de prova?
- c. Construa um gráfico de pontos com os dados dos corpos de prova.

17. Os transdutores de temperatura de um determinado tipo são enviados em lotes de 50. Uma amostra de 60 lotes foi selecionada e o número de transdutores fora das especificações em cada lote foi determinado, resultando nos dados a seguir:

2	1	2	4	0	1	3	2	0	5	3	3	1	3	2	4	7	0	2	3
0	4	2	1	3	1	1	3	4	1	2	3	2	2	8	4	5	1	3	1
5	0	2	3	2	1	0	6	4	2	1	6	0	3	3	3	6	1	2	3

- a. Determine as frequências e frequências relativas dos valores observados de  $x$  = número de transdutores fora das especificações em um lote.
- b. Que proporção de lotes na amostra possui no máximo cinco transdutores fora das especificações? Que proporção tem menos de cinco? Que proporção possui no mínimo cinco unidades fora das especificações?
- c. Desenhe um histograma dos dados, usando a frequência relativa na escala vertical e comente suas características.

18. Em um estudo de produtividade literária (“Lotka’s Test,” *Collection Mgmt.*, 1982, p. 111-118), um grande

número de autores foi classificado de acordo com o número de artigos que tinham publicado durante certo período. Os resultados foram apresentados na distribuição de frequência a seguir:

<i>Número de artigos</i>	1	2	3	4	5	6	7	8
<i>Frequência</i>	784	204	127	50	33	28	19	19

<i>Número de artigos</i>	9	10	11	12	13	14	15	16	17
<i>Frequência</i>	6	7	6	7	4	4	5	3	3

- a. Construa um histograma correspondente a essa distribuição de frequência. Qual é a característica mais interessante do formato da distribuição?
- b. Que proporção desses autores publicou no mínimo cinco artigos? No mínimo 10 artigos? Mais de 10 artigos?
- c. Suponha que os cinco que publicaram 15 artigos, os três que publicaram 16 e os três que publicaram 17 tenham sido agrupados em uma única categoria exibida como “ $\geq 15$ .” Você pode construir um histograma? Explique.
- d. Suponha que, em vez dos valores 15, 16 e 17 relacionados separadamente, eles tenham sido combinados em uma categoria 15-17 com frequência 11. Você pode construir um histograma? Explique.

19. O número de partículas de contaminação de uma pastilha de silício antes de certo processo de limpeza foi determinado para cada pastilha em uma amostra de tamanho 100, resultando nas frequências a seguir:

<i>Número de partículas</i>	0	1	2	3	4	5	6	7
<i>Frequência</i>	1	2	3	12	11	15	18	10

<i>Número de partículas</i>	8	9	10	11	12	13	14
<i>Frequência</i>	12	4	5	3	1	2	1

- a. Que proporção das pastilhas da amostra tinha ao menos uma partícula? Ao menos cinco partículas?
- b. Que proporção das pastilhas da amostra tinha entre cinco e 10 (inclusive) partículas? Estritamente entre cinco e 10 partículas?
- c. Desenhe um histograma usando a frequência relativa no eixo vertical. Como você descreveria o formato do histograma?

20. O artigo “Determination of Most Representative Subdivision” (*J. of Energy Engr.*, 1993, p. 43-55) forneceu dados sobre diversas características de subdivisões que podiam ser usadas na decisão de fornecimento de energia via linhas subterrâneas ou aéreas. Seguem os valores da variável  $x$  = comprimento total das ruas dentro de uma subdivisão:

1280	5320	4390	2100	1240	3060	4770
1050	360	3330	3380	340	1000	960
1320	530	3350	540	3870	1250	2400
960	1120	2120	450	2250	2320	2400
3150	5700	5220	500	1850	2460	5850
2700	2730	1670	100	5770	3150	1890
510	240	396	1419	2109		

- a. Construa um diagrama de caule e folha, usando o dígito de milhares como caule e o de centenas como folha, e comente suas características.
- b. Construa um histograma, usando limites de classe de 0, 1000, 2000, 3000, 4000, 5000 e 6000. Que proporção de subdivisões possui comprimento total inferior a 2000? Entre 2000 e 4000? Como você descreveria o formato do histograma?
21. O artigo citado no Exercício 20 também forneceu os seguintes valores das variáveis  $y$  = números de *culs-de-sac* e  $z$  = número de interseções:

y	1	0	1	0	0	2	0	1	1	1	2	1	0	0	1	1	0	1	1
z	1	8	6	1	1	5	3	0	0	4	4	0	0	1	2	1	4	0	4
y	1	1	0	0	0	1	1	2	0	1	2	2	1	1	0	2	1	1	0
z	0	3	0	1	1	0	1	3	2	4	6	6	0	1	1	8	3	3	5
y	1	5	0	3	0	1	1	0	0										
z	0	5	2	3	1	0	0	0	3										

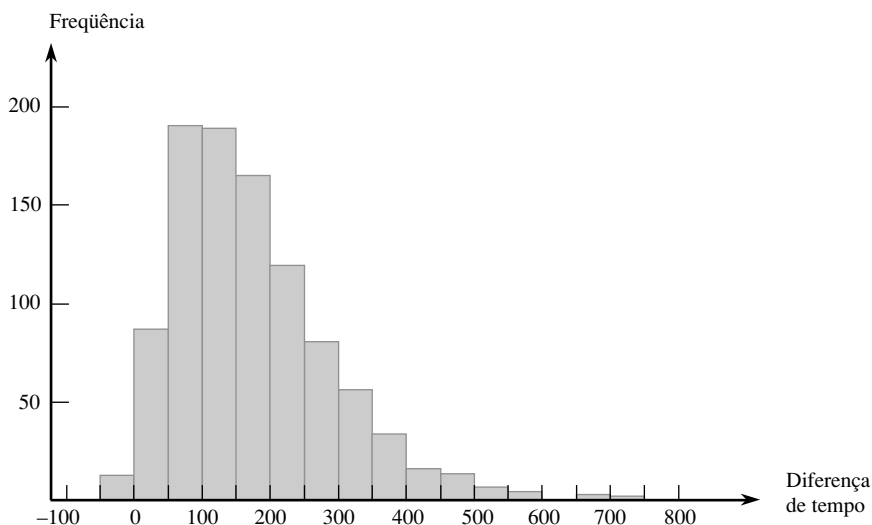
- a. Construa um histograma dos dados de  $y$ . Que proporção dessas subdivisões não possuíam *culs-de-sac*? Ao menos um *cul-de-sac*?
- b. Construa um histograma para os dados  $z$ . Que proporção dessas subdivisões têm no máximo cinco interseções? Menos de cinco interseções?

22. Como varia a velocidade de um corredor no curso de uma maratona (uma distância de 42,195 km)? Considere a determinação do tempo de corrida dos cinco primeiros quilômetros e o tempo de corrida entre os pontos dos kms 35 e 40; subtraia o primeiro tempo do último. Um valor positivo dessa diferença corresponde a um corredor que diminui seu ritmo no final da corrida. O histograma a seguir se baseia nos tempos de corredores que participaram de diversas maratonas diferentes no Japão (“Factors Affecting Runners’ Marathon Performance”, *Chance*, Fall, 1993, p. 24-30). Que características são interessantes nesse histograma? Qual é um valor típico da diferença? Grosso modo, que proporção dos maratonistas correu a última etapa mais rápido que a primeira?
23. Em um estudo de quebras de urdidura durante a tecelagem de tecidos (*Technometrics*, 1982, p. 63), 100 amostras de fios foram testadas. O número de ciclos de esforço para quebra foi determinado para cada amostra de fio, resultando nos dados a seguir:

86	146	251	653	98	249	400	292	131	169
175	176	76	264	15	364	195	262	88	264
157	220	42	321	180	198	38	20	61	121
282	224	149	180	325	250	196	90	229	166
38	337	65	151	341	40	40	135	597	246
211	180	93	315	353	571	124	279	81	186
497	182	423	185	229	400	338	290	398	71
246	185	188	568	55	55	61	244	20	284
393	396	203	829	239	236	286	194	277	143
198	264	105	203	124	137	135	350	193	188

- a. Construa um histograma de frequência relativa com base nos intervalos de classe 0-<100, 100-<200, ... e comente as características do histograma.

Histograma do Exercício 22



b. Construa um histograma com base nos seguintes intervalos de classe: 0–<50, 50–<100, 100–<150, 150–<200, 200–<300, 300–<400, 400–<500, 500–<600 e 600–<900.

c. Se as especificações de tecelagem exigem um esforço de quebra de ao menos 100 ciclos, que proporção das amostras de fio dessa amostra deve ser considerada satisfatória?

24. O conjunto de dados anexo consiste em observações da resistência de corte (lb) de soldas de ponto ultrassônicas feitas sobre um determinado tipo de chapa de Alclad. Construa um histograma de frequência relativa com base em 10 classes de mesma largura com limites 4000, 4200, .... [O histograma coincide com o mostrado em “Comparison of Properties of Joints Prepared by Ultrasonic Welding and Other Means” (*J. of Aircraft*, 1983, p. 552-556).] Comente as características.

5434	4948	4521	4570	4990	5702	5241
5112	5015	4659	4806	4637	5670	4381
4820	5043	4886	4599	5288	5299	4848
5378	5260	5055	5828	5218	4859	4780
5027	5008	4609	4772	5133	5095	4618
4848	5089	5518	5333	5164	5342	5069
4755	4925	5001	4803	4951	5679	5256
5207	5621	4918	5138	4786	4500	5461
5049	4974	4592	4173	5296	4965	5170
4740	5173	4568	5653	5078	4900	4968
5248	5245	4723	5275	5419	5205	4452
5227	5555	5388	5498	4681	5076	4774
4931	4493	5309	5582	4308	4823	4417
5364	5640	5069	5188	5764	5273	5042
5189	4986					

25. A transformação de valores de dados por meio de uma função matemática, como  $\sqrt{x}$  ou  $1/x$ , normalmente resulta em um conjunto de números com “melhores” propriedades estatísticas do que os dados originais. Em particular, é possível encontrar uma função para a qual o histograma dos valores transformados seja mais simétrico (ou, melhor ainda, mais próximo de uma curva normal) do que os dados originais. Como exemplo, o artigo “Time Lapse Cinematographic Analysis of BerylliumLung Fibroblast Interactions” (*Environ. Research*, 1983, p. 34-43) relatou os resultados de experimentos projetados para estudar o comportamento de algumas células que foram expostas ao berílio. Uma característica importante de tal célula individual é seu tempo de interdivisão (IDT). Os IDTs foram determinados para um grande número de células em condições de exposição (tratamento) e não-exposição (controle). Os autores do artigo usaram uma transformação logarítmica, isto é, valor transformado = log (valor original). Considere os seguintes dados representativos de IDT:

IDT	$\log_{10}(\text{IDT})$	IDT	$\log_{10}(\text{IDT})$	IDT	$\log_{10}(\text{IDT})$
28,1	1,45	60,1	1,78	21,0	1,32
31,2	1,49	23,7	1,37	22,3	1,35
13,7	1,14	18,6	1,27	15,5	1,19
46,0	1,66	21,4	1,33	36,3	1,56
25,8	1,41	26,6	1,42	19,1	1,28
16,8	1,23	26,2	1,42	38,4	1,58
34,8	1,54	32,0	1,51	72,8	1,86
62,3	1,79	43,5	1,64	48,9	1,69
28,0	1,45	17,4	1,24	21,4	1,33
17,9	1,25	38,8	1,59	20,7	1,32
19,5	1,29	30,6	1,49	57,3	1,76
21,1	1,32	55,6	1,75	40,9	1,61
31,9	1,50	25,5	1,41		
28,9	1,46	52,1	1,72		

Use os intervalos de classes 10–<20, 20–<30... para construir um histograma dos dados originais. Use os intervalos 1,1–<1,2, 1,2–<1,3, ... para fazer o mesmo para os dados transformados. Qual é o efeito da transformação?

26. O índice de céu claro foi determinado para o céu de Bagdá, compreendendo cada um dos 365 dias de um dado ano (“Contribution to the Study of the Solar Radiation Climate of the Baghdad Environment”, *Solar Energy*, 1990, p. 7-12). A tabela a seguir fornece os resultados.

Classe	Frequência
0,15–<0,25	8
0,25–<0,35	14
0,35–<0,45	28
0,45–<0,50	24
0,50–<0,55	39
0,55–<0,60	51
0,60–<0,65	106
0,65–<0,70	84
0,70–<0,75	11

- a. Determine as frequências relativas e desenhe o histograma correspondente.
- b. Dias nublados são aqueles com o índice de céu limpo inferior a 0,35. Em que porcentagem dos dias o céu esteve nublado?
- c. Dias de céu claro são aqueles para os quais o índice é no mínimo 0,65. Em que porcentagem dos dias o céu esteve limpo?

27. O artigo “Study on the Life Distribution of Microdrills” (*J. of Engr. Manufacture*, 2002: 301-305) relatou as observações a seguir, relacionadas em ordem crescente, da vida útil das brocas (número de furos que uma broca faz antes de quebrar), quando os furos são feitos em uma determinada liga de bronze.

11 14 20 23 31 36 39 44 47 50  
 59 61 65 67 68 71 74 76 78 79  
 81 84 85 89 91 93 96 99 101 104  
 105 105 112 118 123 136 139 141 148 158  
 161 168 184 206 248 263 289 322 388 513

- a. Por que uma distribuição de frequência não pode ter por base os intervalos de classe 0–50, 50–100, 100–150 e assim por diante?
- b. Construa uma distribuição de frequência e um histograma dos dados usando limites de classes 0, 50, 100, ... e então faça comentários sobre as características interessantes.
- c. Construa uma distribuição de frequência e um histograma dos logaritmos naturais relacionados às observações de vida útil e comente as características interessantes.
- d. Que proporção das observações de vida útil dessa amostra é inferior a 100? Que proporção das observações é igual ou maior que 200?

28. Construa um gráfico de pontos para a série de dados anexa. Os dados são mensais e foram obtidos durante o período de 1985-1989. Cada valor é a radiação solar média na faixa 385-530 nm como porcentagem da radiação total (“Global Energy in the Different Spectral Bands at Dhahran, Saudi Arabia,” *J. Solar Energy Engr.*, 1991, p. 290-294). Comente sobre algumas características interessantes dos dados.

20,9 19,6 20,4 20,3 20,8 20,6 20,5 20,4  
 19,9 19,8 19,5 20,2 16,5 18,3 18,7 19,6  
 20,0 20,0 19,5 19,6 19,1 18,8 18,3 17,6  
 17,2 17,8 18,7 19,0 19,0 18,6 18,8 19,0  
 18,5 18,3 17,5 16,9 17,0 17,8 18,1 18,8  
 18,9 18,9 19,1 18,8 18,4 17,8 17,0 16,8  
 17,9 18,4 19,0 19,4 19,7 19,5 19,5 19,5  
 19,0 18,7 18,1 17,9

29. Considere os dados a seguir sobre os tipos de queixas de saúde (J = inflamação de articulações, F = fadiga, B = dor nas costas, M = fadiga muscular, C = tosse, N = irritação nasal/coriza, O = outros) feitas por agricultores. Obtenha as frequências e as frequências relativas das diversas categorias e desenhe um histograma. (Os dados são consistentes com as porcentagens fornecidas no artigo “Physiological Effects of Work Stress and Pesticide Exposure in Tree Planting by British Columbia Silviculture Workers,” *Ergonomics*, 1993, p. 951-961.)

O O N J C F B B F O J O O M  
 O F F O O N O N J F J B O C  
 J O J J F N O B M O J M O B  
 O F J O O B N C O O O M B F  
 J O F N

30. Um **Diagrama de Pareto** é uma variação de um histograma para dados categorizados resultantes de um estudo de controle de qualidade. Cada categoria representa um tipo diferente de não-conformidade de produto ou problema de produção. As categorias são ordenadas de forma que aquela com maior frequência seja exibida na extremidade esquerda, seguida pela categoria com a segunda maior frequência e assim por diante. Suponha que as informações a seguir tenham sido obtidas sobre não-conformidades em pacotes de circuitos: componentes com falha, 126; componentes incorretos, 210; soldas insuficientes, 67; soldas em excesso, 54; falta de componentes, 131. Construa um Diagrama de Pareto.

31. A **frequência acumulada** e a frequência relativa acumulada de um determinado intervalo de classe são a soma das frequências e frequências relativas, respectivamente, desse intervalo e de todos os intervalos abaixo dele. Se, por exemplo, houver quatro intervalos com frequências 9, 16, 13 e 12, as frequências acumuladas serão 9, 25, 38 e 50 e as frequências relativas acumuladas serão 0,18, 0,50, 0,76 e 1,00. Calcule as frequências acumuladas e as frequências relativas acumuladas para os dados do Exercício 24.

32. Uma carga de incêndio (MJ/m<sup>2</sup>) é a energia térmica que pode ser liberada por metro quadrado de área de piso pela combustão de seu conteúdo e da estrutura em si. O artigo “Fire Loads in Office Buildings” (*J. of Structural Engr.*, 1997, p. 365-368) forneceu as seguintes porcentagens acumuladas (lidas de um gráfico) relativas a cargas de incêndio em uma amostra de 388 salas:

Valor	0	150	300	450	600
% Acumulada	0	19,3	37,6	62,7	77,5
Valor	750	900	1050	1200	1350
% Acumulada	87,2	93,8	95,7	98,6	99,1
Valor	1500	1650	1800	1950	
% Acumulada	99,5	99,6	99,8	100,0	

- a. Construa um histograma de frequência relativa e comente as características interessantes.
- b. Que proporção das cargas de incêndio é inferior a 600? Maior ou igual a 1200?
- c. Que proporção das cargas está entre 600 e 1200?

### 1.3 | Medidas de localização

Os resumos visuais de dados são excelentes ferramentas para obter impressões e idéias iniciais. Uma análise mais formal de dados frequentemente exige o cálculo e a interpretação de medidas-resumo numéricas simples. Isto é, a partir dos dados, tentamos extrair diversos números simples, que servem para caracterizar o conjunto de

dados e indicar algumas informações consideráveis. Nossa preocupação principal será com os dados numéricos. Alguns comentários sobre dados categorizados serão apresentados no final da seção.

Suponha, então, que nosso conjunto de dados é do formato  $x_1, x_2, \dots, x_n$ , onde cada  $x_i$  é um número. Que características de tal conjunto de números são de maior interesse e merecem ênfase? Uma característica importante de um conjunto de números é sua localização e, em particular, seu centro. Esta seção apresenta métodos de descrição da localização de um conjunto de dados. Na Seção 1.4, apresentaremos os métodos de medida da dispersão de um conjunto de números.

### A média

Para um determinado conjunto de números  $x_1, x_2, \dots, x_n$ , a medida mais familiar e útil do centro é a *média* do conjunto. Como quase sempre temos os vários  $x_i$  constituindo uma amostra, freqüentemente chamaremos a média aritmética de *média amostral* e a representaremos por  $\bar{x}$ .

**DEFINIÇÃO**

A **média amostral**  $\bar{x}$  das observações  $x_1, x_2, \dots, x_n$ , é dada por

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}$$

O numerador de  $\bar{x}$  pode ser escrito mais informalmente como  $\sum x_i$ , onde a soma se dá sobre todas as observações da amostra.

Para informar  $\bar{x}$ , recomendamos o uso de precisão decimal de um dígito a mais do que a precisão dos  $x_i$ . Dessa forma, se as observações forem distâncias de parada com  $x_1 = 125$ ,  $x_2 = 131$  e assim por diante, podemos ter  $\bar{x} = 127,3$  pés.

#### Exemplo 1.13

As trincas em aço e ferro causadas por fadiga de corrosão cáustica foram estudadas em decorrência de falhas em rebites de caldeiras de aço e em rotores a vapor. Considere as observações a seguir sobre  $x =$  comprimento da trinca ( $\mu\text{m}$ ) como resultado de testes de fadiga por corrosão devido a cargas constantes em amostras de barras de tração lisas durante um período de tempo fixo. (Os dados são consistentes com um histograma e as quantidades-resumo do artigo “On the Role of Phosphorus in the Caustic Stress Corrosion Cracking of Low Alloy Steels”, *Corrosion Science*, 1989: 53-68.)

$x_1 = 16,1$   $x_2 = 9,6$   $x_3 = 24,9$   $x_4 = 20,4$   $x_5 = 12,7$   $x_6 = 21,2$   $x_7 = 30,2$   
 $x_8 = 25,8$   $x_9 = 18,5$   $x_{10} = 10,3$   $x_{11} = 25,3$   $x_{12} = 14,0$   $x_{13} = 27,1$   $x_{14} = 45,0$   
 $x_{15} = 23,3$   $x_{16} = 24,2$   $x_{17} = 14,6$   $x_{18} = 8,9$   $x_{19} = 32,4$   $x_{20} = 11,8$   $x_{21} = 28,5$

A Figura 1.14 mostra um diagrama de caule e folha dos dados. Um comprimento de trinca no início da faixa dos 20 parece ser “típica”.

0H	96	89				
1L	27	03	40	46	18	
1H	61	85				
2L	49	04	12	33	42	
2H	58	53	71	85		
3L	02	24				
3H						
4L						
4H		50				

Caule: dígito das dezenas  
 Folha: dígito das unidades e das dezenas

**Figura 1.14** Um diagrama de caule e folha dos dados dos comprimentos de trincas



Sendo  $\sum x_i = 444,8$ , a média amostral é

$$\bar{x} = \frac{444,8}{21} = 21,18$$

um valor consistente com as informações ilustradas pelo diagrama de caule e folha. ■

Uma interpretação física de  $\bar{x}$  demonstra como ela mede a localização (centro) de uma amostra. Imagine desenhar e definir a escala em um eixo horizontal e depois represente cada observação da amostra por um peso de uma libra colocado no ponto correspondente no eixo. O único ponto em que pode ser colocado um apoio para equilibrar o sistema de pesos é o correspondente ao valor de  $\bar{x}$  (veja a Figura 1.15).

Da mesma forma que  $\bar{x}$  representa o valor médio das observações de uma amostra, a média de todos os valores da população pode ser calculada. Essa média é denominada **média da população** e é representada pela letra grega  $\mu$ . Quando houver  $N$  valores na população (uma população finita),  $\mu = (\text{somatória dos } N \text{ valores da população})/N$ . Nos capítulos 3 e 4, forneceremos uma definição mais geral de  $\mu$  que se aplica a populações finitas e (conceitualmente) infinitas. Da mesma forma que  $\bar{x}$  é uma medida de localização de amostra importante e interessante,  $\mu$  é uma característica interessante e importante (freqüentemente a mais importante) de uma população. Nos capítulos sobre inferência estatística, apresentaremos métodos com base na média amostral para obtenção de conclusões sobre a média de uma população. Por exemplo: podemos usar a média amostral  $\bar{x} = 21,18$  calculada no Exemplo 1.13 como uma *estimativa de ponto* (um único número que é o “melhor” palpite) de  $\mu$ , o comprimento médio verdadeiro de todas as amostras tratadas como descrito.

A média sofre de uma deficiência que a torna uma medida de centro inadequada sob algumas circunstâncias: seu valor pode ser bastante afetado pela presença de um único *outlier* (uma observação incomumente grande ou pequena). No Exemplo 1.13, o valor  $x_{14} = 45,0$  obviamente é um *outlier*. Sem esta observação,  $\bar{x} = 399,8/20 = 19,99$ , o *outlier* aumenta a média em mais de  $1 \mu\text{m}$ . Se a observação  $45,0 \mu\text{m}$  fosse substituída pelo valor catastrófico de  $295,0 \mu\text{m}$ , um *outlier* realmente extremo, então  $\bar{x} = 694,8/21 = 33,09$ , que é maior que todas as observações, exceto uma.

Uma amostra de salários normalmente produz alguns poucos valores aberrantes (dos sortudos que possuem um salário astronômico) e o uso do salário médio como medida de localização freqüentemente será ilusório. Esses exemplos sugerem que procuremos uma medida menos sensível a valores fora da faixa que  $\bar{x}$ , assim, proporemos uma momentaneamente. Entretanto, apesar de  $\bar{x}$  ter essa falha potencial, ela ainda é a medida mais usada, em grande parte porque há muitas populações para as quais um *outlier* extremo na amostra seria altamente improvável. Ao obter uma amostra de uma tal população (a população normal ou em forma de sino, é o exemplo mais importante), a média amostral tenderá a ser estável e muito representativa.



Figura 1.15 A média como ponto de equilíbrio de um sistema de pesos

## A mediana

A palavra *mediana* é sinônimo de “metade” e a mediana amostral é o valor do meio quando as observações são ordenadas da menor para a maior. Quando as observações estiverem representadas por  $x_1, \dots, x_n$ , usaremos o símbolo  $\tilde{x}$  para representar a mediana amostral.

**DEFINIÇÃO**

A **mediana amostral** é obtida pela ordenação das  $n$  observações da menor para a maior (com os valores repetidos incluídos, de forma que cada observação da amostra seja exibida na lista ordenada). Assim,

$$\tilde{x} = \begin{cases} \text{O único} \\ \text{valor} \\ \text{médio se } n \\ \text{for ímpar} & = \left(\frac{n+1}{2}\right)\text{ésimo valor ordenado} \\ \\ \text{A média} \\ \text{dos dois} \\ \text{valores} \\ \text{médios se } n \\ \text{for par} & = \text{média dos valores ordenados } \left(\frac{n}{2}\right) \text{ e } \left(\frac{n}{2} + 1\right) \end{cases}$$

**Exemplo 1.14**

O risco de desenvolvimento de deficiência de ferro é especialmente alto durante a gravidez. O problema na detecção dessa deficiência é que alguns métodos de determinação de nível de ferro podem ser afetados pelo próprio estado de gravidez. Considere os dados a seguir sobre a concentração do receptor de transferrina de uma amostra de mulheres com evidências laboratoriais de uma visível anemia por deficiência de ferro (“Serum Transferrin Receptor for the Detection of Iron Deficiency in Pregnancy,” *Amer. J. of Clinical Nutrition*, 1991: p. 1077-1081):

$$\begin{matrix} x_1 = 15,2 & x_2 = 9,3 & x_3 = 7,6 & x_4 = 11,9 & x_5 = 10,4 & x_6 = 9,7 \\ x_7 = 20,4 & x_8 = 9,4 & x_9 = 11,5 & x_{10} = 16,2 & x_{11} = 9,4 & x_{12} = 8,3 \end{matrix}$$

A lista dos valores ordenados é

$$7,6 \quad 8,3 \quad 9,3 \quad 9,4 \quad 9,4 \quad 9,7 \quad 10,4 \quad 11,5 \quad 11,9 \quad 15,2 \quad 16,2 \quad 20,4$$

Como  $n = 12$  é par, tiramos a média  $n/2 =$  do sexto e sétimo valores ordenados:

$$\text{mediana amostral} = \frac{9,7 + 10,4}{2} = 10,05$$

Observe que, se a maior observação, 20,4, não tivesse aparecido na amostra, a mediana amostral resultante para as  $n = 11$  observações teria sido o único valor médio, 9,7 (o  $(n + 1)/2 =$  sexto valor ordenado). A média amostral é  $\bar{x} = \sum x_i/n = 139,3/12 = 11,61$ , que é um pouco maior que a mediana, por causa dos *outliers*, 15,2, 16,2 e 20,4. ■

Os dados do Exemplo 1.14 ilustram uma propriedade importante de  $\tilde{x}$  em comparação com  $\bar{x}$ : a mediana amostral é muito insensível a muitos valores extremamente pequenos ou extremamente grandes. Se, por exemplo, aumentássemos os dois maiores  $x_i$  de 16,2 e 20,4 para 26,2 e 30,4, respectivamente,  $\tilde{x}$  não seria afetado. Dessa forma, no tratamento de valores de dados fora da faixa,  $\bar{x}$  e  $\tilde{x}$  são extremidades opostas de um espectro:  $\bar{x}$  é sensível mesmo a um único valor, enquanto  $\tilde{x}$  é insensível a um grande número de valores fora da faixa.

Como os valores grandes na amostra do Exemplo 1.14 afetam  $\bar{x}$  mais que  $\tilde{x}$ ,  $\tilde{x} < \bar{x}$  para esses dados. Apesar de  $\bar{x}$  e  $\tilde{x}$  fornecerem uma medida para o centro da amostra em um conjunto de dados, eles em geral não serão iguais, porque enfocam diferentes aspectos da amostra.

De forma análoga,  $\tilde{x}$  como valor médio na amostra é o valor médio da população, a **mediana da população**, representada por  $\tilde{\mu}$ . Como acontece com  $\bar{x}$  e  $\mu$ , podemos considerar o uso da mediana amostral  $\tilde{x}$  para fazer inferências de  $\tilde{\mu}$ . No Exemplo 1.14, podemos usar  $\tilde{x} = 10,05$  como estimativa da concentração da mediana em toda a população a partir da qual a amostra foi selecionada. Uma mediana normalmente é usada para descrever dados de salários ou rendimentos (porque ela não é influenciada por alguns grandes salários). Se a

mediana de uma amostra dos salários de engenheiros fosse  $\tilde{x} = \$ 66.416$ , poderíamos usá-la como base para concluir que o salário mediano dos engenheiros excede \$ 60.000.

A média da população  $\mu$  e a mediana  $\tilde{\mu}$  normalmente não serão idênticas. Se a distribuição da população tiver desvio positivo ou negativo, conforme ilustrado na Figura 1.16, então  $\mu \neq \tilde{\mu}$ . Quando esse for o caso, ao fazer inferências, devemos primeiro decidir quais características das populações são de maior interesse e então proceder de acordo.

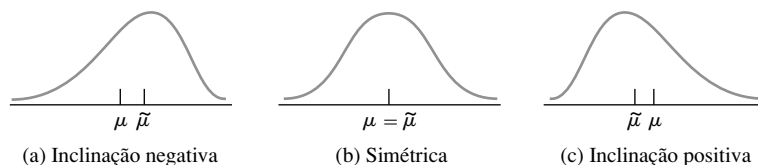


Figura 1.16 Três formatos diferentes para uma distribuição de população

## Outras medidas de localização

### Quartis, Percentis e Médias Aparadas

A mediana (de população ou amostra) divide o conjunto de dados em duas partes de mesmo tamanho. Para obter melhores medidas de localização, podemos dividir os dados em mais de duas partes. Grosso modo, os quartis dividem o conjunto em quatro partes iguais, sendo que as observações acima do terceiro quartil constituem o quarto superior do conjunto de dados, o segundo quartil é idêntico à mediana e o primeiro quartil separa o quarto inferior dos três quartos superiores. De forma similar, um conjunto de dados (amostra ou população) pode ser dividido mais detalhadamente usando percentis; o 99º percentil separa o 1% superior do restante, e assim por diante. A menos que o número de observações seja um múltiplo de 100, recomenda-se cuidado na utilização de percentis. Usaremos percentis no Capítulo 4 com alguns modelos de populações infinitas, de forma que adiaremos a discussão até lá.

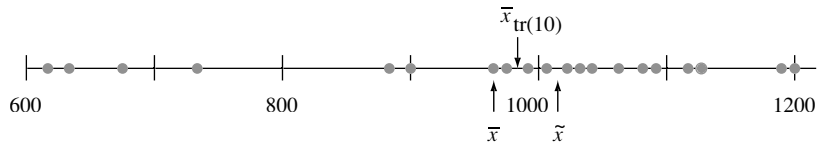
A média amostral e a mediana amostral são influenciadas por valores fora da faixa de uma forma bastante diferente: muito para a média e nada para a mediana. Como o comportamento extremo dos dois valores é indesejável, consideraremos medidas alternativas que não sejam tão sensíveis quanto  $\bar{x}$  e nem tão insensíveis como  $\tilde{x}$ . Para determinar essas alternativas, observe que  $\bar{x}$  e  $\tilde{x}$  são extremidades opostas da mesma “família” de medidas. Após o conjunto de dados ser ordenado,  $\tilde{x}$  é calculado desprezando-se todos os valores possíveis em cada extremidade sem eliminar nada (deixando apenas um ou dois valores centrais) e obtendo a média do que restou. Por outro lado, para calcular  $\bar{x}$ , nada é desprezado antes de se obter a média. Para fazer uma comparação, a média envolve desprezar 0% de cada extremidade da amostra, enquanto, para a mediana, o máximo possível é desprezado de cada extremidade. Uma **média aparada** é algo intermediário entre  $\bar{x}$  e  $\tilde{x}$ . Uma média aparada de 10%, por exemplo, seria calculada eliminando-se os 10% superiores e os 10% inferiores da amostra, obtendo-se, então, a média do restante.

#### Exemplo 1.15

Considere as 20 observações a seguir, ordenadas da menor para a maior, cada uma representando a vida útil (em horas) de um determinado tipo de lâmpada incandescente:

612	623	666	744	883	898	964	970	983	1003
1016	1022	1029	1058	1085	1088	1122	1135	1197	1201

A média das 20 observações é  $\bar{x} = 965,0$  e  $\tilde{x} = 1009,5$ . A média aparada de 10% é obtida pela exclusão das duas menores observações (612 e 623) e as duas maiores (1197 e 1201) seguida do cálculo da média dos 16 valores restantes, para obter  $\bar{x}_{tr(10)} = 979,1$ . O efeito de truncar a média aqui é produzir um “valor central” ligeiramente acima da média ( $\bar{x}$  é trazido para baixo por alguns poucos valores de vida útil) e ainda consideravelmente abaixo da mediana. De forma similar, a média aparada de 20% faz uma média dos 12 valores do meio para obter  $\bar{x}_{tr(20)} = 999,9$ , mais perto ainda da mediana. (Veja a Figura 1.17.)



**Figura 1.17** Gráfico de pontos de vida útil (em horas) de lâmpadas incandescentes ■

Geralmente, o uso da média aparada com proporção de aparagem moderada (entre 5% e 25%) produzirá uma medida que não é nem tão sensível a *outliers* como a média nem tão insensível quanto a mediana. Por esse motivo, as médias truncadas têm sido objeto de crescente atenção dos estatísticos para propósitos descritivos e inferenciais. Mais será dito sobre médias aparadas quando a estimativa por pontos for discutida no Capítulo 6. Finalmente, se a proporção de aparagem for representada por  $\alpha$  e  $n\alpha$  não for inteiro, não será óbvio como calcular a média aparada 100 $\alpha$ %. Por exemplo: se  $\alpha = 0,10$  (10%) e  $n = 22$ , então  $n\alpha = (22)(0,10) = 2,2$  e não é possível aparar 2,2 observações de cada extremidade da amostra ordenada. Nesse caso, a média aparada de 10% seria obtida primeiro com a retirada das duas observações de cada extremidade e pelo cálculo de  $\bar{x}_r$ , seguida pela retirada de três observações de cada extremidade e pelo cálculo de  $\bar{x}_t$ , e então pela interpolação dos dois valores para obtenção de  $\bar{x}_{tr(10)}$ .

### Dados categorizados e proporção de amostras

Quando os dados são categorizados, uma distribuição de freqüência ou distribuição de freqüência relativa fornece um resumo tabular eficiente dos dados. Os indicadores numéricos naturais são, nessa situação, as freqüências individuais e as freqüências relativas. Por exemplo: se for feita uma pesquisa com indivíduos que possuem aparelhos de som para estudar a preferência de marca, cada indivíduo da amostra identificaria a marca do aparelho que possui. A partir disso poderíamos contar as pessoas que possuem aparelhos Sony, Pioneer, Marantz, entre outros. Considere a obtenção de uma amostra de uma população dicotômica, isto é, que consista em apenas duas categorias (votou ou não votou na eleição passada ou possui ou não um aparelho de som etc.). Se fizermos  $x$  representar o número da amostra na categoria 1, o número na categoria 2 será  $n - x$ . A freqüência relativa ou *proporção amostral* da categoria 1 será  $x/n$  e a proporção amostral da categoria 2 será  $1 - x/n$ . Vamos representar uma resposta da categoria 1 por 1 e uma resposta da categoria 2 por 0. Uma amostra de tamanho  $n = 10$  pode então resultar em 1, 1, 0, 1, 1, 1, 0, 0, 1, 1. A média dessa amostra numérica é (já que o número de ocorrências do número 1 =  $x = 7$ )

$$\frac{x_1 + \dots + x_n}{n} = \frac{1 + 1 + 0 + \dots + 1 + 1}{10} = \frac{7}{10} = \frac{x}{n} = \text{proporção amostral}$$

Esse resultado pode ser generalizado e resumido conforme segue: *Se em uma situação de dados categorizados focarmos a atenção em uma determinada categoria e codificarmos os resultados da amostra de forma que 1 seja registrado como um indivíduo da categoria e 0 para um indivíduo fora dela, a proporção amostral de indivíduos da categoria será a média amostral da seqüência de 1s e 0s.* Assim, uma média amostral pode ser usada para resumir os resultados de uma amostra categorizada. Essas observações também se aplicam a situações em que as categorias são definidas por valores agrupados em uma amostra ou população numérica (por exemplo: podemos querer saber se os indivíduos possuem seu automóvel atual há pelo menos cinco anos em vez de estudarmos o tempo exato de posse).

De forma análoga à proporção amostral  $x/n$  de indivíduos que estão em uma determinada categoria, representemos por  $p$  a proporção dos indivíduos da população inteira que pertencem à categoria. Como acontece com  $x/n$ ,  $p$  é uma quantidade entre 0 e 1 e, enquanto  $x/n$  é uma característica da amostra,  $p$  é uma característica da população. A relação entre os dois é semelhante à relação entre  $\tilde{x}$  e  $\tilde{\mu}$  e entre  $\bar{x}$  e  $\mu$ . Em particular, usaremos  $x/n$  para fazer inferências sobre  $p$ . Se, por exemplo, uma amostra de 100 proprietários de carros revelar que 22 possuem seus carros há pelo menos 5 anos, podemos usar  $22/100 = 0,22$  como uma estimativa pontual da proporção

de todos os proprietários que possuem o carro há pelo menos 5 anos. Estudaremos as propriedades de  $x/n$  como um estimador de  $p$  e veremos como  $x/n$  pode ser usado para responder a outras questões inferenciais. Com  $k$  categorias ( $k > 2$ ), podemos usar as proporções amostrais de  $k$  para responder a questões sobre as proporções da população  $p_1, \dots, p_k$ .

## Exercícios | Seção 1.3 (33–43)

**33.** O artigo “The Pedaling Technique of Elite Endurance Cyclists” (*Int. J. of Sport Biomechanics*, 1991, p. 29-53) relatou os dados a seguir sobre a potência de uma única perna de um ciclista em alta carga de trabalho:

244 191 160 187 180 176 174  
205 211 183 211 180 194 200

- Calcule e interprete a média e a mediana amostrais.
- Suponha que a primeira observação tenha sido 204 em vez de 244. Como a média e a mediana seriam afetadas?
- Calcule uma média aparada, eliminando a maior e a menor observações da amostra. Qual é a porcentagem de truncamento correspondente?
- O artigo também relatou valores sobre a potência de uma única perna para uma carga de trabalho baixa. A média amostral de  $n = 13$  observações foi  $\bar{x} = 119,8$  (na verdade 119,7692) e a 14ª observação, um tipo de *outlier*, foi 159. Qual é o valor de  $\bar{x}$  para a amostra toda?

**34.** Considere as observações a seguir sobre resistência ao cisalhamento (MPa) de uma junta soldada de uma determinada forma (de um gráfico no artigo “Diffusion of Silicon Nitride to Austenitic Stainless Steel without Interlayers,” *Metallurgical Trans.*, 1993, p. 1835-1843):

22,2 40,4 16,4 73,7 36,6 109,9  
30,0 4,4 33,1 66,7 81,5

- Determine o valor da média amostral.
- Determine o valor da mediana amostral. Por que esse valor é tão diferente da média?
- Calcule a média aparada, excluindo a menor e a maior observações. Qual é a porcentagem de truncamento correspondente? Como o valor de  $\bar{x}_r$  se compara à média e à mediana?

**35.** A pressão mínima de injeção (psi) em amostras de moldagem por injeção de milho de alta amilose foi determinada para oito amostras diferentes (pressões mais altas correspondem a maior dificuldade de processamento), resultando nas observações a seguir (de “Thermoplastic Starch Blends with a Polyethylene-Co-Vinyl Alcohol: Processability and Physical Properties,” *Polymer Engr. and Science*, 1994, p. 17-23):

15,0 13,0 18,0 14,5 12,0 11,0 8,9 8,0

- Determine os valores da média amostral, mediana amostral e média aparada de 12,5% e os compare.

- Em quanto a menor observação da amostra, atualmente 8,0, pode ser aumentada sem afetar o valor da mediana amostral?

- Suponha que desejemos que os valores da média e da mediana amostrais sejam expressos em quilogramas por polegada quadrada (ksi) em vez de psi. É necessário mudar as unidades de cada observação ou os valores calculados na parte (a) podem ser usados diretamente? *Dica:*  $1\text{kg} = 2,2\text{lb}$ .

**36.** Vinte e seis trabalhadores de plataformas de petróleo *offshore* participaram de um exercício de fuga simulado, resultando nos dados a seguir (em segundos) para concluir a fuga (“Oxygen Consumption and Ventilation During Escape from an Offshore Platform,” *Ergonomics*, 1997, p. 281-292):

389 356 359 363 375 424 325 394 402  
373 373 370 364 366 364 325 339 393  
392 369 374 359 356 403 334 397

- Construa um diagrama de caule e folha dos dados. Como ele sugere que a média e mediana serão comparadas?
- Calcule os valores da média e da mediana amostrais. *Dica:*  $\sum x_i = 9638$ .
- Em quanto o maior tempo, atualmente 424, pode ser aumentado sem afetar o valor da mediana amostral? Em quanto esse valor pode ser diminuído sem afetar o valor da mediana amostral?
- Quais são os valores de  $\bar{x}$  e  $\bar{x}$  quando as observações são reexpressas em minutos?

**37.** O artigo “Snow Cover and Temperature Relationships in North America and Eurasia” (*J. Climate and Applied Meteorology*, 1983, p. 460-469) usou técnicas estatísticas para relacionar a quantidade de cobertura de neve em cada continente com a temperatura média continental. Os dados apresentados incluíram as 10 observações a seguir sobre a cobertura de neve, em outubro, na Eurásia, durante 1970 e 1979 (em milhões de  $\text{km}^2$ ):

6,5 12,0 14,9 10,0 10,7 7,9 21,9 12,5 14,5 9,2

O que você descreveria como valor característico ou representativo da cobertura de neve em outubro para esse período e o que o levou a essa escolha?

**38.** Os valores de pressão sanguínea freqüentemente são informados com aproximação de 5 mmHg (100, 105, 110 etc.). Suponha que os valores reais (sem aproximação) de pressão sanguínea de nove indivíduos selecionados aleatoriamente sejam

118,6 127,4 138,4 130,0 113,7 122,0 108,3 131,5  
133,2

- a. Qual é a mediana dos valores de pressão sanguínea informados?
- b. Suponha que a pressão sanguínea do segundo indivíduo seja 127,6 em vez de 127,4 (uma pequena alteração em um único valor). Como isso afeta a mediana dos valores informados? O que isso diz sobre a sensibilidade da mediana ao arredondamento ou agrupamento dos dados?

39. A propagação de trincas por fadiga em diversas peças de aeronaves tem sido objeto de muitos estudos nos últimos anos. Os dados a seguir consistem dos tempos de propagação (horas de vôo/10<sup>4</sup>) para atingir um determinado tamanho de trinca em furos de fixadores propostos para uso em aeronaves militares (“Statistical Crack Propagation in Fastener Holes under Spectrum Loading,” *J. Aircraft*, 1983, p. 1028-1032):

0,736 0,863 0,865 0,913 0,915 0,937 0,983 1,007  
1,011 1,064 1,109 1,132 1,140 1,153 1,253 1,394

- a. Calcule e compare os valores da média e da mediana amostrais.
  - b. Em quanto a maior observação da amostra pode ser diminuída sem afetar o valor da mediana?
40. Calcule a mediana amostral, a média aparada de 25%, a média aparada de 10% e a média amostral para os dados de concentração fornecidos no Exercício 27 e compare essas medidas.
41. Uma amostra de  $n = 10$  automóveis foi selecionada e cada um deles foi sujeito a um teste de colisão a 5 mph. Representando um carro sem danos visíveis por S (de

sucesso) e um carro com danos por F, os resultados são os seguintes:

S S F S S S F F S S

- a. Qual é o valor da proporção amostral de sucessos  $x/n$ ?
  - b. Substitua cada S por 1 e cada F por 0. Calcule então  $\bar{x}$  para essa amostra codificada numericamente. Como  $\bar{x}$  pode ser comparado a  $x/n$ ?
  - c. Suponha que se decidiu incluir mais 15 carros no experimento. Quantos deles teriam de ser S para fornecer  $x/n = 0,80$  para a amostra de 25 carros?
42. a. Se uma constante  $c$  é adicionada a cada  $x_i$  de uma amostra, resultando em  $y_i = x_i + c$ , como a média e a mediana amostrais dos  $y_i$  se relacionam com a média e a mediana dos  $x_i$ ? Verifique suas hipóteses.
- b. Se cada  $x_i$  é multiplicado por uma constante  $c$ , resultando em  $y_i = cx_i$ , responda à questão da parte (a). Verifique novamente suas hipóteses.
43. Um experimento para estudar a vida útil (em horas) de certo tipo de componente consiste em colocar dez componentes em operação e observá-los por 100 horas. Oito dos componentes apresentaram falhas nesse período e esses valores de vida útil foram registrados. Represente os tempos de vida útil dos dois componentes que ainda funcionam após 100 horas por 100+. As observações resultantes foram

48 79 100+ 35 92 86 57 100+ 17 29

Que medidas de tendência central discutidas nesta seção podem ser calculadas e quais são os valores dessas medidas? (Nota: Os dados deste experimento são “censurados pela direita”.)

## 1.4 | Medidas de Dispersão

Informar apenas a medida de tendência central fornece apenas informações parciais sobre um conjunto de dados ou uma distribuição. Diferentes amostras ou populações podem ter medidas de tendência central idênticas e apresentar diferenças entre si em outros aspectos importantes. A Figura 1.18 apresenta gráficos de pontos de três amostras com a mesma média e a mesma mediana, mas com dispersões diferentes ao redor do centro. A primeira amostra é a que apresenta maior dispersão, a terceira a menor e a segunda é intermediária em relação a elas.

### Medidas de dispersão para dados amostrais

A medida de dispersão mais simples de uma amostra é a **amplitude**, a diferença entre o maior e o menor valores da amostra. Observe que o valor da amplitude da amostra 1 da Figura 1.18 é muito maior do que o da amostra 3, o que reflete maior dispersão na primeira amostra do que na terceira. Um defeito da amplitude, entretanto, é que ela depende apenas das duas observações mais extremas e não considera as posições dos  $n - 2$  valores restantes. As amostras 1 e 2 na Figura 1.18 possuem amplitudes idênticas mas, se levarmos em conta as observações entre os dois extremos, há muito menos dispersão na segunda amostra do que na primeira.

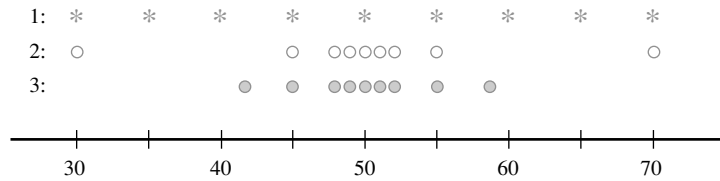


Figura 1.18 Amostras com medidas centrais idênticas mas com variabilidade diferente

Nossa principal medida de dispersão envolve os **desvios em relação à média**,  $x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_n - \bar{x}$ . Ou seja, os desvios da média são obtidos pela subtração de  $\bar{x}$  de cada uma das  $n$  observações da amostra. Um desvio será positivo se a observação for maior que a média (à direita da média no eixo das medidas) e negativo se a observação for menor que a média. Se todos os desvios forem pequenos em magnitude, todos os  $x_i$  estarão próximos à média e haverá pouca dispersão. Por outro lado, se alguns desvios forem grandes, alguns  $x_i$  estarão distantes de  $\bar{x}$ , indicando maior dispersão. Uma forma simples de combinar os desvios em uma única quantidade é calcular a sua média (somá-los e dividi-los por  $n$ ). Infelizmente, há um problema grave com essa sugestão:

$$\text{somatória dos desvios} = \sum_{i=1}^n (x_i - \bar{x}) = 0$$

de forma que o desvio será sempre zero. A demonstração usa diversas regras-padrão de somatória e o fato de que  $\sum \bar{x} = \bar{x} + \bar{x} + \dots + \bar{x} = n\bar{x}$ :

$$\sum (x_i - \bar{x}) = \sum x_i - \sum \bar{x} = \sum x_i - n\bar{x} = \sum x_i - n\left(\frac{1}{n} \sum x_i\right) = 0$$

Como podemos alterar os desvios para quantidades não-negativas de forma que os desvios positivos e negativos não se cancelem quando forem somados? Uma possibilidade é trabalhar com os valores absolutos e calcular o desvio médio absoluto  $\sum |x_i - \bar{x}|/n$ . Como a operação em valor absoluto conduz a diversas dificuldades teóricas, considere os quadrados dos desvios  $(x_1 - \bar{x})^2, (x_2 - \bar{x})^2, \dots, (x_n - \bar{x})^2$ . Em vez de usar o quadrado do desvio médio  $\sum (x_i - \bar{x})^2/n$ , por diversos motivos, dividiremos a soma dos quadrados dos desvios por  $n - 1$  em vez de  $n$ .

DEFINIÇÃO

A **variância amostral**, representada por  $s^2$ , é dada por

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1} = \frac{S_{xx}}{n - 1}$$

O **desvio padrão amostral**, representado por  $s$ , é a raiz quadrada (positiva) da variância:

$$s = \sqrt{s^2}$$

A unidade de  $s$  é a mesma de cada um dos  $x_i$ . Se, por exemplo, as observações forem consumo de combustível em milhas por galão, podemos ter  $s = 2,0$  milhas/galão. Uma interpretação grosseira do desvio padrão da amostra é que ele é o tamanho de um desvio típico ou representativo da média amostral para a amostra selecionada. Dessa forma, se  $s = 2,0$  milhas/galão, então alguns  $x_i$  da amostra estão dentro do intervalo  $\bar{x}$  mais de que 2,0, enquanto outros estão mais distantes; 2,0 é um desvio representativo (ou “padrão”) do consumo médio de combustível. Se  $s = 3,0$  para uma segunda amostra de carros de outro tipo, um desvio típico dessa amostra é cerca de 1,5 vez maior do que na primeira amostra, indicando maior dispersão da segunda.

**Exemplo 1.16**

A resistência é uma importante característica de materiais usados em casas pré-fabricadas. Cada um dos  $n = 11$  elementos de placas pré-fabricadas foi submetido a um teste de tensão severo e a largura máxima (mm) das trincas resultantes foi registrada. Os dados fornecidos (Tabela 1.3) foram relatados no artigo “Prefabricated Ferrocement Ribbed Elements for Low-Cost Housing” (*J. Ferrocement*, 1984, p. 347-364). Conseqüências do arredondamento influenciam para a soma dos desvios não ser exatamente nula. O numerador de  $s^2$  é 11,9359, portanto  $s^2 = 11,9359/(11 - 1) = 11,9359/10 = 1,19359$  e  $s = \sqrt{1,19359} = 1,0925$  mm. ■

**Tabela 1.3** Dados do Exemplo 1.16

$x_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
0,684	-0,9841	0,9685
2,540	0,8719	0,7602
0,924	-0,7441	0,5537
3,130	1,4619	2,1372
1,038	-,6301	0,3970
0,598	-1,0701	1,1451
0,483	-1,1851	1,4045
3,520	1,8519	3,4295
1,285	-0,3831	0,1468
2,650	0,9819	0,9641
1,497	-0,1711	0,0293
$\sum x_i = 18,349$	$\sum (x_i - \bar{x}) = -0,0001$	$S_{xx} = \sum (x_i - \bar{x})^2 = 11,9359$
$\bar{x} = \frac{18,349}{11} = 1,6681$		

**Dedução de  $s^2$**

Para explicar por que  $s^2$  deve ser usado em vez do quadrado do desvio médio para medir a dispersão, observe primeiro que, enquanto  $s^2$  mede a dispersão da amostra, há uma medida de dispersão da população denominada *variância da população*. Usaremos  $\sigma^2$  (o quadrado da letra grega minúscula sigma) para representar a variância da população e  $\sigma$  para representar o desvio padrão da população (a raiz quadrada de  $\sigma^2$ ). Quando a população é finita e consiste de  $N$  valores,

$$\sigma^2 = \sum_{i=1}^N (x_i - \mu)^2 / N$$

que é a média de todos os quadrados dos desvios da média da população (no caso de população, o divisor é  $N$  e não  $N - 1$ ). Definições mais gerais de  $\sigma^2$  são apresentadas nos capítulos 3 e 4.

Da mesma forma que  $\bar{x}$  é usado para fazer inferências sobre a média da população  $\mu$ , devemos definir a variância da amostra de maneira que possa ser usada para fazer inferências sobre  $\sigma^2$ . Observe que  $\sigma^2$  envolve os quadrados dos desvios em torno da média da população  $\mu$ . Se conhecêssemos o valor real de  $\mu$ , poderíamos então definir a variância da amostra como o quadrado do desvio médio da amostra  $x_i$  em torno de  $\mu$ . Entretanto, o valor de  $\mu$  quase nunca é conhecido, de modo que deve ser usada a soma dos quadrados dos desvios em torno de  $\bar{x}$ . Todavia, os  $x_i$  tendem a estar mais próximos de sua média do que da média da população  $\mu$ , assim, para compensar tal fato, é usado o divisor  $n - 1$  em vez de  $n$ . Em outras palavras, se usássemos o divisor  $n$  na fórmula da variância da amostra, quantidade resultante tenderia a subestimar  $\sigma^2$  (gerar valores na média muito pequenos para a estimativa), enquanto a divisão pelo valor, ligeiramente menor,  $n - 1$  corrige a subestimativa.



É costume se referir a  $s^2$  como tendo  $n - 1$  **graus de liberdade** (gl) como base. Essa terminologia resulta do fato de que, apesar de  $s^2$  ter como base as  $n$  quantidades  $x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_n - \bar{x}$ , sua soma é nula. Assim, especificar os valores de qualquer  $n - 1$  das quantidades determina o valor restante. Por exemplo: se  $n = 4$  e  $x_1 - \bar{x} = 8, x_2 - \bar{x} = -6$  e  $x_4 - \bar{x} = -4$ , automaticamente  $x_3 - \bar{x} = 2$ , de forma que apenas três dos quatro valores de  $x_i - \bar{x}$  serão determinados livremente (3 gl).

### Uma fórmula para o cálculo de $s^2$

O cálculo dos quadrados dos desvios é tedioso, especialmente se for usada para  $\bar{x}$  precisão decimal suficiente para prevenir os efeitos de arredondamento. Uma fórmula alternativa para o numerador de  $s^2$  evita a necessidade de todas as subtrações para obter os desvios. A fórmula envolve  $(\sum x_i)^2$ , somando e depois obtendo os quadrados e  $\sum x_i^2$ , obtendo os quadrados e depois somando.

Uma expressão alternativa para o numerador de  $s^2$  é

$$S_{xx} = \sum (x_i - \bar{x})^2 = \sum x_i^2 - \frac{(\sum x_i)^2}{n}$$

**Demonstração** Como  $\bar{x} = \sum x_i/n$ ,  $n\bar{x}^2 = (\sum x_i)^2/n$ . Então,

$$\begin{aligned} \sum (x_i - \bar{x})^2 &= \sum (x_i^2 - 2\bar{x} \cdot x_i + \bar{x}^2) = \sum x_i^2 - 2\bar{x} \sum x_i + \sum (\bar{x})^2 \\ &= \sum x_i^2 - 2\bar{x} \cdot n\bar{x} + n(\bar{x})^2 = \sum x_i^2 - n(\bar{x})^2 \end{aligned}$$

#### Exemplo 1.17

A quantidade de reflexão de luz pelas folhas foi usada para diversos propósitos, incluindo a avaliação da cor da plantação de grama, a estimativa do *status* de nitrogênio e a medida de biomassa. O artigo “Leaf Reflectance-Nitrogen-Chlorophyll Relations in Buffel-Grass” (*Photogrammetric Engr. and Remote Sensing*, 1985, p. 463-466) forneceu as observações a seguir, obtidas usando espectrofotogrametria, sobre a reflexão das folhas sob condições experimentais especificadas.

Observação	$x_i$	$x_i^2$	Observação	$x_i$	$x_i^2$
1	15,2	231,04	9	12,7	161,29
2	16,8	282,24	10	15,8	249,64
3	12,6	158,76	11	19,2	368,64
4	13,2	174,24	12	12,7	161,29
5	12,8	163,84	13	15,6	243,36
6	13,8	190,44	14	13,5	182,25
7	16,3	265,69	15	12,9	166,41
8	13,0	169,00			
				$\sum x_i = 216,1$	$\sum x_i^2 = 3168,13$

A fórmula de cálculo fornece

$$\begin{aligned} S_{xx} &= \sum x_i^2 - \frac{(\sum x_i)^2}{n} = 3168,13 - \frac{(216,1)^2}{15} \\ &= 3168,13 - 3113,28 = 54,85 \end{aligned}$$

de onde  $s^2 = S_{xx}/(n - 1) = 54,85/14 = 3,92$  e  $s = 1,98$ . ■

O método alternativo pode resultar em valores de  $s^2$  e  $s$  diferentes dos valores calculados, usando as definições. Essas diferenças se devem a efeitos de arredondamento e não serão importantes na maioria das amostras. Para minimizar os efeitos de arredondamento ao usar a fórmula alternativa, os cálculos intermediários devem ser feitos usando-se vários dígitos significativos a mais do que serão mantidos na resposta final. Como o numerador de  $s^2$  é a soma de quantidades não-negativas (quadrados desvios),  $s^2$  é com certeza não-negativo. Ainda assim, se for usada alternativa, particularmente com dados de pouca dispersão, um pequeno erro numérico pode resultar em um numerador negativo ( $\sum x_i^2$  menor que  $(\sum x_i)^2/n$ ). Se o valor de  $s^2$  for negativo, foi cometido um erro de cálculo.

Diversas outras propriedades de  $s^2$  podem facilitar esse cálculo.

#### PROPOSIÇÃO

Considere  $x_1, x_2, \dots, x_n$  como uma amostra e  $c$  como uma constante qualquer diferente de zero.

1. Se  $y_1 = x_1 + c, y_2 = x_2 + c, \dots, y_n = x_n + c$ , então  $s_y^2 = s_x^2$  e
2. Se  $y_1 = cx_1, \dots, y_n = cx_n$ , então  $s_y^2 = c^2 s_x^2, s_y = |c| s_x$ ,

onde  $s_x^2$  é a variância amostral dos  $x$  e  $s_y^2$  é a variância amostral dos  $y$ .

Em palavras, o resultado 1 informa que, se uma constante  $c$  for adicionada a (ou subtraída de) cada valor dos dados, a variância não será alterada. Isso é intuitivo, já que a adição ou a subtração de  $c$  altera a localização do conjunto de dados, mas mantém as distâncias entre os valores inalteradas. De acordo com o resultado 2, a multiplicação de cada  $x_i$  por  $c$  resulta em  $s^2$ , sendo multiplicado por um fator de  $c^2$ . Essas propriedades podem ser demonstradas pela observação de que  $\bar{y} = \bar{x} + c$  no resultado 1 e de que  $\bar{y} = c\bar{x}$  no resultado 2.

### Boxplots

Diagramas de caule e folha e histogramas conduzem a impressões gerais sobre um conjunto de dados, enquanto um único valor como a média ou o desvio padrão enfoca apenas um aspecto dos dados. Nos últimos anos, um resumo esquemático denominado *boxplot* vem sendo usado para descrever as características mais proeminentes de conjuntos de dados. Essas características incluem (1) centro, (2) dispersão, (3) a extensão e a natureza de qualquer desvio em relação à simetria e (4) a identificação de *outliers*, observações que normalmente estão distantes da maior parte dos dados. Como apenas um *outlier* pode afetar drasticamente os valores de  $\bar{x}$  e  $s$ , um *boxplot* é baseado em medidas “resistentes” à presença de alguns *outliers*: a mediana e uma medida de dispersão denominada *dispersão entre os quartos*.

#### DEFINIÇÃO

Ordene as  $n$  observações da menor para a maior e então separe a metade menor da maior. A mediana  $\tilde{x}$  estará incluída em ambas as partes se  $n$  for ímpar. Então o **quarto inferior** será a mediana da metade menor e o **quarto superior** será a mediana da metade maior. Uma medida de dispersão resistente a *outliers* é a **dispersão entre os quartos**  $f_s$ , dada por

$$f_s = \text{quarto superior} - \text{quarto inferior}$$

*Grosso modo*, a dispersão entre os quartos não será alterada pelas posições das observações nos menores 25% ou nos maiores 25% dos dados.

O *boxplot* mais simples tem base no seguinte resumo de cinco números:

menor  $x_i$       quarto inferior      mediana      quarto superior      maior  $x_i$

Primeiro, desenhe um eixo de medida horizontal. Então, coloque um retângulo sobre o eixo; a extremidade esquerda do retângulo estará no quarto inferior e a extremidade direita estará no quarto superior (de forma que a largura da caixa =  $f_s$ ). Trace um segmento de reta vertical ou outro símbolo dentro do retângulo na posição da mediana. A posição do símbolo da mediana em relação às duas extremidades indica informações sobre o desvio nos 50% centrais dos dados. Por fim, desenhe “bigodes” saindo de cada extremidade do retângulo para as observações maior e menor. Um *boxplot* com uma orientação vertical também pode ser desenhado, fazendo-se modificações óbvias no processo de construção.

### Exemplo 1.18

O ultra-som foi usado para obter informações sobre dados de corrosão na espessura da chapa do assoalho de um reservatório elevado usado para armazenar óleo bruto (“Statistical Analysis of UT Corrosion Data from Floor Plates of a Crude Oil Aboveground Storage Tank,” *Materials Eval.*, 1994, p. 846-849). Cada observação é a maior profundidade do orifício na placa, expressa em milipolegadas.

40 52 55 60 70 75 85 85 90 90 92 94 94 95 98 100 115 125 125

O resumo de cinco números segue:

menor  $x_i = 40$     quarto inferior = 72,5     $\bar{x} = 90$     quarto superior = 96,5  
maior  $x_i = 125$

A Figura 1.19 exibe o *boxplot* resultante. A extremidade direita da caixa está muito mais próxima da mediana do que a esquerda, indicando um desvio substancial na parte central dos dados. A largura da caixa ( $f_s$ ) também é relativamente grande em relação à amplitude dos dados (distância entre as pontas dos bigodes).

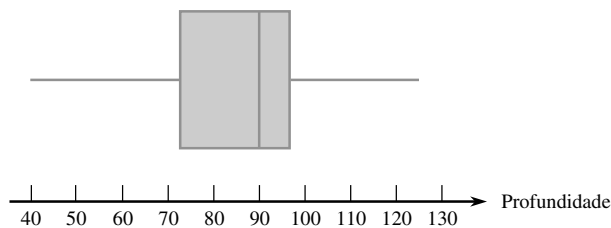


Figura 1.19 Um *boxplot* dos dados de corrosão

A Figura 1.20 mostra o resultado do MINITAB de uma solicitação de descrição dos dados de corrosão. A média aparada é a média das 17 observações que permanecem após a exclusão do maior e do menor valores (porcentagem de truncamento  $\approx 5\%$ ). Q1 e Q3 são os quartis inferior e superior. Eles são similares aos quartos, mas calculados de uma forma ligeiramente diferente. A Média SE é  $s/\sqrt{n}$ , quantidade que será importante no trabalho subsequente em inferências de  $\mu$ .

Variável	N	Média	Mediana	Média Ap	Desv Padrão	Média SE
profundidade	19	86,32	90,00	86,76	23,32	5,35
Variável	Mínimo	Máximo	Q1	Q3		
profundidade	40,00	125,00	70,00	98,00		

Figura 1.20 Descrição em MINITAB dos dados de profundidade do orifício

### Boxplots com Outliers

Um *boxplot* pode ser formatado para indicar explicitamente a presença de *outliers*.

**DEFINIÇÃO** Qualquer observação distante mais do que  $1,5f_s$  do quarto mais próximo é um *outlier*. Um *outlier* é **extremo** se estiver a mais de  $3f_s$  do quarto mais próximo. Caso contrário, é **moderado**.

Muitos procedimentos inferenciais se baseiam na suposição de que a amostra é proveniente de uma distribuição normal. Mesmo um único *outlier* extremo na amostra avisa o investigador de que tais procedimentos não devem ser usados e a presença de diversos *outliers* moderados passa a mesma mensagem.

Modifiquemos agora nossa construção anterior de um *boxplot*, desenhando um “bigode” em cada extremidade da “caixa”, para que a menor e a maior observações não sejam *outliers*. Cada *outlier* moderado é representado por um círculo cheio e cada *outlier* extremo é representado por um círculo vazio. Alguns softwares estatísticos não fazem distinções entre *outliers* extremos e moderados.

**Exemplo 1.19**

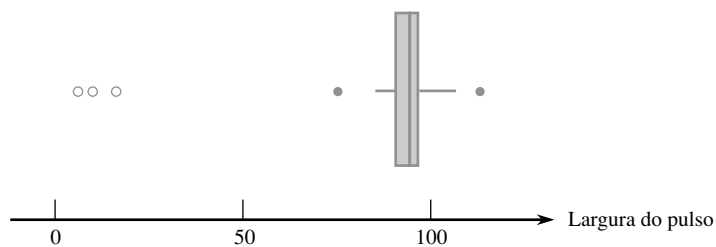
Os efeitos de descargas parciais na degradação de materiais de cavidades isolante têm importantes implicações na vida útil de componentes de alta voltagem. Consideremos a seguinte amostra de  $n = 25$  larguras de pulso de descargas lentas em uma cavidade cilíndrica de polietileno. (Esses dados são consistentes com um histograma de 250 observações no artigo “Assessment of Dielectric Degradation by Ultrawide-band PD Detection,” *IEEE Trans. on Dielectrics and Elec. Insul.*, 1995, p. 744-760.) O autor do artigo nota o impacto de diversas ferramentas estatísticas na interpretação dos dados de descarga.

5,3    8,2    13,8    74,1    85,3    88,0    90,2    91,5    92,4    92,9    93,6    94,3    94,8  
 94,9    95,5    95,8    95,9    96,6    96,7    98,1    99,0    101,4    103,7    106,0    113,5

Os indicadores relevantes são

$$\begin{array}{lll} \bar{x} = 94,8 & \text{quarto inferior} = 90,2 & \text{quarto superior} = 96,7 \\ f_s = 6,5 & 1,5f_s = 9,75 & 3f_s = 19,50 \end{array}$$

Dessa forma, qualquer observação menor que  $90,2 - 9,75 = 80,45$  ou maior que  $96,7 + 9,75 = 106,45$  é um *outlier*. Há um *outlier* na extremidade superior da amostra e quatro *outliers* na inferior. Como  $90,2 - 19,5 = 70,7$ , as três observações: 5,3, 8,2 e 13,8 são *outliers* extremos. Os outros dois *outliers* são moderados. Os “bigodes” se estendem até 85,3 e 106,0, as observações mais extremas que não são *outliers*. O *boxplot* resultante está na Figura 1.21. Há desvio negativo grande na parte central da amostra, assim como na amostra inteira.



**Figura 1.21** Um *boxplot* dos dados de largura de pulso que exibem *outliers* extremos e moderados ■

**Boxplots Comparativos**

Um *boxplot* comparativo ou lado a lado é uma forma muito eficiente de revelar semelhanças e diferenças entre dois ou mais conjuntos de dados consistindo de observações da mesma variável.

**Exemplo 1.20**

Nos últimos anos, algumas evidências sugerem que a alta concentração de radônio em ambientes fechados pode estar ligada ao desenvolvimento de cânceres infantis, mas muitos profissionais da saúde não estão convencidos. Um artigo recente (“Indoor Radon and Childhood Cancer”, *The Lancet*, 1991, p. 1537-1538) apresentou os seguintes dados sobre a concentração de radônio (Bq/m<sup>3</sup>) em duas amostras diferentes de casas. A primeira amostra consistia de casas em que havia residido uma criança com diagnóstico de câncer. As casas da segunda amostra não possuíam casos registrados de câncer infantil. A Figura 1.22 apresenta um diagrama de caule e folha dos dedos.

1. Câncer		2. Ausência de câncer	
	9683795		0 95768397678993
	86071815066815233150		1 12271713114
	12302731		2 99494191
	8349		3 839
	5		4
	7		5 55
			6
			7
HI: 210			8 5

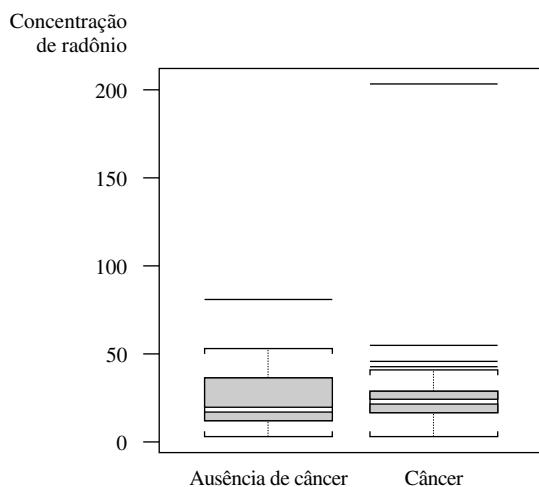
Caule: Dígito das dezenas  
 Folha: Dígito das unidades

**Figura 1.22** Diagrama de caule e folha do Exemplo 1.20

As quantidades numéricas resumidas são as seguintes:

	$\bar{x}$	$\tilde{x}$	$s$	$f_s$
Câncer	22,8	16,0	31,7	11,0
Sem câncer	19,2	12,0	17,0	18,0

Os valores da média e da mediana sugerem que a amostra de câncer está centrada um pouco à direita da amostra de ausência de câncer na escala de medida. A média, entretanto, exagera a grandeza do desvio, principalmente devido ao valor 210 na amostra de câncer. Os valores de  $s$  sugerem dispersão maior na amostra de câncer do que na amostra de ausência de câncer, mas essa impressão é contrariada pela quarta dispersão. Novamente, a observação 210, um *outlier* extremo, é a culpada. A Figura 1.23 mostra um *boxplot* comparativo do pacote de software S-Plus. A caixa dos indivíduos com ausência de câncer é mais extensa se comparada à caixa de câncer ( $f_s = 18$  vs.  $f_s = 11$ ), além disso, as posições das retas das medianas mostram desvio maior na metade da amostra de ausência de câncer do que na amostra de câncer. Os *outliers* são representados por segmentos de reta horizontais e não há distinção entre *outliers* moderados e extremos.



**Figura 1.23** Um *boxplot* dos dados do Exemplo 1.20, no S-Plus

## Exercícios | Seção 1.4 (44–61)

44. O artigo “Oxygen Consumption During Fire Suppression: Error of Heart Rate Estimation” (*Ergonomics*, 1991, p. 1469-1474) informou os dados a seguir sobre consumo de oxigênio (mL/kg/min) para uma amostra de 10 bombeiros em uma simulação de supressão de incêndio:  
29,5 49,3 30,6 28,2 28,0 26,3 33,9 29,4 23,5 31,6  
Calcule:
- A amplitude amostral.
  - A variância amostral  $s^2$  pela definição (ou seja, primeiro calculando os desvios e depois obtendo os quadrados etc.).
  - O desvio padrão amostral.
  - $s^2$  usando o método alternativo.
45. O valor do módulo de Young (GPa) foi determinado para chapas fundidas feitas de algumas substâncias metálicas, resultando nas observações a seguir (“Strength and Modulus of a Molybdenum-Coated Ti-25Al-10Nb-3U-1Mo Intermetallic,” *J. of Materials Engr. and Performance*, 1997, p. 46-50):  
116,4 115,9 114,6 115,2 115,8
- Calcule  $\bar{x}$  e os desvios em relação à média.
  - Use os desvios calculados na parte (a) para obter a variância amostral e o desvio padrão amostral.
  - Calcule  $s^2$  usando a fórmula do numerador  $S_{xx}$ .
  - Subtraia 100 de cada observação para obter uma amostra de valores transformados. Agora calcule a variância amostral desses valores transformados e a compare ao  $s^2$  dos dados originais.
46. As observações a seguir da viscosidade estabilizada (cP) para amostras de certo tipo de asfalto com 18% de borracha adicionada são do artigo “Viscosity Characteristics of Rubber-Modified Asphalts” (*J. of Materials in Civil Engr.*, 1996, p. 153-156):  
2781 2900 3013 2856 2888
- Quais são os valores da média amostral e da mediana amostral?
  - Calcule a variância amostral usando a fórmula. (Dica: Primeiro subtraia um valor conveniente de cada observação.)
47. Calcule e interprete os valores da mediana amostral, da média amostral e do desvio padrão amostral das observações a seguir da resistência à ruptura (MPa, lidas de um gráfico de “Heat-Resistant Active Brazing of Silicon Nitride: Mechanical Evaluation of Braze Joints,” *Welding J.*, August, 1997):  
87 93 96 98 105 114 128 131 142 168
48. O Exercício 36 na Seção 1.3 apresentou uma amostra de 26 tempos de fuga dos trabalhadores de uma plataforma de petróleo em uma simulação de fuga. Calcule e interprete o desvio padrão amostral. (Dica:  $\sum x_i = 9638$  e  $\sum x_i^2 = 3.587.566$ .)
49. Um estudo da relação entre idade e diversas funções visuais (como precisão e percepção de profundidade) informou as seguintes observações da área de lâmina escleral (mm<sup>2</sup>) nas extremidades do nervo óptico humano (“Morphometry of Nerve Fiber Bundle Pores in the Optic Nerve Head of the Human,” *Experimental Eye Research*, 1988, p. 559-568):  
2,75 2,62 2,74 3,85 2,34 2,74 3,93 4,21 3,88  
4,33 3,46 4,52 2,43 3,65 2,78 3,56 3,01
- Calcule  $\sum x_i$  e  $\sum x_i^2$ .
  - Use os valores calculados na parte (a) para obter a variância amostral  $s^2$  e o desvio padrão amostral  $s$ .
50. Em 1997, uma mulher processou um fabricante de teclados de computadores, sob a acusação de lesões por esforços repetitivos causados pelo teclado (*Genessy v. Digital Equipment Corp.*). O pleito era de cerca de 3,5 milhões de dólares por danos físicos, mas a corte negou esse valor pois julgou a indenização exagerada. Ao fazer essa determinação, a corte identificou um grupo “normativo” de 27 casos similares e especificou como razoável uma indenização limitada por dois desvios padrão em relação à média das indenizações dos 27 casos. As 27 indenizações foram (em milhares de dólares) 37, 60, 75, 115, 135, 140, 149, 150, 238, 290, 340, 410, 600, 750, 750, 750, 1050, 1100, 1139, 1150, 1200, 1200, 1250, 1576, 1700, 1825 e 2000, das quais  $\sum x_i = 20.179$ ,  $\sum x_i^2 = 24.657.511$ . Qual é o valor máximo que pode ser indenizado pela regra de dois desvios padrão?
51. O artigo “A Thin-Film Oxygen Uptake Test for the Evaluation of Automotive Crankcase Lubricants” (*Lubric. Engr.*, 1984, p. 75-83) informou os seguintes dados sobre tempo de oxidação-indução (min) para diversos óleos comerciais:  
87 103 130 160 180 195 132 145 211 105 145  
153 152 138 87 99 93 119 129
- Calcule a variância e o desvio padrão amostrais.
  - Se as observações fossem especificadas em horas, quais seriam os valores resultantes para a variância e para o desvio padrão amostrais? Responda sem reescrever os valores.
52. Os primeiros quatro desvios em relação à média de uma amostra de  $n = 5$  tempos de reação foram 0,3, 0,9, 1,0 e 1,3. Qual é o quinto desvio em relação à média? Forneça uma amostra para a qual esses são os cinco desvios em relação à média.
53. Reconsidere os dados sobre a área de lâmina escleral do Exercício 49.
- Determine os quartos inferior e superior.

- b. Calcule o valor da dispersão entre os quartos.
- c. Se os dois maiores valores da amostra, 4,33 e 4,52, fossem 5,33 e 5,52, como  $f_s$  seria afetado? Explique.
- d. Em quanto a observação 2,34 pode ser aumentada sem afetar  $f_s$ ? Explique.
- e. Se uma 18ª observação,  $x_{18} = 4,60$ , fosse adicionada à amostra, qual seria o valor de  $f_s$ ?

54. Reconsidere as observações de resistência ao cisalhamento (MPa) apresentadas no Exercício 34 deste capítulo:

22,2 40,4 16,4 73,7 36,6 109,9  
30,0 4,4 33,1 66,7 81,5

- a. Quais são os valores dos quartos e qual é o valor de  $f_s$ ?
- b. Construa um *boxplot* com base no resumo de cinco valores e comente suas características.
- c. Quão grande ou pequena deve ser uma observação para se qualificar como um *outlier*? E como um *outlier* extremo?
- d. Em quanto a maior observação pode ser diminuída sem afetar  $f_s$ ?

55. Segue um diagrama de caule e folha dos dados de tempos de fuga apresentados no Exercício 36 deste capítulo.

32	55
33	49
34	
35	6699
36	34469
37	03345
38	9
39	2347
40	23
41	
42	4

- a. Calcule o valor da dispersão entre os quartos.
- b. Há algum *outlier* na amostra? Algum *outlier* extremo?
- c. Construa um *boxplot* e comente suas características.
- d. Em quanto a maior observação, 424, pode ser diminuída sem afetar o valor da dispersão entre os quartos?

56. A quantidade de contaminação por alumínio (ppm) em certo tipo de plástico foi determinada para uma amostra de 26 espécimes de plástico, resultando nos dados a seguir (“The Lognormal Distribution for Modeling Quality Data when the Mean Is Near Zero,” *J. of Quality Technology*, 1990, p. 105-110):

30	30	60	63	70	79	87	90	101
102	115	118	119	119	120	125	140	145
172	182	183	191	222	244	291	511	

Construa um *boxplot* que mostre *outliers* e comente suas características.

57. Uma amostra de 20 garrafas de certo tipo de vidro foi selecionada e a resistência à pressão interna de cada garrafa foi determinada. Considere as seguintes informações parciais da amostra:

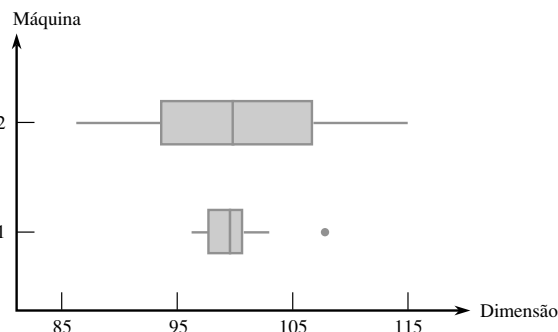
mediana = 202,2    quarto inferior = 196,0  
quarto superior = 216,8

Três observações menores 125,8 188,1 193,7  
Três observações maiores 221,3 230,5 250,2

- a. Há algum *outlier* na amostra? Algum *outlier* extremo?
- b. Construa um *boxplot* que mostre *outliers* e comente suas características.

58. Uma empresa usa duas máquinas diferentes para fabricar certo tipo de peça. Durante um turno, uma amostra de  $n = 20$  peças produzidas por cada máquina é selecionada e o valor de uma importante dimensão de cada peça é determinado. O *boxplot* comparativo da figura a seguir foi construído a partir dos dados resultantes. Compare e destaque as diferenças entre as duas amostras.

Boxplot comparativo do Exercício 58



59. A concentração de cocaína no sangue (mg/L) foi determinada para uma amostra de indivíduos que morreram de delírio induzido por cocaína (ED) e para uma amostra de indivíduos que morreram de overdose de cocaína sem delírio. O tempo de sobrevivência das pessoas em ambos os grupos foi de, no máximo, 6 horas. Os dados a seguir foram obtidos de um *boxplot* comparativo do artigo “Fatal Excited Delirium Following Cocaine Use” (*J. of Forensic Sciences*, 1997, p. 25-31).

ED            0 0 0 0 0,1 0,1 0,1 0,1 0,2 0,2 0,3 0,3  
0,3 0,4 0,5 0,7 0,8 1,0 1,5 2,7 2,8  
3,5 4,0 8,9 9,2 11,7 21,0

Não-ED      0 0 0 0 0 0,1 0,1 0,1 0,1 0,2 0,2 0,2  
0,3 0,3 0,3 0,4 0,5 0,5 0,6 0,8 0,9 1,0  
1,2 1,4 1,5 1,7 2,0 3,2 3,5 4,1  
4,3 4,8 5,0 5,6 5,9 6,0 6,4 7,9  
8,3 8,7 9,1 9,6 9,9 11,0 11,5  
12,2 12,7 14,0 16,6 17,8

- a. Determine as medianas, quartos e quartas dispersões das duas amostras.
- b. Há algum *outlier* nas amostras? Algum *outlier* extremo?

c. Construa um *boxplot* comparativo e use-o como base para comparar e destacar as diferenças das amostras ED e não-ED.

60. Foram obtidas observações sobre a resistência à explosão (lb/in<sup>2</sup>) de soldas de fechamento de bocais de teste e soldas de bocais de canisters de produção (“Proper Procedures Are the Key to Welding Radioactive Waste Cannisters,” *Welding J.*, Aug. 1997, p. 61-67).

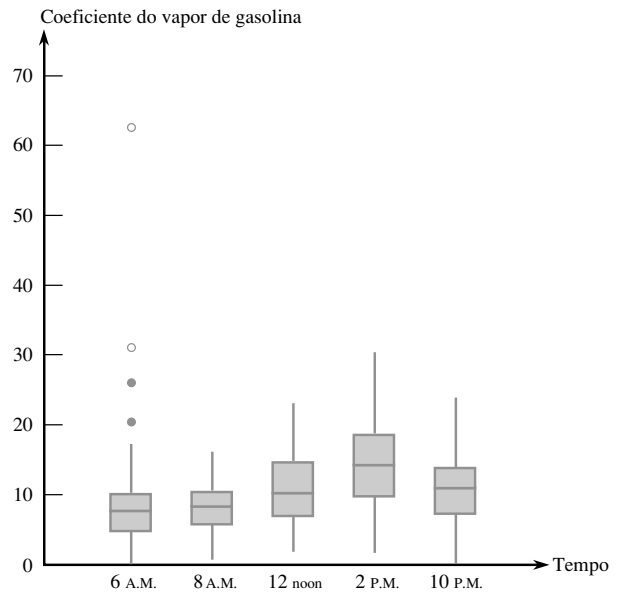
Teste 7200 6100 7300 7300 8000 7400  
7300 7300 8000 6700 8300

Canister 5250 5625 5900 5900 5700 6050  
5800 6000 5875 6100 5850 6600

Construa um *boxplot* comparativo e faça comentários sobre os aspectos interessantes (o artigo citado não inclui a figura, mas os autores comentaram que ela foi vista).

61. O seguinte *boxplot* comparativo sobre coeficientes de vapor de gasolina para veículos em Detroit foi exibido no artigo “Receptor Modeling Approach to VOC Emission Inventory Validation” (*J. of Envir. Engr.*, 1995, p. 483-490). Comente as características interessantes.

Boxplot comparativo do Exercício 61



Exercícios Suplementares (62–83)

62. Considere as seguintes informações do limite de resistência à tração (lb/in) de uma amostra de  $n = 4$  espécimes de fio rígido de cobre e zircônio (de “Characterization Methods for Fine Copper Wire” *Wire J. Intl.*, ago., 1997, p. 74-80):

$\bar{x} = 76831$   $s = 180$  menor  $x_i = 76683$   
maior  $x_i = 77048$

Determine os valores das duas observações da metade da amostra (e não o faça adivinhando!).

63. Três valores diferentes de vazão  $C_2F_6$  (SCCM) foram consideradas em um experimento para investigar o efeito da vazão na uniformidade (%) da extremidade de um *wafer* de silício usado na fabricação de circuitos integrados, resultando nos dados a seguir:

Vazão

125 2,6 2,7 3,0 3,2 3,8 4,6  
160 3,6 4,2 4,2 4,6 4,9 5,0  
200 2,9 3,4 3,5 4,1 4,6 5,1

Compare e destaque as observações de uniformidade resultantes dessas diferentes vazões.

64. A quantidade de radiação recebida em uma estufa tem importante papel na determinação da taxa de fotossíntese. As observações a seguir sobre radiação solar foram lidas de um gráfico no artigo “Radiation Components over Bare and Planted Soils in a Greenhouse” (*Solar Energy*, 1990, p. 1011-1016).

6,3 6,4 7,7 8,4 8,5 8,8 8,9  
9,0 9,1 10,0 10,1 10,2 10,6 10,6  
10,7 10,7 10,8 10,9 11,1 11,2 11,2  
11,4 11,9 11,9 12,2 13,1

Use alguns dos métodos discutidos neste capítulo para descrever e resumir os dados.

65. Os seguintes dados sobre emissões de HC e CO para um determinado veículo foram fornecidos na introdução do capítulo.

HC (gm/mi) 13,8 18,3 32,2 32,5  
CO (gm/mi) 118 149 232 236

- a. Calcule os desvios padrão amostrais das observações de HC e CO. A crença parece ter fundamento?  
b. O coeficiente de dispersão amostral  $s/\bar{x}$  (ou  $100 s/\bar{x}$ ) avalia o valor da dispersão em relação à média. Valores desses coeficientes de diversos conjuntos de dados podem ser comparados para determinar quais conjuntos de dados exibem mais ou menos dispersão. Faça uma comparação dos dados fornecidos.

66. A distribuição de frequência a seguir, de observações sobre resistência à ruptura (MPa) de barras cerâmicas tratadas em um determinado forno, foi explicada no artigo “Evaluating Tunnel Kiln Performance” (*Amer. Ceramic Soc. Bull.*, ago. 1997, p. 59-63).



Classe	81– <83	83– <85	85– <87	87– <89	89– <91
Frequência	6	7	17	30	43
Classe	91– <93	93– <95	95– <97	97– <99	
Frequência	28	22	13	3	

- a. Construa um histograma com base nas frequências relativas e comente as características interessantes.
- b. Que proporção das observações de resistência são iguais ou maiores que 85? Inferiores a 95?
- c. *Grosso modo*, que proporção das observações é inferior a 90?

67. Quinze amostras de ar de certa região foram obtidas e para cada uma delas foi determinada a concentração de monóxido de carbono. Os resultados (em ppm) foram:  
 9,3 10,7 8,5 9,6 12,2 15,6 9,2 10,5  
 9,0 13,2 11,0 8,8 13,7 12,1 9,8

Usando o método de interpolação sugerido na Seção 1.3, calcule a média aparada 10%.

- 68. a. Para que valor de  $c$  a quantidade  $\sum(x_i - c)^2$  é minimizada? (*Dica*: Obtenha a derivada em relação a  $c$ , iguale a 0 e resolva a equação).
- b. Usando o resultado da parte (a), qual das duas quantidades  $\sum(x_i - \bar{x})^2$  e  $\sum(x_i - \mu)^2$  será menor que a outra (assumindo que  $\bar{x} \neq \mu$ )?
- 69. a. Considere  $a$  e  $b$  como constantes e  $y_i = ax_i + b$  para  $i = 1, 2, \dots, n$ . Quais são as relações entre  $\bar{x}$  e  $\bar{y}$  e entre  $s_x^2$  e  $s_y^2$ ?
- b. Uma amostra de temperaturas iniciais de uma determinada reação química resultou em uma média amostral (°C) de 87,3 e um desvio padrão amostral de 1,04. Quais são a média e o desvio padrão em °F? (*Dica*:  $F = \frac{9}{5} C + 32.$ )

70. O elevado consumo de energia durante o exercício continua após o fim do treino. Como as calorias queimadas após o exercício contribuem para a perda de peso e têm outras conseqüências, é importante entender esse processo. O estudo “Effect of Weight Training Exercise and Treadmill Exercise on Post-Exercise Oxygen Consumption” (*Medicine and Science in Sports and Exercise*, 1998, p. 518-522) relatou os dados a seguir de um estudo em que o consumo de oxigênio (litros) foi medido continuamente por 30 minutos para cada um dos 15 indivíduos após um exercício de levantamento de peso e após um exercício em esteira rolante.

Indivíduo	1	2	3	4	5	6	7	8	9
	10	11	12	13	14	15			
Levantamento (x) de peso	14,6	14,4	19,5	24,3	16,3	22,1			
		23,0	18,7	19,0	17,0	19,1	19,6		
		23,2	18,5	15,9					
Esteira (y) rolante	11,3	5,3	9,1	15,2	10,1	19,6			
	20,8	10,3	10,3	2,6	16,6	22,4			
	23,6	12,6	4,4						

- a. Construa um *boxplot* comparativo das observações de levantamento de peso e de esteira rolante e comente as características encontradas.
- b. Como os dados estão na forma de pares  $(x, y)$ , com as medidas  $x$  e  $y$  da mesma variável em duas condições diferentes, é natural focar as diferenças entre os pares:  $d_1 = x_1 - y_1, \dots, d_n = x_n - y_n$ . Construa um *boxplot* das diferenças das amostras. O que ele sugere?

71. Segue uma descrição do MINITAB dos dados de resistência fornecidos no Exercício 13.

Variável	N	Média	Mediana	Média Ap	DesvPad	Média SE
resistência	153	135,39	135,40	135,41	4,59	0,37
Variável	Mínimo	Máximo	Q1	Q3		
resistência	122,20	147,70	132,95	138,25		

- a. Comente as características interessantes (os quartis e quartos são praticamente idênticos aqui).
- b. Construa um *boxplot* dos dados com base nos quartis e comente suas características.

72. Distúrbios causados por ansiedade e seus sintomas normalmente podem ser tratados com medicamentos à base de benzodiazepina. É sabido que animais expostos a estresse apresentam redução da capacidade de absorção de receptores de benzodiazepina no córtex frontal. O artigo “Decreased Benzodiazepine Receptor Binding in Prefrontal Cortex in Combat-Related Post-traumatic Stress Disorder” (*Amer. J. of Psychiatry*, 2000, p. 1120-1126) descreveu o primeiro estudo da redução da capacidade de absorção de receptores de benzodiazepina em indivíduos que sofriam de DSPT. Os dados a seguir, relacionados a uma medida da capacidade de absorção (distribuição ajustada do volume), foram tirados de um gráfico do artigo.

DSPT: 10, 20, 25, 28, 31, 35, 37, 38, 38, 39, 39, 42, 46  
 Saudável: 23, 39, 40, 41, 43, 47, 51, 58, 63, 66, 67, 69, 72

Use os diversos métodos deste capítulo para descrever e resumir os dados.

73. O artigo “Can We Really Walk Straight?” (*Amer. J. of Physical Anthropology*, 1992, p. 19-27) relatou um experimento em que cada um de 20 homens saudáveis foi solicitado a caminhar da forma mais reta possível até um alvo a 60m de distância com velocidade normal. Considere as observações a seguir sobre a cadência (número de passos por segundo):

0,95 0,85 0,92 0,95 0,93 0,86 1,00 0,92 0,85 0,81  
 0,78 0,93 0,93 1,05 0,93 1,06 1,06 0,96 0,81 0,96

Use os métodos desenvolvidos neste capítulo para resumir os dados e inclua interpretações ou discussões, se apropriado. (*Nota*: O autor do artigo usou uma análise estatística sofisticada para concluir que as pessoas não podem andar em linha reta e sugeriu diversas explicações para isso.)

74. A **moda** de um conjunto de dados numéricos é o valor que ocorre mais freqüentemente no conjunto.

- a. Determine a moda dos dados de cadência fornecidos no Exercício 73.
- b. Para uma amostra categorizada, como você definiria a categoria modal?

75. Foram selecionadas amostras de três tipos de corda e o limite de fadiga (MPa) foi determinado para cada amostra, resultando os dados a seguir.

Tipo 1 350 350 350 358 370 370 370 371  
371 372 372 384 391 391 392

Tipo 2 350 354 359 363 365 368 369 371  
373 374 376 380 383 388 392

Tipo 3 350 361 362 364 364 365 366 371  
377 377 377 379 380 380 392

- a. Construa um boxplot comparativo e comente as semelhanças e diferenças.
- b. Construa um gráfico de pontos (*dotplot*) comparativo (um *dotplot* para cada amostra com uma escala comum). Comente as semelhanças e diferenças.
- c. O boxplot comparativo da parte (a) fornece uma avaliação informativa das semelhanças e diferenças? Explique seu raciocínio.

76. As três medidas de tendência central apresentadas neste capítulo são a média, a mediana e a média aparada. Duas medidas de tendência central adicionais, usadas ocasionalmente, são o centro da amplitude, que é a média entre a menor e a maior observações e a média dos quartos, que é a média dos dois quartos. Quais destas cinco medidas de tendência central são indiferentes aos efeitos de *outliers* e quais não são? Explique seu raciocínio.

77. Considere os dados a seguir sobre tempo de conserto ativo (horas) para uma amostra de  $n = 46$  rádios de comunicação aéreos:

0,2 0,3 0,5 0,5 0,5 0,6 0,6 0,7 0,7 0,7 0,8 0,8  
0,8 1,0 1,0 1,0 1,0 1,1 1,3 1,5 1,5 1,5 1,5 2,0  
2,0 2,2 2,5 2,7 3,0 3,0 3,3 3,3 4,0 4,0 4,5 4,7  
5,0 5,4 5,4 7,0 7,5 8,8 9,0 10,3 22,0 24,5

Construa:

- a. Um diagrama de caule e folha em que os dois maiores valores são exibidos separadamente em uma linha denominada HI;
- b. Um histograma com base em seis intervalos de classe com 0 como limite inferior do primeiro intervalo e as larguras de intervalo de 2, 2, 2, 4, 10 e 10, respectivamente.

78. Considere uma amostra  $x_1, x_2, \dots, x_n$  e suponha que os valores de  $\bar{x}$ ,  $s^2$  e  $s$  tenham sido calculados.

- a. Assuma  $y_i = x_i - \bar{x}$  para  $i = 1, \dots, n$ . Como os valores de  $s^2$  e  $s$  dos  $y_i$  se comparam aos valores correspondentes dos  $x_i$ ? Explique.

- b. Assuma  $z_i = (x_i - \bar{x})/s$  para  $i = 1, \dots, n$ . Quais são os valores da variância amostral e do desvio padrão amostral dos  $z_i$ ?

79. Representemos por  $\bar{x}_n$  e  $s_n^2$  a média e a variância amostral  $\bar{x}_{n+1}$  e  $s_{n+1}^2$ , as quantidades quando uma observação adicional  $x_{n+1}$  é adicionada à amostra:

- a. Mostre como  $\bar{x}_{n+1}$  pode ser calculado a partir de  $\bar{x}_n$  e  $x_{n+1}$ .
- b. Mostre que

$$ns_{n+1}^2 = (n - 1)s_n^2 + \frac{n}{n + 1} (x_{n+1} - \bar{x}_n)^2$$

de forma que  $s_{n+1}^2$  possa ser calculado a partir de  $x_{n+1}$ ,  $\bar{x}_n$ , e  $s_n^2$ .

- c. Suponha que uma amostra de 15 fios de tapeçaria tenha uma média de alongamento de 12,58mm e um desvio padrão de 0,512mm. O 16º fio resulta em um valor de alongamento de 11,8. Quais são os valores da média amostral e do desvio padrão das 16 observações de alongamento?

80. As distâncias das rotas de ônibus em qualquer sistema de trânsito normalmente variam de um trajeto para outro. O artigo “Planning of City Bus Routes” (*J. of the Institution of Engineers*, 1995, p. 211-215) fornece as informações a seguir sobre as distâncias (km) de um determinado sistema:

<i>Extensão</i>	6–	8–	10–	12–	14–
	<8	<10	<12	<14	<16
<i>Freqüência</i>	6	23	30	35	32
<i>Extensão</i>	16–	18–	20–	22–	24–
	<18	<20	<22	<24	<26
<i>Freqüência</i>	48	42	40	28	27
<i>Extensão</i>	26–	28–	30–	35–	40–
	<28	<30	<35	<40	<45
<i>Freqüência</i>	26	14	27	11	2

- a. Desenhe um histograma correspondente a estas freqüências.
- b. Que proporção destas distâncias são inferiores a 20? Que proporção das rotas possui distância ao menos de 30?
- c. *Grosso modo*, qual é o valor do 90º percentil da distribuição de distâncias de rotas?
- d. *Grosso modo*, qual é a mediana da distância de rota?

81. Um estudo realizado para investigar a distribuição do tempo total de frenagem (tempo de reação mais tempo do movimento entre acelerador e freio, em ms) em condições reais de dirigibilidade a 60km/h forneceu as informações resumidas sobre a distribuição dos tempos (“A Field Study on Braking Responses during Driving,” *Ergonomics*, 1995: 1903-1910):

média = 535    mediana = 500    moda = 500  
dp = 96    mínimo = 220    máximo = 925  
5º percentil = 400    10º percentil = 430  
90º percentil = 640    95º percentil = 720

O que você pode concluir sobre o formato do histograma desses dados? Explique seu raciocínio.

82. Os dados de amostra  $x_1, x_2, \dots, x_n$  algumas vezes representam uma **série temporal**, onde  $x_t$  = valor observado de uma variável de resposta  $x$  no momento  $t$ . Frequentemente, a série observada mostra grande variação aleatória, o que dificulta o estudo do comportamento de longo prazo. Nessas situações, é desejável produzir uma versão suavizada da série. Uma das técnicas usadas é a suavização **exponencial**. O valor de uma constante de suavização  $\alpha$  é escolhido ( $0 < \alpha < 1$ ). Então, com  $\bar{x}_t$  = valor ajustado no instante  $t$ , definimos  $\bar{x}_t = x_t$  e para  $t = 2, 3, \dots, n$ ,  $\bar{x}_t = \alpha x_t + (1 - \alpha)\bar{x}_{t-1}$ .

- Considere a seguinte série de tempo em que  $x_t$  = temperatura (°F) de efluente em uma estação de tratamento de esgoto no dia  $t$ : 47, 54, 53, 50, 46, 46, 47, 50, 51, 50, 46, 52, 50, 50. Plote cada coordenada  $x_t$  em relação a  $t$  em um sistema bidimensional (um gráfico de série de tempo). Parece haver algum padrão?
- Calcule o  $\bar{x}_t$  usando  $\alpha = 0,1$ . Repita, para  $\alpha = 0,5$ . Que valor de  $\alpha$  fornece uma série  $\bar{x}_t$  mais suavizada?
- Substitua  $\bar{x}_{t-1} = \alpha x_{t-1} + (1 - \alpha)\bar{x}_{t-2}$  no lado direito da expressão de  $\bar{x}_t$  e então substitua  $\bar{x}_{t-2}$  em termos de  $x_{t-2}$ ,  $\bar{x}_{t-3}$  e assim por diante. De quantos valores  $x_p$ ,  $x_{t-1}$ , ...,  $x_1$ ,  $\bar{x}_t$  depende? O que acontece ao coeficiente de  $x_{t-k}$  quando  $k$  aumenta?
- Consulte a parte (c). Se  $t$  for grande, qual é a sensibilidade de  $\bar{x}_t$  em relação ao valor inicial  $\bar{x}_t = x_1$ ? Explique.

(Nota: Uma referência relevante é o artigo "Simple Statistics for Interpreting Environmental Data," *Water Pollution Control Fed. J.*, 1981, p. 167-175.)

83. Considere as observações numéricas  $x_1, \dots, x_n$ . Frequentemente é de interesse saber se os  $x_i$  estão (ao menos aproximadamente) distribuídos simetricamente ao redor de algum valor. Se  $n$  é pelo menos moderadamente grande, a extensão da simetria pode ser avaliada por meio de um diagrama de caule e folha ou de um histograma. Entretanto, se  $n$  não for muito grande, esses gráficos não serão muito informativos. Considere a alternativa a seguir: seja  $y_1$  o menor  $x_i$ ,  $y_2$  o segundo menor  $x_i$  e assim por diante. Faça então o gráfico dos seguintes pares como pontos em um sistema de coordenadas bidimensional:  $(y_n - \bar{x}, \bar{x} - y_1)$ ,  $(y_{n-1} - \bar{x}, \bar{x} - y_2)$ ,  $(y_{n-2} - \bar{x}, \bar{x} - y_3)$ , ... Há  $n/2$  pontos em que  $n$  é par e  $(n - 1)/2$  em que  $n$  é ímpar.

- Qual é a aparência do gráfico quando há simetria perfeita dos dados? Qual sua aparência quando as observações se estendem mais acima do que abaixo da mediana (uma cauda superior longa)?
- Os dados a seguir sobre precipitação (acre-pé) de 26 nuvens esparsas foram obtidos do artigo "A Bayesian Analysis of a Multiplicative Treatment Effect in Weather Modification" (*Technometrics*, 1975, p. 161-166). Construa o gráfico e comente sobre a extensão da simetria ou a natureza do seu desvio.

4,1	7,7	17,5	31,4	32,7	40,6	92,4
115,3	118,3	119,0	129,6	198,6	200,7	242,5
255,0	274,7	274,7	302,8	334,1	430,0	489,1
703,4	978,0	1656,0	1697,8	2745,6		

## Bibliografia

CHAMBERS, John. CLEVELAND, William, KLEINER, Beat e TUKEY, Paul. *Graphical Methods for Data Analysis*. Brooks/Cole, Pacific Grove, CA, 1983. Uma apresentação altamente recomendada sobre metodologia gráfica antiga e mais recente da estatística.

DEVORE, Jay e PECK, Roxy. *Statistics: The Exploration and Analysis of Data* (4ª ed.). Duxbury Press, Pacific Grove, CA, 2001. Os primeiros capítulos fornecem uma pesquisa não-matemática de métodos para descrição e resumo de dados.

FREEDMAN, David, PISANI, Robert e PURVES, Roger. *Statistics* (3ª ed.). Norton, Nova York, 1998. Uma excelente pesquisa não-matemática de raciocínio e metodologia básicos da estatística.

HOAGLIN, David, MOSTELLER, Frederick e TUKEY, John. *Understanding Robust and Exploratory Data Analysis*.

Wiley, Nova York, 1983. Discute por que e como os métodos exploratórios devem ser empregados. É bastante útil em detalhes de diagramas de caule e folha e *boxplots*.

HOAGLIN, David e VELLEMAN, Paul. *Applications, Basics, and Computing of Exploratory Data Analysis*. Duxbury Press, Boston, 1980. Uma boa discussão de alguns métodos exploratórios básicos.

MOORE, David. *Statistics: Concepts and Controversies* (5ª ed.). Freeman, San Francisco, 2001. Um estudo extremamente agradável e de fácil leitura que contém uma discussão intuitiva de problemas relacionados à amostragem e experimentos projetados.

TANUR, Judith et al. (eds.). *Statistics: A Guide to the Unknown* (3ª ed.). Duxbury Press, Belmont, CA, 1988. Contém vários artigos não-técnicos sobre diversas aplicações da estatística.

# Probabilidade

## Introdução

O termo **probabilidade** se refere ao estudo da aleatoriedade e da incerteza.

Em qualquer situação em que ocorrem diversos resultados, a teoria da probabilidade oferece métodos de quantificação das chances ou possibilidades de ocorrência associadas aos diversos resultados. A linguagem da probabilidade é constantemente usada de maneira informal em contextos escritos e falados. Exemplos incluem declarações como: “é provável que a média Dow-Jones cresça no final do ano”, “há 50% de chance de o titular buscar a reeleição”, “provavelmente pelo menos uma seção deste curso será oferecida no ano que vem”, “as chances favorecem um acordo rápido para o fim da greve” e “espera-se que pelo menos 20 mil ingressos sejam vendidos para o show”. Neste capítulo, apresentaremos alguns conceitos básicos de probabilidade, indicaremos como as probabilidades podem ser interpretadas e como suas regras podem ser aplicadas para calcular as possibilidades de ocorrência de muitos eventos. A metodologia da probabilidade permitirá que expressemos, em linguagem precisa, declarações informais como as fornecidas acima.

O estudo da probabilidade como ramo da matemática data de mais de 300 anos e teve sua gênese relacionada a questões que envolviam jogos de azar. Muitos livros se dedicam exclusivamente à probabilidade, mas nosso objetivo é cobrir apenas a parte da disciplina que tem maior ligação com os problemas de inferência estatística.

---

## 2.1 | Espaços amostrais e eventos

Um **experimento** é qualquer ação ou processo cujo resultado está sujeito à incerteza. Apesar de o termo *experimento* geralmente sugerir uma situação altamente controlada de teste em laboratório, aqui será usado em um sentido muito mais amplo. Dessa forma, os experimentos que são de interesse incluem jogar uma moeda uma ou diversas vezes, selecionar uma ou várias cartas de um baralho, pesar um pedaço de pão, determinar o tempo de

viagem de casa ao trabalho em uma dada manhã, obter tipos sanguíneos de um grupo de indivíduos ou medir as resistências de compressão de diversas vigas metálicas.

## O espaço amostral de um experimento

### DEFINIÇÃO

O **espaço amostral** de um experimento, representado por  $S$ , é o conjunto de todos os resultados possíveis desse experimento.

### Exemplo 2.1

O experimento mais simples a que a probabilidade se aplica é aquele com dois resultados finais possíveis. Um exemplo de experimento desse tipo é examinar um único fusível para ver se funciona. O espaço amostral desse experimento é expresso como  $\mathcal{S} = \{N, D\}$ , onde  $N$  representa sem defeito,  $D$  representa com defeito e as chaves são usadas para englobar os elementos do conjunto. Outro experimento desse tipo envolve jogar um percevejo e observar se sua ponta cai para cima ou para baixo, com o espaço amostral  $\mathcal{S} = \{C, B\}$  e um terceiro consiste em observar o sexo da próxima criança nascida no hospital local, com  $\mathcal{S} = \{M, F\}$ .

### Exemplo 2.2

Se examinarmos três fusíveis em seqüência e anotarmos o resultado de cada exame, o resultado do experimento é qualquer seqüência de  $N$  e  $D$  de 3 elementos, de forma que

$$\mathcal{S} = \{NNN, NND, NDN, NDD, DNN, DND, DDN, DDD\}$$

Se tivéssemos jogado um percevejo três vezes, o espaço amostral seria obtido substituindo  $N$  por  $C$  no  $\mathcal{S}$  acima, e com uma alteração similar na notação do espaço amostral para o experimento em que o sexo de três crianças recém-nascidas é observado. ■

### Exemplo 2.3

Dois postos de gasolina estão localizados em uma determinada interseção. Cada um possui seis bombas. Considere o experimento em que o número de bombas em uso em determinada hora do dia é determinado para cada posto. Um resultado experimental especifica quantas bombas estão em uso no primeiro posto e quantas são usadas no segundo. Um resultado possível é (2, 2), outro é (4, 1) e outro ainda é (1, 4). Os 49 resultados em  $\mathcal{S}$  são exibidos na tabela a seguir. O espaço amostral do experimento em que um dado de seis lados é lançado duas vezes é obtido excluindo a linha 0 e a coluna 0 da tabela, fornecendo 36 resultados.

		<i>Segundo Posto</i>						
		0	1	2	3	4	5	6
<i>Primeiro posto</i>	0	(0, 0)	(0, 1)	(0, 2)	(0, 3)	(0, 4)	(0, 5)	(0, 6)
	1	(1, 0)	(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)
	2	(2, 0)	(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
	3	(3, 0)	(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)
	4	(4, 0)	(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	(4, 6)
	5	(5, 0)	(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)	(5, 6)
	6	(6, 0)	(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

### Exemplo 2.4

Se uma bateria de lanterna nova, tipo D, tiver uma voltagem fora de certos limites, será classificada como falha ( $F$ ); se a voltagem estiver dentro dos limites especificados, será classificada como sucesso ( $S$ ). Suponha que um experimento consista em testar cada bateria quando sai de uma linha de montagem até que seja observado um sucesso. Apesar de não ser muito provável, um resultado possível é que as primeiras 10 (ou 100 ou 1000 ou ...) sejam  $F$  e a próxima seja  $S$ . Isto é, para qualquer inteiro  $n$  possível, teremos de examinar  $n$  baterias antes de obter o primeiro  $S$ . O espaço amostral é  $\mathcal{S} = \{S, FS, FFS, FFFS, \dots\}$ , que contém um número infinito de resultados possíveis. A mesma forma abreviada de espaço amostral é apropriada para um experimento em que, é registrado o sexo de cada criança recém-nascida, em um horário especificado, até que seja observado o nascimento de uma criança do sexo masculino. ■

### Eventos

No estudo de probabilidade, estaremos interessados não apenas nos resultados individuais de  $\mathcal{S}$ , como também em qualquer grupo de resultados de  $\mathcal{S}$ .

#### DEFINIÇÃO

**Evento** é qualquer grupo (subconjunto) de resultados contidos no espaço amostral  $\mathcal{S}$ . O evento é denominado **simples** se consistir um único resultado e **composto** se consistir em mais de um resultado.

Quando um experimento é realizado, determinado evento  $A$  ocorre se o resultado experimental estiver contido em  $A$ . Em geral, ocorrerá exatamente um evento simples, mas diversos eventos compostos também podem ocorrer simultaneamente.

### Exemplo 2.5

Considere um experimento em que cada um de três veículos que trafeguem em uma determinada estrada siga pela saída à esquerda ( $E$ ) ou à direita ( $D$ ) no final da rampa de saída. Os oito resultados possíveis que compõem o espaço amostral são:  $EEE, DEE, EDE, EED, EDD, DED, DDE$  e  $DDD$ . Dessa forma, há oito eventos simples, dentre os quais estão  $E_1 = \{EEE\}$  e  $E_5 = \{EDD\}$ . Os eventos compostos incluem:

$A = \{DEE, EDE, EED\}$  = o evento em que exatamente um dos três veículos vira à direita;

$B = \{EEE, DEE, EDE, EED\}$  = o evento em que no máximo um dos veículos vira à direita;

$C = \{EEE, DDD\}$  = o evento em que os três veículos viram na mesma direção.

Suponha que, quando o experimento é executado, o resultado seja  $EEE$ . Então, o evento simples  $E_1$  terá ocorrido, da mesma forma que os eventos  $B$  e  $C$  (mas não  $A$ ). ■

### Exemplo 2.6 (continuação do Exemplo 2.3)

Quando o número de bombas em uso em cada um dos dois postos de seis bombas for observado, haverá 49 resultados possíveis, de forma que haverá 49 eventos simples:  $E_1 = \{(0, 0)\}$ ,  $E_2 = \{(0, 1)\}$ , ...,  $E_{49} = \{(6, 6)\}$ . Exemplos de eventos compostos são

$A = \{(0, 0), (1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)\}$  = o evento em que o número de bombas em uso é o mesmo nos dois postos;

$B = \{(0, 4), (1, 3), (2, 2), (3, 1), (4, 0)\}$  = o evento em que o número total de bombas em uso é quatro;

$C = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$  = o evento em que no máximo uma bomba está em uso em cada posto. ■

**Exemplo 2.7** (continuação do Exemplo 2.4)

O espaço amostral do experimento de exame das baterias contém um número infinito de resultados, portanto há um número infinito de eventos simples. Os eventos compostos incluem:

$A = \{S, FS, FFS\}$  = o evento em que no máximo três baterias são examinadas;

$E = \{FS, FFFS, FFFFFS, \dots\}$  = o evento em que um número par de baterias é examinado. ■

**Algumas relações sobre a teoria dos conjuntos**

Um evento é essencialmente um conjunto, de forma que as relações e resultados da teoria elementar dos conjuntos podem ser usados para o estudo dos eventos. As operações a seguir serão usadas para construção de novos eventos, a partir de eventos conhecidos.

**DEFINIÇÃO**

1. A **união** de dois eventos  $A$  e  $B$ , representada por  $A \cup B$  e lida “ $A$  união  $B$ ”, é o evento que consiste em todos os resultados que estão no evento  $A$  ou no  $B$  ou em ambos (de forma que a união inclui resultados em que ocorram  $A$  e  $B$ , bem como aqueles em que exatamente um ocorre), isto é, todos os resultados em ao menos um dos eventos.
2. A **interseção** dos dois eventos  $A$  e  $B$ , representada por  $A \cap B$  e lida “ $A$  interseção  $B$ ”, é o evento que consiste de todos os resultados que estão em *ambos*  $A$  e  $B$ .
3. O **complemento** de um evento  $A$ , representado por  $A'$ , é o conjunto de todos os resultados em  $\mathcal{S}$  que não estão contidos em  $A$ .

**Exemplo 2.8** (continuação do Exemplo 2.3)

Para o experimento em que é observado o número de bombas em uso em um posto de gasolina de seis bombas, assuma  $A = \{0, 1, 2, 3, 4\}$ ,  $B = \{3, 4, 5, 6\}$  e  $C = \{1, 3, 5\}$ . Então

$$A \cup B = \{0, 1, 2, 3, 4, 5, 6\} = \mathcal{S}; A \cup C = \{0, 1, 2, 3, 4, 5\},$$

$$A \cap B = \{3, 4\}, A \cap C = \{1, 3\}, A' = \{5, 6\}, \{A \cup C\}' = \{6\}$$

**Exemplo 2.9** (continuação do Exemplo 2.4)

No experimento da bateria, defina  $A$ ,  $B$  e  $C$  por

$$A = \{S, FS, FFS\}$$

$$B = \{S, FFS, FFFFFS\}$$

e

$$C = \{FS, FFFS, FFFFFS, \dots\}$$

Então

$$A \cup B = \{S, FS, FFS, FFFFFS\}$$

$$A \cap B = \{S, FFS\}$$

$$A' = \{FFFS, FFFFFS, FFFFFS, \dots\}$$

e

$$C' = \{S, FFS, FFFFFS, \dots\} = \{\text{evento em que é inspecionado um número ímpar de baterias}\} \quad \blacksquare$$

Algumas vezes,  $A$  e  $B$  não possuem resultados em comum, de modo que a interseção de  $A$  e  $B$  não contém nenhum resultado.

**DEFINIÇÃO**

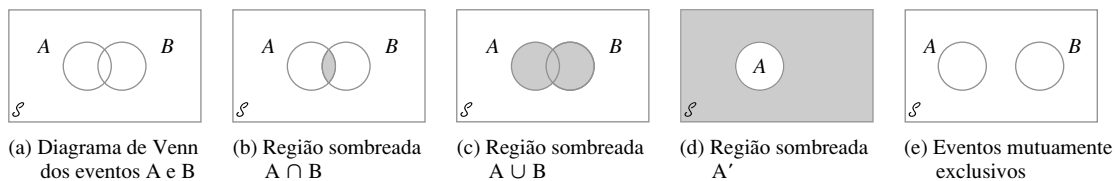
Quando  $A$  e  $B$  não possuem resultados em comum, são chamados eventos **mutuamente exclusivos** ou **disjuntos**.

**Exemplo 2.10**

Uma cidade pequena possui três revendedores de automóveis: um revendedor GM que vende Chevrolets, Pontiacs e Buicks, um revendedor Ford que vende Fords e Mercurys e um revendedor Chrysler que vende Plymouths e Chryslers. Se um experimento consistir em observar a marca do próximo carro vendido, os eventos  $A = \{\text{Chevrolet, Pontiac, Buick}\}$  e  $B = \{\text{Ford, Mercury}\}$  são mutuamente exclusivos porque o próximo carro vendido não pode ser um produto GM e Ford ao mesmo tempo. ■

As operações de união e interseção podem ser estendidas a mais de dois eventos. Para quaisquer três eventos  $A$ ,  $B$  e  $C$ , o evento  $A \cup B \cup C$  é o conjunto de resultados contidos em no mínimo um dos três eventos, enquanto  $A \cap B \cap C$  é o conjunto de resultados contidos nos três eventos. Dados os eventos  $A_1, A_2, A_3, \dots$ , eles serão denominados mutuamente exclusivos (ou disjuntos em pares) se dois deles não tiverem resultados em comum.

Uma representação gráfica de eventos e manipulações de eventos é obtida pelo uso de diagramas de Venn. Para construir um desses diagramas, desenhe um retângulo cujo interior representará o espaço amostral  $\mathcal{S}$ . Então, qualquer evento  $A$  é representado como o interior de uma curva fechada (normalmente um círculo) contido em  $\mathcal{S}$ . A Figura 2.1 mostra exemplos de diagramas de Venn.



**Figura 2.1** Diagramas de Venn

**Exercícios** | Seção 2.1 (1–10)

1. Quatro universidades —1, 2, 3 e 4— estão participando de um torneio de basquete. Na primeira *etapa*, 1 jogará com 2 e 3 com 4. Os dois vencedores disputarão o campeonato e os dois perdedores também jogarão. Um resultado possível pode ser representado por 1324 (1 ganha de 2 e 3 ganha de 4 nos jogos da primeira *etapa* e 1 ganha de 3 e 2 ganha de 4).
  - a. Relacione todos os resultados de  $\mathcal{S}$ .
  - b. Represente por  $A$  o evento em que 1 ganha o torneio. Relacione os resultados de  $A$ .
  - c. Represente por  $B$  o evento em que 2 seja um dos finalistas do campeonato. Relacione os resultados de  $B$ .
  - d. Quais são os resultados de  $A \cup B$  e de  $A \cap B$ ? Quais são os resultados de  $A'$ ?
2. Suponha que os veículos que trafegam em uma determinada estrada possam tomar uma saída à direita ( $D$ ), à esquerda ( $E$ ) ou ir em frente ( $F$ ). Observe a direção de cada um de três veículos sucessivamente.
  - a. Relacione todos os resultados do evento  $A$  em que os três veículos seguem na mesma direção.
  - b. Relacione todos os resultados do evento  $B$  em que os três veículos tomam diferentes direções.
  - c. Relacione os resultados do evento  $C$  em que exatamente dois dos três veículos viram à direita.
  - d. Relacione todos os resultados do evento  $D$  em que exatamente dois veículos seguem na mesma direção.
  - e. Relacione os resultados em  $D'$ ,  $C \cup D$ , e  $C \cap D$ .
3. Três componentes estão conectados para formar um sistema conforme exibido no diagrama a seguir. Como os componentes no subsistema 2-3 estão conectados em paralelo, esse subsistema funcionará se ao menos um dos dois componentes individuais funcionar. Para que todo o sistema funcione, o componente 1 deve funcionar, bem como o subsistema 2-3.



- O experimento consiste em determinar a condição de cada componente [ $S$  (sucesso) para um componente que funciona bem e  $F$  (falha) para um componente que não funciona].
- Que resultados estão contidos no evento  $A$  para que exatamente dois dos três componentes funcionem?
  - Que resultados estão contidos no evento  $B$  para que ao menos dois componentes funcionem?
  - Que resultados estão contidos no evento  $C$  para que o sistema funcione?
  - Relacione os resultados de  $C'$ ,  $A \cup C$ ,  $A \cap C$ ,  $B \cup C$ , e  $B \cap C$ .
- Cada item de uma amostra de quatro hipotecas está classificado como taxa fixa ( $F$ ) ou taxa variável ( $V$ ).
    - Quais são os 16 resultados de  $\mathcal{S}$ ?
    - Que resultados do evento possuem exatamente três hipotecas de taxa fixa?
    - Que resultados do evento possuem quatro hipotecas de taxa fixa?
    - Que resultados pertencem ao evento em que as quatro hipotecas são de taxa variável?
    - Qual é a união dos eventos das partes (c) e (d) e qual a interseção desses dois eventos?
    - Quais são a união e a interseção dos dois eventos das partes (b) e (c)?
  - Uma família de três pessoas  $-A, B$  e  $C-$  tem plano de uma clínica médica que sempre possui um médico em cada posto 1, 2, e 3. Durante determinada semana, cada membro da família visita a clínica uma vez e é designada aleatoriamente a um posto. O experimento consiste em registrar o número do posto para cada membro. Um resultado é  $(1, 2, 1)$  para  $A$  no posto 1,  $B$  no posto 2 e  $C$  no posto 1.
    - Relacione os 27 resultados do espaço amostral.
    - Relacione todos os resultados do evento em que os três membros seguem para o mesmo posto.
    - Relacione todos os resultados do evento em que todos os membros seguem para postos diferentes.
    - Relacione todos os resultados do evento em que ninguém segue para o posto 2.
  - A biblioteca de uma faculdade possui cinco exemplares de certo livro-texto de reserva. Duas cópias (1 e 2) são primeiras edições e as outras três (3, 4 e 5) são segundas edições. Um aluno examina esses livros em ordem aleatória, parando apenas quando uma segunda edição é selecionada. Um resultado possível é 5 e outro é 213.
    - Relacione todos os resultados de  $\mathcal{S}$ .
    - Represente por  $A$  o evento em que exatamente um livro deve ser examinado. Quais são os resultados em  $A$ ?
    - Represente por  $B$  o evento em que o livro 5 é o selecionado. Quais resultados pertencem a  $B$ ?
    - Represente por  $C$  o evento em que o livro 1 não é examinado. Quais resultados pertencem a  $C$ ?
  - Um departamento acadêmico terminou a votação secreta para sua chefia. A urna contém quatro cédulas com votos para o candidato  $A$  e três com votos para o candidato  $B$ . Suponha que essas cédulas sejam removidas da urna uma a uma.
    - Relacione todos os resultados possíveis.
    - Suponha que uma contagem seja feita conforme as cédulas são removidas. Para que resultados  $A$  se mantém na frente de  $B$  na contagem?
  - Uma empresa de engenharia civil está trabalhando atualmente em usinas de energia em três locais diferentes. Represente por  $A_i$  o evento em que a usina no local  $i$  é completada na data do contrato. Use as operações de união, interseção e complemento para descrever cada um dos eventos a seguir nos termos de  $A_1, A_2$  e  $A_3$ , desenhe um diagrama de Venn e sombreie a região correspondente a cada uma.
    - Ao menos uma usina é concluída até a data do contrato.
    - Todas as usinas são concluídas até a data do contrato.
    - Apenas a usina do local 1 é concluída até a data do contrato.
    - Exatamente uma usina é concluída até a data do contrato.
    - Apenas a usina do local 1 ou as duas outras são concluídas até a data do contrato.
  - Use diagramas de Venn para provar as duas relações a seguir para eventos  $A$  e  $B$  (denominadas leis de De Morgan) para quaisquer eventos  $A$  e  $B$ :
    - $(A \cup B)' = A' \cap B'$
    - $(A \cap B)' = A' \cup B'$
  - No Exemplo 2.10, identifique três eventos que não são mutuamente exclusivos.
    - Suponha que não haja um resultado comum a todos os três eventos  $A, B$  e  $C$ . Esses eventos são necessariamente mutuamente exclusivos? Se sua resposta for sim, explique. Se for não, dê um exemplo contrário, usando o experimento do Exemplo 2.10.

## 2.2

## Axiomas, Interpretações e Propriedades da Probabilidade

Dados um experimento e um espaço amostral  $\mathcal{S}$ , o objetivo da probabilidade é atribuir a cada evento  $A$  um número  $P(A)$ , denominado probabilidade do evento  $A$ , que fornecerá uma medida precisa da chance de

ocorrência de  $A$ . Para assegurar que as atribuições de probabilidade sejam consistentes com nossas noções intuitivas de probabilidade, todas as atribuições devem satisfazer os axiomas a seguir (propriedades básicas) de probabilidade.

AXIOMA 1	Para qualquer evento, $A$ , $P(A) \geq 0$ .
AXIOMA 2	$P(\mathcal{S}) = 1$ .
AXIOMA 3	<p>a. Se <math>A_1, A_2, \dots, A_k</math> for um conjunto finito de eventos mutuamente exclusivos, então</p> $P(A_1 \cup A_2 \cup \dots \cup A_k) = \sum_{i=1}^k P(A_i)$ <p>b. Se <math>A_1, A_2, A_3, \dots</math> for um conjunto infinito de eventos mutuamente exclusivos, então</p> $P(A_1 \cup A_2 \cup A_3 \cup \dots) = \sum_{i=1}^{\infty} P(A_i)$

O Axioma 1 reflete a noção intuitiva de que a chance de ocorrência de  $A$  deve ser não-negativa. O espaço amostral é, por definição, o evento que deve ocorrer quando o experimento é realizado ( $\mathcal{S}$  contém todos os resultados possíveis), de forma que o Axioma 2 diz que a maior probabilidade possível de ser atribuída a  $\mathcal{S}$  é 1. O terceiro axioma formaliza a idéia de que, se desejarmos que a probabilidade de ao menos um de diversos eventos ocorra e que dois eventos não ocorram simultaneamente, a chance de pelo menos um ocorrer será a soma das chances dos eventos individuais.

### Exemplo 2.11

No experimento em que uma única moeda é lançada, o espaço amostral é  $\mathcal{S} = \{H, T\}$ . Os axiomas especificam que  $P(\mathcal{S}) = 1$ , de forma que, para completar a atribuição de probabilidade, falta apenas determinar  $P(H)$  e  $P(T)$ . Já que  $H$  e  $T$  são eventos disjuntos e  $H \cup T = \mathcal{S}$ , o Axioma 3 implica que

$$1 = P(\mathcal{S}) = P(H) + P(T)$$

Essa expressão implica que  $P(T) = 1 - P(H)$ . A única liberdade permitida pelos axiomas nesses experimentos é a probabilidade atribuída a  $H$ . Uma possível atribuição de probabilidades é  $P(H) = 0,5$ ,  $P(T) = 0,5$ , enquanto outra atribuição possível é  $P(H) = 0,75$ ,  $P(T) = 0,25$ . De fato, representar  $p$  por qualquer número fixo entre 0 e 1,  $P(H) = p$  e  $P(T) = 1 - p$  é uma atribuição consistente com os axiomas. ■

### Exemplo 2.12

Considere o experimento do Exemplo 2.4, em que as baterias que saem de uma linha de montagem são testadas uma a uma até que seja encontrada uma com voltagem dentro dos limites especificados. Esses eventos simples são  $E_1 = \{S\}$ ,  $E_2 = \{FS\}$ ,  $E_3 = \{FFS\}$ ,  $E_4 = \{FFFS\}$ , ... Suponha que a probabilidade de uma bateria ser satisfatória seja de 0,99. Então, pode ser demonstrado que  $P(E_1) = 0,99$ ,  $P(E_2) = (0,01)(0,99)$ ,  $P(E_3) = (0,01)^2(0,99)$ ... é uma atribuição de probabilidades a eventos simples que satisfazem os axiomas. Em particular, porque os  $E_i$  são disjuntos e  $\mathcal{S} = E_1 \cup E_2 \cup E_3 \cup \dots$ , deve ser o caso de que

$$\begin{aligned} 1 = P(\mathcal{S}) &= P(E_1) + P(E_2) + P(E_3) + \dots \\ &= 0,99[1 + 0,01 + (0,01)^2 + (0,01)^3 + \dots] \end{aligned}$$

A validade dessa igualdade é uma conseqüência de um resultado matemático referente à soma de uma série geométrica.

Entretanto, outra atribuição de probabilidade legítima (de acordo com os axiomas) do mesmo tipo “geométrico” é obtida pela substituição de 0,99 por qualquer número  $p$  entre 0 e 1 e 0,01 por  $(1 - p)$ . ■

## Interpretando probabilidade

Os Exemplos 2.11 e 2.12 mostram que os axiomas não determinam completamente uma atribuição de probabilidades a eventos. Os axiomas servem apenas para cortar atribuições inconsistentes com as noções intuitivas de probabilidade. No experimento de lançamento de uma moeda do Exemplo 2.11, duas atribuições particulares foram sugeridas. A atribuição correta ou apropriada depende da forma pela qual o experimento é executado e também da interpretação da probabilidade. A interpretação usada com mais frequência e mais compreendida baseia-se na noção de frequências relativas.

Considere um experimento que pode ser executado repetidamente de uma forma idêntica e independente e represente por  $A$  um evento que consista de um conjunto fixo de resultados do experimento. Exemplos simples de experimentos passíveis de repetição incluem o lançamento de dados e moedas discutidos anteriormente. Se o experimento for executado  $n$  vezes, em algumas delas o evento  $A$  ocorrerá (o resultado será o conjunto  $A$ ) e em outras, não. Represente por  $n(A)$  o número de repetições em que  $A$  ocorre. A relação  $n(A)/n$  é denominada *frequência relativa* de ocorrência do evento  $A$  na sequência de  $n$  repetições. Evidências empíricas, com base nos resultados de várias dessas sequências de experimentos repetidos, indicam que, conforme  $n$  aumenta, a frequência relativa  $n(A)/n$  se estabiliza, de acordo com ilustração da Figura 2.2. Isto é, conforme  $n$  torna-se arbitrariamente grande, a frequência relativa se aproxima de um valor-limite denominado *frequência relativa-limite* do evento  $A$ . A interpretação objetiva da probabilidade identifica essa frequência relativa-limite com  $P(A)$ .

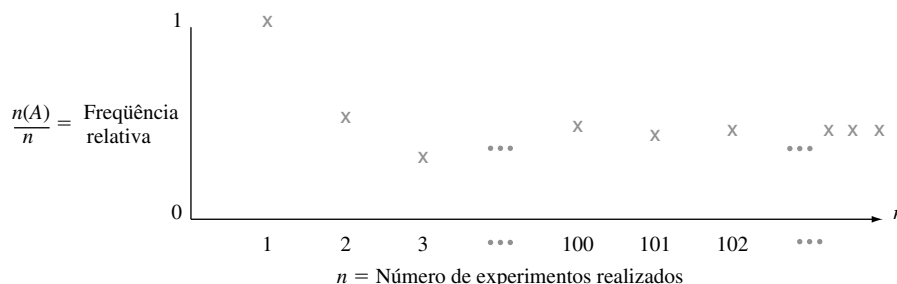


Figura 2.2 Estabilização da frequência relativa

Se forem atribuídas probabilidades a eventos de acordo com suas frequências relativas-limite, podemos interpretar a afirmativa “a probabilidade de dar ‘cara’ em um lançamento de moeda é de 0,5” para indicar que, em grande número de lançamentos, a “cara” será exibida em cerca de metade deles e “coroa”, na outra metade.

A interpretação de frequência relativa da probabilidade é tida como objetiva porque toma como base uma propriedade do experimento, em vez de qualquer indivíduo preocupado com o experimento. Por exemplo: dois observadores de uma sequência de lançamentos de moeda devem usar as mesmas atribuições de probabilidade, já que eles não influenciam a frequência relativa-limite. Na prática, essa interpretação não é tão objetiva quanto parece, já que a frequência relativa-limite de um evento não será conhecida. Portanto, temos de atribuir probabilidades com base em nossa crença sobre a frequência relativa-limite dos eventos em estudo. Felizmente, há muitos experimentos para os quais há consenso no que diz respeito a atribuições de probabilidade. Quando falamos de uma moeda justa, queremos dizer que  $P(H) = P(T) = 0,5$ ; e um dado justo é aquele em que as frequências relativas-limite dos seis resultados são todas  $\frac{1}{6}$ , sugerindo atribuições de probabilidade  $P(\{1\}) = \dots = P(\{6\}) = \frac{1}{6}$ .

Como a interpretação objetiva da probabilidade se baseia na noção de frequência-limite, sua aplicabilidade é limitada a situações experimentais que podem ser repetidas. Ainda assim, a linguagem de probabilidade é normalmente usada em relação a situações que não podem ser repetidas. Os exemplos incluem: “há boas chances de um acordo de paz”, “provavelmente nossa empresa fechará o contrato” e “como o melhor atacante deles está machucado, espero que eles não marquem mais de 10 contra nós”. Nessas situações, desejamos, como antes, atribuir probabilidades numéricas a diversos resultados e eventos (por exemplo: a probabilidade de fecharmos o contrato é de 0,9). Devemos, portanto, adotar uma interpretação alternativa dessas probabilidades. Como

diferentes observadores podem ter diversas informações prévias e opiniões a respeito de situações experimentais, as atribuições de probabilidade podem ser diferentes de indivíduo para indivíduo. As interpretações nestas situações são denominadas *subjetivas*. O livro de Robert Winkler relacionado nas referências do capítulo fornece uma boa pesquisa sobre diversas interpretações subjetivas.

## Propriedades de probabilidade

### PROPOSIÇÃO

Para qualquer evento,  $A$ ,  $P(A) = 1 - P(A')$ .

### Demonstração

No Axioma 3a, assumamos  $k = 2$ ,  $A_1 = A$ , e  $A_2 = A'$ . Já que por definição  $A'$ ,  $A \cup A' = \mathcal{S}$  enquanto  $A$  e  $A'$  são disjuntos,  $1 = P(\mathcal{S}) = P(A \cup A') = P(A) + P(A')$ , de onde segue o resultado. ■

Essa proposição é surpreendentemente útil, porque há várias situações em que  $P(A')$  é mais facilmente obtido pelos métodos diretos do que  $P(A)$ .

### Exemplo 2.13

Considere um sistema de cinco componentes idênticos ligados em série, conforme ilustrado na Figura 2.3.

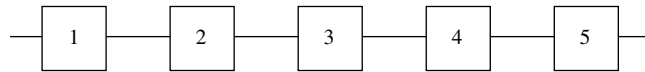


Figura 2.3 Um sistema de cinco componentes ligados em série

Represente um componente que falha por  $F$  e o que não falha por  $S$  (de sucesso). Represente por  $A$  o evento em que o sistema falha. Para que  $A$  ocorra, ao menos um dos componentes individuais deve falhar. Os resultados de  $A$  incluem  $SSFSS$  (1, 2, 4 e 5 funcionam, mas 3 não),  $FFSSS$  e assim por diante. Na verdade há 31 diferentes resultados em  $A$ . Entretanto,  $A'$ , o evento em que o sistema funciona, consiste em um único resultado  $SSSSS$ . Na Seção 2.5 veremos que se 90% de todos esses componentes não apresentarem falhas e se componentes diferentes apresentarem falhas independentemente um do outro,  $P(A') = P(SSSSS) = 0,9^5 = 0,59$ . Portanto,  $P(A) = 1 - 0,59 = 0,41$ . Dessa forma, em um grande número de tais sistemas, cerca de 41% apresentarão falhas. ■

Em geral, a proposição acima é útil quando o evento de interesse pode ser expresso como “ao menos...”, já que o complemento “menos que...” pode ser mais fácil de se lidar (em alguns problemas, “mais que...” é mais fácil de lidar que “no máximo...”). Se você tiver dificuldade para calcular  $P(A)$  diretamente, pense em determinar  $P(A')$ .

### PROPOSIÇÃO

Se  $A$  e  $B$  forem mutuamente exclusivos, então  $P(A \cap B) = 0$ .

### Demonstração

Visto que  $A \cap B$  não contém resultados,  $(A \cap B)' = \mathcal{S}$ . Portanto,  $1 = P[(A \cap B)'] = 1 - P(A \cap B)$ , o que implica  $P(A \cap B) = 1 - 1 = 0$ . ■

Quando os eventos  $A$  e  $B$  são mutuamente exclusivos, o Axioma 3 fornece  $P(A \cup B) = P(A) + P(B)$ . Quando  $A$  e  $B$  não são mutuamente exclusivos, a probabilidade da união é obtida do resultado a seguir.

### PROPOSIÇÃO

Para quaisquer dois eventos  $A$  e  $B$ ,

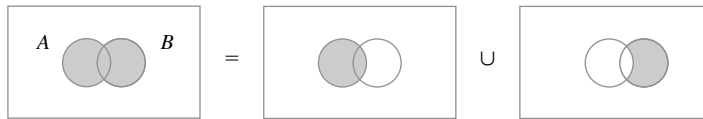
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Observe que a proposição é válida mesmo se  $A$  e  $B$  forem mutuamente exclusivos, já que  $P(A \cap B) = 0$ . A idéia central é que, ao adicionar  $P(A)$  e  $P(B)$ , a probabilidade da interseção  $A \cap B$  será contada duas vezes, de forma que  $P(A \cap B)$  deve ser subtraído.

**Demonstração**

Observe primeiro que  $A \cup B = A \cup (B \cap A')$ , conforme ilustrado na Figura 2.4. Visto que  $A$  e  $(B \cap A')$  são mutuamente exclusivos,  $P(A \cup B) = P(A) + P(B \cap A')$ . Mas  $B = (B \cap A) \cup (B \cap A')$  (a união da parte de  $B$  contida em  $A$  e a parte de  $B$  não contida em  $A$ ), sendo  $(B \cap A)$  e  $(B \cap A')$  mutuamente exclusivos, de forma que  $P(B) = P(B \cap A) + P(B \cap A')$ . Combinando esses resultados, temos

$$\begin{aligned} P(A \cup B) &= P(A) + P(B \cap A') = P(A) + [P(B) - P(A \cap B)] \\ &= P(A) + P(B) - P(A \cap B) \end{aligned}$$



**Figura 2.4** Representação de  $A \cup B$  como união de eventos disjuntos

**Exemplo 2.14**

Em um determinado bairro residencial, 60% de todos os lares assinam o jornal metropolitano publicado em uma cidade próxima, 80% assinam o jornal local e 50% de todos os lares assinam os dois. Se um lar for selecionado aleatoriamente, qual será a probabilidade de ele assinar (1) ao menos um dos jornais e (2) exatamente um dos jornais?

Fazendo  $A = \{\text{assinaturas do jornal metropolitano}\}$  e  $B = \{\text{assinaturas do jornal local}\}$ , as informações fornecidas implicam  $P(A) = 0,6$ ,  $P(B) = 0,8$ , e  $P(A \cap B) = 0,5$ . A proposição anterior se aplica a dar

$P(\text{assinaturas de ao menos um dos jornais})$

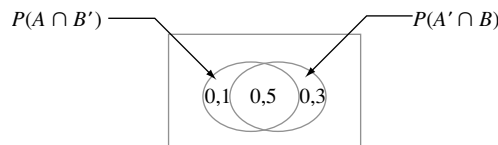
$$= P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0,6 + 0,8 - 0,5 = 0,9$$

O evento de um lar assinar apenas o jornal local pode ser escrito como  $A' \cap B$  [(não-metropolitano) e local]. A Figura 2.4 implica que

$$0,9 = P(A \cup B) = P(A) + P(A' \cap B) = 0,6 + P(A' \cap B)$$

pela qual  $P(A' \cap B) = 0,3$ . De forma similar,  $P(A \cap B') = P(A \cup B) - P(B) = 0,1$ . Como ilustrado na Figura 2.5, onde se vê que

$$P(\text{exatamente um}) = P(A \cap B') + P(A' \cap B) = 0,1 + 0,3 = 0,4$$



**Figura 2.5** Probabilidades do Exemplo 2.14

A probabilidade de uma união de mais de dois eventos pode ser calculada de forma análoga. Para os três eventos  $A$ ,  $B$  e  $C$ , o resultado é

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) \\ &\quad - P(B \cap C) + P(A \cap B \cap C) \end{aligned}$$

Como vemos no diagrama de Venn de  $A \cup B \cup C$ , exibido na Figura 2.6, quando  $P(A)$ ,  $P(B)$  e  $P(C)$  são adicionados, certas interseções são contadas duas vezes, de forma que devem ser subtraídas, mas isso resulta em  $P(A \cap B \cap C)$  sendo subtraída mais de uma vez.

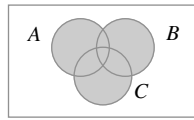


Figura 2.6  $A \cup B \cup C$

### Determinando as sistemáticas de probabilidades

Quando o número de resultados possíveis (eventos simples) é grande, haverá muitos eventos compostos. Uma forma simples de determinar as probabilidades desses eventos, que evita a violação dos axiomas e as propriedades derivadas, é primeiro determinar as probabilidades  $P(E_i)$  para todos os eventos simples. Isso deve satisfazer  $P(E_i) \geq 0$  e  $\sum_{\text{todos } i} P(E_i) = 1$ . Dessa forma, a probabilidade de qualquer evento composto  $A$  é calculada pela adição das  $P(E_i)$  para todos os  $E_i$  em  $A$ :

$$P(A) = \sum_{\text{todos os } E_i \text{ em } A} P(E_i)$$

#### Exemplo 2.15

Represente os seis eventos básicos  $\{1\}, \dots, \{6\}$  associados ao lançamento de um dado de seis lados uma vez por  $E_1, \dots, E_6$ . Se o dado for construído de tal forma que qualquer um dos três resultados pares tenha o dobro de probabilidade de ocorrer em relação aos ímpares, uma atribuição apropriada de probabilidades a eventos elementares é  $P(E_1) = P(E_3) = P(E_5) = \frac{1}{9}$ ,  $P(E_2) = P(E_4) = P(E_6) = \frac{2}{9}$ . Dessa forma, para o evento  $A = \{\text{resultado par}\} = E_2 \cup E_4 \cup E_6$ ,  $P(A) = P(E_2) + P(E_4) + P(E_6) = \frac{6}{9} = \frac{2}{3}$ ; para  $B = \{\text{resultado} \leq 3\} = E_1 \cup E_2 \cup E_3$ ,  $P(B) = \frac{1}{9} + \frac{2}{9} + \frac{1}{9} = \frac{4}{9}$ . ■

### Resultados igualmente prováveis

Em muitos experimentos que consistem em  $N$  resultados, é razoável atribuir probabilidades iguais a todos os  $N$  eventos simples. Tais eventos incluem exemplos óbvios, como lançamento de uma moeda; ou um dado não viciado uma ou duas vezes (ou qualquer número fixo de vezes); ou selecionar uma ou diversas cartas de um baralho de 52 cartas bem embaralhado. Com  $p = P(E_i)$  para cada  $i$ ,

$$1 = \sum_{i=1}^N P(E_i) = \sum_{i=1}^N p = p \cdot N \quad \text{então } p = \frac{1}{N}$$

Isto é, se houver  $N$  resultados possíveis, a probabilidade atribuída a cada um será  $1/N$ . Consideremos um evento  $A$ , com  $N(A)$  representando o número de resultados contidos em  $A$ . Então

$$P(A) = \sum_{E_i \text{ em } A} P(E_i) = \sum_{E_i \text{ em } A} \frac{1}{N} = \frac{N(A)}{N}$$

Depois de contarmos o número  $N$  de resultados no espaço amostral para calcular a probabilidade de qualquer evento, devemos contar os resultados nele contidos e obter a razão entre os dois números. Dessa forma, quando os resultados forem igualmente prováveis, o cálculo das probabilidades se resume na contagem.

## Exemplo 2.16

Quando dois dados são lançados separadamente, há  $N = 36$  resultados (excluir as primeiras linha e coluna da tabela no Exemplo 2.3). Se os dois dados forem justos, todos os 36 resultados serão igualmente prováveis, então  $P(E_i) = \frac{1}{36}$ . Dessa forma, o evento  $A = \{\text{soma dos dois números} = 7\}$  consistirá em seis resultados (1, 6), (2, 5), (3, 4), (4, 3), (5, 2) e (6, 1). Assim,

$$P(A) = \frac{N(A)}{N} = \frac{6}{36} = \frac{1}{6}$$

■

## Exercícios Seção 2.2 (11–28)

11. Uma empresa de fundos mútuos oferece a seus clientes diversos fundos: um de mercado, três de títulos diferentes, (curto, médio e longo prazos), dois fundos de ações (moderado e de alto risco) e um misto. Dentre os usuários que possuem cotas em apenas um fundo, seguem as porcentagens de clientes nos diferentes fundos.
- |                      |     |                        |     |
|----------------------|-----|------------------------|-----|
| Mercado              | 20% | Ações de alto risco    | 18% |
| Título curto prazo   | 15% | Ação de risco moderado | 25% |
| Título intermediário | 10% | Misto                  | 7%  |
| Título longo prazo   | 5%  |                        |     |
- Um cliente que possui cotas em apenas um fundo é selecionado aleatoriamente.
- Qual é a probabilidade de o indivíduo selecionado possuir cotas do fundo misto?
  - Qual é a probabilidade de o indivíduo selecionado possuir cotas em um fundo de títulos?
  - Qual é a probabilidade de o indivíduo selecionado não possuir cotas em um fundo de ações?
12. Selecione aleatoriamente um estudante em uma determinada universidade e represente por  $A$  o evento de ele possuir um cartão de crédito Visa e por  $B$  o evento análogo para um MasterCard. Suponha que  $P(A) = 0,5$ ,  $P(B) = 0,4$  e  $P(A \cap B) = 0,25$ .
- Calcule a probabilidade de que o indivíduo selecionado tenha pelo menos um dos dois tipos de cartão (ou seja, a probabilidade do evento  $A \cup B$ ).
  - Qual é a probabilidade de o indivíduo selecionado não ter nenhum dos tipos de cartão?
  - Descreva, em termos de  $A$  e  $B$ , o evento em que o estudante selecionado possui um cartão Visa, mas não um MasterCard. Calcule a probabilidade desse evento.
13. Uma empresa de consultoria em informática apresenta suas propostas de três projetos. Represente por  $A_i = \{\text{projeto } i \text{ fechado}\}$ , para  $i = 1, 2, 3$  e suponha que  $P(A_1) = 0,22$ ,  $P(A_2) = 0,25$ ,  $P(A_3) = 0,28$ ,  $P(A_1 \cap A_2) = 0,11$ ,  $P(A_1 \cap A_3) = 0,05$ ,  $P(A_2 \cap A_3) = 0,07$ ,  $P(A_1 \cap A_2 \cap A_3) = 0,01$ . Expresse em palavras cada evento a seguir e calcule sua probabilidade:
- $A_1 \cup A_2$
  - $A_1' \cap A_2'$  [Sugestão:  $(A_1 \cup A_2)' = A_1' \cap A_2'$ ]
  - $A_1 \cup A_2 \cup A_3$
  - $A_1' \cap A_2' \cap A_3'$
  - $A_1' \cap A_2' \cap A_3$
  - $(A_1' \cap A_2') \cup A_3$
14. Uma empresa de eletricidade oferece uma taxa vitalícia de energia a qualquer lar cuja utilização de energia esteja abaixo de 240 kWh durante um determinado mês. Represente por  $A$  o evento de um lar selecionado aleatoriamente em uma comunidade que não excede a utilização da taxa vitalícia em janeiro e por  $B$  o evento análogo para o mês de julho ( $A$  e  $B$  se referem ao mesmo lar). Suponha que  $P(A) = 0,8$ ,  $P(B) = 0,7$  e  $P(A \cup B) = 0,9$ . Calcule:
- $P(A \cap B)$ .
  - A probabilidade de a quantia da taxa vitalícia ser excedida em exatamente um dos dois meses. Descreva esse evento em termos de  $A$  e  $B$ .
15. Considere o tipo de secadora de roupas (a gás ou elétricas) comprada por cinco clientes diferentes em uma loja.
- Se a probabilidade de no máximo um desses clientes fazer uma compra de uma secadora elétrica for 0,428, qual será a probabilidade de ao menos dois clientes comprarem uma secadora elétrica?
  - Se  $P(\text{os cinco comprarem a gás}) = 0,116$  e  $P(\text{os cinco comprarem elétricas}) = 0,005$ , qual será a probabilidade de haver uma compra de ao menos uma de cada tipo?
16. Um indivíduo recebe três copos de refrigerante diferentes rotulados de  $C$ ,  $D$  e  $P$ . Solicita-se que ele experimente os três e os relacione em ordem de preferência. Suponha que o mesmo refrigerante tenha sido colocado nos três copos.
- Quais são os eventos simples nesse experimento de classificação e qual probabilidade você atribuiria a cada um deles?
  - Qual é a probabilidade de  $C$  ser classificado como o primeiro?
  - Qual é a probabilidade de  $C$  ser classificado como o primeiro e  $D$  como o último?
17. Represente por  $A$  o evento de que a próxima solicitação de assistência de um consultor de um software estatístico seja relacionada ao pacote SPSS e por  $B$  o evento

de a próxima solicitação de ajuda ser relacionada ao pacote SAS. Suponha que  $P(A) = 0,30$  e  $P(B) = 0,50$ .

- a. Por que não é o caso de  $P(A) + P(B) = 1$ ?
- b. Calcule  $P(A')$ .
- c. Calcule  $P(A \cup B)$ .
- d. Calcule  $P(A' \cap B')$ .

18. Uma caixa contém quatro lâmpadas de 40-W, cinco de 60-W e seis de 75-W. Se as lâmpadas forem selecionadas uma a uma em ordem aleatória, qual é a probabilidade de ao menos duas serem selecionadas para obter uma de 75 W?

19. A inspeção visual de juntas de solda em placas de circuitos impressos pode ser bastante subjetiva. Parte do problema se origina dos diversos tipos de defeitos de soldas (por exemplo, falta solda em pontos variados) e até da quantidade de um ou mais desses defeitos. Conseqüentemente, até mesmo inspetores altamente treinados podem discordar sobre a disposição de uma junta. Em um lote de 10.000 juntas, o inspetor A encontrou 724 que julgou defeituosas, o inspetor B encontrou 751 e 1159 foram encontradas por ao menos um dos inspetores. Suponha que uma dessas 10.000 juntas seja selecionada aleatoriamente.

- a. Qual é a probabilidade de que a junta selecionada não seja julgada defeituosa por nenhum dos dois inspetores?
- b. Qual é a probabilidade de que a junta selecionada seja julgada defeituosa pelo inspetor B, mas não pelo inspetor A?

20. Uma determinada fábrica opera em três turnos diferentes. No ano anterior, ocorreram 200 acidentes na fábrica. Alguns deles podem ser atribuídos em parte a condições de trabalho inseguras, enquanto os outros não estão relacionados a condições de trabalho. A tabela a seguir fornece as porcentagens de acidentes que se encaixam em cada categoria de turno de trabalho.

	Condições Inseguras	Não relacionado às condições
<i>Dia</i>	10%	35%
<b>Turno Alternado</b>	8%	20%
<i>Noite</i>	5%	22%

Suponha que um dos 200 relatórios de acidente seja selecionado aleatoriamente de um arquivo de relatórios e sejam determinados o tipo de acidente e o turno.

- a. Quais são os eventos simples?
- b. Qual é a probabilidade de que o acidente selecionado seja atribuído a condições inseguras?
- c. Qual é a probabilidade de que o acidente selecionado não tenha ocorrido no turno do dia?

21. Uma empresa de seguros oferece quatro níveis de dedução – nenhum, baixo, médio e alto – para os possuidores de apólices de seguros residenciais e três níveis diferentes – baixo, médio e alto – para os possuidores de apólices de seguros de automóveis. A tabela a seguir

fornece as proporções das diversas categorias de segurados que possuem ambos os tipos de seguros. Por exemplo: a proporção de indivíduos com baixa dedução de seguro residencial e baixa dedução de seguro de automóvel é 0,06 (6% de todos os indivíduos).

Automóvel	Residencial			
	N	B	M	A
<b>B</b>	0,04	0,06	0,05	0,03
<b>M</b>	0,07	0,10	0,20	0,10
<b>A</b>	0,02	0,03	0,15	0,15

Suponha que um indivíduo que possua ambos os tipos de apólices seja selecionado aleatoriamente.

- a. Qual é a probabilidade de que o indivíduo tenha dedução média de automóvel e alta de residência?
- b. Qual é a probabilidade de que o indivíduo tenha uma dedução baixa de automóvel? Uma dedução baixa de residência?
- c. Qual é a probabilidade de que um indivíduo esteja na mesma categoria para deduções de automóvel e residência?
- d. Com base na resposta da parte (c), qual é a probabilidade de que as duas categorias sejam diferentes?
- e. Qual é a probabilidade de que o indivíduo tenha ao menos um nível baixo de dedução?
- f. Usando a resposta da parte (e), qual é a probabilidade de que nenhum nível de dedução seja baixo?

22. A rota usada por um motorista que vai ao trabalho contém dois cruzamentos com semáforos. A probabilidade de que ele tenha de parar no primeiro semáforo é 0,4, a probabilidade análoga para o segundo semáforo é 0,5 e a probabilidade de que ele tenha de parar em pelo menos um dos dois semáforos é 0,6. Qual é a probabilidade de ele ter de parar:

- a. Nos dois semáforos?
- b. No primeiro semáforo mas não no segundo?
- c. Em exatamente um semáforo?

23. Os computadores de seis membros do corpo docente de um determinado departamento serão substituídos. Dois deles escolheram *laptops* e os outros quatro *desktops*. Suponha que apenas duas configurações podem ser feitas em um determinado dia e que os dois computadores a serem configurados sejam selecionados aleatoriamente entre os seis (implicando 15 resultados igualmente prováveis se os computadores estiverem numerados 1, 2, ..., 6, de forma que um resultado consistirá nos computadores 1 e 2, outro nos computadores 1 e 3 e assim por diante).

- a. Qual é a probabilidade de que as duas configurações selecionadas sejam de *laptops*?
- b. Qual é a probabilidade de que as configurações selecionadas sejam de *desktops*?
- c. Qual é a probabilidade de que ao menos uma das configurações selecionadas seja de um computador *desktop*?



- d. Qual é a probabilidade de que ao menos um computador de cada tipo seja escolhido para a configuração?
24. Use os axiomas para mostrar que, se um evento  $A$  estiver contido em outro evento  $B$  (ou seja,  $A$  é um subconjunto de  $B$ ), então  $P(A) \leq P(B)$ . [Sugestão: neste caso,  $A$  e  $B$ ,  $A$  e  $B \cap A'$  são disjuntos e  $B = A \cup (B \cap A')$ , como pode ser visto em um diagrama de Venn.] Para os  $A$  e  $B$  genéricos, o que podemos deduzir sobre as relações  $P(A \cap B)$ ,  $P(A)$ , e  $P(A \cup B)$ ?
25. Os três principais itens opcionais de certo tipo de carro novo são transmissão automática ( $A$ ), teto solar ( $B$ ) e rádio com CD-player ( $C$ ). Se 70% de todos os compradores solicitarem  $A$ , 80% solicitarem  $B$ , 75% solicitarem  $C$ , 85% solicitarem  $A$  ou  $B$ , 90% solicitarem  $A$  ou  $C$ , 95% solicitarem  $B$  ou  $C$  e 98% solicitarem  $A$  ou  $B$  ou  $C$ , calcule as probabilidades dos eventos a seguir. [Sugestão: “ $A$  ou  $B$ ” é o evento em que no mínimo dois opcionais são solicitados.] Desenhe um diagrama de Venn e identifique todas as regiões.
- O comprador seguinte solicita um dos três opcionais.
  - O comprador seguinte não solicita nenhum opcional.
  - O comprador seguinte solicita apenas transmissão automática e nenhum dos outros dois opcionais.
  - O comprador seguinte solicita exatamente um opcional.
26. Um determinado sistema pode ter três tipos diferentes de defeitos. Represente por  $A_i$  ( $i = 1, 2, 3$ ) o evento em que o sistema apresenta um defeito do tipo  $i$ . Suponha que
- $$P(A_1) = 0,12 \quad P(A_2) = 0,07 \quad P(A_3) = 0,05$$
- $$P(A_1 \cup A_2) = 0,13 \quad P(A_1 \cup A_3) = 0,14$$
- $$P(A_2 \cup A_3) = 0,10 \quad P(A_1 \cap A_2 \cap A_3) = 0,01$$
- Qual é a probabilidade de o sistema não ter um defeito do tipo 1?
  - Qual é a probabilidade de o sistema ter defeitos dos tipos 1 e 2?
  - Qual é a probabilidade de o sistema ter defeitos dos tipos 1 e 2, mas não do tipo 3?
  - Qual é a probabilidade de o sistema ter no máximo dois desses defeitos?
27. Um departamento acadêmico com cinco membros do corpo docente — Anderson, Box, Cox, Cramer e Fisher — deve selecionar dois deles para servir a um comitê de avaliação de pessoal. Como o trabalho levará muito tempo, ninguém deseja servir, de forma que foi decidido que os representantes serão selecionados por sorteio, colocando-se cinco tiras de papel em uma caixa, misturando-as e selecionando duas.
- Qual é a probabilidade de Anderson e Box serem selecionados? (Sugestão: Relacione os resultados igualmente prováveis.)
  - Qual é a probabilidade de que pelo menos um dos dois membros cujos nomes sejam iniciados por C seja selecionado?
  - Se os cinco membros tiverem 3, 6, 7, 10 e 14 anos de ensino na universidade, respectivamente, qual é a probabilidade de que os dois representantes escolhidos tenham no mínimo 15 anos de ensino na universidade (tempos de ensino somados)?
28. No Exercício 5, suponha que qualquer indivíduo tenha possibilidade igualmente provável de ser designado para qualquer um dos três postos independentemente de para onde os outros sejam designados. Qual é a possibilidade de:
- Os três membros da família serem enviados ao mesmo posto?
  - No máximo dois membros da família serem enviados ao mesmo posto?
  - Cada membro da família ser enviado a um posto diferente?

## 2.3 | Técnicas de contagem

Quando os diversos resultados de um experimento são igualmente prováveis (a mesma probabilidade é atribuída a cada evento simples), a tarefa de calcular probabilidades se reduz a uma contagem. Em particular, se  $N$  for a quantidade de resultados de um espaço amostral e  $N(A)$  for a quantidade de resultados contidos em um evento  $A$ , então

$$P(A) = \frac{N(A)}{N} \quad (2.1)$$

Se uma lista dos resultados estiver disponível ou for de fácil construção, e  $N$  for pequeno, então o numerador e o denominador da Equação (2.1) podem ser obtidos sem o uso de quaisquer princípios de contagem geral.

Há, entretanto, muitos experimentos para os quais o esforço despendido na elaboração de tal lista é proibitivo porque  $N$  é muito grande. Explorando algumas regras gerais de contagem, é possível calcular

probabilidades da forma (2.1) sem relacionar os resultados. Essas regras também são úteis em vários problemas que envolvem resultados que não sejam igualmente prováveis. Várias das regras desenvolvidas aqui serão usadas no estudo das distribuições de probabilidades no próximo capítulo.

## A regra do produto para pares ordenados

Nossa primeira regra se aplica a qualquer situação em que um conjunto (evento) consiste em pares ordenados de objetos e desejamos contar o número desses pares. Entendemos por par ordenado que, se  $O_1$  e  $O_2$  forem objetos, o par  $(O_1, O_2)$  será diferente do par  $(O_2, O_1)$ . Por exemplo: se um indivíduo seleciona uma linha aérea para uma viagem de Los Angeles a Chicago e (após fazer negócios em Chicago) uma segunda linha para seguir até Nova York, uma possibilidade é (American, United), outra é (United, American) e outra, (United, United).

### PROPOSIÇÃO

Se o primeiro elemento ou objeto de um par ordenado puder ser selecionado de  $n_1$  formas e para cada uma dessas  $n_1$  formas, o segundo elemento do par pode ser selecionado de  $n_2$  formas, o que faz com que o número de pares seja  $n_1 n_2$ .

### Exemplo 2.17

O proprietário de uma casa em reforma solicita os serviços de um encanador e de um eletricista. Se houver 12 empreiteiros de encanamento e 9 empreiteiros eletricistas disponíveis na área, de quantas formas os empreiteiros podem ser contratados? Se representarmos os encanadores por  $P_1 \dots P_{12}$  e os eletricistas por  $Q_1 \dots Q_9$ , desejamos que os pares sejam da forma  $(P_i, Q_j)$ . Com  $n_1 = 12$  e  $n_2 = 9$ , a regra do produto fornece  $N = (12)(9) = 108$  formas possíveis de escolher dois tipos de empreiteiros. ■

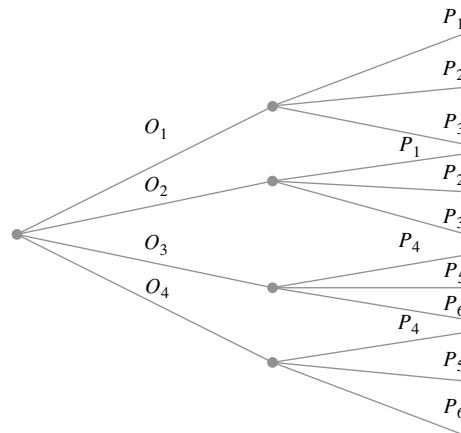
No Exemplo 2.17, a escolha do segundo elemento do par não depende de o primeiro ter sido escolhido ou ter ocorrido. Enquanto houver o mesmo número de escolhas do segundo elemento para cada primeiro, a regra do produto será válida, mesmo quando o conjunto de segundos elementos possíveis depender do primeiro.

### Exemplo 2.18

Uma família se mudou para uma cidade e precisa dos serviços de um obstetra e de um pediatra. Há duas clínicas de fácil acesso e cada uma tem dois obstetras e três pediatras. A família obterá benefícios máximos de seguro de saúde se escolher uma clínica e selecionar os dois especialistas. De quantas formas isso pode ser feito? Represente os obstetras por  $O_1, O_2, O_3,$  e  $O_4$  e os pediatras por  $P_1, \dots P_6$ . Desejamos então determinar o número de pares  $(O_i, P_j)$  para os quais  $O_i$  e  $P_j$  estão associados à mesma clínica. Como há quatro obstetras,  $n_1 = 4$  e para cada um há três escolhas de pediatra,  $n_2 = 3$ , a aplicação da regra do produto fornece  $N = n_1 n_2 = 12$  escolhas possíveis. ■

## Diagramas de árvore

Em diversos problemas de contagem e probabilidade, uma configuração denominada *diagrama de árvore* pode ser usada para representar graficamente todas as possibilidades. O diagrama de árvore associado ao Exemplo 2.18 é exibido na Figura 2.7. Partindo de um ponto do lado esquerdo do diagrama, para cada primeiro elemento possível de um par origina-se um segmento de reta direcionado para a direita. Cada uma dessas retas é denominada ramo de primeira geração. Para cada ramo de primeira geração criamos outro segmento de reta originando-se na extremidade do ramo, para cada escolha possível do segundo elemento do par. Cada segmento de reta é um ramo de segunda geração. Como há quatro obstetras, há quatro ramos de primeira geração e três pediatras para cada obstetra resultam em três ramos de segunda geração, originando-se de cada ramo de primeira geração.



**Figura 2.7** Diagrama de árvore do Exemplo 2.18

Suponha que, genericamente, haja  $n_1$  ramos de primeira geração e para cada um deles haja  $n_2$  ramos de segunda geração. O número total de ramos de segunda geração será então  $n_1 n_2$ . Já que a extremidade de cada ramo de segunda geração corresponde a exatamente um par possível (escolher um primeiro elemento e então um segundo nos coloca exatamente na extremidade de um ramo de segunda geração), há pares  $n_1 n_2$ , o que comprova a regra do produto.

A construção de um diagrama de árvore não depende de ter o mesmo número de ramos de segunda geração originando-se de cada ramo de primeira geração. Se a segunda clínica tivesse quatro pediatras, haveria apenas três ramos partindo de dois dos ramos de primeira geração e quatro partindo de cada um dos dois outros ramos de primeira geração. Um diagrama de árvore pode então ser usado para representar graficamente experimentos diferentes daqueles a que se aplica a regra do produto.

### Uma regra do produto mais abrangente

Se um dado de seis lados é lançado cinco vezes consecutivas em vez de apenas duas, cada resultado possível é um conjunto de cinco números como (1, 3, 1, 2, 4) ou (6, 5, 2, 2, 2). Denominaremos um conjunto ordenado de  $k$  objetos uma ***k*-tupla** (assim um par é uma 2-tupla e um trio é uma 3-tupla). Cada resultado de experimento de lançamento de dados é uma 5-tupla.

#### Regra do produto para $k$ -tuplas

Suponha que um conjunto consista em conjuntos ordenados de  $k$  elementos ( $k$ -tuplas) e que haja  $n_1$  escolhas possíveis para o primeiro elemento. Para cada escolha do primeiro elemento, há  $n_2$  escolhas possíveis do segundo elemento . . . ; para cada escolha possível dos primeiros  $k - 1$  elementos, há  $n_k$  escolhas do  $k^{\circ}$  elemento. Então há  $n_1 n_2 \cdots n_k$  resultados possíveis  $k$ -tuplas.

Essa regra mais geral também pode ser ilustrada por um diagrama de árvore. Apenas construa um diagrama mais elaborado, adicionando ramos de terceira geração partindo de cada extremidade de um ramo de segunda geração e assim por diante, até que os ramos de  $k^{\circ}$  geração sejam adicionados.

**Exemplo 2.19** (continuação do Exemplo 2.17)

Suponha que o trabalho de reforma da casa envolva a compra de diversos utensílios de cozinha. Eles serão comprados do mesmo fornecedor e há cinco deles disponíveis na área. Com os fornecedores representados por  $D_1, \dots, D_5$ , há  $N = n_1 n_2 n_3 = (5)(12)(9) = 540$  3-tuplas da forma  $(D_i, P_j, Q_k)$ , havendo 540 formas de escolher primeiro um fornecedor de eletrodomésticos, depois um encanador e por fim um electricista. ■

**Exemplo 2.20** (continuação do Exemplo 2.18)

Se cada clínica tiver três especialistas em medicina interna e dois cirurgiões gerais, haverá  $n_1 n_2 n_3 n_4 = (4)(3)(3)(2) = 72$  formas de selecionar um médico de cada tipo, de maneira que todos os médicos sejam da mesma clínica. ■

## Arranjos e Permutações

Até agora foram selecionados os elementos sucessivos de uma  $k$ -tupla a partir de conjuntos totalmente diferentes (por exemplo: fornecedores de utensílios de cozinha, encanadores e eletricitistas). Em diversos lançamentos de um dado, o conjunto a partir do qual os elementos sucessivos são escolhidos é sempre  $\{1, 2, 3, 4, 5, 6\}$ , mas as escolhas são feitas “com reposição”, de forma que o mesmo elemento pode aparecer mais de uma vez. Consideremos um conjunto determinado que consiste em  $n$  elementos diferentes e suponhamos que um  $k$ -tupla seja formado pela seleção sucessiva a partir do conjunto *sem reposição*, de modo que um elemento possa aparecer uma única vez nas  $k$  posições possíveis.

### DEFINIÇÃO

Qualquer seqüência ordenada de  $k$  objetos selecionados de um conjunto de  $n$  objetos distintos é denominada **arranjo** de tamanho  $k$  dos objetos. O número de arranjos de tamanho  $k$  que pode ser criado a partir dos  $n$  objetos é representado por  $A_{k,n}$ .

O número de arranjos de tamanho  $k$  é obtido imediatamente pela regra geral do produto. O primeiro elemento pode ser escolhido de  $n$  formas diversas; para cada uma das  $n$  formas, o segundo elemento pode ser escolhido de  $n - 1$  formas e assim por diante. Finalmente, para cada forma de seleção dos primeiros  $k - 1$  elementos, o  $k$ -ésimo elemento pode ser escolhido de  $n - (k - 1) = n - k + 1$  formas. Então

$$A_{k,n} = n(n - 1)(n - 2) \cdots (n - k + 2)(n - k + 1)$$

**Exemplo 2.21**

Há 10 assistentes de professores disponíveis para correção de provas em um determinado curso. O primeiro exame consiste em quatro questões e o professor deseja selecionar um assistente diferente para corrigir cada uma (apenas um assistente por questão). De quantas formas diferentes os assistentes podem ser escolhidos para a correção? Aqui,  $n =$  número de assistentes  $= 10$  e  $k =$  número de questões  $= 4$ . O número de atribuições diferentes de correção então será  $A_{4,10} = (10)(9)(8)(7) = 5040$ . ■

A utilização da notação fatorial permite que  $A_{k,n}$  seja expresso de forma mais compacta.

### DEFINIÇÃO

Para qualquer inteiro positivo  $m$ ,  $m!$  é lido “fatorial de  $m$ ” e definido por  $m! = m(m - 1) \cdots (2)(1)$ . Além disso,  $0! = 1$ .

Usando a notação fatorial,  $(10)(9)(8)(7) = (10)(9)(8)(7)(6!)/6! = 10!/6!$ . De forma mais geral,

$$\begin{aligned} A_{k,n} &= n(n - 1) \cdots (n - k + 1) \\ &= \frac{n(n - 1) \cdots (n - k + 1)(n - k)(n - k - 1) \cdots (2)(1)}{(n - k)(n - k - 1) \cdots (2)(1)} \end{aligned}$$

que se transforma em

$$P_{k,n} = \frac{n!}{(n - k)!}$$

Por exemplo:  $A_{3,9} = 9!/(9 - 3)! = 9!/6! = 9 \cdot 8 \cdot 7 \cdot 6!/6! = 9 \cdot 8 \cdot 7$ . Observe também que, como  $0! = 1$ ,  $0! = 1$ ,  $P_{n,n} = n!/(n - n)! = n!/0! = n!/1 = n!$ , conforme mencionado.

## Combinações

Há muitos problemas de contagem em que é fornecido um conjunto de  $n$  objetos distintos e se deseja contar o número de subconjuntos não-ordenados de tamanho  $k$ . Por exemplo: no *bridge* importam apenas as 13 cartas da mão e não a ordem em que elas são dadas. Na formação de um comitê, a ordem em que os membros são relacionados normalmente não tem importância.

### DEFINIÇÃO

Dado um conjunto de  $n$  objetos diferentes, qualquer subconjunto não-ordenado de tamanho  $k$  é denominado **combinação**. O número de combinações de tamanho  $k$  que podem ser formadas a partir de  $n$  objetos distintos é representado por  $\binom{n}{k}$ . (Essa notação é mais comum na probabilidade do que  $C_{k,n}$  que seria análogo à notação de permutações.)

O número de combinações de tamanho  $k$  de um determinado conjunto é menor que o número de arranjos porque, quando a ordem não importa, diversos arranjos correspondem à mesma combinação. Considere, por exemplo, o conjunto  $\{A, B, C, D, E\}$ , consistindo em cinco elementos. Há  $5!/(5-3)! = 60$  arranjos de tamanho 3. Há seis arranjos de tamanho 3 que consistem nos elementos  $A, B$  e  $C$ , já que esses três elementos podem ser ordenados de  $3 \cdot 2 \cdot 1 = 3! = 6$  formas:  $(A, B, C)$ ,  $(A, C, B)$ ,  $(B, A, C)$ ,  $(B, C, A)$ ,  $(C, A, B)$  e  $(C, B, A)$ . Esses seis arranjos são equivalentes a uma única combinação  $\{A, B, C\}$ . De forma similar, para qualquer outra combinação de tamanho 3, há  $3!$  arranjos a cada uma obtida pela ordenação dos três objetos. Dessa forma,

$$60 = A_{3,5} = \binom{5}{3} \cdot 3!; \quad \text{então} \quad \binom{5}{3} = \frac{60}{3!} = 10$$

Essas dez combinações são

$\{A, B, C\}$   $\{A, B, D\}$   $\{A, B, E\}$   $\{A, C, D\}$   $\{A, C, E\}$   $\{A, D, E\}$ ,  $\{B, C, D\}$   
 $\{B, C, E\}$   $\{B, D, E\}$   $\{C, D, E\}$

Quando há  $n$  objetos diferentes, qualquer arranjo de tamanho  $k$  é obtido pela ordenação dos  $k$  objetos não-ordenados de uma combinação em uma de  $k!$  formas, de modo que a quantidade de arranjos seja o produto de  $k!$  e do número de combinações. Isso fornece

$$\binom{n}{k} = \frac{P_{k,n}}{k!} = \frac{n!}{k!(n-k)!}$$

Observe que  $\binom{n}{n} = 1$  e  $\binom{n}{0} = 1$ , já que há apenas uma forma de escolhermos um conjunto de (todos)  $n$  elementos ou de nenhum elemento, e  $\binom{n}{1} = n$ , já que há  $n$  subconjuntos de tamanho 1.

### Exemplo 2.22

Uma mão de *bridge* consiste em quaisquer 13 cartas selecionadas de um baralho de 52 cartas sem importar a ordem. Há  $\binom{52}{13} = 52!/13!39!$  diferentes mãos de *bridge*, o que dá cerca de 635 bilhões. Como há 13 cartas em cada naipe, o número de mãos formadas unicamente por paus e/ou espadas (sem cartas vermelhas) é  $\binom{26}{13} = 26!/13!13! = 10,400,600$ . Uma dessas  $\binom{26}{13}$  mãos é formada totalmente por espadas e outra por paus, de forma que há  $[\binom{26}{13} - 2]$  mãos formadas totalmente por espadas e paus na mão. Suponha que uma mão de *bridge* seja dada de um baralho bem embaralhado (ou seja, 13 cartas selecionadas aleatoriamente entre as 52 possibilidades) e seja

$A = \{ \text{a mão formada inteiramente por espadas e paus com os dois naipes} \}$

$B = \{ \text{a mão formada por exatamente dois naipes} \}.$

Os  $N = \binom{52}{13}$  resultados possíveis são igualmente prováveis, de forma que

$$P(A) = \frac{N(A)}{N} = \frac{\binom{26}{13} - 2}{\binom{52}{13}} = 0,0000164$$

Como há  $\binom{4}{2} = 6$  combinações formadas por dois naipes, das quais espadas e paus são tal combinação,

$$P(B) = \frac{6 \left[ \binom{26}{13} - 2 \right]}{\binom{52}{13}} = 0,0000983$$

Ou seja, uma mão totalmente formada por cartas de exatamente dois dos quatro naipes ocorre cerca de uma vez a cada 10.000. Se você joga *bridge* apenas uma vez por mês, é provável que nunca receba uma dessas mãos. ■

### Exemplo 2.23

Um armazém de universidade recebeu uma entrega de 25 impressoras, das quais 10 são impressoras a laser e 15 são a jato de tinta. Se 6 das 25 forem selecionadas aleatoriamente para serem verificadas por um técnico, qual será a probabilidade de que exatamente 3 delas sejam a laser (sendo as outras 3 a jato de tinta)?

Represente por  $D_3 = \{\text{exatamente 3 das 6 selecionadas são impressoras a jato de tinta}\}$ . Assumindo que qualquer conjunto específico de 6 impressoras tem a mesma chance de ser escolhido em relação a qualquer outro conjunto de 6, temos resultados igualmente prováveis, de forma que  $P(D_3) = N(D_3)/N$ , onde  $N$  é o número de formas de escolha das 6 impressoras dentre as 25 e  $N(D_3)$  é o número de formas de seleção de 3 impressoras a laser e 3 a jato de tinta. Assim,  $N = \binom{25}{6}$ . Para obter  $N(D_3)$ , pense em primeiro escolher 3 dos 15 modelos a jato de tinta e então 3 impressoras a laser. Há  $\binom{15}{3}$  formas de escolha de 3 modelos a jato de tinta e  $\binom{10}{3}$  formas de escolha de 3 impressoras a laser. Nesse caso,  $N(D_3)$  é o produto desses dois números (visualize um diagrama de árvore – estamos realmente usando a regra do produto), de forma que

$$P(D_3) = \frac{N(D_3)}{N} = \frac{\binom{15}{3}\binom{10}{3}}{\binom{25}{6}} = \frac{15!}{3!12!} \cdot \frac{10!}{3!7!} = \frac{25!}{6!19!} = 0,3083$$

Represente por  $D_4 = \{\text{exatamente 4 das 6 impressoras selecionadas são modelos a jato de tinta}\}$  e defina  $D_5$  e  $D_6$  de forma análoga. Então a probabilidade de ao menos 3 impressoras a jato de tinta serem selecionadas é

$$\begin{aligned} P(D_3 \cup D_4 \cup D_5 \cup D_6) &= P(D_3) + P(D_4) + P(D_5) + P(D_6) \\ &= \frac{\binom{15}{3}\binom{10}{3}}{\binom{25}{6}} + \frac{\binom{15}{4}\binom{10}{2}}{\binom{25}{6}} + \frac{\binom{15}{5}\binom{10}{1}}{\binom{25}{6}} + \frac{\binom{15}{6}\binom{10}{0}}{\binom{25}{6}} = 0,8530 \quad \blacksquare \end{aligned}$$

---

**Exercícios** | Seção 2.3 (29–44)

- 29.** O conselho de estudantes de engenharia de certa faculdade possui um aluno representante de cada uma das áreas de engenharia (civil, elétrica, produção de materiais e mecânica). De quantas formas é possível:
- Selecionar um presidente e um vice-presidente?
  - Selecionar um presidente, um vice-presidente e um secretário?
  - Selecionar dois membros para o conselho da reitoria?
- 30.** Um amigo meu vai oferecer um jantar. Sua adega inclui 8 garrafas de zinfandel, 10 de merlot e 12 de cabernet (ele só toma vinho tinto), todos de vinícolas diferentes.
- Se ele quiser servir 3 garrafas de zinfandel e a ordem para servir for importante, de quantas formas pode fazê-lo?
  - Se forem selecionadas 6 garrafas de vinho aleatoriamente entre as 30 disponíveis para servir, quantas formas há de selecioná-las?
  - Se forem selecionadas 6 garrafas aleatoriamente, de quantas formas será possível selecionar duas garrafas de cada variedade?
  - Se forem selecionadas 6 garrafas aleatoriamente, qual a probabilidade de serem escolhidas duas garrafas de cada variedade?
  - Se forem selecionadas 6 garrafas aleatoriamente, qual a probabilidade de todas serem do mesmo tipo?
- 31. a.** Beethoven escreveu 9 sinfonias e Mozart escreveu 27 concertos para piano. Se a estação de rádio de uma universidade desejar tocar primeiro uma sinfonia de Beethoven e depois um concerto de Mozart, de quantas formas pode ser feita a escolha?
- b.** O gerente da rádio decide que em cada noite sucessiva (7 dias por semana) será tocada uma sinfonia de Beethoven, seguida por um concerto de piano de Mozart, seguido por um quarteto de cordas de Schubert (de um total de 15). Por cerca de quantos anos essa política pode continuar, antes que exatamente o mesmo programa seja repetido?
- 32.** Uma cadeia de lojas de aparelhos de som oferece um preço especial para o conjunto completo de componentes (aparelho de som, CD-player, alto-falantes, tocafitas). O comprador tem a seguinte opção de fabricante para cada componente:
- 
- Aparelho de som: Kenwood, Onkyo, Pioneer, Sony, Sherwood  
 CD-player: Onkyo, Pioneer, Sony, Technics  
 Alto-falantes: Boston, Infinity, Polk  
 Toca-fitas: Onkyo, Sony, Teac, Technics
- 
- Um painel de ligação da loja permite ao cliente conectar qualquer seleção de componentes (um de cada tipo). Use a regra do produto para responder as seguintes perguntas:
- De quantas formas um componente de cada tipo pode ser selecionado?
  - De quantas formas os componentes podem ser selecionados se o aparelho de som e o CD-player forem Sony?
  - De quantas formas os componentes podem ser selecionados se nenhum for Sony?
  - De quantas formas pode ser feita uma seleção se ao menos um componente Sony for incluído?
  - Se os interruptores forem ligados de forma totalmente aleatória, qual será a probabilidade de o sistema selecionado conter ao menos um componente Sony? E exatamente um componente Sony?
- 33.** Logo após terem sido colocados em serviço, alguns ônibus fabricados por uma determinada empresa apresentaram trincas na parte inferior do chassi. Suponha que uma cidade tenha 25 desses ônibus e que haja trincas em 8 deles.
- Há quantas formas de selecionar uma amostra de 5 ônibus dos 25 para uma inspeção completa?
  - De quantas formas uma amostra de 5 ônibus pode conter exatamente 4 com trincas visíveis?
  - Se uma amostra de 5 ônibus for selecionada ao acaso, qual é a probabilidade de exatamente 4 dos 5 apresentarem trincas visíveis?
  - Se os ônibus forem selecionados como na parte (c), qual é a probabilidade de que pelo menos 4 dos ônibus selecionados tenham trincas visíveis?
- 34.** Uma instalação de produção emprega 20 operários no turno diurno, 15 operários no noturno e 10 operários no da madrugada. Um consultor de controle de qualidade deve selecionar 6 desses operários para entrevistas detalhadas. Suponha que a seleção seja feita de tal forma que qualquer grupo específico de 6 operários tenha a mesma possibilidade de ser selecionado que qualquer outro grupo (escolha de 6 nomes sem reposição entre 45).
- Quantas escolhas contêm 6 operários do turno diurno? Qual é a probabilidade de os 6 operários selecionados serem do turno diurno?
  - Qual é a probabilidade de os 6 operários selecionados serem do mesmo turno?
  - Qual é a probabilidade de pelo menos dois turnos diferentes terem representantes entre os operários selecionados?
  - Qual é a probabilidade de pelo menos um dos turnos não ter representante na amostra de operários?
- 35.** Um departamento acadêmico com cinco membros docentes restringiu sua escolha para chefe do departamento ao candidato *A* ou candidato *B*. Cada membro então votou em um dos candidatos em uma cédula.

- Suponha que haja três votos para o candidato  $A$  e dois para o candidato  $B$ . Se as cédulas forem selecionadas para contagem em ordem aleatória, qual é a probabilidade de  $A$  se manter à frente de  $B$  na contagem dos votos (por exemplo: esse evento ocorre se a ordem selecionada for  $AABAB$ , mas não se for  $ABBAA$ )?
- 36.** Um pesquisador está estudando os efeitos da temperatura, da pressão e do tipo de catalisador no resultado de certa reação química. Três temperaturas diferentes, quatro pressões diferentes e cinco catalisadores diferentes são considerados.
- Se cada teste envolver a utilização de uma única temperatura, pressão e catalisador, quantos testes serão possíveis?
  - Quantos testes envolvem o uso da menor temperatura e de duas menores pressões?
- 37.** Refira-se ao o exercício 36 e suponha que cinco testes diferentes sejam feitos no primeiro dia de experiência. Se os cinco forem selecionados aleatoriamente dentre todas as possibilidades, de forma que qualquer grupo de cinco tenha a mesma probabilidade de escolha, qual a probabilidade de que um catalisador diferente seja usado em cada teste?
- 38.** Uma caixa em um depósito contém quatro lâmpadas de 40W, cinco de 60W e seis de 75W. Suponha que três lâmpadas sejam selecionadas aleatoriamente.
- Qual a probabilidade de que exatamente duas das lâmpadas selecionadas sejam de 75W?
  - Qual a probabilidade de que as três lâmpadas selecionadas tenham a mesma potência?
  - Qual a probabilidade de que uma lâmpada de cada tipo seja selecionada?
  - Suponha que as lâmpadas sejam selecionadas uma a uma até que seja encontrada uma de 75W. Qual a probabilidade de que seja necessário examinar pelo menos seis lâmpadas?
- 39.** Foram recebidos quinze telefones em uma assistência técnica autorizada. Cinco deles são celulares, cinco são sem fio e os outros cinco são telefones com fio. Suponha que esses componentes sejam numerados aleatoriamente de 1, 2, ..., a 15 para definir a ordem em que serão consertados.
- Qual é a probabilidade de que todos os telefones sem fio estejam entre os primeiros 10 a serem consertados?
  - Qual é a probabilidade de que após consertar 10 telefones, somente dois dos três tipos ainda estejam com conserto pendente?
  - Qual é a probabilidade de que dois telefones de cada tipo estejam entre os primeiros seis a serem consertados?
- 40.** Três moléculas do tipo  $A$ , três do tipo  $B$ , três do tipo  $C$  e três do tipo  $D$  serão vinculadas uma à outra para formar uma cadeia molecular. Uma molécula desse tipo é  $ABCDABCDABCD$  e outra é  $BCDDAAABDBCC$ .
- Quantas moléculas desse tipo podem ser formadas? (*Sugestão:* se as três moléculas  $A$  forem diferentes uma da outra  $-A_1, A_2$  e  $A_3$  – assim como as  $B, C$  e  $D$ , quantas moléculas haverá? Como esse número será reduzido se não houver distinção entre as moléculas  $A$ ?)
  - Suponha que uma molécula composta do tipo descrito seja selecionada aleatoriamente. Qual é a probabilidade de que todas as moléculas de cada tipo estejam juntas (como em  $BBBAAADDCC$ )?
- 41.** Uma professora de matemática deseja marcar uma reunião com cada um de seus oito assistentes, quatro homens e quatro mulheres, para discutir o curso de cálculo. Suponha que todas as seqüências de reuniões sejam igualmente prováveis de serem escolhidas.
- Qual é a probabilidade de que pelo menos uma assistente mulher esteja entre as primeiras três pessoas com quem a professora se reunirá?
  - Qual é a probabilidade de que após as primeiras cinco reuniões ela tenha se reunido com todas as assistentes mulheres?
  - Suponha que a professora tenha os mesmos oito assistentes no semestre seguinte e marque reuniões sem considerar a ordem seguida no primeiro semestre. Qual a probabilidade de que a ordem dos compromissos seja diferente?
- 42.** Três casais compraram ingressos de teatro e estão sentados em uma fileira que consiste em apenas seis assentos. Se eles se sentarem de uma forma totalmente aleatória (ordem aleatória), qual será a probabilidade de Jim e Paula (marido e mulher) se sentarem nos dois assentos da esquerda? Qual é a probabilidade de Jim e Paula se sentarem um ao lado do outro? Qual é a probabilidade de que ao menos uma das esposas se sente ao lado de seu respectivo marido?
- 43.** No pôquer de cinco cartas, um *straight* consiste em cinco cartas com denominações adjacentes (por exemplo: 9 de paus, 10 de copas, valete de copas, dama de espadas e rei de paus). Assumindo que os ases podem ficar nas duas pontas, se você receber um jogo de cinco cartas, qual é a probabilidade de que ele será um *straight* com a carta 10 alta? Qual é a probabilidade de se receber um *straight*? Qual é a probabilidade de ser um *straight flush* (todas as cartas do mesmo naipe)?
- 44.** Mostre que  $\binom{n}{k} = \binom{n}{n-k}$ . Forneça uma interpretação envolvendo subconjuntos.



## 2.4 | Probabilidade Condicional

As probabilidades atribuídas a diversos eventos dependem do que seja conhecido sobre a situação experimental quando a atribuição é feita. Depois da atribuição inicial, informações parciais sobre ou relevantes para o resultado do experimento podem estar disponíveis. Essas informações podem fazer com que revisemos algumas atribuições de probabilidade. Para um determinado evento  $A$ , usamos  $P(A)$  para representar a probabilidade atribuída a  $A$ . Agora interpretamos  $P(A)$  como a probabilidade original ou incondicional do evento  $A$ . Nesta seção, examinamos como a informação “ocorreu um evento  $B$ ” afeta a probabilidade atribuída a  $A$ . Por exemplo:  $A$  pode se referir a um indivíduo que tenha determinada doença pela presença de alguns sintomas. Se for feito um exame de sangue no indivíduo e o resultado for negativo ( $B =$  exame de sangue negativo), então a probabilidade de ter a doença será alterada (ela deve diminuir, mas normalmente não para zero, já que os exames de sangue não são infalíveis). Usaremos a notação  $P(A | B)$  para representar a **probabilidade condicional de  $A$  dado que ocorreu o evento  $B$** .

### Exemplo 2.24

Componentes complexos são montados em uma fábrica que usa duas linhas de montagem diferentes:  $A$  e  $A'$ . A linha  $A$  usa equipamentos mais antigos que  $A'$ , de forma que é mais lenta e um pouco menos confiável. Suponha que em determinado dia, a linha  $A$  tenha montado 8 componentes, dos quais 2 foram identificados como defeituosos ( $B$ ) e 6 como não-defeituosos ( $B'$ ), ao passo que a linha  $A'$  produziu 1 defeituoso e 9 não-defeituosos. Essas informações estão resumidas na tabela a seguir.

		Condição	
		$B$	$B'$
Linha	$A$	2	6
	$A'$	1	9

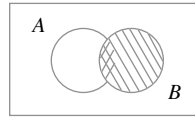
Não tendo essas informações, o gerente de vendas seleciona aleatoriamente 1 desses 18 componentes para uma demonstração. Antes da demonstração

$$P(\text{componente selecionado da linha } A) = P(A) = \frac{N(A)}{N} = \frac{8}{18} = 0,44$$

Entretanto, se o componente escolhido tiver defeito, o evento  $B$  terá ocorrido, de forma que o componente deve ser 1 dos 3 na coluna  $B$  da tabela. Como esses 3 componentes são igualmente prováveis entre si após a ocorrência de  $B$ ,

$$P(A | B) = \frac{2}{3} = \frac{\frac{2}{18}}{\frac{3}{18}} = \frac{P(A \cap B)}{P(B)} \quad (2.2) \quad \blacksquare$$

Na Equação (2.2), a probabilidade condicional é expressa como uma relação entre probabilidades não-condicionais: o numerador é a probabilidade de interseção dos dois eventos, enquanto o denominador é a probabilidade do evento de condição  $B$ . Um diagrama de Venn esclarece a relação (veja a Figura 2.8).



**Figura 2.8** Demonstração da definição de probabilidade condicional

Dado que  $B$  ocorreu, o espaço amostral relevante não é mais  $\mathcal{S}$ , mas consiste em resultados contidos em  $B$ ;  $A$  ocorreu se, e somente se, um dos resultados da interseção tiver ocorrido, de forma que a probabilidade condicional de  $A$  dado  $B$  é proporcional a  $P(A \cap B)$ . A proporção constante  $1/P(B)$  é utilizada para assegurar que a probabilidade  $P(B|B)$  do novo espaço amostral de  $B$  seja igual a 1.

### Definição de probabilidade condicional

O Exemplo 2.24 demonstra que, quando os resultados são igualmente prováveis, o cálculo de probabilidades condicionais pode se basear na intuição. Quando os experimentos são mais complicados, a intuição pode nos enganar, portanto queremos uma definição geral de probabilidade condicional que forneça respostas intuitivas em problemas simples. O diagrama de Venn e a Equação (2.2) indicam a definição apropriada.

DEFINIÇÃO

Para quaisquer dois eventos  $A$  e  $B$  com  $P(B) > 0$ , a **probabilidade condicional de  $A$  dado que ocorreu  $B$**  é definida por

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{2.3}$$

#### Exemplo 2.25

Suponha que, de todos os indivíduos que compram uma determinada câmera digital, 60% incluem um cartão de memória opcional na compra, 40% incluem uma pilha extra e 30% incluem um cartão e uma pilha. Considere a seleção aleatória de um comprador e sejam  $A = \{\text{compra de cartão de memória}\}$  e  $B = \{\text{compra de pilha}\}$ . Dessa forma,  $P(A) = 0,60$ ,  $P(B) = 0,40$  e  $P(\text{compra de ambos}) = P(A \cap B) = 0,30$ . Dado que o indivíduo selecionado comprou uma pilha extra, a probabilidade de compra de um cartão opcional é

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0,30}{0,40} = 0,75$$

Isto é, de todos os que compraram uma pilha extra, 75% compraram um cartão de memória extra. De forma análoga,

$$P(\text{pilha} | \text{cartão de memória}) = P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{0,30}{0,60} = 0,50$$

Observe que  $P(A|B) \neq P(A)$  e  $P(B|A) \neq P(B)$ . ■

#### Exemplo 2.26

Uma revista publica três colunas, intituladas “Arte” ( $A$ ), “Livros” ( $B$ ) e “Cinema” ( $C$ ). Os hábitos de leitura de um leitor selecionado aleatoriamente em relação a essas colunas são:

<i>Lê regularmente</i>	$A$	$B$	$C$	$A \cap B$	$A \cap C$	$B \cap C$	$A \cap B \cap C$
<i>Probabilidade</i>	0,14	0,23	0,37	0,08	0,09	0,13	0,05

(Veja a Figura 2.9.)

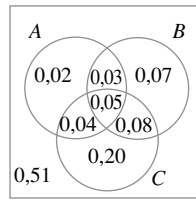


Figura 2.9 Diagrama de Venn do Exemplo 2.26

Assim, temos:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0,08}{0,23} = 0,348$$

$$P(A|B \cup C) = \frac{P(A \cap (B \cup C))}{P(B \cup C)} = \frac{0,04 + 0,05 + 0,03}{0,47} = \frac{0,12}{0,47} = 0,255$$

$$P(A | \text{pelo menos uma}) = P(A|A \cup B \cup C) = \frac{P(A \cap (A \cup B \cup C))}{P(A \cup B \cup C)}$$

$$= \frac{P(A)}{P(A \cup B \cup C)} = \frac{0,14}{0,49} = 0,286$$

e

$$P(A \cup B|C) = \frac{P((A \cup B) \cap C)}{P(C)} = \frac{0,04 + 0,05 + 0,08}{0,37} = 0,459$$

### A regra da multiplicação para $P(A \cap B)$

A definição de probabilidade condicional fornece o resultado a seguir, obtido pela multiplicação de ambos os lados da equação (2.3) por  $P(B)$ .

A Regra da multiplicação

$$P(A \cap B) = P(A|B) \cdot P(B)$$

Essa regra é importante porque freqüentemente se deseja  $P(A \cap B)$ , ao passo que  $P(B)$  e  $P(A|B)$  podem ser especificados pela descrição do problema. Considerando a expressão  $P(B|A)$  obtém-se  $P(A \cap B) = P(B|A) \cdot P(A)$ .

#### Exemplo 2.27

Quatro indivíduos responderam a uma solicitação de um banco de sangue para doação. Nenhum deles doou sangue antes, de forma que seus tipos sanguíneos são desconhecidos. Suponha que apenas o tipo O+ seja desejado e apenas um dos quatro indivíduos tenha esse tipo sanguíneo. Se os doadores potenciais forem selecionados em ordem aleatória para determinação do tipo sanguíneo, qual será a probabilidade de que pelo menos três indivíduos tenham de ser testados para obtenção do tipo desejado?

Fazendo  $B = \{\text{primeiro tipo não O+}\}$  e  $A = \{\text{segundo tipo não O+}\}$ ,  $P(B) = \frac{3}{4}$ . Dado que o primeiro tipo não é O+, dois dos três indivíduos restantes não serão O+, de forma que  $P(A|B) = \frac{2}{3}$ . A regra de multiplicação agora fornece:

$$\begin{aligned}
 P(\text{pelo menos três indivíduos foram testados}) &= P(A \cap B) \\
 &= P(A | B) \cdot P(B) \\
 &= \frac{2}{3} \cdot \frac{3}{4} = \frac{6}{12} \\
 &= 0,5
 \end{aligned}$$

■

A regra da multiplicação é mais útil quando o experimento consiste em diversas etapas consecutivas. O evento condicional  $B$  então descreve o resultado da primeira etapa e  $A$  descreve o resultado da segunda, de forma que  $P(A | B)$  – condicionado ao que ocorre primeiro – normalmente é conhecida. A regra é facilmente estendida a experimentos que envolvem mais de duas etapas. Por exemplo:

$$\begin{aligned}
 P(A_1 \cap A_2 \cap A_3) &= P(A_3 | A_1 \cap A_2) \cdot P(A_1 \cap A_2) \\
 &= P(A_3 | A_1 \cap A_2) \cdot P(A_2 | A_1) \cdot P(A_1)
 \end{aligned} \tag{2.4}$$

onde  $A_1$  ocorre primeiro, seguido por  $A_2$  e finalmente por  $A_3$ .

### Exemplo 2.28

Para o experimento de definição de tipo sanguíneo do Exemplo 2.27,

$$\begin{aligned}
 P(\text{terceiro tipo é O+}) &= P(\text{terceiro é } | \text{primeiro não é} \cap \text{segundo não é}) \\
 &\quad \cdot P(\text{segundo não é} | \text{primeiro não é}) \cdot P(\text{primeiro não é}) \\
 &= \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} = \frac{1}{4} = 0,25
 \end{aligned}$$

■

Quando o experimento consistir uma seqüência de diversas etapas, é conveniente representá-las em um diagrama de árvore. Depois que tivermos um diagrama de árvore apropriado, as probabilidades e as probabilidades condicionais são inseridas nos diversos ramos; o que é feito com o uso da regra da multiplicação de forma direta.

### Exemplo 2.29

Uma cadeia de lojas de vídeo vende três marcas diferentes de videocassetes. Dessas vendas, 50% são da marca 1 (a mais barata), 30% são da marca 2 e 20% são da marca 3. Cada fabricante oferece um ano de garantia para peças e mão-de-obra. É sabido que 25% dos videocassetes da marca 1 necessitam de reparos de garantia, enquanto os percentuais correspondentes para as marcas 2 e 3 são 20% e 10%, respectivamente.

1. Qual é a probabilidade de que um comprador selecionado aleatoriamente compre um videocassete da marca 1 que precise de reparo durante a garantia?
2. Qual é a probabilidade de que um comprador selecionado aleatoriamente possua um aparelho que necessite de reparos durante a garantia?
3. Se um cliente voltar à loja com um videocassete que precise de reparos em garantia, qual é a probabilidade de ele ser da marca 1? E da marca 2? E da marca 3?

A primeira etapa do problema envolve um cliente que escolhe uma de três marcas de videocassete. Seja  $A_i = \{\text{compra da marca } i\}$ , para  $i = 1, 2$  e  $3$ . Então  $P(A_1) = 0,50$ ,  $P(A_2) = 0,30$  e  $P(A_3) = 0,20$ . Depois que a marca do videocassete for selecionada, a segunda etapa envolve a observação da necessidade de reparo na garantia. Com  $B = \{\text{precisa de reparo}\}$  e  $B' = \{\text{não precisa de reparo}\}$ , as informações fornecidas implicam  $P(B | A_1) = 0,25$ ,  $P(B | A_2) = 0,20$ , e  $P(B | A_3) = 0,10$ .

O diagrama de árvore que representa essa situação experimental é exibido na Figura 2.10. Os ramos iniciais correspondem a diferentes marcas de videocassetes e há ramos de segunda geração que partem da extremidade de cada ramo inicial: um para “precisa de reparo” e outro para “não precisa de reparo”. A probabilidade

$P(A_i)$  é exibida no ramo inicial, enquanto as probabilidades condicionais  $P(B|A_i)$  e  $P(B'|A_i)$  são exibidas nos ramos de segunda geração. À direita de cada ramo de segunda geração correspondente à ocorrência de  $B$  exibimos o produto das probabilidades nos ramos que vão até o ponto. Isto é simplesmente a regra de multiplicação em ação. A resposta à pergunta proposta em 1, portanto, é  $P(A_1 \cap B) = P(B|A_1) \cdot P(A_1) = 0,125$ . A resposta à pergunta 2 é

$$\begin{aligned} P(B) &= P[(\text{marca 1 e reparo}) \text{ ou } (\text{marca 2 e reparo}) \text{ ou } (\text{marca 3 e reparo})] \\ &= P(A_1 \cap B) + P(A_2 \cap B) + P(A_3 \cap B) \\ &= 0,125 + 0,060 + 0,020 = 0,205 \end{aligned}$$

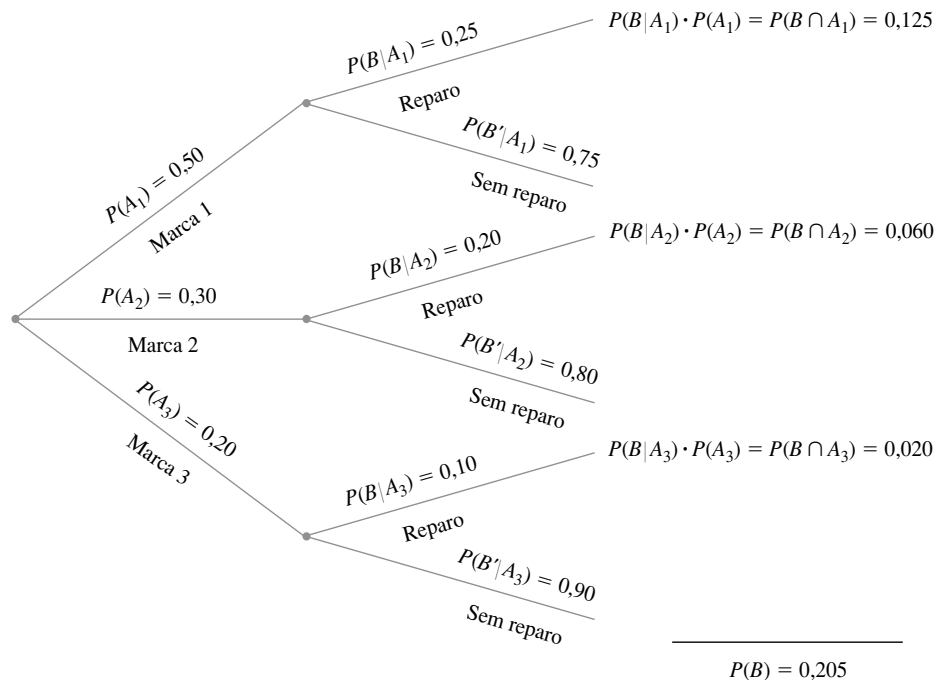


Figure 2.10 Diagrama de árvore do Exemplo 2.29

Finalmente,

$$P(A_1|B) = \frac{P(A_1 \cap B)}{P(B)} = \frac{0,125}{0,205} = 0,61$$

$$P(A_2|B) = \frac{P(A_2 \cap B)}{P(B)} = \frac{0,060}{0,205} = 0,29$$

e

$$P(A_3|B) = 1 - P(A_1|B) - P(A_2|B) = 0,10$$

Observe que a *probabilidade anterior* ou inicial da marca 1 é 0,50, enquanto, depois que se sabe que o videocassete selecionado precisa de reparos, a *probabilidade posterior* da marca 1 aumenta para 0,61. Isso ocorre porque os videocassetes da marca 1 são igualmente mais prováveis de precisar de reparos do que os das outras marcas. A probabilidade posterior da marca 3 é  $P(A_3|B) = 0,10$ , que é muito menor que a probabilidade anterior  $P(A_3) = 0,20$ . ■

## Teorema de Bayes

O cálculo de uma probabilidade posterior  $P(A_j|B)$  a partir de probabilidades anteriores dadas  $P(A_i)$  e das probabilidades condicionais  $P(B|A_i)$  ocupa uma posição central em teoria elementar de probabilidade. A regra geral

desses cálculos, que na verdade é uma aplicação simples da regra de multiplicação, remete ao reverendo Thomas Bayes, que viveu no século XVIII. Para expressá-la, precisamos de outro resultado. Lembre-se de que os eventos  $A_1, \dots, A_k$  são mutuamente exclusivos se nenhum par deles tiver resultados comuns. Os eventos são *exaustivos* se um  $A_i$  tiver de ocorrer para que  $A_1 \cup \dots \cup A_k = \mathcal{S}$ .

#### Lei ou Teorema da probabilidade total

Sejam  $A_1, \dots, A_k$  eventos mutuamente exclusivos e exaustivos. Então, para qualquer outro evento  $B$ ,

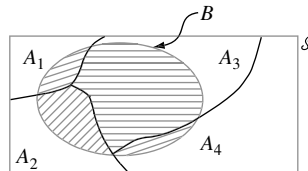
$$\begin{aligned} P(B) &= P(B|A_1)P(A_1) + \dots + P(B|A_k)P(A_k) \\ &= \sum_{i=1}^k P(B|A_i)P(A_i) \end{aligned} \quad (2.5)$$

#### Demonstração

Visto que os eventos  $A_i$  são mutuamente exclusivos e exaustivos, se  $B$  ocorrer será com exatamente um dos  $A_i$ , isto é,  $B = (A_1 \text{ e } B) \text{ ou } \dots \text{ ou } (A_k \text{ e } B) = (A_1 \cap B) \cup \dots \cup (A_k \cap B)$ , onde os eventos  $(A_i \cap B)$  são mutuamente exclusivos. Essa “partição de  $B$ ” é ilustrada na Figura 2.11. Assim,

$$P(B) = \sum_{i=1}^k P(A_i \cap B) = \sum_{i=1}^k P(B|A_i)P(A_i)$$

conforme desejado.



**Figura 2.11** Partição de  $B$  pelos eventos mutuamente exclusivos e exaustivos  $A_i$  ■

Um exemplo do uso da Equação (2.5) foi mostrado na resposta à pergunta 2 do Exemplo 2.29, onde  $A_1 = \{\text{marca 1}\}$ ,  $A_2 = \{\text{marca 2}\}$ ,  $A_3 = \{\text{marca 3}\}$  e  $B = \{\text{reparo}\}$ .

#### Teorema de Bayes

Sejam  $A_1, A_2, \dots, A_k$  uma coleção de  $k$  eventos mutuamente exclusivos e exaustivos com  $P(A_i) > 0$  para  $i = 1, \dots, k$ . Então, para qualquer outro evento  $B$  em que  $P(B) > 0$ ,

$$P(A_j|B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^k P(B|A_i) \cdot P(A_i)} \quad j = 1, \dots, k \quad (2.6)$$

A transição da segunda para a terceira expressão em (2.6) baseia-se na aplicação da regra da multiplicação no numerador e da lei da probabilidade total no denominador.

A proliferação de eventos e índices na expressão (2.6) pode intimidar usuários com pouca experiência em probabilidade. Desde que haja relativamente poucos eventos na partição, pode-se usar um gráfico de árvore

(como no Exemplo 2.29) como base para calcular as probabilidades posteriores sem referência explícita ao teorema de Bayes.

### Exemplo 2.30

*Incidência de doença rara.* Apenas 1 em 1000 adultos é acometido por uma doença rara para a qual foi desenvolvido um teste de diagnóstico. O teste funciona de tal forma que, se o indivíduo tiver a doença, o resultado do teste será positivo em 99% das vezes e, se não a tiver, será positivo em apenas 2% das vezes. Se um indivíduo selecionado aleatoriamente for testado e o resultado for positivo, qual é a probabilidade de ele ter a doença?

Para usar o teorema de Bayes, considere por  $A_1 = \{\text{indivíduo tem a doença}\}$ ,  $A_2 = \{\text{indivíduo não tem a doença}\}$  e  $B = \{\text{resultado do teste positivo}\}$ . Então,  $P(A_1) = 0,001$ ,  $P(A_2) = 0,999$ ,  $P(B|A_1) = 0,99$ , e  $P(B|A_2) = 0,02$ . O diagrama de árvore deste problema está na Figura 2.12.

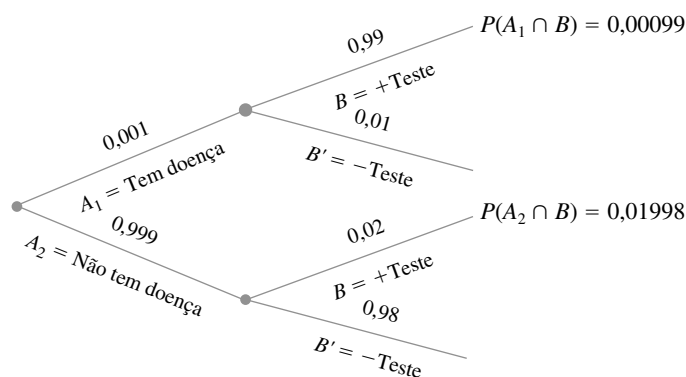


Figura 2.12 Diagrama de árvore para o problema da doença rara

Próximo a cada ramo correspondente a um resultado positivo, a regra de multiplicação indica as probabilidades registradas. Portanto,  $P(B) = 0,00099 + 0,01998 = 0,02097$ , pelo qual obtemos

$$P(A_1|B) = \frac{P(A_1 \cap B)}{P(B)} = \frac{0,00099}{0,02097} = 0,047$$

O resultado parece ir de encontro à intuição. O resultado do teste parece tão preciso que esperamos que alguém com um resultado positivo tenha alta probabilidade de ter a doença, enquanto a probabilidade condicional calculada é de apenas 0,047. Entretanto, como a doença é rara e o teste é moderadamente confiável, a maior parte dos resultados positivos dos testes provém de erros e não de indivíduos doentes. A probabilidade de ter a doença foi aumentada por um fator multiplicativo 47 (da probabilidade anterior 0,001 para a posterior 0,047); mas, para se obter maior aumento na probabilidade posterior, é necessário um teste de diagnóstico com menores taxas de erros. Se a doença não fosse tão rara (por exemplo, 25% de incidência na população), as taxas de erros do presente teste forneceriam um bom diagnóstico. ■

## Exercícios | Seção 2.4 (45–67)

45. A população de um país consiste em três grupos étnicos. Cada indivíduo pertence a um de quatro grupos sanguíneos principais. A *tabela de probabilidade conjunta* fornece as proporções de indivíduos nas diversas combinações de grupos étnicos-grupos sanguíneos.

		Grupo sanguíneo			
		O	A	B	AB
Grupo étnico	1	0,082	0,106	0,008	0,004
	2	0,135	0,141	0,018	0,006
	3	0,215	0,200	0,065	0,020

Suponha que um indivíduo seja selecionado aleatoriamente da população e defina os eventos  $A = \{\text{tipo A selecionado}\}$ ,  $B = \{\text{tipo B selecionado}\}$  e  $C = \{\text{grupo étnico 3 selecionado}\}$ .

- a. Calcule  $P(A)$ ,  $P(C)$ , e  $P(A \cap C)$ .
- b. Calcule  $P(A|C)$  e  $P(C|A)$  e explique, no contexto, o que significa cada uma dessas probabilidades.
- c. Se o indivíduo selecionado não tiver o tipo sanguíneo B, qual é a probabilidade de que ele ou ela seja do grupo étnico 1?

**46.** Suponha que um indivíduo seja selecionado aleatoriamente da população de todos os homens adultos que vivem nos Estados Unidos. Sejam  $A$  o evento em que o indivíduo selecionado tem mais de 1,80m de altura e  $B$  o evento em que o indivíduo selecionado é um jogador profissional de basquete. Qual você pensa ser maior:  $P(A|B)$  ou  $P(B|A)$ ? Por quê?

**47.** Volte ao cenário dos cartões de crédito do Exercício 12 (Seção 2.2), onde  $A = \{\text{Visa}\}$ ,  $B = \{\text{MasterCard}\}$ ,  $P(A) = 0,5$ ,  $P(B) = 0,4$  e  $P(A \cap B) = 0,25$ . Calcule e interprete cada uma das probabilidades a seguir (um diagrama de Venn pode ser útil).

- a.  $P(B|A)$       b.  $P(B'|A)$
- c.  $P(A|B)$       d.  $P(A'|B)$
- e. Dado que o indivíduo selecionado tem ao menos um cartão de crédito, qual é a probabilidade de ele ter um cartão Visa?

**48.** Reconsidere a situação de defeitos descrita no Exercício 26 (Seção 2.2).

- a. Dado que o sistema tem um defeito do tipo 1, qual é a probabilidade de ele ter um defeito do tipo 2?
- b. Dado que o sistema tem um defeito do tipo 1, qual é a probabilidade de ele ter os três tipos de defeitos?
- c. Dado que o sistema tem ao menos um tipo de defeito, qual é a probabilidade de ele ter exatamente um tipo de defeito?
- d. Dado que o sistema tem os dois primeiros tipos de defeitos, qual é a probabilidade de ele não ter o terceiro tipo de defeito?

**49.** Se duas lâmpadas forem selecionadas aleatoriamente da caixa do Exercício 38 (Seção 2.3) e pelo menos uma delas for de 75W, qual é a probabilidade de que ambas gerem 75W? Dado que pelo menos uma das lâmpadas não seja de 75W, qual é a probabilidade de ambas as lâmpadas selecionadas terem a mesma potência?

**50.** Uma loja de departamentos vende camisas esportivas em três tamanhos (pequeno, médio e grande), três padrões (xadrez, estampado e listrado) e dois comprimentos de manga (curta e longa). As tabelas a seguir fornecem as proporções de camisas vendidas nas diversas combinações de categorias.

**Manga curta**

Tamanho	Padrão		
	Xadrez	Estampado	Listrado
<b>P</b>	0,04	0,02	0,05
<b>M</b>	0,08	0,07	0,12
<b>G</b>	0,03	0,07	0,08

**Manga longa**

Tamanho	Padrão		
	Xadrez	Estampado	Listrado
<b>P</b>	0,03	0,02	0,03
<b>M</b>	0,10	0,05	0,07
<b>G</b>	0,04	0,02	0,08

- a. Qual é a probabilidade de a próxima camisa vendida ser média, de mangas longas e estampada?
- b. Qual é a probabilidade de a próxima camisa vendida ser uma camisa média, estampada?
- c. Qual é a probabilidade de a próxima camisa vendida ser de mangas curtas? E de mangas longas?
- d. Qual é a probabilidade de a próxima camisa vendida ser média? E de ser estampada?
- e. Dado que a última camisa vendida foi uma xadrez de mangas curtas, qual é a probabilidade de o tamanho ser médio?
- f. Dado que a última camisa vendida foi uma média e xadrez, qual é a probabilidade de as mangas serem curtas? E de serem longas?

**51.** Uma caixa contém seis bolas vermelhas e três verdes e uma segunda caixa contém sete bolas vermelhas e três verdes. Uma bola é retirada da primeira caixa e colocada na segunda. Então uma bola é retirada da segunda caixa e colocada na primeira.

- a. Qual é a probabilidade de uma bola vermelha ser selecionada na primeira caixa e outra bola vermelha na segunda?
- b. No fim do processo de seleção, qual é a probabilidade de o número de bolas vermelhas e verdes da primeira e da segunda caixas ser idêntico ao do início?

**52.** Um sistema consiste em duas bombas idênticas, nº 1 e nº 2. Se uma bomba falha, o sistema continua em operação. Entretanto, devido ao esforço adicional, a bomba restante passa a estar mais suscetível a falhas do que originalmente. Isto é,  $r = P(\text{nº 2 falha} | \text{nº 1 falha}) > P(\text{nº 2 falha}) = q$ . Se pelo menos uma das bombas falhar durante a vida útil do sistema 7% das vezes e ambas as bombas falharem durante aquele período em apenas 1% das vezes, qual é a probabilidade de a bomba nº 1 falhar durante a vida útil do sistema?

**53.** Certa loja faz reparos em componentes de áudio e vídeo. Represente por  $A$  o evento em que o próximo



componente trazido para conserto seja de áudio e por  $B$  o evento em que o próximo componente seja um CD-player (de forma que o evento  $B$  está contido em  $A$ ). Suponha que  $P(A) = 0,6$  e  $P(B) = 0,05$ . Qual é  $P(B|A)$ ?

54. No Exercício 13,  $A_i = \{\text{projeto escolhido } i\}$ , para  $i = 1, 2, 3$ . Use as probabilidades fornecidas para calcular as probabilidades a seguir:

- $P(A_2|A_1)$
- $P(A_2 \cap A_3|A_1)$
- $P(A_2 \cup A_3|A_1)$
- $P(A_1 \cap A_2 \cap A_3|A_1 \cup A_2 \cup A_3)$ .

Expresse em palavras a probabilidade calculada.

55. No Exercício 42, seis pessoas (três casais) escolheram assentos aleatoriamente em uma fileira formada por seis assentos.

- Use a regra da multiplicação para calcular a probabilidade de Jim e Paula se sentarem juntos à esquerda (evento  $A$ ) e de John e Mary Lou (marido e mulher) se sentarem juntos no meio (evento  $B$ ).
- Dado que John e Mary Lou sentam-se juntos no meio, qual é a probabilidade de os outros dois maridos se sentarem junto a suas esposas?
- Dado que John e Mary Lou sentam-se juntos, qual é a probabilidade de todos os maridos se sentarem junto a suas esposas?

56. Se  $P(B|A) > P(B)$ , mostre que  $P(B'|A) < P(B')$ . (*Sugestão:* Adicione  $P(B'|A)$  aos dois lados da inequação fornecida e então use o resultado do Exercício 57.)

57. Para quaisquer eventos  $A$  e  $B$  com  $P(B) > 0$ , mostre que  $P(A|B) + P(A'|B) = 1$ .

58. Demonstre que para quaisquer três eventos  $A$ ,  $B$ , e  $C$  com  $P(C) > 0$ ,  $P(A \cup B|C) = P(A|C) + P(B|C) - P(A \cap B|C)$ .

59. Em determinado posto de gasolina, 40% dos clientes usam gasolina comum ( $A_1$ ), 35% usam gasolina aditivada ( $A_2$ ) e 25% usam gasolina *premium* ( $A_3$ ). Dos clientes que usam gasolina comum, apenas 30% enchem o tanque (evento  $B$ ). Dos clientes que usam gasolina aditivada e *premium*, 60% enchem o tanque, enquanto dentre os que usam *premium*, 50% enchem o tanque.

- Qual é a probabilidade de o próximo cliente pedir gasolina aditivada e encher o tanque ( $A_2 \cap B$ )?
- Qual é a probabilidade de o próximo cliente encher o tanque?
- Se o próximo cliente encher o tanque, qual é a probabilidade de pedir gasolina comum? E gasolina aditivada? E gasolina *premium*?

60. Setenta por cento das aeronaves leves que desaparecem em vôo em certo país são localizadas posteriormente. Das aeronaves localizadas, 60% possuem localizador de emergência, enquanto 90% das aeronaves não localizadas não possuem esse dispositivo. Suponha que uma aeronave leve tenha desaparecido.

- Se ela tiver localizador de emergência, qual é a probabilidade de não ser localizada?
- Se ela não tiver localizador de emergência, qual é a probabilidade de ser localizada?

61. Componentes de determinado tipo são enviados a um fornecedor em lotes de 10. Suponha que 50% desses lotes não contenham componentes defeituosos, 30% contenham um componente defeituoso e 20% contenham dois componentes defeituosos. Dois componentes de um lote são selecionados aleatoriamente e testados. Quais são as probabilidades associadas a 0, 1 e 2 componentes defeituosos estarem no lote em cada uma das condições a seguir?

- Nenhum componente testado apresenta defeito.
- Um de dois componentes testados apresenta defeito. (*Sugestão:* Desenhe um diagrama de árvore com três ramos de primeira geração para os três tipos diferentes de lotes.)

62. Uma empresa que fabrica câmeras de vídeo produz um modelo básico e um modelo luxo. No ano passado, 40% das câmeras vendidas foram do modelo básico. Dos clientes que compraram o modelo básico, 30% compraram uma garantia estendida, enquanto 50% dos compradores do modelo luxo a compraram. Se você souber que um cliente selecionado aleatoriamente possui uma garantia estendida, qual a probabilidade de ele ter o modelo básico?

63. Considere os eventos a seguir para clientes que compram um conjunto de pneus em uma determinada loja, considere os eventos

$A = \{\text{os pneus comprados foram fabricados nos Estados Unidos}\}$

$B = \{\text{o comprador fez o balanceamento dos pneus imediatamente}\}$

$C = \{\text{o comprador solicita alinhamento dianteiro e traseiro}\}$

em conjunto com  $A'$ ,  $B'$ , e  $C'$ . Assuma as probabilidades condicionais e não-condicionais a seguir:

$$P(A) = 0,75 \quad P(B|A) = 0,9 \quad P(B|A') = 0,8$$

$$P(C|A \cap B) = 0,8 \quad P(C|A \cap B') = 0,6$$

$$P(C|A' \cap B) = 0,7 \quad P(C|A' \cap B') = 0,3$$

- Construa um diagrama de árvore formado por ramos de primeira, segunda e terceira gerações, identifique cada evento e a probabilidade apropriada próximo a cada ramo.
- Calcule  $P(A \cap B \cap C)$ .
- Calcule  $P(B \cap C)$ .
- Calcule  $P(C)$ .
- Calcule  $P(A|B \cap C)$ , a probabilidade de uma compra de pneus americanos, dado que são solicitados alinhamento e balanceamento.

64. No Exemplo 2.30, suponha que a taxa de incidência da doença seja de 1 em 25, em vez de 1 em 1000. Qual

será então a probabilidade de um resultado positivo do teste? Dado que o resultado é positivo, qual a probabilidade de o indivíduo ter a doença? Dado um resultado negativo, qual a probabilidade de o indivíduo não ter a doença?

65. Em uma grande universidade, na pesquisa por um livro-texto satisfatório, o departamento de estatística testou um livro diferente durante os últimos três trimestres. No trimestre do outono, 500 alunos usaram o livro do professor Média; durante o inverno, 300 alunos usaram o livro do professor Mediana e durante a primavera, 200 alunos usaram o livro do professor Moda. Uma pesquisa no final de cada trimestre mostrou que 200 estavam satisfeitos com o livro do Média, 150 estavam satisfeitos com o livro do Mediana e 160 estavam satisfeitos com o livro do Moda. Se um aluno que estudou estatística durante um dos trimestres for selecionado aleatoriamente e admitir estar satisfeito com o livro-texto, ele terá mais probabilidade de ter usado o livro do Média, do Mediana ou do Moda? Quem é o autor igualmente menos provável? (*Sugestão*: Desenhe um diagrama de árvore ou use o teorema de Bayes.)
66. Uma amiga que vive em Los Angeles faz viagens frequentes de consultoria para Washington, D.C. Cinquenta por cento das vezes ela viaja pela linha aérea nº 1, 30 % das vezes pela nº 2 e os 20 % restantes das vezes pela nº 3. Na linha nº 1, os vôos para Washington atrasam 30% das vezes e para L.A., 10% das vezes. Na linha aérea nº 2, essas porcentagens são 25% e 20%, enquanto na linha aérea nº 3 são 40% e 25%. Se soubermos que em uma viagem ela chega atrasada em exatamente um dos destinos, quais são as probabilidades posteriores de ela ter voado nas linhas nº 1, nº 2 e nº 3? (*Sugestão*: Na extremidade de cada ramo de primeira geração em um diagrama de árvore, desenhe três ramos de segunda geração denominados 0 atrasado, 1 atrasado e 2 atrasado.)
67. No Exercício 59, considere as informações adicionais a seguir sobre a utilização de cartões de crédito:
- 70% de todos os clientes da gasolina comum que enchem o tanque usam um cartão de crédito.
  - 50% de todos os clientes da gasolina comum que não enchem o tanque usam um cartão de crédito.
  - 60% de todos os clientes da gasolina extra que enchem o tanque usam um cartão de crédito.
  - 50% de todos os clientes da gasolina extra que não enchem o tanque usam um cartão de crédito.
  - 50% de todos os clientes da gasolina *premium* que enchem o tanque usam um cartão de crédito.
  - 40% de todos os clientes da gasolina *premium* que não enchem o tanque usam um cartão de crédito.
- Calcule a probabilidade de cada um dos eventos a seguir para o próximo cliente a ser atendido (um diagrama de árvore pode ajudar).
- a. {extra, tanque cheio e cartão de crédito}
  - b. {*premium*, não tanque cheio e cartão de crédito}
  - c. {*premium* e cartão de crédito}
  - d. {tanque cheio e cartão de crédito}
  - e. {cartão de crédito}
  - f. se o próximo cliente usar cartão de crédito, qual a probabilidade de que seja pedida a gasolina *premium*?

## 2.5 | Independência

A definição de probabilidade condicional nos permite revisar a probabilidade  $P(A)$  originalmente atribuída a  $A$  quando somos informados posteriormente da ocorrência de outro evento  $B$ . A nova probabilidade de  $A$  é  $P(A|B)$ . Em nossos exemplos, normalmente ocorria de  $P(A|B)$  ser diferente da probabilidade não-condicional  $P(A)$ , indicando que a informação “ $B$  ocorreu” resultou em uma alteração da possibilidade de ocorrência de  $A$ . Há outras situações em que a possibilidade de ocorrência de  $A$  não é afetada pelo conhecimento da ocorrência de  $B$ , de forma que  $P(A|B) = P(A)$ . É natural, então, pensar em  $A$  e  $B$  como eventos independentes, indicando que a ocorrência ou não-ocorrência desses eventos não afeta a possibilidade de ocorrer o outro evento.

### DEFINIÇÃO

Dois eventos  $A$  e  $B$  são **independentes** se  $P(A|B) = P(A)$  e **dependentes** em caso contrário.

A definição de independência pode parecer “assimétrica” porque não se exige também que  $P(B|A) = P(B)$ . Entretanto, usando a definição de probabilidade condicional e a regra da multiplicação,

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A|B)P(B)}{P(A)} \quad (2.7)$$

O lado direito da Equação (2.7) é  $P(B)$  se, e somente se,  $P(A|B) = P(A)$  (independência), de forma que a igualdade da definição implica outra igualdade (e vice-versa). Também é fácil demonstrar que, se  $A$  e  $B$  forem independentes, os seguintes pares de eventos também o serão: (1)  $A'$  e  $B$ , (2)  $A$  e  $B'$ , e (3)  $A'$  e  $B'$ .

### Exemplo 2.31

Considere o lançamento de um dado uma vez e defina os eventos  $A = \{2, 4, 6\}$ ,  $B = \{1, 2, 3\}$ , e  $C = \{1, 2, 3, 4\}$ . Temos então  $P(A) = \frac{1}{2}$ ,  $P(A|B) = \frac{1}{3}$  e  $P(A|C) = \frac{1}{2}$ . Isto é, os eventos  $A$  e  $B$  são dependentes, enquanto os eventos  $A$  e  $C$  são independentes. Intuitivamente, se esse dado for lançado e formos informados de que o resultado foi 1, 2, 3 ou 4 ( $C$  ocorreu), a probabilidade de  $A$  ter ocorrido é  $\frac{1}{2}$ , conforme era inicialmente, já que dois dos quatro resultados relevantes são pares e são igualmente prováveis. ■

### Exemplo 2.32

Sejam  $A$  e  $B$  dois eventos mutuamente exclusivos com  $P(A) > 0$ . Por exemplo: para um automóvel escolhido aleatoriamente, seja  $A = \{\text{o carro possui quatro cilindros}\}$  e  $B = \{\text{o carro possui seis cilindros}\}$ . Como os eventos são mutuamente exclusivos, se  $B$  ocorrer,  $A$  não pode ter ocorrido, de forma que  $P(A|B) = 0 \neq P(A)$ . A mensagem aqui é: *se dois eventos forem mutuamente exclusivos, não podem ser independentes*. Quando  $A$  e  $B$  são mutuamente exclusivos, a informação da ocorrência de  $A$  indica algo sobre  $B$  (não pode ter ocorrido), de forma que a independência é excluída. ■

## $P(A \cap B)$ Quando os eventos são independentes

Freqüentemente a natureza de um experimento indica que dois eventos  $A$  e  $B$  devem ser considerados independentes. Este é o caso, por exemplo, de um fabricante que recebe uma placa de circuito de cada um de dois fornecedores diferentes; cada placa é testada em sua chegada e  $A = \{\text{a primeira tem defeito}\}$  e  $B = \{\text{a segunda tem defeito}\}$ . Se  $P(A) = 0,1$ , também deve ser o caso de  $P(A|B) = 0,1$ , sabendo que a condição da segunda placa não fornece informações sobre as condições da primeira. Nosso próximo resultado mostra como calcular  $P(A \cap B)$  quando os eventos são independentes.

#### PROPOSIÇÃO

$A$  e  $B$  são independentes se, e somente se,

$$P(A \cap B) = P(A) \cdot P(B) \quad (2.8)$$

Parafraseando a proposição,  $A$  e  $B$  são eventos independentes se, e somente se, a probabilidade de ambos ocorrerem ( $A \cap B$ ) for o produto de duas probabilidades individuais. Segue a demonstração:

$$P(A \cap B) = P(A|B) \cdot P(B) = P(A) \cdot P(B) \quad (2.9)$$

onde a segunda igualdade da Equação (2.9) é válida se, e somente se,  $A$  e  $B$  forem independentes. Por causa da equivalência de independência pela Equação (2.8), esta última pode ser usada como definição de independência.

**Exemplo 2.33**

Sabe-se que 30% das lavadoras de roupa de uma determinada empresa requerem manutenção enquanto estiverem na garantia e somente 10% das secadoras de roupa precisam de manutenção. Se alguém comprar uma lavadora e uma secadora de roupas feitas por essa empresa, qual é a probabilidade de que ambas as máquinas precisem de conserto?

Seja  $A$  o evento em que a lavadora precisa de manutenção em garantia e  $B$  o evento análogo para a secadora. Então  $P(A) = 0,30$  e  $P(B) = 0,10$ . Assumindo que as máquinas funcionem de forma independente uma da outra, a probabilidade desejada será

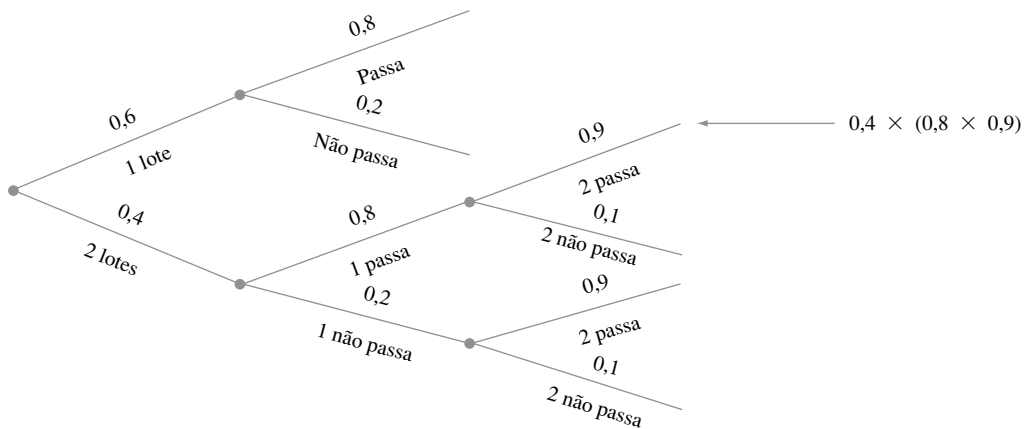
$$P(A \cap B) = P(A) \cdot P(B) = (0,30)(0,10) = 0,03$$

A probabilidade de que nenhuma das máquinas precise de serviço é

$$P(A' \cap B') = P(A') \cdot P(B') = (0,70)(0,90) = 0,63$$

**Exemplo 2.34**

Cada dia, de segunda a sexta-feira, chega em uma determinada instalação de inspeção um lote de componentes enviado por um fornecedor. Duas vezes por semana, chega um lote de um segundo fornecedor. Oitenta por cento de todos os lotes do fornecedor 1 passam na inspeção, assim como 90% dos lotes do fornecedor 2. Qual é a probabilidade de, em um dia selecionado aleatoriamente, dois lotes passarem na inspeção? Responderemos a essa questão assumindo que, em dias em que os dois lotes são testados, os fatos de o primeiro ou de o segundo lote passarem ou não na inspeção são independentes. A Figura 2.13 exibe as informações relevantes.



**Figura 2.13** Diagrama de árvore do Exemplo 2.34

$$\begin{aligned} P(\text{ambos passam}) &= P(\text{dois recebidos} \cap \text{ambos passam}) \\ &= P(\text{ambos passam} \mid \text{dois recebidos}) \cdot P(\text{dois recebidos}) \\ &= [(0,8)(0,9)](0,4) = 0,288 \end{aligned}$$

**Independência de mais de dois eventos**

A noção de independência de dois eventos pode ser estendida para coleções de mais de dois eventos. Apesar de ser possível estender a definição para dois eventos independentes trabalhando em termos de probabilidades condicionais e não-condicionais, é mais direto e menos trabalhoso seguir a linha da última proposição.

**DEFINIÇÃO**

Os eventos  $A_1, \dots, A_n$  são **mutuamente independentes** se para cada  $k$  ( $k = 2, 3, \dots, n$ ) e cada subconjunto de índices  $i_1, i_2, \dots, i_k$ ,

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \dots \cdot P(A_{i_k}).$$

Parafraseando a definição, os eventos serão mutuamente independentes se a probabilidade da interseção de qualquer subconjunto dos  $n$  eventos for igual ao produto das probabilidades individuais. Como foi o caso de dois eventos, freqüentemente especificamos a independência de certos eventos no início do problema. A definição pode então ser usada para calcular a probabilidade de uma interseção.

### Exemplo 2.35

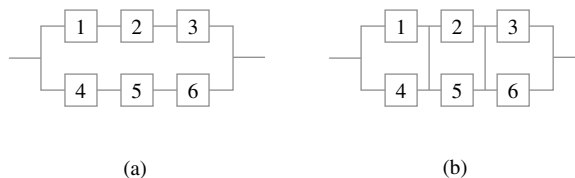
O artigo “Reliability Evaluation of Solar Photovoltaic Arrays” (*Solar Energy*, 2002, p. 129-141) apresenta diversas configurações de matrizes fotovoltaicas solares formados por células solares de silício cristalino. Considere o sistema ilustrado na Figura 2.14 (a). Há dois subsistemas ligados em paralelo, cada um com três células. Para que o sistema funcione, pelo menos dois subsistemas paralelos devem funcionar. Em cada subsistema, as três células estão ligadas em série, de forma que um subsistema só funciona se todas as suas células estiverem funcionando. Considere a vida útil  $t_0$  e suponha que se queira determinar a probabilidade da vida útil do sistema exceder  $t_0$ . Seja  $A_i$  o evento vida útil da célula  $i$  excede  $t_0$  ( $i = 1, 2, \dots, 6$ ). Assumimos que os eventos  $A_i$  sejam independentes (se alguma célula específica durar mais de  $t_0$  horas não tem nenhuma relação com a duração das outras células) e que  $P(A_i) = 0,9$  para cada  $i$ , já que as células são idênticas. Então

$$\begin{aligned} P(\text{vida útil do sistema excede } t_0) &= P[(A_1 \cap A_2 \cap A_3) \cup (A_4 \cap A_5 \cap A_6)] \\ &= P(A_1 \cap A_2 \cap A_3) + P(A_4 \cap A_5 \cap A_6) \\ &\quad - P[(A_1 \cap A_2 \cap A_3) \cap (A_4 \cap A_5 \cap A_6)] \\ &= (0,9)(0,9)(0,9) + (0,9)(0,9)(0,9) - (0,9)(0,9)(0,9)(0,9)(0,9)(0,9) = 0,927 \end{aligned}$$

De forma alternativa,

$$\begin{aligned} P(\text{vida útil do sistema excede } t_0) &= 1 - P(\text{a vida útil de ambos os subsistemas são } \leq t_0) \\ &= 1 - [P(\text{vida útil do subsistema é } \leq t_0)]^2 \\ &= 1 - [1 - P(\text{vida útil do subsistema é } > t_0)]^2 \\ &= 1 - [1 - (0,9)^3]^2 = 0,927 \end{aligned}$$

Em seguida, considere o sistema cruzado da Figura 2.14(b), obtido do arranjo série-paralelo pela ligação das colunas de junção.



**Figura 2.14** Configurações de sistema do Exemplo 2.35: (a) série-paralelo; (b) cruzado

Neste caso o sistema falhará se houver falha de uma coluna, e a vida útil do sistema excederá  $t_0$  apenas se a vida útil de cada coluna também exceder essa configuração,

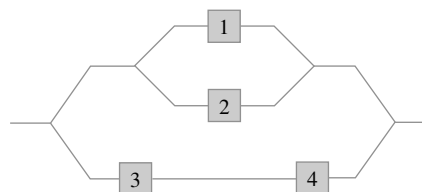
$$\begin{aligned} P(\text{vida útil do sistema é no mínimo } t_0) &= [P(\text{vida útil da coluna excede } t_0)]^3 \\ &= [1 - P(\text{vida útil da coluna é } \leq t_0)]^3 \\ &= [1 - P(\text{ambas as células da coluna têm vida útil } \leq t_0)]^3 \\ &= [1 - (1 - 0,9)^2]^3 = 0,970 \end{aligned}$$

**Exercícios** | Seção 2.5 (68–87)

68. Reconsidere o cenário de cartões de crédito do Exercício 47 (Seção 2.4) e mostre que  $A$  e  $B$  são dependentes primeiro usando a definição de independência e depois demonstrando que não se aplica a regra da multiplicação.
69. Uma empresa de exploração de petróleo possui dois projetos ativos, um na Ásia e outro na Europa. Sejam por  $A$  o evento em que o projeto da Ásia tem sucesso e  $B$  o evento em que o projeto da Europa tem sucesso. Suponha que  $A$  e  $B$  sejam eventos independentes com  $P(A) = 0,4$  e  $P(B) = 0,7$ .
- Se o projeto da Ásia não obtiver sucesso, qual é a probabilidade de o projeto da Europa também não obtê-lo? Explique seu raciocínio.
  - Qual é a probabilidade de pelo menos um dos dois projetos ter sucesso?
  - Dado que pelo menos um dos dois projetos obteve sucesso, qual é a probabilidade de apenas o projeto da Ásia ter sucesso?
70. No Exercício 13, há algum  $A_i$  independente de qualquer outro  $A_j$ ? Responda, usando a propriedade da multiplicação para eventos independentes.
71. Se  $A$  e  $B$  forem eventos independentes, mostre que  $A'$  e  $B$  também são independentes. [*Sugestão*: Primeiro defina a relação entre  $P(A' \cap B)$ ,  $P(B)$ , e  $P(A \cap B)$ .]
72. Suponha que as proporções de fenótipos sanguíneos em uma população sejam as seguintes:
- | A    | B    | AB   | O    |
|------|------|------|------|
| 0,42 | 0,10 | 0,04 | 0,44 |
- Assumindo que os fenótipos de dois indivíduos selecionados aleatoriamente sejam independentes um do outro, qual é a probabilidade de que ambos os fenótipos sejam O? Qual é a probabilidade de que os fenótipos de dois indivíduos selecionados aleatoriamente sejam iguais?
73. Uma das hipóteses básicas da teoria de gráficos de controle (consulte o Capítulo 16) é a que pontos locados sucessivamente sejam independentes uns dos outros. Cada ponto locado indica que o processo de fabricação está operando corretamente ou que há algum tipo de problema. Mesmo quando o processo está funcionando bem, há uma pequena probabilidade de que um determinado ponto indique um problema no processo. Suponha que essa probabilidade seja 0,05. Qual é a probabilidade de que pelo menos um de 10 pontos sucessivos indique um problema quando na verdade o processo opera corretamente? Responda esta pergunta para 25 pontos sucessivos.
74. A probabilidade de um aluno errar a marcação de uma pergunta específica num exame de múltipla escolha é 0,1. Se houver 10 perguntas e elas forem marcadas de

forma independente, qual é a probabilidade de não ser cometido nenhum erro? E de ser cometido pelo menos um erro? Se houver  $n$  perguntas e a probabilidade de um erro de marcação for  $p$  em vez de 0,1, forneça expressões para as duas probabilidades.

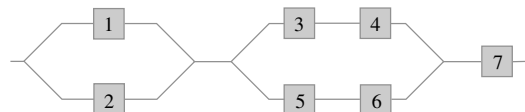
75. A junta de uma aeronave requer 25 rebites. A junta terá que ser refeita se qualquer um dos rebites estiver defeituoso. Suponha que os defeitos dos rebites sejam independentes um do outro e tenham a mesma probabilidade.
- Se 20% de todas as juntas tiverem que ser refeitas, qual será a probabilidade de um rebite ter defeito?
  - Quão pequena deve ser a probabilidade de um rebite ter defeito para garantir que apenas 10% das juntas tenham que ser refeitas?
76. Uma caldeira tem cinco válvulas de alívio idênticas. A probabilidade de uma válvula específica ser aberta sob demanda é de 0,95. Assumindo operação independente das válvulas, calcule  $P$ (pelo menos uma válvula é aberta) e  $P$ (pelo menos uma válvula tem falha ao abrir).
77. Duas bombas ligadas em paralelo falham independentemente uma da outra em um determinado dia. A probabilidade de que apenas a bomba mais antiga apresente falha é de 0,10 e a probabilidade de que apenas a bomba mais nova o faça é de 0,05. Qual é a probabilidade de que o sistema de bombas apresente uma falha em um determinado dia (que acontece se houver falha em ambas as bombas)?
78. Considere o sistema de componentes ligados como na figura a seguir. Os componentes 1 e 2 estão ligados em paralelo, de forma que o subsistema funciona se e, somente se, 1 ou 2 funcionar. Como 3 e 4 estão ligados em série, o subsistema funcionará se, e somente se, 3 e 4 funcionarem. Se os componentes funcionarem independentemente um do outro e  $P$ (componente funciona) = 0,9, calcule  $P$ (sistema funciona).



79. Refira-se à configuração de sistema série-paralelo apresentada no Exemplo 2.35 e suponha que haja apenas duas células em vez de três em cada subsistema paralelo [na Figura 2.14(a), elimine as células 3 e 6 e renuncie as células 4 e 5 como 3 e 4]. Usando  $P(A_i) = 0,9$ , a probabilidade de que a vida útil do sistema exceda  $t_0$  é facilmente vista ser 0,9639. Para que valor 0,9 teria de ser alterado para aumentar a confiabilidade

da vida útil do sistema de 0,9639 para 0,99? (*Sugestão*: Seja  $P(A_i) = p$ , expresse a confiabilidade do sistema em termos de  $p$  e assuma  $x = p^2$ ).

- 80.** Considere o lançamento independente de dois dados, um verde e um vermelho. Sejam  $A$  o evento em que o dado vermelho mostra 3 pontos,  $B$  o evento em que o dado verde mostra 4 pontos e  $C$  o evento em que o número total de pontos mostrados nos dois dados é 7. Esses eventos são independentes aos pares (ou seja,  $A$  e  $B$  são eventos independentes,  $A$  e  $C$  também e  $B$  e  $C$  também)? Os três eventos são mutuamente independentes?
- 81.** Os componentes que chegam em um distribuidor são verificados em busca de defeitos por dois inspetores diferentes (cada componente é verificado por ambos os inspetores). O primeiro inspetor detecta 90% de todos os componentes defeituosos, assim como o segundo inspetor. Pelo menos um inspetor não detecta 20% de todos os componentes defeituosos. Qual é a probabilidade de ocorrência dos itens a seguir?
- Um componente com defeito ser detectado apenas pelo primeiro inspetor? Por exatamente um dos inspetores?
  - Os três componentes defeituosos de um lote passarem pelos dois inspetores sem detecção de defeito (assumindo que as inspeções de diferentes componentes sejam independentes uma da outra)?
- 82.** Setenta por cento de todos os veículos inspecionados em um determinado posto de inspeção de emissões passam na inspeção. Assumindo que veículos sucessivos passam ou não passam independentemente um do outro, calcule as seguintes probabilidades:
- $P$ (os três próximos veículos inspecionados passam)
  - $P$ (pelo menos um dos três próximos veículos inspecionados falha)
  - $P$ (exatamente um dos três próximos veículos inspecionados passa)
  - $P$ (no máximo um dos três próximos veículos passa)
  - Dado que ao menos um dos três próximos veículos passa na inspeção, qual é a probabilidade de que todos os três passem (uma probabilidade condicional)?
- 83.** Um inspetor de controle de qualidade pesquisa defeitos em itens produzidos. O inspetor pesquisa as falhas do item numa série de fixações independentes, cada uma com duração fixa. Dado que existe uma falha, seja  $p$  a probabilidade de a falha ser detectada durante qualquer fixação (esse modelo é discutido em "Human Performance in Sampling Inspection," *Human Factors*, 1979, p. 99-105).
- Assumindo que um item tenha uma falha, qual é a probabilidade de ele ser detectado até o final da segunda fixação (depois da detecção de uma falha, as fixações são interrompidas)?
  - Forneça uma expressão da probabilidade de que uma falha será detectada até o final da  $n$ ésima fixação.
  - Se, quando uma falha não for detectada em três fixações, o item passar, qual será a probabilidade de que um item com falha passe na inspeção?
- Suponha que 10% de todos os itens contenham uma falha [ $P$ (item escolhido aleatoriamente apresenta falha) = 0,1]. Com a suposição da parte (c), qual é a probabilidade de um item com falha passar na inspeção (ele passará automaticamente se não tiver falha, mas também pode passar se tiver)?
  - Dado que um item passou na inspeção (nenhuma falha em três fixações), qual é a probabilidade de ele possuir uma falha? Calcule para  $p = 0,5$ .
- 84. a.** Uma serraria recebe um lote de 10.000 tábuas  $2 \times 4$ . Suponha que 20% dessas tábuas (2.000) na verdade estejam muito verdes para serem usadas em construção de primeira qualidade. Duas tábuas são selecionadas aleatoriamente, uma após a outra. Sejam  $A = \{a \text{ primeira tábua está verde}\}$  e  $B = \{a \text{ segunda tábua está verde}\}$ . Calcule  $P(A)$ ,  $P(B)$  e  $P(A \cap B)$  (um diagrama de árvore é útil).  $A$  e  $B$  são independentes?
- Com  $A$  e  $B$  independentes e  $P(A) = P(B) = 0,2$ , qual é  $P(A \cap B)$ ? Qual é a diferença entre essa resposta e  $P(A \cap B)$  na parte (a)? Para fins de cálculo de  $P(A \cap B)$ , podemos assumir que  $A$  e  $B$  da parte (a) sejam independentes para obter unicamente a probabilidade correta?
  - Suponha que o lote consista em 10 tábuas, das quais duas estão verdes. A hipótese de independência agora apresenta aproximadamente a resposta correta para  $P(A \cap B)$ ? Qual é a principal diferença entre a situação descrita aqui e a da parte (a)? Quando você considera que a hipótese de independência seria válida para a obtenção de uma resposta aproximadamente correta para  $P(A \cap B)$ ?
- 85.** Consulte as hipóteses do Exercício 78 e responda à questão proposta para o sistema da figura abaixo. Como a probabilidade mudaria se este fosse um subsistema ligado em paralelo ao subsistema ilustrado na Figura 2.14 (a)?



- 86.** O professor Stan der Deviation pode usar uma de duas rotas em seu caminho de volta do trabalho para casa. Na primeira rota, há quatro cruzamentos com ferrovias. A probabilidade de ele ter de parar em qualquer um dos cruzamentos é 0,1 e os trens operam de forma independente nos quatro cruzamentos. A outra rota é mais longa, mas tem apenas dois cruzamentos, independentes um do outro, com a mesma probabilidade de parada da primeira rota. Um dia, o professor Deviation tem um compromisso em casa com hora marcada. Em qualquer rota ele estima que chegará atrasado se for parado por trens em pelo menos metade dos cruzamentos encontrados.
- Que rota deve seguir para minimizar a probabilidade de chegar atrasado?

- b. Se jogar uma moeda para decidir que rota tomar e chegar atrasado, qual será a probabilidade de ele ter usado a rota com quatro cruzamentos?
87. Suponha que etiquetas idênticas tenham sido colocadas na orelha direita e na esquerda de uma raposa. A raposa é libertada por um período de tempo. Considere os dois eventos  $C_1 = \{\text{etiqueta da orelha esquerda é perdida}\}$  e  $C_2 = \{\text{etiqueta da orelha direita é per-$

dida). Seja  $\pi = P(C_1) = P(C_2)$  e assuma  $C_1$  e  $C_2$  como eventos independentes. Deduza uma expressão (envolvendo  $\pi$ ) para a probabilidade de exatamente uma etiqueta ser perdida (“Ear Tag Loss in Red Foxes,” *J. Wildlife Mgmt.*, 1976, p. 164-167). (Sugestão: Desenhe um diagrama de árvore em que os dois ramos iniciais refiram-se ao evento etiqueta da orelha esquerda é perdida).

Exercícios Suplementares (88–111)

88. Uma pequena empresa de fabricação começará a operar no turno da noite. Há 20 maquinistas empregados.
- Se o turno da noite tiver 3 maquinistas, quantas equipes podem ser formadas?
  - Se os maquinistas forem classificados de 1, 2, ..., 20 em ordem de competência, quantas equipes não teriam o melhor maquinista?
  - Quantas equipes teriam pelo menos um dos 10 melhores maquinistas?
  - Se uma dessas equipes for selecionada aleatoriamente em uma noite específica, qual é a probabilidade de o melhor maquinista não trabalhar naquela noite?
89. Uma empresa usa três linhas de produção para fabricar certo tipo de latas. A tabela fornece as porcentagens de latas fora da conformidade, categorizadas por tipo de não-conformidade, para cada uma das três linhas durante um período de tempo.

	Linha 1	Linha 2	Linha 3
<b>Mancha</b>	15	12	20
<b>Trinca</b>	50	44	40
<b>Problema em pull-tab</b>	21	28	24
<b>Defeito na superfície</b>	10	8	15
<b>Outro</b>	4	8	2

Durante esse período, a linha 1 produziu 500 latas fora de conformidade, a linha 2 produziu 400 e a linha 3, 600. Suponha que uma dessas 1500 latas seja selecionada aleatoriamente.

- Qual é a probabilidade de ela ter sido produzida pela linha 1? E que o motivo da não-conformidade seja uma trinca?
  - Se a lata selecionada tiver vindo da linha 1, qual é a probabilidade de ela ter uma mancha?
  - Dado que a lata selecionada possui um defeito na superfície, qual é a probabilidade de ter vindo da linha 1?
90. Um funcionário do escritório de registros de uma universidade possui 10 formulários em sua mesa esperando processamento. Seis deles são petições de saída e os outros quatro são solicitações de substituição de curso.

- Se ele selecionar aleatoriamente seis formulários para entregar a um subordinado, qual é a probabilidade de apenas um dos dois tipos de formulários permanecer na mesa?
  - Suponha que ele tenha tempo para processar apenas quatro desses formulários até o fim do dia. Se esses quatro forem selecionados aleatoriamente um a um, qual é a probabilidade de que cada formulário sucessivo seja de um tipo diferente de seu predecessor?
91. Um satélite tem seu lançamento previsto do cabo Canaveral, na Flórida, e o outro da Base Aérea de Vandenberg, na Califórnia. Sejam  $A$  o evento do lançamento de Vandenberg ser feito na data e  $B$  o evento análogo do cabo Canaveral. Se  $A$  e  $B$  forem eventos independentes com  $P(A) > P(B)$  e  $P(A \cup B) = 0,626$ ,  $P(A \cap B) = 0,144$ , determine os valores de  $P(A)$  e  $P(B)$ .

92. Um transmissor está enviando uma mensagem usando código binário, ou seja, uma seqüência de 0s e 1s. Cada bit transmitido (0 ou 1) deve passar por três relés para chegar ao receptor. Em cada relé há uma probabilidade de 0,20 de que o bit enviado seja feito de forma diferente do bit recebido (uma reversão). Assuma que os relés operem de forma independente um do outro.

Transmissor → Relé 1 → Relé 2 → Relé 3

→ Receptor

- Se um 1 for enviado pelo transmissor, qual será a probabilidade de um 1 ser enviado por todos os relés?
  - Se um 1 for enviado pelo transmissor, qual é a probabilidade de um 1 ser recebido pelo receptor? (Sugestão: Os oito resultados experimentais podem ser exibidos em um diagrama de árvore com três gerações de ramos, um para cada relé).
  - Suponha que 70% de todos os bits enviados do transmissor sejam 1. Se um 1 for recebido pelo receptor, qual é a probabilidade de um 1 ter sido enviado?
93. O indivíduo A tem um círculo de cinco amigos íntimos (B, C, D, E e F). A ouviu um boato de fora do círculo e convidou os cinco amigos para uma festa para transmitir o boato. Para iniciar, A seleciona um dos cinco aleatoriamente e conta o boato para ele. Esse indivíduo



então seleciona um dos quatro restantes e repete o que ouviu. Continuando, uma nova pessoa é selecionada dentre os que não ouviram o boato e assim sucessivamente até que todos saibam.

- a. Qual é a probabilidade de o boato ser repetido na ordem B, C, D, E e F?
  - b. Qual é a probabilidade de F ser a terceira pessoa a saber do boato?
  - c. Qual é a probabilidade de F ser a última pessoa a ouvir o boato?
94. Refira-se ao Exercício 93. Se em cada etapa a pessoa que “tem” o boato no momento não souber quem já ouviu e selecionar o próximo receptor aleatoriamente dentre as cinco possibilidades, qual é a probabilidade de F ainda não ter ouvido o boato depois ter sido contado 10 vezes na festa?
95. Um engenheiro químico está interessado em determinar se traços de certa impureza estão presentes em um produto. Um experimento tem probabilidade 0,80 de detectar a presença da impureza. A probabilidade de não detectar a impureza se ela não estiver presente é de 0,90. As probabilidades anteriores de a impureza estar presente e estar ausente são de 0,40 e 0,60, respectivamente. Três experimentos separados resultam em apenas duas detecções. Qual é a probabilidade posterior de a impureza estar presente?
96. Cada competidor de um programa de perguntas e respostas é solicitado a especificar uma de seis possíveis categorias que servirão de tema para as perguntas feitas. Suponha que  $P$  (competidor escolhe a categoria  $i$ ) =  $\frac{1}{6}$  e os competidores sucessivos escolham suas categorias independentemente um do outro. Se houver três competidores em cada programa e os três escolherem categorias diferentes, qual é a probabilidade de que exatamente um tenha selecionado a categoria 1?
97. Os fixadores usados na fabricação de aviões são ligeiramente tortos para evitar afrouxarem com a vibração. Suponha que 95% de todos os fixadores passem na inspeção inicial. Dos 5% com falhas, 20% possuem defeitos tão sérios que devem ser sucateados. Os fixadores restantes são enviados para retrabalho, onde 40% não podem ser salvos e são descartados. Os outros 60% são corrigidos pelo processo de ondulação e, depois, passam na inspeção.
- a. Qual é a probabilidade de um fixador selecionado aleatoriamente passar na inspeção inicialmente ou após o retrabalho?
  - b. Dado que um fixador tenha passado na inspeção, qual é a probabilidade de ele ter passado na inspeção inicial e não ter precisado de retrabalho?
98. Um por cento de todos os indivíduos de certa população são portadores de determinada doença. Um teste de diagnóstico para essa doença tem taxa de detecção de 90% para portadores e 5% para não-portadores. Suponha que o teste seja aplicado independentemente a duas amostras de sangue do mesmo indivíduo selecionado aleatoriamente.
- a. Qual é a probabilidade de ambos os testes terem o mesmo resultado?
  - b. Se ambos os testes forem positivos, qual será a probabilidade de o indivíduo selecionado ser um portador?
99. Um sistema consiste em dois componentes. A probabilidade de o segundo componente funcionar de forma satisfatória durante a vida útil do projeto é 0,9, a probabilidade de pelo menos um dos dois componentes funcionar é de 0,96 e a de ambos os componentes funcionarem é de 0,75. Dado que o primeiro componente funciona de forma satisfatória por toda a vida útil do projeto, qual é a probabilidade de o segundo também funcionar?
100. Certa empresa envia 40% de sua correspondência expressa noturna pelo serviço expresso  $E_1$ . Dessas remessas, 2% chegam após a hora de entrega garantida (represente o evento “entrega atrasada” por  $L$ ). Se um registro de um envio expresso noturno for selecionado aleatoriamente no arquivo da empresa, qual será a probabilidade de a remessa ter sido enviada via  $E_1$  e estar atrasada?
101. Refira-se ao Exercício 100. Suponha que 50% das remessas expressas sejam enviadas via serviço expresso  $E_2$  e as 10% restantes sejam enviadas por  $E_3$ . Dos enviados por  $E_2$ , apenas 1% chega com atraso, enquanto 5% das parcelas enviadas por  $E_3$  chegam com atraso.
- a. Qual é a probabilidade de uma parcela selecionada aleatoriamente chegar com atraso?
  - b. Se uma parcela selecionada aleatoriamente chegar no horário, qual será a probabilidade de ela não ter sido enviada por  $E_1$ ?
102. Uma empresa usa três linhas de montagem diferentes:  $A_1$ ,  $A_2$  e  $A_3$  para fabricar certo componente. Dos componentes fabricados pela linha  $A_1$ , 5% exigem retrabalho para corrigir um defeito, enquanto 8% dos componentes da linha  $A_2$  exigem retrabalho, assim como 10% de  $A_3$ . Suponha que 50% de todos os componentes sejam produzidos pela linha  $A_1$ , 30% por  $A_2$  e 20% por  $A_3$ . Se um componente selecionado aleatoriamente exigir retrabalho, qual será a probabilidade de ele ser proveniente da linha  $A_1$ ? Da linha  $A_2$ ? E da linha  $A_3$ ?
103. Desprezando a possibilidade de aniversário no dia 29 de fevereiro, suponha que um indivíduo selecionado aleatoriamente tenha iguais probabilidades de ter nascido em qualquer um dos outros 365 dias.
- a. Se 10 pessoas forem selecionadas aleatoriamente, qual será a probabilidade de todas terem datas de aniversário diferentes? Que ao menos duas tenham a mesma data de aniversário?
  - b. Com  $k$  no lugar de 10 na parte (a), qual é o menor  $k$  para o qual há possibilidade de 50% de duas ou mais pessoas terem a mesma data de aniversário?

- c. Se 10 pessoas forem selecionadas aleatoriamente, qual será a probabilidade de que pelo menos duas tenham a mesma data de aniversário ou que pelo menos duas tenham os mesmos três últimos dígitos do número de seguro social (ao todo são 9 dígitos)? [Nota: O artigo “Methods for Studying Coincidences” (F. Mosteller and P. Diaconis, *J. Amer. Stat. Assoc.*, 1989, p. 853-861) discute problemas desse tipo].

**104.** Um método usado para distinguir rochas graníticas ( $G$ ) e basálticas ( $B$ ) é examinar uma parte do espectro infravermelho da energia solar refletida na superfície da rocha. Sejam  $R_1, R_2$  e  $R_3$  as intensidades de espectro medidas em três diferentes comprimentos de onda. Normalmente, para granito,  $R_1 < R_2 < R_3$ , enquanto para basalto  $R_3 < R_1 < R_2$ . Quando as medidas são feitas de forma remota (usando aeronaves), podem surgir diversas ordenações de  $R_i$ s, se a rocha for basalto ou granito. Vôos sobre regiões de composição conhecida resultaram nas seguintes informações:

	Granito	Basalto
$R_1 < R_2 < R_3$	60%	10%
$R_1 < R_3 < R_2$	25%	20%
$R_3 < R_1 < R_2$	15%	70%

Suponha que, para uma rocha selecionada aleatoriamente em determinada região,  $P(\text{granito}) = 0,25$  e  $P(\text{basalto}) = 0,75$ .

- a. Mostre que  $P(\text{granito} | R_1 < R_2 < R_3) > P(\text{basalto} | R_1 < R_2 < R_3)$ . Se as medidas fornecerem  $R_1 < R_2 < R_3$ , você classificaria a rocha como granito ou como basalto?
- b. Se as medidas fornecerem  $R_1 < R_3 < R_2$ , como você classificaria a rocha? Responda à mesma questão para  $R_3 < R_1 < R_2$ .
- c. Usando as regras de classificação indicadas nas partes (a) e (b), ao selecionar uma rocha dessa região, qual é a probabilidade de uma classificação errada? [Sugestão:  $G$  pode ser classificado como  $B$  ou  $B$  como  $C$  e  $P(B)$  e  $P(G)$  são conhecidos].
- d. Se  $P(\text{granito}) = p$  em vez de 0,25, há valores de  $p$  (diferentes de 1) para os quais uma pessoa sempre classificaria uma rocha como granito?

**105.** Uma seqüência de vislumbres é fornecida para que um alvo possa ser detectado. Seja  $G_i =$  (o alvo é detectado no  $i$ -ésimo vislumbre), com  $p_i = P(G_i)$ . Suponha que os  $G_i$ s sejam eventos independentes e escreva uma expressão da probabilidade de que o alvo foi detectado no final do  $i$ -ésimo vislumbre. (Nota: Esse modelo é discutido em “Predicting Aircraft Detectability,” *Human Factors*, 1979, p. 277-291.)

**106.** Num jogo de beisebol da Liga Pequena, o lançador do time A lança *strikes* em 50% das vezes e lança *ball* (bolas ruins) 50% das vezes, os arremessos são inde-

pendentes um do outro e o lançador nunca atinge o rebatedor. Sabendo disso, o treinador do time B instruiu o rebatedor para não rebater nunca. Calcule a probabilidade de

- a. o rebatedor ganhar a base no quarto arremesso (4 *balls* consecutivos);
- b. o rebatedor ganhar a base no sexto arremesso (de forma que dois dos cinco primeiros devem ter sido *strikes*) usando um argumento de contagem ou construindo um diagrama de árvore;
- c. o rebatedor ganhar a base (4 *balls* antes de 3 *strikes*);
- d. o primeiro rebatedor marcar sem ninguém ter sido eliminado (assumindo que cada rebatedor use a estratégia de não rebater).

**107.** Quatro engenheiros, A, B, C e D, tiveram entrevistas marcadas para as 10 da manhã na sexta-feira, 13 de janeiro, na Random Sampling, Inc. O gerente de RH marcou as quatro entrevistas para as salas 1, 2, 3 e 4, respectivamente. Entretanto, a secretária do gerente não sabe disso e os manda a quatro salas, de forma totalmente aleatória (o que mais a fazer!). Qual é a probabilidade de:

- a. Os quatro acabarem nas salas certas?
- b. Nenhum dos quatro ir para a sala certa?

**108.** Determinada companhia aérea tem vôos às 10 da manhã de Chicago para Nova York, Atlanta e Los Angeles. Seja  $A$  o evento em que o vôo para Nova York está cheio e defina os eventos  $B$  e  $C$  de forma análoga para os outros vôos. Suponha que  $P(A) = 0,6$ ,  $P(B) = 0,5$ ,  $P(C) = 0,4$  e que os três eventos sejam independentes. Qual é a probabilidade de:

- a. Os três vôos estarem cheios? Ao menos um dos vôos não estar cheio?
- b. Apenas o vôo para Nova York estar cheio? Exatamente um dos três vôos estar cheio?

**109.** Um gerente de pessoal entrevistará quatro candidatos a um cargo. Eles estão classificados como 1, 2, 3 e 4, em ordem de preferência, e serão entrevistados em ordem aleatória. Entretanto, no final de cada entrevista, o gerente saberá apenas como o candidato atual se compara em relação aos candidatos já entrevistados. Por exemplo: a ordem de entrevistas 3, 4, 1, 2 não fornece informações após a primeira entrevista, mostra que o segundo candidato é pior que o primeiro e que o terceiro é melhor que os dois primeiros. Entretanto, a ordem 3, 4, 2, 1 geraria as mesmas informações após cada uma das três primeiras entrevistas. O gerente deseja contratar o melhor candidato, mas deve tomar uma decisão irrevogável após cada entrevista. Considere a estratégia a seguir: rejeitar automaticamente os primeiros  $s$  candidatos e então contratar o primeiro candidato subsequente que é o melhor dentre os entrevistados (se não houver tal candidato, contratar o último entrevistado).

Por exemplo, com  $s = 2$ , a ordem 3, 4, 1, 2 resultaria na contratação do melhor candidato, enquanto a

ordem 3, 1, 2, 4 não o faria. Dos quatro valores possíveis (0, 1, 2 e 3), qual maximiza  $P$  (contratação do melhor)? (*Sugestão*: Escreva as 24 ordens de igual probabilidade:  $s = 0$  significa que o primeiro candidato é contratado automaticamente).

**110.** Considere quatro eventos independentes  $A_1, A_2, A_3$  e  $A_4$  e seja  $p_i = P(A_i)$  para  $i = 1, 2, 3, 4$ . Expresse a probabilidade de que pelo menos um desses quatro eventos ocorra em termos dos  $p_i$  e faça o mesmo para a probabilidade de que pelo menos dois eventos ocorram.

**111.** Uma caixa contém quatro pedaços de papel exatamente de mesmas dimensões: (1) ganha o prêmio 1; (2) ganha o prêmio 2, (3) ganha o prêmio 3 e (4) ganha os prêmios 1, 2 e 3. Um pedaço é selecionado aleatoriamente. Seja  $A_1 = \{\text{ganha o prêmio 1}\}$ ;  $A_2 = \{\text{ganha o prêmio 2}\}$  e  $A_3 = \{\text{ganha o prêmio 3}\}$ . Mostre que  $A_1$  e  $A_2$  são independentes, que  $A_1$  e  $A_3$  são independentes, assim como  $A_2$  e  $A_3$  (isso é independência *aos pares*). Mostre, entretanto, que  $P(A_1 \cap A_2 \cap A_3) \neq P(A_1) \cdot P(A_2) \cdot P(A_3)$ , de forma que esses eventos *não* sejam mutuamente independentes.

---

## Bibliografia

DURRETT, Richard. *The Essentials of Probability*. Duxbury Press, Belmont, CA, 1993. Uma apresentação concisa, em um nível ligeiramente superior ao deste livro.

MOSTELLER, Frederick, ROURKE, Robert e THOMAS, George, *Probability with Statistical Applications* (2ª ed.), Addison-Wesley, Reading, MA, 1970. Uma introdução muito boa à probabilidade, com muitos exemplos divertidos; bom especialmente em regras de contagem e suas aplicações.

OLKIN, Ingram, DERMAN, Cyrus e GLENER, Leon, *Probability Models and Application* (2ª ed.), Macmil-

lan, Nova York, 1994. Uma abrangente introdução à probabilidade, escrita em um nível ligeiramente mais difícil do que o deste livro, mas com muitos exemplos interessantes.

ROSS, Sheldon, *A First Course in Probability* (6ª ed.), Macmillan, Nova York, 2002. De leitura mais difícil e mais sofisticado matematicamente, esse livro contém muitos exemplos e exercícios interessantes.

WINKLER, Robert, *Introduction to Bayesian Inference and Decision*, Holt, Rinehart & Winston. Nova York, 1972. Excelente introdução à probabilidade subjetiva.

# Variáveis Aleatórias Discretas e Distribuições de Probabilidade

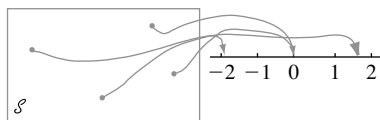
## Introdução

Independentemente de um experimento gerar resultados quantitativos ou qualitativos, os métodos de análise estatística requerem que sejam enfocados certos aspectos numéricos dos dados (como proporção de amostra  $x/n$ , média  $\bar{x}$  ou desvio padrão  $s$ ). O conceito de variável aleatória nos permite passar dos resultados do experimento propriamente ditos para uma função numérica dos resultados. Há dois tipos fundamentalmente diferentes de variáveis aleatórias: variáveis aleatórias discretas e variáveis aleatórias contínuas. Neste capítulo examinaremos as propriedades básicas e discutiremos os exemplos mais importantes das variáveis discretas. O Capítulo 4 enfocará as variáveis aleatórias contínuas.

## 3.1 Variáveis Aleatórias

Em qualquer experimento há diversas características que podem ser observadas ou medidas, mas na maioria dos casos o experimento enfocará um ou alguns aspectos específicos da amostra. Por exemplo: em um estudo de padrões de transporte de uma área metropolitana, cada indivíduo pode ser consultado sobre distância, sobre o número de pessoas que usam o mesmo veículo, mas não sobre o QI, renda, tamanho da família e outras características do tipo. Como alternativa, um pesquisador pode testar uma amostra de componentes e registrar apenas quantos apresentaram falhas dentro de mil horas, em vez de manter o registro das falhas individuais.

Em geral, cada resultado de um experimento é associado a um número, especificando-se uma regra de associação (por exemplo: a quantidade de componentes que apresentam falha em um período de mil horas em uma amostra de 10 componentes ou o peso total de bagagem para uma amostra de 25 passageiros de uma companhia aérea). Tal regra de associação é denominada **variável aleatória**. Variável porque é possível obter diferentes valores numéricos, e aleatória porque o valor observado depende de qual dos resultados possíveis do experimento é obtido (Veja a Figura 3.1).



**Figura 3.1** Uma variável aleatória

#### DEFINIÇÃO

Para um dado espaço amostral  $\mathcal{S}$  de um experimento, uma **variável aleatória (va)** é qualquer regra que associe um valor a cada resultado de  $\mathcal{S}$ . Em termos matemáticos, uma variável aleatória é uma função cujo domínio é o espaço amostral e o contra-domínio é um conjunto de números reais.

As variáveis aleatórias normalmente são representadas por letras maiúsculas, como  $X$  e  $Y$ , próximas ao final do alfabeto. Em contraste com a utilização anterior de uma letra maiúscula, como  $X$ , para representar uma variável, usaremos letras minúsculas para representar um valor específico da variável aleatória correspondente. A notação  $X(s) = x$  significa que  $x$  é o valor associado ao resultado  $s$  pela va  $X$ .

#### Exemplo 3.1

Quando um estudante tenta acessar um computador em um sistema de compartilhamento de tempo, todas as portas estão ocupadas ( $F$ ), caso em que o aluno não terá sucesso, ou haverá ao menos uma porta livre ( $S$ ), caso em que o estudante conseguirá acessar o sistema. Com  $\mathcal{S} = \{S, F\}$ , defina uma va  $X$  como

$$X(S) = 1 \quad X(F) = 0$$

A va  $X$  indica se o estudante pode (1) ou não (0) acessá-lo. ■

No Exemplo 3.1, a va  $X$  foi especificada, relacionando-se explicitamente cada elemento de  $\mathcal{S}$  e o número associado. Se  $\mathcal{S}$  tiver muitos resultados, a tarefa levará muito tempo para ser feita, mas pode ser omitida frequentemente.

#### Exemplo 3.2

Considere o experimento em que um número de telefone em um determinado código de área é discado por meio de discador aleatório (dispositivos muito usados por empresas de pesquisa) e defina uma va  $Y$  como

$$Y = \begin{cases} 1 & \text{se o número selecionado não estiver na lista telefônica} \\ 0 & \text{se o número selecionado estiver na lista telefônica} \end{cases}$$

Por exemplo: se 5528-2966 estiver na lista telefônica —  $Y(5528-2966) = 0$  — enquanto  $Y(7772-7350) = 1$  nos diz que o número 7772-7350 não está na lista. A descrição desse tipo de evento em palavras é mais econômica do que uma listagem completa, de forma que a usaremos sempre que possível. ■

Nos exemplos 3.1 e 3.2, os únicos valores possíveis da variável aleatória eram 0 e 1. Tais variáveis aleatórias aparecem tão frequentemente que receberam um nome especial em homenagem ao primeiro indivíduo que as estudou.

#### DEFINIÇÃO

Qualquer variável aleatória cujos únicos valores possíveis são 0 e 1 é denominada **Variável Aleatória de Bernoulli**.

Frequentemente, precisamos definir e estudar diversas variáveis aleatórias diferentes do mesmo espaço amostral.

**Exemplo 3.3**

O Exemplo 2.3 descreveu um experimento no qual foi determinado o número de bombas em uso de dois postos de gasolina. Defina as vas  $X$ ,  $Y$  e  $U$  por

$X$  = o número total de bombas em uso nos dois postos

$Y$  = a diferença entre o número de bombas em uso no posto 1 e o número em uso no posto 2

$U$  = o máximo de bombas em uso nos dois postos

Se o experimento for realizado e o resultado for  $s = (2, 3)$ , então  $X((2, 3)) = 2 + 3 = 5$ , podendo-se dizer que o valor observado de  $X$  foi  $x = 5$ . De forma similar, o valor observado de  $Y$  seria  $y = 2 - 3 = -1$  e o valor observado de  $U$  seria  $u = \text{máx.}(2, 3) = 3$ . ■

Cada variável aleatória dos exemplos 3.1–3.3 pode assumir apenas um número finito de valores possíveis. Esse não é obrigatoriamente o caso.

**Exemplo 3.4**

No Exemplo 2.4, consideramos o experimento em que diversas baterias eram examinadas até a obtenção de uma em bom estado ( $S$ ). O espaço amostral foi  $\mathcal{S} = \{S, FS, FFS, \dots\}$ . Defina uma va  $X$  como

$X$  = o número de baterias examinadas antes do fim do experimento

Então  $X(S) = 1$ ,  $X(FS) = 2$ ,  $X(FFS) = 3$ ,  $\dots$ ,  $X(FFFFFFFS) = 7$  e assim por diante. Qualquer inteiro positivo é um valor possível de  $X$ , de modo que o conjunto de valores possíveis é infinito. ■

**Exemplo 3.5**

Suponha que um local (latitude e longitude) nos Estados Unidos continental seja selecionado de forma aleatória. Defina uma va  $Y$  por

$Y$  = a altura acima do nível do mar do local selecionado

Por exemplo: se o local selecionado foi  $(39^\circ 50'N, 98^\circ 35'O)$ , podemos ter  $Y((39^\circ 50'N, 98^\circ 35'O)) = 1748,26$  pés. O maior valor possível de  $Y$  é 14.494 pés (monte Whitney) e o menor valor possível é  $-282$  pés (Vale da Morte). O conjunto de todos os valores possíveis de  $Y$  é o conjunto de todos os números no intervalo entre  $-282$  e 14.494, isto é,

$$\{y: y \text{ é um número, } -282 \leq y \leq 14.494\}$$

e há um número infinito de valores nesse intervalo. ■

**Dois tipos de variáveis aleatórias**

Na Seção 1.2 foi feita uma distinção entre dois tipos de variáveis numéricas, as *discretas* e as *contínuas*. A mesma distinção é feita para as variáveis aleatórias.

**DEFINIÇÃO**

Uma variável aleatória **discreta** é uma variável, cujos valores possíveis constituem um conjunto finito ou podem ser relacionados em uma seqüência infinita na qual haja um primeiro elemento, um segundo e assim por diante. Uma variável aleatória é **contínua** se seu conjunto de valores possíveis consiste em um intervalo completo da reta de números (Reta Real).

Apesar de qualquer intervalo da reta de números conter um número infinito de valores, podemos demonstrar que não há uma forma de criarmos uma lista infinita de todos esses valores, pois há muitos deles.

## Exemplo 3.6

Todas as variáveis aleatórias mostradas nos Exemplos 3.1–3.4 são discretas. Como outro exemplo, suponha que selecionemos casais aleatoriamente e façamos um teste sanguíneo em cada pessoa, até encontrarmos um marido e mulher com o mesmo fator Rh. Sendo  $X$  = o número de testes sanguíneos a serem realizados, os valores possíveis de  $X$  são  $D = \{2, 4, 6, 8, \dots\}$ . Como os valores possíveis foram relacionados em seqüência,  $X$  é uma va discreta. ■

Para estudar as propriedades básicas das vas discretas, são requeridas apenas as ferramentas da matemática discreta: soma e subtração. O estudo das variáveis contínuas requer a matemática contínua do cálculo: integrais e derivadas.

## Exercícios | Seção 3.1 (1–10)

1. Uma viga de concreto pode apresentar falha por cisalhamento ( $C$ ) ou flexão ( $F$ ). Suponha que três vigas com defeito sejam selecionadas aleatoriamente e o tipo de falha seja determinado para cada uma delas. Seja  $X$  = número de vigas entre as três selecionadas que falharam por cisalhamento. Relacione cada resultado no espaço amostral juntamente com o valor de  $X$  associado.
2. Dê três exemplos de va de Bernoulli (diferentes dos do livro).
3. Usando o experimento do Exemplo 3.3, defina mais duas variáveis aleatórias e relacione os valores possíveis de cada uma.
4. Seja  $X$  = número de dígitos não-nulos de um código CEP selecionado aleatoriamente. Quais são os valores possíveis de  $X$ ? Forneça três resultados possíveis e seus valores associados  $X$ .
5. Se o espaço amostral é um conjunto infinito, implica necessariamente que qualquer va  $X$  definida em  $\mathcal{S}$  terá um conjunto infinito de valores possíveis? Caso a resposta seja sim, explique o porquê. Caso contrário, dê um exemplo.
6. A partir de certo horário fixo, cada carro que cruza uma interseção é observado para obtenção de sua direção: à esquerda ( $E$ ), à direita ( $D$ ) ou segue em frente ( $F$ ). O experimento termina assim que um carro vira à esquerda. Seja  $X$  = o número de carros observados. Quais são os valores possíveis de  $X$ ? Relacione cinco resultados e seus valores associados  $X$ .
7. Para cada variável aleatória definida a seguir, descreva o conjunto de valores possíveis da variável e diga se é discreta.
  - a.  $X$  = número de ovos não quebrados em uma caixa de ovos padrão selecionada aleatoriamente
  - b.  $Y$  = número de estudantes ausentes no primeiro dia de aula em uma lista de chamada de um determinado curso
  - c.  $U$  = número de vezes que um jogador balança o taco de golfe antes de atingir a bola
  - d.  $X$  = comprimento de uma cascavel selecionada aleatoriamente
  - e.  $Z$  = quantia ganha em *royalties* pela venda de uma primeira edição de 10.000 livros-texto
  - f.  $Y$  = pH de uma amostra de solo selecionada aleatoriamente
  - g.  $X$  = tensão (psi) no encordoamento de uma raquete de tênis selecionada aleatoriamente
  - h.  $X$  = número total de lançamentos de moeda exigidos para que três indivíduos obtenham resultados iguais (*cara, cara, cara* ou *coroa, coroa, coroa*)
8. Cada vez que um componente é testado, o teste resulta em sucesso ( $S$ ) ou falha ( $F$ ). Suponha que o componente seja testado repetidamente até que ocorra sucesso em três tentativas *consecutivas*. Seja  $Y$  o número de tentativas necessárias para atingir esse objetivo. Relacione todos os resultados correspondentes aos cinco menores valores de  $Y$  e diga que valor de  $Y$  está associado a cada um.
9. Um indivíduo chamado Claudius está localizado no ponto 0 do diagrama a seguir.

Usando um dispositivo de aleatoriedade adequado (como um dado tetraédrico, com quatro lados), Claudius primeiro vai para um dos quatro locais  $B_1, B_2, B_3, B_4$ . Uma vez que ele esteja em um desses locais, outro dispositivo aleatório é usado para decidir se ele deve retornar a 0 ou visitar um dos dois outros pontos adjacentes. O processo continua; após cada movimento, outro movimento para um dos (novos) pontos adjacentes é determinado pelo lançamento do dado ou moeda apropriado.

- a. Seja  $X$  = número de movimentos que Claudius faz antes de retornar a 0. Quais são os valores possíveis de  $X$ ?  $X$  é discreta ou contínua?
  - b. Se forem permitidos movimentos ao longo da diagonal ligando 0 a  $A_1, A_2, A_3$  e  $A_4$ , respectivamente, responda às perguntas da parte (a).
10. O número de bombas em uso em dois postos de gasolina, um com seis bombas e outro com quatro, será determinado. Forneça os valores possíveis de cada uma das variáveis aleatórias a seguir:
- a.  $T$  = número total de bombas em uso
  - b.  $X$  = diferença entre as quantidades em uso nos postos 1 e 2
  - c.  $U$  = número máximo de bombas em uso em cada posto
  - d.  $Z$  = número de postos com exatamente duas bombas em uso

### 3.2 Distribuições de Probabilidade para Variáveis Aleatórias Discretas

Quando são atribuídas probabilidades a diversos resultados de  $\mathcal{S}$ , elas, por sua vez, determinam probabilidades associadas aos valores de qualquer  $va$   $X$  em particular. A *distribuição de probabilidade de  $X$*  expressa como a probabilidade total 1 é distribuída (alocada a) entre os diversos valores possíveis de  $X$ .

#### Exemplo 3.7

Seis lotes de componentes estão prontos para embarque em um fornecedor. O número de componentes com defeito em cada lote é mostrado a seguir:

Lote	1	2	3	4	5	6
Número de peças com defeito	0	2	0	1	2	0

Um desses lotes será selecionado aleatoriamente para embarque a um cliente específico. Seja  $X$  o número de peças com defeito no lote selecionado. Os três valores possíveis de  $X$  são 0, 1 e 2. Dos seis eventos igualmente prováveis, três resultam em  $X = 0$ , um em  $X = 1$  e os outros dois em  $X = 2$ . Sejam  $p(0)$  a probabilidade de  $X = 0$  e  $p(1)$  e  $p(2)$  as probabilidades dos outros dois valores possíveis de  $X$ . Então

$$p(0) = P(X = 0) = P(\text{lote 1 ou 3 ou 6 é enviado}) = \frac{3}{6} = 0,500$$

$$p(1) = P(X = 1) = P(\text{lote 4 é enviado}) = \frac{1}{6} = 0,167$$

$$p(2) = P(X = 2) = P(\text{lote 2 ou 5 é enviado}) = \frac{2}{6} = 0,333$$

Ou seja, a probabilidade 0,500 é atribuída para o valor 0 de  $X$ , a probabilidade 0,167 é alocada para o valor 1 de  $X$  e a probabilidade restante, 0,333, é associada ao valor 2 de  $X$ . Os valores de  $X$  juntamente com suas probabilidades especificam coletivamente a distribuição de probabilidade ou *função de massa da probabilidade de  $X$* . Se o experimento fosse repetido diversas vezes no longo prazo,  $X = 0$  ocorreria na metade das vezes,  $X = 1$  em um sexto das vezes e  $X = 2$  em um terço das vezes. ■

#### DEFINIÇÃO

A **função distribuição de probabilidade** ou **função de massa de probabilidade** (fmp) de uma  $va$  discreta é definida para cada número  $x$  por  $p(x) = P(X = x) = P(\text{todos os } s \in \mathcal{S}: X(s) = x)$ .<sup>1</sup>

<sup>1</sup>  $P(X = x)$  lê-se “a probabilidade de a  $va$   $X$  assumir o valor  $x$ ”. Por exemplo:  $P(X = 2)$  indica a probabilidade de que o valor resultante de  $X$  seja 2.



Expondo em palavras, para cada valor possível  $x$  da variável aleatória, a fmp especifica a probabilidade de observar aquele valor quando o experimento for realizado. As condições  $p(x) \geq 0$  e  $\sum_{\text{todos os } x \text{ possíveis}} p(x) = 1$  são obrigatórias para qualquer fmp.

### Exemplo 3.8

Suponha que visitemos a livraria de uma universidade durante a primeira semana de aulas, e observemos se a próxima pessoa a comprar um computador comprará um *laptop* ou um *desktop*. Seja

$$X = \begin{cases} 1 & \text{se o cliente comprar um } \textit{laptop} \\ 0 & \text{se o cliente comprar um } \textit{desktop} \end{cases}$$

Se 20% de todos os compradores durante aquela semana selecionarem um *laptop*, a fmp de  $X$  será

$$p(0) = P(X = 0) = P(\text{próximo cliente compra um } \textit{desktop}) = 0,8$$

$$p(1) = P(X = 1) = P(\text{próximo cliente compra um } \textit{laptop}) = 0,2$$

$$p(x) = P(X = x) = 0 \quad \text{para } x \neq 0 \text{ ou } 1$$

Uma descrição equivalente é

$$p(x) = \begin{cases} 0,8 & \text{se } x = 0 \\ 0,2 & \text{se } x = 1 \\ 0,0 & \text{se } x \neq 0 \text{ ou } 1 \end{cases}$$

A Figura 3.2 é uma ilustração dessa fmp, denominada *gráfico de linhas*.

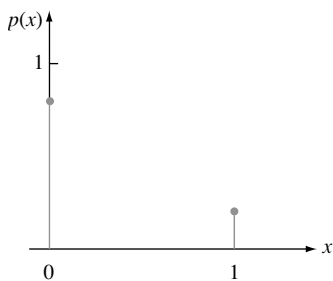


Figura 3.2 O gráfico de linhas da fmp do Exemplo 3.8

### Exemplo 3.9

Considere um grupo de cinco doadores de sangue potenciais: A, B, C, D e E. Desses, apenas A e B possuem o tipo O+. Cinco amostras de sangue, uma de cada indivíduo, serão testadas em ordem aleatória até que seja identificado um indivíduo O+. Seja a va  $Y =$  número de testes necessários para identificar um indivíduo O+. Então a fmp de  $Y$  é

$$p(1) = P(Y = 1) = P(\text{A ou B identificados primeiro}) = \frac{2}{5} = 0,4$$

$$p(2) = P(Y = 2) = P(\text{C, D, ou E primeiro e então A ou B})$$

$$= P(\text{C, D, ou E primeiro}) \cdot P(\text{A ou B depois} \mid \text{C, D ou E primeiro}) = \frac{3}{5} \cdot \frac{2}{4} = 0,3$$

$$p(3) = P(Y = 3) = P(\text{C, D, ou E primeiro e segundo, então A ou B})$$

$$= \left(\frac{3}{5}\right)\left(\frac{2}{4}\right)\left(\frac{2}{3}\right) = 0,2$$

$$p(4) = P(Y = 4) = P(C, D, e E, todos feitos primeiro) = \left(\frac{3}{5}\right)\left(\frac{2}{4}\right)\left(\frac{1}{3}\right) = 0,1$$

$$p(y) = 0 \quad \text{se } y \neq 1, 2, 3, 4$$

A fmp pode ser melhor apresentada na forma tabular:

$y$	1	2	3	4
$p(y)$	0,4	0,3	0,2	0,1

onde qualquer valor  $y$  não-relacionado recebe probabilidade zero. Essa fmp também pode ser exibida em um gráfico de linhas (Veja a Figura 3.3).

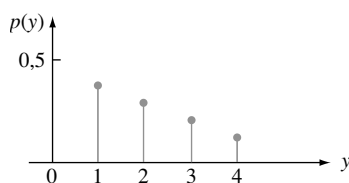


Figura 3.3 O gráfico de linhas da fmp do Exemplo 3.9

O nome “função de massa de probabilidade” é sugerido por um modelo usado em física para um sistema de “pontos de massa”. Nesse modelo, as massas são distribuídas nos diversos locais  $x$  ao longo de um eixo unidimensional. Nossa fmp descreve como a massa total das probabilidades igual a 1 é distribuída nos diversos pontos ao longo dos eixos dos possíveis valores da variável aleatória (onde e quanto de massa em cada  $x$ ).

Outra representação ilustrativa útil de uma fmp, denominada histograma de probabilidades, é similar ao histograma descrito no Capítulo 1. Acima de cada  $y$  com  $p(y) > 0$ , construa um retângulo centrado em  $y$ . A altura de cada retângulo é proporcional a  $p(y)$  e a base é a mesma para todos os retângulos. Quando os valores possíveis estão igualmente espaçados, a base é frequentemente escolhida como sendo a distância entre os valores sucessivos de  $y$  (apesar de poder ser menor). A Figura 3.4 mostra dois histogramas de probabilidades.

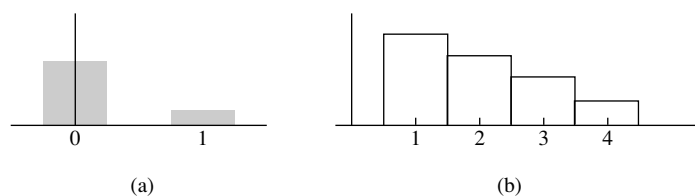


Figura 3.4 Histogramas de probabilidade: (a) Exemplo 3.8; (b) Exemplo 3.9

### Parâmetro de uma distribuição de probabilidades

No Exemplo 3.8, tínhamos  $p(0) = 0,8$  e  $p(1) = 0,2$  porque 20% de todos os compradores selecionaram computadores portáteis. Em outra livraria, pode ocorrer que  $p(0) = 0,9$  e  $p(1) = 0,1$ . De forma mais geral, a fmp de qualquer va Bernoulli pode ser expressa na forma  $p(1) = \alpha$  e  $p(0) = 1 - \alpha$ , onde  $0 < \alpha < 1$ . Como a fmp depende do valor específico de  $\alpha$ , normalmente escrevemos  $p(x; \alpha)$  em vez de apenas  $p(x)$ :

$$p(x; \alpha) = \begin{cases} 1 - \alpha & \text{se } x = 0 \\ \alpha & \text{se } x = 1 \\ 0 & \text{caso contrário} \end{cases} \quad (3.1)$$

Cada escolha de  $\alpha$  na Expressão (3.1) resulta em uma fmp diferente.

## DEFINIÇÃO

Suponha que  $p(x)$  dependa de uma quantidade que pode ser atribuída a qualquer um de diversos valores possíveis, em que cada valor diferente define uma distribuição de probabilidade diferente. Tal quantidade é denominada **parâmetro** da distribuição. A coleção de todas as distribuições de probabilidade dos diferentes valores do parâmetro é denominada uma **família** de distribuições de probabilidade.

A quantidade  $\alpha$  da Expressão (3.1) é um parâmetro. Cada número  $\alpha$  entre 0 e 1 determina um membro diferente de uma família para distribuições. Dois desses membros são

$$p(x; 0,6) = \begin{cases} 0,4 & \text{se } x = 0 \\ 0,6 & \text{se } x = 1 \\ 0 & \text{caso contrário} \end{cases} \quad \text{e} \quad p(x; 0,5) = \begin{cases} 0,5 & \text{se } x = 0 \\ 0,5 & \text{se } x = 1 \\ 0 & \text{caso contrário} \end{cases}$$

Cada distribuição de probabilidade de uma va Bernoulli tem a forma da Expressão (3.1), denominada a *família de distribuições de Bernoulli*.

**Exemplo 3.10**

Iniciando em um horário fixo, observamos o sexo de cada criança nascida em um determinado hospital até que nasça um menino ( $B$ ). Seja  $p = P(B)$ , assuma que nascimentos sucessivos sejam independentes e defina a va como  $X =$  número de nascimentos observados. Então

$$\begin{aligned} p(1) &= P(X = 1) = P(B) = p \\ p(2) &= P(X = 2) = P(GB) = P(G) \cdot P(B) = (1 - p)p \end{aligned}$$

e

$$p(3) = P(X = 3) = P(GGB) = P(G) \cdot P(G) \cdot P(B) = (1 - p)^2 p$$

Continuando dessa forma, define-se a fórmula geral:

$$p(x) = \begin{cases} (1 - p)^{x-1} p & x = 1, 2, 3, \dots \\ 0 & \text{caso contrário} \end{cases} \quad (3.2)$$

A quantidade  $p$  na Expressão (3.2) representa um número entre 0 e 1 e é um parâmetro da distribuição de probabilidades. No exemplo do sexo das crianças,  $p = 0,51$  pode ser apropriado, mas, se estivermos buscando a primeira criança com sangue com fator Rh positivo, podemos assumir  $p = 0,85$ . ■

**A função distribuição acumulada**

Para um valor fixo  $x$ , normalmente desejamos computar a probabilidade de o valor observado de  $X$  ser no máximo  $x$ . Por exemplo: a fmp do Exemplo 3.7 foi

$$p(x) = \begin{cases} 0,500 & x = 0 \\ 0,167 & x = 1 \\ 0,333 & x = 2 \\ 0 & \text{caso contrário} \end{cases}$$

A probabilidade de  $X$  ser no máximo 1 é então

$$P(X \leq 1) = p(0) + p(1) = 0,500 + 0,167 = 0,667$$

Nesse exemplo,  $X \leq 1,5$  se e somente se  $X \leq 1$ , de forma que  $P(X \leq 1,5) = P(X \leq 1) = 0,667$ . De forma similar,  $P(X \leq 0) = P(X = 0) = 0,5$ , e  $P(X \leq 0,75) = 0,5$  também. Como 0 é o menor valor possível de  $X$ ,  $P(X \leq -1,7) = 0$ ,  $P(X \leq -0,0001) = 0$  e assim por diante. O maior valor possível de  $X$  é 2, de forma que  $P(X \leq 2) = 1$  e, se  $x$  for um número maior que 2,  $P(X \leq x) = 1$ ; isto é,  $P(X \leq 5) = 1$ ,  $P(X \leq 10,23) = 1$  e assim por diante. Observe que  $P(X < 1) = 0,5 \neq P(X \leq 1)$ , pois a probabilidade de  $X$  com valor 1 está incluída na última probabilidade, mas não na anterior. Quando  $X$  é uma variável aleatória discreta e  $x$  é um valor possível de  $X$ ,  $P(X < x) < P(X \leq x)$ .

**DEFINIÇÃO**

A **função de distribuição acumulada** (FDA)  $F(x)$  de uma variável aleatória discreta  $X$  com fmp  $p(x)$  é definida para cada valor de  $x$  por

$$F(x) = P(X \leq x) = \sum_{y: y \leq x} p(y)$$

Para qualquer valor  $x$ ,  $F(x)$  é a probabilidade de o valor  $X$  observado ser no máximo  $x$ .

**Exemplo 3.11**

A fmp de  $Y$  (quantidade de classificações sanguíneas) no Exemplo 3.9 foi

$y$	1	2	3	4
$p(y)$	0,4	0,3	0,2	0,1

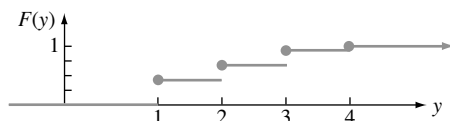
Primeiro determinamos  $F(y)$  para cada valor do conjunto  $\{1, 2, 3, 4\}$  de valores possíveis:

$$\begin{aligned} F(1) &= P(Y \leq 1) = P(Y = 1) = p(1) = 0,4 \\ F(2) &= P(Y \leq 2) = P(Y = 1 \text{ ou } 2) = p(1) + p(2) = 0,7 \\ F(3) &= P(Y \leq 3) = P(Y = 1 \text{ ou } 2 \text{ ou } 3) = p(1) + p(2) + p(3) = 0,9 \\ F(4) &= P(Y \leq 4) = P(Y = 1 \text{ ou } 2 \text{ ou } 3 \text{ ou } 4) = 1 \end{aligned}$$

Assim sendo, para qualquer outro valor de  $y$ ,  $F(y)$  será igual ao valor de  $F$  mais próximo possível de  $Y$  à esquerda de  $y$ . Por exemplo:  $F(2,7) = P(Y \leq 2,7) = P(Y \leq 2) = 0,7$ , e  $F(3,999) = F(3) = 0,9$ . A FDA é, portanto

$$F(y) = \begin{cases} 0 & \text{se } y < 1 \\ 0,4 & \text{se } 1 \leq y < 2 \\ 0,7 & \text{se } 2 \leq y < 3 \\ 0,9 & \text{se } 3 \leq y < 4 \\ 1 & \text{se } 4 \leq y \end{cases}$$

O gráfico de  $F(y)$  é mostrado na Figura 3.5.



**Figura 3.5** O gráfico da FDA do Exemplo 3.11

Para uma va discreta  $X$ , o gráfico de  $F(x)$  terá um salto em cada valor possível de  $X$  e será uma reta entre os valores possíveis. Esse gráfico é denominado **função degrau**.

**Exemplo 3.12**

No Exemplo 3.10, qualquer inteiro positivo tinha um valor  $X$  possível e a fmp era

$$p(x) = \begin{cases} (1 - p)^{x-1}p & x = 1, 2, 3, \dots \\ 0 & \text{caso contrário} \end{cases}$$

Para qualquer inteiro positivo  $x$ ,

$$F(x) = \sum_{y \leq x} p(y) = \sum_{y=1}^x (1 - p)^{y-1} p = p \sum_{y=0}^{x-1} (1 - p)^y \tag{3.4}$$

Para calcular essa soma, usamos o fato de que a soma parcial de uma série geométrica é

$$\sum_{y=0}^k a^y = \frac{1 - a^{k+1}}{1 - a}$$

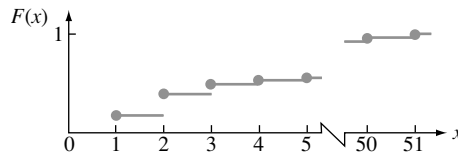
Usando essa expressão na Equação (3.4), com  $a = 1 - p$  e  $k = x - 1$ , temos

$$F(x) = p \cdot \frac{1 - (1 - p)^x}{1 - (1 - p)} = 1 - (1 - p)^x \quad \text{sendo } x \text{ um inteiro positivo}$$

Como  $F$  é constante entre inteiros positivos,

$$F(x) = \begin{cases} 0 & x < 1 \\ 1 - (1 - p)^{[x]} & x \geq 1 \end{cases} \tag{3.5}$$

onde  $[x]$  é o maior inteiro  $\leq x$  (por exemplo,  $[2,7] = 2$ ). Portanto, se  $p = 0,51$ , como no exemplo dos nascimentos, a probabilidade de se ter de examinar no máximo cinco nascimentos para se ver o primeiro menino é de  $F(5) = 1 - (0,49)^5 = 1 - 0,0282 = 0,9718$ , enquanto  $F(10) \approx 1,0000$ . Essa FDA está representada na Figura 3.6.



**Figura 3.6** Gráfico de  $F(x)$  do Exemplo 3.12

Nos exemplos até agora, a FDA origina-se da fmp. Tal processo pode ser revertido para se obter a fmp da FDA sempre que esta última função estiver disponível. Suponha, por exemplo, que  $X$  represente o número de componentes com defeito em um lote de seis componentes, de forma que os valores possíveis de  $X$  são 0, 1, ..., 6. Então

$$\begin{aligned} p(3) &= P(X = 3) \\ &= [p(0) + p(1) + p(2) + p(3)] - [p(0) + p(1) + p(2)] \\ &= P(X \leq 3) - P(X \leq 2) \\ &= F(3) - F(2) \end{aligned}$$

De forma mais geral, a probabilidade de  $X$  estar em um intervalo especificado é obtida facilmente a partir da FDA. Por exemplo:

$$\begin{aligned} P(2 \leq X \leq 4) &= p(2) + p(3) + p(4) \\ &= [p(0) + \dots + p(4)] - [p(0) + p(1)] \\ &= P(X \leq 4) - P(X \leq 1) \\ &= F(4) - F(1) \end{aligned}$$

Observe que  $P(2 \leq X \leq 4) \neq F(4) - F(2)$ . O fato ocorre porque o valor 2 de  $X$  está incluído em  $2 \leq X \leq 4$ , de forma que não queremos desprezar essa probabilidade. Entretanto,  $P(2 < X \leq 4) = F(4) - F(2)$  porque  $X = 2$  não está incluído no intervalo  $2 < X \leq 4$ .

**PROPOSIÇÃO**

Para quaisquer dois números  $a$  e  $b$  com  $a \leq b$ ,

$$P(a \leq X \leq b) = F(b) - F(a-)$$

onde “ $a-$ ” representa o maior valor possível de  $X$  estritamente menor que  $a$ . Em particular, se os únicos valores possíveis forem inteiros e, se  $a$  e  $b$  forem inteiros, então

$$\begin{aligned} P(a \leq X \leq b) &= P(X = a \text{ ou } a + 1 \text{ ou } \dots \text{ ou } b) \\ &= F(b) - F(a - 1) \end{aligned}$$

Considerando  $a = b$  resulta que  $P(X = a) = F(a) - F(a - 1)$ , nesse caso.

O motivo de subtrair  $F(a-)$  em vez de  $F(a)$  é que queremos incluir  $P(X = a)$ ;  $F(b) - F(a)$  fornece  $P(a < X \leq b)$ . Essa proposição será usada extensivamente no cálculo da probabilidade binomial e de Poisson nas Seções 3.4 e 3.6.

**Exemplo 3.13**

Seja  $X =$  o número de dias de licença por doença de um funcionário de uma grande empresa, selecionado aleatoriamente em certo ano. Se o número máximo de dias permitidos por ano for 14, os valores possíveis de  $X$  são 0, 1, ..., 14. Com  $F(0) = 0,58$ ,  $F(1) = 0,72$ ,  $F(2) = 0,76$ ,  $F(3) = 0,81$ ,  $F(4) = 0,88$ , e  $F(5) = 0,94$ ,

$$P(2 \leq X \leq 5) = P(X = 2, 3, 4, \text{ ou } 5) = F(5) - F(1) = 0,22$$

e

$$P(X = 3) = F(3) - F(2) = 0,05$$



**Outra visão das funções de massa de probabilidade**

Freqüentemente é útil imaginar a fmp como especificação de um modelo matemático para população discreta.

**Exemplo 3.14**

Considere a escolha aleatória de um estudante entre os 15.000 matriculados no semestre corrente na Mega University. Seja  $X =$  número de cursos em que o estudante selecionado está matriculado e suponha que  $X$  tenha a fmp mostrada a seguir.

$x$	1	2	3	4	5	6	7
$p(x)$	0,01	0,03	0,13	0,25	0,39	0,17	0,02

Uma forma de interpretar tal situação é imaginar a população formada por 15.000 indivíduos, cada um com seu próprio valor  $X$ . A proporção de cada valor  $X$  é dada por  $p(x)$ . Um ponto de vista alternativo é esquecer dos alunos e pensar na população em si como formada pelos valores  $X$ : há alguns 1s na população, 2s e alguns 7s. A população então consiste dos números 1, 2, ..., 7 (portanto, é discreta) e  $p(x)$  fornece um modelo da distribuição dos valores da população. ■

Uma vez definido o modelo da população, podemos usá-lo para calcular os valores das características da população (por exemplo, a média  $\mu$ ) e fazer inferências sobre tais características.

## Exercícios | Seção 3.2 (11–27)

11. Uma instalação de condicionamento de automóveis especializada em regulagem de motores sabe que 45% de todas as regulagens são feitas em automóveis de quatro cilindros, 40% em automóveis de seis cilindros e 15% em automóveis de oito cilindros. Seja  $X$  = número de cilindros do próximo carro a ser preparado.

- Qual é a fmp de  $X$ ?
- Desenhe um gráfico de linhas e um histograma de probabilidade da fmp da parte (a).
- Qual é a probabilidade de o próximo carro a ser regulado ter no mínimo seis cilindros? Mais de seis cilindros?

12. As empresas aéreas algumas vezes fazem *overbook* de vôos. Suponha que, para um avião de 50 lugares, tenham sido vendidas 55 passagens. Defina a variável aleatória  $Y$  como o número de passageiros com passagens que compareceram ao vôo. A função massa de probabilidade de  $Y$  é exibida na tabela a seguir.

$y$	45	46	47	48	49	50	51	52	53	54	55
$p(y)$	0,050	0,100	0,120	0,140	0,250	0,170	0,060	0,050	0,030	0,020	0,010

- Qual é a probabilidade de que o vôo acomode todos os passageiros que comparecerem?
- Qual é a probabilidade de nem todos os passageiros que comparecerem serem acomodados?
- Se você for a primeira pessoa da lista de espera (o que significa que será o primeiro a subir no avião, se houver assentos disponíveis após todos os passageiros terem sido acomodados), qual será a probabilidade de estar no vôo? Qual é a probabilidade, se você for a terceira pessoa da lista de espera?

13. Uma empresa que fornece computadores pelo correio tem seis linhas telefônicas. Seja  $X$  o número de linhas em uso em determinado horário. Suponha que a fmp de  $X$  seja conforme a tabela a seguir.

$x$	0	1	2	3	4	5	6
$p(x)$	0,10	0,15	0,20	0,25	0,20	0,06	0,04

Calcule a probabilidade de cada um dos seguintes eventos.

- {no máximo três linhas estão em uso}
- {menos de três linhas estão em uso}
- {pelo menos três linhas estão em uso}
- {entre duas e cinco linhas, inclusive, estão em uso}
- {entre duas e quatro linhas, inclusive, não estão em uso}
- {pelo menos quatro linhas não estão em uso}

14. Um empreiteiro é solicitado pelo departamento de planejamento de uma cidade a enviar um, dois, três,

quatro ou cinco formulários (dependendo da natureza do projeto) quando requer um alvará de construção. Seja  $Y$  = o número de formulários requeridos do próximo empreiteiro. Sabe-se que a probabilidade de  $y$  formulários serem exigidos é proporcional a  $y$ , isto é,  $p(y) = ky$  para  $y = 1, \dots, 5$ .

- Qual é o valor de  $k$ ? [Sugestão:  $\sum_{y=1}^5 p(y) = 1$ ]
- Qual é a probabilidade de no máximo três formulários serem exigidos?
- Qual é a probabilidade de serem requeridos entre dois e quatro formulários (inclusive)?
- Poderia  $p(y) = y^2/50$  para  $y = 1, \dots, 5$  ser a fmp de  $Y$ ?

15. Muitos fabricantes possuem programas de controle de qualidade que incluem a inspeção de defeitos no recebimento dos materiais. Suponha que um fabricante de computadores receba placas em lotes de cinco. Duas placas em cada lote são selecionadas para inspeção. Podemos representar os resultados possíveis do processo pela seleção de pares. Por exemplo: o par (1, 2) representa a seleção das placas 1 e 2 para inspeção.

- Relacione os 10 resultados diferentes possíveis.
- Suponha que as placas 1 e 2 sejam as únicas com defeito em um lote de cinco. Duas placas serão escolhidas aleatoriamente. Defina  $X$  como o número observado de placas com defeito entre as inspecionadas. Determine a distribuição de probabilidades de  $X$ .
- Seja  $F(x)$  a FDA de  $X$ . Primeiro determine  $F(0) = P(X \leq 0)$ ,  $F(1)$ , e  $F(2)$ , e depois obtenha  $F(x)$  para todos os outros  $x$ .

16. Algumas partes da Califórnia são particularmente propensas a terremotos. Suponha que em tal área, 30% de todos os moradores tenham seguro contra danos por terremotos. Quatro moradores são selecionados aleatoriamente. Seja  $X$  o número dos que possuem seguro contra danos por terremotos, entre os quatro.

- Determine a distribuição de probabilidades de  $X$ . [Sugestão: seja  $S$  o morador que tem seguro e  $N$  o que não o tem. Uma possibilidade será  $SNSS$ , com probabilidade  $(0,3)(0,7)(0,3)(0,3)$  e valor associado  $X = 3$ . Há outros 15 resultados].
- Desenhe o histograma de probabilidade correspondente.
- Qual é o valor mais provável de  $X$ ?
- Qual é a probabilidade de que ao menos dois dos quatro moradores selecionados tenham seguro?

17. A voltagem de uma pilha nova pode ser aceitável ( $A$ ) ou inaceitável ( $I$ ). Uma lanterna específica exige duas pilhas, que serão selecionadas e testadas independentemente até que sejam encontradas duas aceitáveis. Suponha que 90% de todas as pilhas tenham voltagem aceitável. Seja  $Y$  o número de pilhas que devem ser testadas.

- a. Qual é o valor de  $p(2)$ , isto é,  $P(Y = 2)$ ?
- b. Qual é o valor de  $p(3)$ ? (*Sugestão*: Há dois resultados diferentes que dão  $Y = 3$ .)
- c. Para ter  $Y = 5$ , o que deve ser verdadeiro para a quinta pilha selecionada? Relacione os quatro resultados possíveis para os quais  $Y = 5$  e depois determine  $p(5)$ .
- d. Use o padrão de suas respostas para as partes (a)–(c) para obter a fórmula geral de  $p(y)$ .

18. Dois dados de seis lados são lançados independentemente. Seja  $M =$  o máximo dos dois lançamentos (então  $M(1, 5) = 5$ ,  $M(3, 3) = 3$  etc.).

- a. Qual é a fmp de  $M$ ? [*Sugestão*: Determine primeiro  $p(1)$ , depois  $p(2)$  e assim por diante].
- b. Determine a FDA de  $M$  e desenhe o gráfico.

19. No Exemplo 3.9, suponha que haja apenas quatro doadores potenciais de sangue dos quais apenas um tem sangue tipo O+. Calcule a fmp de  $Y$ .

20. Uma biblioteca assina duas revistas semanais diferentes, que supostamente chegam na correspondência de quarta-feira. Na verdade, elas podem chegar na quarta, quinta, sexta ou sábado. Suponha que as duas cheguem independentemente uma da outra e para cada uma  $P(\text{qua.}) = 0,3$ ,  $P(\text{qui.}) = 0,4$ ,  $P(\text{sex.}) = 0,2$  e  $P(\text{sáb.}) = 0,1$ . Seja  $Y =$  número de dias, após a quarta-feira, que as revistas levam para chegar (os valores possíveis de  $Y$  são 0, 1, 2 ou 3). Calcule a fmp de  $Y$ . [*Sugestão*: Há 16 resultados possíveis;  $Y(\text{qua.}, \text{qua.}) = 0$ ,  $Y(\text{qua.}, \text{qui.}) = 2$  e assim por diante].

21. Refira-se ao Exercício 13, calcule e desenhe o gráfico da FDA  $F(x)$ . Utilize-a para calcular a probabilidade dos eventos descritos nas partes (a)–(d) do problema.

22. Uma organização de consumidores que avalia automóveis novos relata costumeiramente o número de defeitos graves em cada carro examinado. Seja  $X$  o número de defeitos graves em um carro de determinado tipo selecionado aleatoriamente. A FDA de  $X$  é como abaixo:

$$F(x) = \begin{array}{ll} 0 & x < 0 \\ 0,06 & 0 \leq x < 1 \\ 0,19 & 1 \leq x < 2 \\ 0,39 & 2 \leq x < 3 \\ 0,67 & 3 \leq x < 4 \\ 0,92 & 4 \leq x < 5 \\ 0,97 & 5 \leq x < 6 \\ 1 & 6 < x \end{array}$$

Calcule as seguintes probabilidades diretamente pela FDA:

- a.  $p(2)$ , isto é,  $P(X = 2)$
- b.  $P(X > 3)$
- c.  $P(2 \leq X \leq 5)$
- d.  $P(2 < X < 5)$

23. Uma empresa de seguros oferece aos seus segurados diferentes opções de pagamento *premium*. Para um segurado selecionado aleatoriamente, seja  $X =$  número

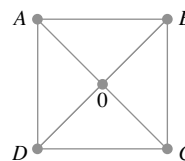
de meses entre pagamentos sucessivos. A FDA de  $X$  é como segue:

$$F(x) = \begin{array}{ll} 0 & x < 1 \\ 0,30 & 1 \leq x < 3 \\ 0,40 & 3 \leq x < 4 \\ 0,45 & 4 \leq x < 6 \\ 0,60 & 6 \leq x < 12 \\ 1 & 12 \leq x \end{array}$$

- a. Qual é a fmp de  $X$ ?
- b. Usando apenas a FDA, calcule  $P(3 \leq X \leq 6)$  e  $P(4 \leq X)$ .

24. No Exemplo 3.10, seja  $Y =$  o número de meninas nascidas antes do final do experimento. Com  $p = P(B)$  e  $1 - p = P(G)$ , qual é a fmp de  $Y$ ? (*Sugestão*: Primeiro relacione os valores possíveis de  $Y$ , começando pelo menor, e continue até obter uma fórmula geral).

25. Alvie Singer vive no ponto 0 do diagrama abaixo e tem quatro amigos que vivem em  $A, B, C$  e  $D$ . Um dia, Alvie decide visitá-los e lança uma moeda duas vezes para decidir quais dos quatro visitará. Depois de chegar na casa de um amigo, ele pode retornar para casa ou continuar a visita a uma das duas casas adjacentes (como 0,  $A$  ou  $C$  quando em  $B$ ), tendo cada uma das três possibilidades com probabilidade  $\frac{1}{3}$ . Dessa forma, ele continua a visitar amigos até voltar para casa.



- a. Seja  $X =$  o número de vezes que Alvie visita um amigo. Determine a fmp em relação a  $X$ .
- b. Seja  $Y =$  o número de segmentos de reta que Alvie percorre (incluindo os que chegam e saem de 0). Qual é a fmp de  $Y$ ?
- c. Suponha que amigas vivam em  $A$  e  $C$  e amigos em  $B$  e  $D$ . Se  $Z =$  o número de visitas a amigas, qual é a fmp de  $Z$ ?

26. Depois que todos os estudantes saírem da sala de aula, um professor de estatística observa que quatro cópias do livro foram deixadas sob as mesas. No começo da próxima aula, o professor distribuiu os quatro livros de forma completamente aleatória para cada um de quatro alunos (1, 2, 3 e 4) que dizem ter esquecido os livros. Um resultado possível é que 1 receba o livro de 2, 2 receba o livro de 4, 3 receba o seu livro e 4 receba o livro de 1. Esse resultado pode ser abreviado como (2, 4, 3, 1).

- a. Relacione os outros 23 resultados possíveis.
- b. Seja  $X =$  o número de estudantes que recebem o próprio livro. Determine a fmp de  $X$ .

27. Demonstre que a FDA  $F(x)$  é uma função não-decrescente; isto é,  $x_1 < x_2$  implica  $F(x_1) \leq F(x_2)$ . Em que condições  $F(x_1) = F(x_2)$ ?



### 3.3 | Valores Esperados de Variáveis Aleatórias Discretas

No Exemplo 3.14, consideramos uma universidade com 15.000 alunos e consideramos  $X =$  o número de cursos em que um aluno selecionado aleatoriamente estava matriculado. A fmp de  $X$  é mostrada a seguir. Como  $p(1) = 0,01$ , sabemos que  $(0,01) \cdot (15.000) = 150$  alunos estão matriculados em um curso e assim sucessivamente para os outros  $x$  valores.

$x$	1	2	3	4	5	6	7	
$p(x)$	0,01	0,03	0,13	0,25	0,39	0,17	0,02	(3.6)
Número de alunos matriculados	150	450	1950	3750	5850	2550	300	

Para calcular o número médio de cursos por aluno ou o valor médio de  $X$  na população, devemos calcular o número total de cursos e dividir pelo número total de alunos. Como cada um dos 150 alunos está matriculado em um curso, esses 150 contribuem com 150 cursos no total. De forma similar, 450 alunos contribuem com  $2(450)$  cursos e assim por diante. O valor médio da população  $X$  é

$$\frac{1(150) + 2(450) + 3(1950) + \cdots + 7(300)}{15.000} = 4,57 \quad (3.7)$$

Como  $150/15.000 = 0,01 = p(1)$ ,  $450/15.000 = 0,03 = p(2)$ , e assim por diante, uma expressão alternativa para (3,7) é

$$1 \cdot p(1) + 2 \cdot p(2) + \cdots + 7 \cdot p(7) \quad (3.8)$$

A Expressão (3.8) mostra que, para calcular o valor médio da população de  $X$ , precisamos apenas dos valores possíveis de  $X$  juntamente com suas probabilidades (proporções). Em particular, o tamanho da população é irrelevante, desde que a fmp seja dada por (3,6). A média ou valor médio de  $X$  é, portanto, uma média *ponderada* dos valores possíveis 1, ..., 7, em que os pesos são as probabilidades dos valores.

#### Valor esperado de $X$

##### DEFINIÇÃO

Seja  $X$  como uma va discreta com conjunto de valores possíveis  $D$  e fmp  $p(x)$ . O **valor esperado** ou **valor médio** de  $X$ , denotado por  $E(X)$  ou  $\mu_x$ , é

$$E(X) = \mu_x = \sum_{x \in D} x \cdot p(x)$$

Quando está claro a que  $X$  o valor esperado se refere,  $\mu$  é usado no lugar de  $\mu_x$ .

##### Exemplo 3.15

Para a fmp em (3.6),

$$\begin{aligned} \mu &= 1 \cdot p(1) + 2 \cdot p(2) + \cdots + 7 \cdot p(7) \\ &= (1)(0,01) + 2(0,03) + \cdots + (7)(0,02) \\ &= 0,01 + 0,06 + 0,39 + 1,00 + 1,95 + 1,02 + 0,14 = 4,57 \end{aligned}$$

Se imaginarmos que a população consiste dos  $X$  valores 1, 2, ..., 7, então  $\mu = 4,57$  é a média da população. Na seqüência, freqüentemente nos referiremos a  $\mu$  como a *média da população* em vez de média de  $X$  na população. ■

No Exemplo 3.15, o valor esperado de  $\mu$  era 4,57, que não é um valor possível de  $X$ . A palavra *esperado* deve ser interpretada com cautela porque uma pessoa não espera ver um valor de  $X = 4,57$  quando um único estudante é selecionado.

### Exemplo 3.16

Após cada nascimento, os bebês são classificados de acordo com uma escala denominada Apgar. As classificações possíveis são 0, 1, ..., 10, com a classificação do bebê determinada por cor, tônus muscular, esforço respiratório, batimentos cardíacos e irritabilidade reflexas (a melhor pontuação possível é 10). Seja  $X$  o escore Apgar de uma criança selecionada aleatoriamente em um determinado hospital no próximo ano e suponha que a fmp de  $X$  seja

$x$	0	1	2	3	4	5	6	7	8	9	10
$p(x)$	0,002	0,001	0,002	0,005	0,02	0,04	0,18	0,37	0,25	0,12	0,01

O valor médio de  $X$  será

$$\begin{aligned} E(X) = \mu &= 0(0,002) + 1(0,001) + 2(0,002) \\ &\quad + \cdots + 8(0,25) + 9(0,12) + 10(0,01) \\ &= 7,15 \end{aligned}$$

Novamente,  $\mu$  não é um valor possível da variável  $X$ . Além disso, como a variável se refere a uma futura criança, não há uma população existente, concreta, a que se refira. Em vez disso, imaginamos a fmp como um modelo para uma população conceitual consistindo dos valores 0, 1, 2, ..., 10. O valor médio dessa população conceitual é  $\mu = 7,15$ . ■

### Exemplo 3.17

Seja  $X = 1$  se um componente selecionado aleatoriamente precisar de reparo em garantia e  $= 0$  caso contrário. Então  $X$  será uma va Bernoulli com fmp

$$p(x) = \begin{cases} 1 - p & x = 0 \\ p & x = 1 \\ 0 & x \neq 0, 1 \end{cases}$$

pela qual  $E(X) = 0 \cdot p(0) + 1 \cdot p(1) = 0(1 - p) + 1(p) = p$ . Isto é, o valor esperado de  $X$  é apenas a probabilidade de  $X$  ser igual ao valor 1. Se conceitualizarmos uma população formada por zeros na proporção  $1 - p$  e de 1s na proporção  $p$ , então a média da população é  $\mu = p$ . ■

### Exemplo 3.18

A forma geral da fmp de  $X =$  número de crianças nascidas até o nascimento do primeiro menino (inclusive) é

$$p(x) = \begin{cases} p(1 - p)^{x-1} & x = 1, 2, 3, \dots \\ 0 & \text{caso contrário} \end{cases}$$

Pela definição,

$$E(X) = \sum_D x \cdot p(x) = \sum_{x=1}^{\infty} xp(1 - p)^{x-1} = p \sum_{x=1}^{\infty} \left[ -\frac{d}{dp} (1 - p)^x \right] \quad (3.9)$$

Se trocarmos a ordem de cálculo da derivada e da soma, a soma será a de uma série geométrica. Depois de calcular a soma, calcula-se a derivada e o resultado final é  $E(X) = 1/p$ . Se  $p$  estiver próximo de 1, espera-se ver um menino em breve, enquanto, se  $p$  estiver próximo de 0, esperam-se muitos nascimentos antes do primeiro menino. Para  $p = 0,5$ ,  $E(X) = 2$ . ■

Há outra interpretação de  $\mu$  usada freqüentemente. Considere a fmp

$$p(x) = \begin{cases} (0,5) \cdot (0,5)^{x-1} & \text{se } x = 1, 2, 3, \dots \\ 0 & \text{caso contrário} \end{cases}$$

Essa é a fmp de  $X =$  número de lançamentos necessários de uma moeda para obter a primeira “cara” (um caso especial do Exemplo 3.18). Suponha que observemos um valor  $x$  dessa fmp (lançar uma moeda até aparecer uma “cara”), e então observemos outro valor independentemente (continue lançando), e depois outro e assim sucessivamente. Se depois de observarmos um grande número de valores  $x$  calcularmos a média deles, a média amostral resultante será muito próxima de  $\mu = 2$ . Isto é,  $\mu$  pode ser interpretado como o valor observado da média de longo prazo de  $X$  quando o experimento é executado repetidamente.

### Exemplo 3.19

Seja  $X$  o número de entrevistas pelas quais um estudante passa antes de conseguir um emprego, com fmp

$$p(x) = \begin{cases} k/x^2 & x = 1, 2, 3, \dots \\ 0 & \text{caso contrário} \end{cases}$$

onde  $k$  é escolhido de forma que  $\sum_{x=1}^{\infty} (k/x^2) = 1$ . (Nos cursos de matemática sobre séries infinitas, é demonstrado que  $\sum_{x=1}^{\infty} (1/x^2) < \infty$ , que implica que  $k$  existe, mas seu valor exato não nos preocupa). O valor esperado de  $X$  é

$$\mu = E(X) = \sum_{x=1}^{\infty} x \cdot \frac{k}{x^2} = k \sum_{x=1}^{\infty} \frac{1}{x} \quad (3.10)$$

A soma à direita da Equação (3.10) é a famosa série harmônica da matemática e pode ser demonstrado que é igual a  $\infty$ .  $E(X)$  não é finito, neste caso, porque  $p(x)$  não decresce tão rapidamente quanto  $x$  aumenta. Os estatísticos dizem que a distribuição de probabilidade de  $X$  possui “uma cauda longa”. Se for escolhida uma sequência de valores de  $X$  usando essa distribuição, a média amostral não terá um valor finito, mas tenderá a crescer sem limite.

Os estatísticos usam a expressão “caudas longas” em relação a quaisquer distribuições que tenham valor de probabilidade grande distante de  $\mu$ , (portanto caudas longas não requerem  $\mu = \infty$ ). Caudas longas dificultam as inferências sobre  $\mu$ . ■

## Valor esperado de uma função

Freqüentemente estamos interessados no valor esperado de alguma função  $h(X)$  em vez de  $X$  em si mesmo.

### Exemplo 3.20

Suponha que uma livraria compre 10 cópias de um livro a US\$ 6,00 cada para vendê-las a US\$ 12,00, sabendo que ao fim de um período de 3 meses os livros não vendidos podem ser devolvidos por US\$ 2,00. Se  $X =$  ao número de cópias vendidas, então a receita líquida será  $= h(X) = 12X + 2(10 - X) - 60 = 10X - 40$ . ■

Uma forma fácil de calcular o valor esperado de  $h(X)$  é sugerida pelo exemplo a seguir.

### Exemplo 3.21

Seja  $X =$  o número de cilindros do motor do próximo carro a ser regulado em certa oficina. O custo de uma regulagem é relacionado a  $X$  por  $h(X) = 20 + 3X + 0,5X^2$ . Como  $X$  é uma variável aleatória,  $h(X)$  também o é. Denote essa última va por  $Y$ . As fmps de  $X$  e  $Y$  são:

$x$	4	6	8	$y$	40	56	76
$p(x)$	0,5	0,3	0,2	$p(y)$	0,5	0,3	0,2

Denotando os valores possíveis de  $Y$  por  $D^*$ , temos

$$\begin{aligned} E(Y) = E[h(X)] &= \sum_{D^*} y \cdot p(y) \\ &= (40)(0,5) + (56)(0,3) + (76)(0,2) \\ &= h(4) \cdot (0,5) + h(6) \cdot (0,3) + h(8) \cdot (0,2) \\ &= \sum_D h(x) \cdot p(x) \end{aligned} \quad (3.11)$$

De acordo com a Equação (3.11), não é necessário determinar a fmp de  $Y$  para obter  $E(Y)$ . O valor esperado é a média ponderada dos valores possíveis de  $h(x)$  (em vez de  $x$ ). ■

#### PROPOSIÇÃO

Se a va  $X$  tiver um conjunto de valores possíveis  $D$  e fmp  $p(x)$ , o valor esperado de qualquer função  $h(X)$ , expresso por  $E[h(X)]$  ou  $\mu_{h(X)}$ , é calculado por

$$E[h(X)] = \sum_D h(x) \cdot p(x)$$

De acordo com a proposição,  $E[h(X)]$  é calculada da mesma forma que  $E(X)$  em si, exceto que  $h(x)$  é usado no lugar de  $x$ .

#### Exemplo 3.22

Uma loja de computadores comprou três computadores de certo tipo a US\$ 500 cada. Eles serão vendidos a US\$ 1.000 cada. O fabricante concordou em aceitar a devolução dos computadores não vendidos, após um período especificado, por US\$ 200 cada. Seja  $X$  o número de computadores vendidos e suponha que  $p(0) = 0,1$ ,  $p(1) = 0,2$ ,  $p(2) = 0,3$ , e  $p(3) = 0,4$ . Definindo como  $h(X)$  o lucro associado à venda de  $X$  unidades, as informações fornecidas implicam que  $h(X) = \text{receita} - \text{custo} = 1000X + 200(3 - X) - 1500 = 800X - 900$ . O lucro esperado será então

$$\begin{aligned} E[h(X)] &= h(0) \cdot p(0) + h(1) \cdot p(1) + h(2) \cdot p(2) + h(3) \cdot p(3) \\ &= (-900)(0,1) + (-100)(0,2) + (700)(0,3) + (1500)(0,4) \\ &= \$ 700 \end{aligned} \quad \blacksquare$$

### Regras do valor Esperado

A função  $h(X)$  freqüentemente é uma função linear  $aX + b$ . Nesse caso,  $E[h(X)]$  é facilmente calculado pela  $E(X)$ .

#### PROPOSIÇÃO

$$E(aX + b) = a \cdot E(X) + b$$

(Ou, usando notação alternativa,  $\mu_{aX+b} = a \cdot \mu_X + b$ .)

Em palavras, o valor esperado de uma função linear é igual à função linear calculada com o valor esperado  $E(X)$ . Como  $h(X)$  no Exemplo 3.22 é linear e  $E(X) = 2$ ,  $E[h(X)] = 800(2) - 900 = \text{US\$ } 700$ , como antes.

#### Demonstração

$$\begin{aligned} E(aX + b) &= \sum_D (ax + b) \cdot p(x) = a \sum_D x \cdot p(x) + b \sum_D p(x) \\ &= aE(X) + b \end{aligned} \quad \blacksquare$$

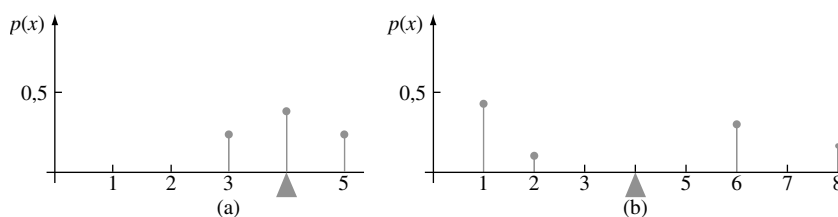
Dois casos especiais da proposição fornecem duas regras importantes do valor esperado.

1. Para qualquer constante  $a$ ,  $E(aX) = a \cdot E(X)$  (assumir  $b = 0$ ). (3.12)
2. Para qualquer constante  $b$ ,  $E(X + b) = E(X) + b$  (assumir  $a = 1$ ).

A multiplicação de  $X$  por uma constante  $a$  altera a unidade de medida (de dólares para centavos, onde  $a = 100$ , de polegadas para centímetros, onde  $a = 2,54$  etc.). A regra 1 diz que o valor esperado nas novas unidades é igual ao valor das antigas, multiplicado pelo fator de conversão  $a$ . De forma similar, se uma constante  $b$  for adicionada a cada valor possível de  $X$ , então o valor esperado será deslocado do valor da constante.

## Variância de $X$

O valor esperado de  $X$  descreve onde a distribuição de probabilidades está centrada. Usando a analogia física de colocação de pontos de massa  $p(x)$  no valor  $x$  de um eixo unidimensional, se o eixo estiver apoiado em um fulcro posicionado em  $\mu$ , não há tendência de inclinação do eixo. O fato é ilustrado para duas distribuições diferentes na Figura 3.7.



**Figura 3.7** Duas distribuições de probabilidade diferentes com  $\mu = 4$

Apesar de ambas as distribuições ilustradas na Figura 3.7 terem o mesmo centro  $\mu$ , a distribuição da Figura 3.7(b) tem maior dispersão ou variabilidade do que a da Figura 3.7(a). Usaremos a variância de  $X$  para avaliar o valor da variabilidade de (a distribuição de)  $X$ , do mesmo modo que  $s^2$  foi usado no Capítulo 1 para medir a variabilidade de uma amostra.

### DEFINIÇÃO

Seja  $X$  com fmp  $p(x)$  e o valor esperado  $\mu$ . A **variância** de  $X$ , denotada por  $V(X)$  ou  $\sigma_x^2$ , ou apenas  $\sigma^2$ , é

$$V(X) = \sum_D (x - \mu)^2 \cdot p(x) = E[(X - \mu)^2]$$

O **desvio padrão** (DP) de  $X$  é

$$\sigma_x = \sqrt{\sigma_x^2}$$

A quantidade  $h(X) = (X - \mu)^2$  é o quadrado do desvio de  $X$  em relação à sua média e  $\sigma^2$  é o quadrado do desvio esperado. Se a maior parte da distribuição de probabilidade estiver próxima de  $\mu$ ,  $\sigma^2$  será relativamente pequena. Entretanto, se houver valores de  $x$  distantes de  $\mu$  com  $p(x)$  grande,  $\sigma^2$  será bem grande.

### Exemplo 3.23

Se  $X$  é o número de cilindros do próximo carro a ser regulado em uma oficina, com a fmp do Exemplo 3.21 [ $p(4) = 0,5$ ,  $p(6) = 0,3$ ,  $p(8) = 0,2$ , de onde  $\mu = 5,4$ ], então

$$\begin{aligned} V(X) &= \sigma^2 = \sum_{x=4}^8 (x - 5,4)^2 \cdot p(x) \\ &= (4 - 5,4)^2(0,5) + (6 - 5,4)^2(0,3) + (8 - 5,4)^2(0,2) = 2,44 \end{aligned}$$

O desvio padrão de  $X$  é  $\sigma = \sqrt{2,44} = 1,562$ . ■

Quando a fmp  $p(x)$  especifica um modelo matemático para a distribuição dos valores da população,  $\sigma^2$  e  $\sigma$  medem a dispersão dos valores da população:  $\sigma^2$  é a variância da população e  $\sigma$  é o desvio padrão da população.

### Fórmula Alternativa de $\sigma^2$

A quantidade de operações aritméticas necessárias para calcular  $\sigma^2$  pode ser reduzida usando-se uma fórmula alternativa de cálculo.

PROPOSIÇÃO

$$V(X) = \sigma^2 = \left[ \sum_D x^2 \cdot p(x) \right] - \mu^2 = E(X^2) - [E(X)]^2$$

Ao usar essa fórmula,  $E(X^2)$  é calculado primeiro sem necessidade de subtração e então  $E(X)$  é calculado, elevado ao quadrado e subtraído (uma vez) de  $E(X^2)$ .

#### Exemplo 3.24

A fmp do número de cilindros  $X$  do próximo carro a ser regulado em uma oficina foi fornecida no Exemplo 3.23 como  $p(4) = 0,5$ ,  $p(6) = 0,3$ , e  $p(8) = 0,2$ , de onde  $\mu = 5,4$  e

$$E(X^2) = (4^2)(0,5) + (6^2)(0,3) + (8^2)(0,2) = 31,6$$

Portanto  $\sigma^2 = 31,6 - (5,4)^2 = 2,44$ , conforme o Exemplo 3.23. ■

### Demonstração da Fórmula Alternativa

Desenvolva  $(x - \mu)^2$  da definição de  $\sigma^2$  para obter  $x^2 - 2\mu x + \mu^2$ , e depois calcule a somatória ( $\sum$ ) de cada um dos três termos:

$$\begin{aligned} \sigma^2 &= \sum_D x^2 \cdot p(x) - 2\mu \cdot \sum_D x \cdot p(x) + \mu^2 \sum_D p(x) \\ &= E(X^2) - 2\mu \cdot \mu + \mu^2 = E(X^2) - \mu^2 \end{aligned}$$

### Regras da Variância

A variância de  $h(X)$  é o valor esperado do quadrado da diferença entre  $h(X)$  e seu valor esperado:

$$V[h(X)] = \sigma_{h(X)}^2 = \sum_D \{h(x) - E[h(X)]\}^2 \cdot p(x) \quad (3.13)$$

Quando  $h(X)$  for uma função linear,  $V[h(X)]$  é facilmente relacionada a  $V(X)$ .

PROPOSIÇÃO

$$V(aX + b) = \sigma_{aX+b}^2 = a^2 \cdot \sigma_X^2 \quad \text{e} \quad \sigma_{aX+b} = |a| \cdot \sigma_X$$

Esse resultado nos diz que a adição da constante  $b$  não afeta a variância, o que é intuitivo, porque a adição de  $b$  muda a localização (valor da média) mas não a dispersão dos valores. Em particular,

$$\begin{aligned} 1. \quad &\sigma_{aX}^2 = a^2 \cdot \sigma_X^2, \quad \sigma_{aX} = |a| \cdot \sigma_X \\ 2. \quad &\sigma_{X+b}^2 = \sigma_X^2 \end{aligned} \quad (3.14)$$

O motivo do valor absoluto na expressão de  $\sigma_{aX}$  é que  $a$  pode ser negativo, ao passo que o desvio padrão não pode.  $a^2$  resulta quando  $a$  é eliminado do termo que se eleva ao quadrado na Equação (3.13).

## Exemplo 3.25

No problema da venda de computador do Exemplo 3.22,  $E(X) = 2$  e

$$E(X^2) = (0)^2(0,1) + (1)^2(0,2) + (2)^2(0,3) + (3)^2(0,4) = 5$$

então  $V(X) = 5 - (2)^2 = 1$ . A função lucro  $h(X) = 800X - 900$  tem variância  $(800)^2 \cdot V(X) = (640.000)(1) = 640.000$  e desvio padrão 800. ■

## Exercícios | Seção 3.3 (28–43)

28. A fmp de  $X =$  o número de defeitos graves em um eletrodoméstico selecionado aleatoriamente é

$x$	0	1	2	3	4
$p(x)$	0,08	0,15	0,45	0,27	0,05

Calcule os dados a seguir:

- $E(X)$
- $V(X)$  diretamente pela definição
- O desvio padrão de  $X$
- $V(X)$  usando a fórmula alternativa.

29. Um indivíduo que possui um seguro de automóvel de uma determinada empresa é selecionado aleatoriamente. Seja  $Y$  o número de infrações ao código de trânsito para as quais o indivíduo foi reincidente nos últimos 3 anos. A fmp de  $Y$  é

$y$	0	1	2	3
$p(y)$	0,60	0,25	0,10	0,05

- Calcule  $E(Y)$ .
- Suponha que um indivíduo com  $Y$  infrações reincidentes incorra em uma multa de US\$  $100Y^2$ . Calcule o valor esperado da multa.

30. Refira-se ao Exercício 12 e calcule  $V(Y)$  e  $\sigma_y$ . Determine, então, a probabilidade de  $Y$  estar dentro do intervalo de 1 desvio padrão em torno da média.

31. Uma loja de eletrodomésticos vende três modelos diferentes de *freezers* verticais com 13,5, 15,9 e 19,1 pés cúbicos de espaço, respectivamente. Seja  $X =$  volume de armazenagem comprado pelo próximo cliente a comprar um *freezer*. Suponha que a fmp de  $X$  seja

$x$	13,5	15,9	19,1
$p(x)$	0,2	0,5	0,3

- Calcule  $E(X)$ ,  $E(X^2)$  e  $V(X)$ .
- Se o preço de um *freezer* com  $X$  pés cúbicos de capacidade for  $25X - 8,5$ , qual será o preço esperado pago pelo próximo cliente a comprar um *freezer*?
- Qual é a variância do preço  $25X - 8,5$  pago pelo próximo cliente?
- Suponha que, apesar de a capacidade nominal de um *freezer* ser  $X$ , a capacidade real seja  $h(X) = X - 0,01X^2$ . Qual é a capacidade real esperada do *freezer* comprado pelo próximo cliente?

32. Seja  $X$  uma va de Bernoulli com fmp como no Exemplo 3.17.

- Calcule  $E(X^2)$ .
- Demonstre que  $V(X) = p(1 - p)$ .
- Calcule  $E(X^{79})$ .

33. Suponha que o número de certo tipo de planta encontrada em uma região retangular (denominada *quadrat* pelos ecologistas) em uma determinada área geográfica seja uma va  $X$  com fmp

$$p(x) = \begin{cases} c/x^3 & x = 1, 2, 3, \dots \\ 0 & \text{caso contrário} \end{cases}$$

$E(X)$  é finito? Justifique sua resposta (essa é outra distribuição que os estatísticos denominam “cauda longa”).

34. Uma pequena drogaria solicita cópias de uma revista para seu *display* todas as semanas. Represente por  $X =$  demanda da revista, com fmp

$x$	1	2	3	4	5	6
$p(x)$	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{3}{15}$	$\frac{4}{15}$	$\frac{3}{15}$	$\frac{2}{15}$

Suponha que o proprietário da loja pague US\$ 1,00 por cópia e o preço de venda seja US\$ 2,00. Se as revistas não vendidas até o fim da semana não têm valor, é melhor comprar três ou quatro cópias? (*Sugestão*: Expresse a receita líquida em função da demanda  $X$  de três e quatro cópias da revista e depois calcule a receita esperada).

35. Seja  $X$  o dano (em valor monetário) incorrido por um determinado tipo de acidente em um ano. Os valores possíveis de  $X$  são 0, 1000, 5000 e 10000, com probabilidades 0,8, 0,1, 0,08 e 0,02 respectivamente. Uma empresa oferece uma apólice dedutível de US\$ 500. Se quiser que o seu lucro seja de US\$ 100, que valor de prêmio deve cobrar?

36. Os  $n$  candidatos a uma vaga foram classificados como 1, 2, 3, ...,  $n$ . Seja  $X =$  a classificação de um candidato selecionado aleatoriamente, de forma que  $X$  tenha fmp

$$p(x) = \begin{cases} 1/n & x = 1, 2, 3, \dots, n \\ 0 & \text{caso contrário} \end{cases}$$

(tal função é denominada *distribuição uniforme discreta*). Calcule  $E(X)$  e  $V(X)$ , usando a fórmula alternativa. [Sugestão: a soma dos primeiros  $n$  inteiros positivos é  $n(n + 1)/2$ , enquanto a soma de seus quadrados é  $n(n + 1)(2n + 1)/6$ .]

37. Seja  $X$  = resultado de um dado lançado uma vez. Se antes de lançar o dado, oferecessem a você (1/3,5) dólares ou  $h(X) = 1/X$  dólares, você aceitaria a quantia garantida ou faria a aposta? [Nota: Normalmente não é verdade que  $1/E(X) = E(1/X)$ .]
38. Uma empresa de fornecimento de materiais químicos atualmente possui em seu estoque 100 libras de um determinado produto, vendido aos clientes em lotes de 5 libras. Seja  $X$  = número de lotes pedidos por um cliente selecionado aleatoriamente e suponha que  $X$  tenha fmp

$x$	1	2	3	4
$p(x)$	0,2	0,4	0,3	0,1

Calcule  $E(X)$  e  $V(X)$  e depois o número esperado de libras restantes, após o envio do pedido do próximo cliente, e a variância do número de libras restantes. (Sugestão: o número de libras restantes é uma função linear de  $X$ ).

39. a. Desenhe um gráfico de linhas da fmp de  $X$  do Exercício 34. Determine, então, a fmp de  $-X$  e desenhe seu gráfico de linhas. Considerando as duas situações, o que você pode dizer sobre  $V(X)$  e  $V(-X)$ ?
- b. Use a proposição que envolve  $V(aX + b)$  para definir uma relação geral entre  $V(X)$  e  $V(-X)$ .

40. Use a definição da Expressão (3.13) para demonstrar que  $V(aX + b) = a^2 \cdot \sigma_X^2$ . [Sugestão: com  $h(X) = aX + b$ ,  $E[h(X)] = a\mu + b$  onde  $\mu = E(X)$ .]

41. Suponha que  $E(X) = 5$  e  $E[X(X - 1)] = 27,5$ . Qual é
- a.  $E(X^2)$ ? [Sugestão:  $E[X(X - 1)] = E[X^2 - X] = E(X^2) - E(X)$ ]
- b.  $V(X)$ ?
- c. A relação geral entre as quantidades  $E(X)$ ,  $E[X(X - 1)]$  e  $V(X)$ ?

42. Escreva a regra geral para  $E(X - c)$  onde  $c$  é uma constante. O que acontece quando você assume  $c = \mu$ , o valor esperado de  $X$ ?

43. O resultado denominado **desigualdade de Chebyshev** diz que, para qualquer distribuição de probabilidade de uma va  $X$  e qualquer número  $k$  ao menos igual a 1,  $P(|X - \mu| \geq k\sigma) \leq 1/k^2$ . Em palavras, a probabilidade do valor  $X$  estar a pelo menos  $k$  desvios padrão da média é de no máximo  $1/k^2$ .

- a. Qual é o valor do limite superior para  $k = 2$ ?  $k = 3$ ?  $k = 4$ ?  $k = 5$ ?  $k = 10$ ?
- b. Calcule  $\mu$  e  $\sigma$  da distribuição do Exercício 13. Depois calcule  $P(|X - \mu| \geq k\sigma)$  para os valores de  $k$  fornecidos na parte (a). O que isto sugere sobre o limite superior em relação à probabilidade correspondente?
- c. Assuma que  $X$  possa ter três valores:  $-1$ ,  $0$ , e  $1$ , com as probabilidades  $\frac{1}{18}$ ,  $\frac{8}{9}$ , e  $\frac{1}{18}$ , respectivamente. Qual é  $P(|X - \mu| \geq 3\sigma)$ , e como ela se compara ao limite correspondente?
- d. Forneça uma distribuição para a qual  $P(|X - \mu| \geq 5\sigma) = 0,04$ .

### 3.4 Distribuição de probabilidade binomial

Há diversos experimentos que satisfazem exatamente ou aproximadamente a seguinte lista de requisitos:

1. O experimento consiste em uma seqüência de  $n$  experimentos menores denominados *tentativas*, onde  $n$  é estabelecido antes do experimento.
2. Cada tentativa pode resultar em um de dois resultados possíveis (tentativas dicotômicas), chamados de sucesso ( $S$ ) ou falha ( $F$ ).
3. As tentativas são independentes, de forma que o resultado de qualquer tentativa particular não influencia o resultado de qualquer outra tentativa.
4. A probabilidade de sucesso é constante de uma tentativa para a outra. Denominamos essa probabilidade  $p$ .

**DEFINIÇÃO**

Um experimento para o qual as Condições 1–4 são satisfeitas é denominado **experimento binomial**.



**Exemplo 3.26**

A mesma moeda é lançada sucessiva e independentemente  $n$  vezes. Usamos  $S$  arbitrariamente para representar o resultado  $H$  (cara) e  $F$  para representar o resultado  $T$  (coroa). Esse experimento satisfaz, então, as Condições 1–4. Lançar um percevejo  $n$  vezes, com  $S =$  ponta para cima e  $F =$  ponta para baixo, também resulta um experimento binomial. ■

Muitos experimentos envolvem uma seqüência de tentativas independentes para as quais há mais de dois resultados possíveis em qualquer tentativa. Um experimento binomial pode então ser criado, dividindo-se os resultados possíveis em dois grupos.

**Exemplo 3.27**

A cor das sementes de ervilhas é determinada por um único *locus* genético. Se dois alelos desse *locus* são  $AA$  ou  $Aa$  (o genótipo), então a ervilha será amarela (o fenótipo) e, se o alelo for  $aa$ , será verde. Suponha que organizemos 20 sementes  $Aa$  aos pares e cruzemos as duas ervilhas de cada par para obtenção de 10 novos genótipos. Cada novo genótipo será um sucesso  $S$  se for  $aa$  e uma falha, caso contrário. Então, com o identificador  $S$  ou  $F$ , o experimento será binomial com  $n = 10$  e  $p = P(\text{genótipo } aa)$ . Se cada membro do par for igualmente provável de contribuir com a ou  $A$ , então  $p = P(a) \cdot P(a) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{4}$ . ■

**Exemplo 3.28**

Suponha que certa cidade tenha 50 restaurantes licenciados, dos quais atualmente 15 possuem pelo menos uma violação grave do código de saúde e os outros 35 não possuam violações graves. Há cinco inspetores, cada um dos quais inspeciona um restaurante por semana. O nome de cada restaurante é escrito em um pedaço de papel e, após serem misturados, cada inspetor retira um dos pedaços de papel *sem reposição*. A  $i$ -ésima tentativa será um sucesso se o  $i$ -ésimo restaurante selecionado ( $i = 1, \dots, 5$ ) não tiver violações graves. Então

$$P(S \text{ na primeira tentativa}) = \frac{35}{50} = 0,70$$

e

$$\begin{aligned} P(S \text{ na segunda tentativa}) &= P(SS) + P(FS) \\ &= P(\text{segunda } S \mid \text{primeira } S) P(\text{primeira } S) \\ &\quad + P(\text{segunda } S \mid \text{primeira } F) P(\text{primeira } F) \\ &= \frac{34}{49} \cdot \frac{35}{50} + \frac{35}{49} \cdot \frac{15}{50} = \frac{35}{50} \left( \frac{34}{49} + \frac{15}{49} \right) = \frac{35}{50} = 0,70 \end{aligned}$$

De forma similar, é possível mostrar que  $P(S \text{ na } i\text{-ésima tentativa}) = 0,70$  para  $i = 3, 4, 5$ . Entretanto,

$$P(S \text{ na quinta tentativa} \mid SSSS) = \frac{31}{46} = 0,67$$

enquanto

$$P(S \text{ na quinta tentativa} \mid FFFF) = \frac{35}{46} = 0,76$$

O experimento não é binomial porque as tentativas não são independentes. Em geral, se a amostragem for feita sem reposição, o experimento não terá tentativas independentes. Se cada pedaço de papel fosse repostado após sua retirada, as tentativas seriam independentes, o que podia fazer com que o mesmo restaurante fosse inspecionado por mais de um inspetor. ■

**Exemplo 3.29**

Suponha que certo estado tenha 500.000 motoristas cadastrados, dos quais 400.000 possuem seguro. Uma amostra de 10 motoristas é selecionada sem reposição. A  $i$ -ésima tentativa será  $S$  se o  $i$ -ésimo motorista tiver

seguro. Apesar desta situação parecer idêntica à do Exemplo 3.28, a diferença importante é que o tamanho da população é muito grande em relação ao tamanho da amostra. Neste caso

$$P(S \text{ em } 2 \mid S \text{ em } 1) = \frac{399,999}{499,999} = 0,80000$$

e

$$P(S \text{ em } 10 \mid S \text{ no primeiro } 9) = \frac{399,991}{499,991} = 0,799996 \approx 0,80000$$

Esses cálculos sugerem que, apesar das tentativas serem exatamente independentes, as probabilidades condicionais diferem tão pouco uma da outra que, para fins práticos, as tentativas podem ser consideradas independentes com  $P(S) = 0,8$  constante. Assim, com uma boa aproximação, o experimento é binomial com  $n = 10$  e  $p = 0,8$ . ■

Usaremos a seguinte regra prática para decidir se um experimento “sem reposição” pode ser tratado como experimento binomial.

REGRA

Considere a amostragem sem reposição de uma população dicotômica de tamanho  $N$ . Se o tamanho da amostra (número de tentativas)  $n$  for no máximo 5% do tamanho da população, o experimento pode ser analisado como se fosse exatamente binomial.

Por “analisado”, queremos dizer que as probabilidades baseadas nas hipóteses de experimento binomial estarão muito próximas das probabilidades reais “sem reposição”, que normalmente são mais difíceis de calcular. No Exemplo 3.28,  $n/N = 5/50 = 0,1 > 0,05$ , de forma que o experimento binomial não é uma boa aproximação, mas, no Exemplo 3.29,  $n/N = 10/500.000 < 0,05$ .

## Variável aleatória binomial e sua distribuição

Na maioria dos experimentos binomiais, interessa o número total de  $S$  e não o conhecimento de exatamente quais tentativas resultaram em  $S$ .

DEFINIÇÃO

Dado um experimento binomial consistindo de  $n$  tentativas, a **variável aleatória binomial**  $X$  a ele associada é definida como

$$X = \text{quantidade de } S \text{ nas } n \text{ tentativas}$$

Suponha, por exemplo, que  $n = 3$ . Haverá, então, oito resultados possíveis para o experimento:

*SSS SSF SFS SFF FSS FSF FFS FFF*

A partir da definição de  $X$ ,  $X(SSS) = 3$ ,  $X(SFF) = 1$  e assim por diante. Os valores possíveis de  $X$  em um experimento de  $n$  tentativas são  $x = 0, 1, 2, \dots, n$ . Frequentemente usaremos  $X \sim \text{Bin}(n, p)$  para indicar que  $X$  é uma va binomial baseada em  $n$  tentativas com probabilidade de sucesso  $p$ .

NOTAÇÃO

Como a fmp de uma va binomial depende de dois parâmetros  $n$  e  $p$ , representamos a fmp por  $b(x; n, p)$ .

Considere primeiro o caso de  $n = 4$  para o qual cada resultado, sua probabilidade e o valor de  $x$  correspondente estão relacionados na Tabela 3.1. Por exemplo:

$$\begin{aligned}
 P(SSFS) &= P(S) \cdot P(S) \cdot P(F) \cdot P(S) && \text{(tentativas independentes)} \\
 &= p \cdot p \cdot (1 - p) \cdot p && \text{(} P(S) \text{ constante)} \\
 &= p^3 \cdot (1 - p)
 \end{aligned}$$

**Tabela 3.1** Resultados e probabilidades de um experimento binominal com quatro tentativas

Resultado	$x$	Probabilidade	Resultado	$x$	Probabilidade
SSSS	4	$p^4$	FSSS	3	$p^3(1 - p)$
SSSF	3	$p^3(1 - p)$	FSSF	2	$p^2(1 - p)^2$
SSFS	3	$p^3(1 - p)$	FSFS	2	$p^2(1 - p)^2$
SSFF	2	$p^2(1 - p)^2$	FSFF	1	$p(1 - p)^3$
SFSS	3	$p^3(1 - p)$	FFSS	2	$p^2(1 - p)^2$
SFSF	2	$p^2(1 - p)^2$	FFSF	1	$p(1 - p)^3$
SFFS	2	$p^2(1 - p)^2$	FFFS	1	$p(1 - p)^3$
SFFF	1	$p(1 - p)^3$	FFFF	0	$(1 - p)^4$

Neste caso especial, queremos  $b(x; 4, p)$  para  $x = 0, 1, 2, 3$  e  $4$ . Para  $b(3; 4, p)$ , identificamos quais dos 16 resultados fornecem um valor de  $x$  igual a 3 e somamos as probabilidades associadas a cada resultado:

$$b(3; 4, p) = P(FSSS) + P(SFSS) + P(SSFS) + P(SSSF) = 4p^3(1 - p)$$

Há quatro resultados com  $x = 3$ , cada um com probabilidade  $p^3(1 - p)$  (a ordem de  $S$  e  $F$  não é importante, apenas a quantidade  $S$ ), então

$$b(3; 4, p) = \left\{ \begin{array}{l} \text{número de resultados} \\ \text{com } X = 3 \end{array} \right\} \cdot \left\{ \begin{array}{l} \text{probabilidade de qualquer resultado} \\ \text{particular com } X = 3 \end{array} \right\}$$

De forma similar,  $b(2; 4, p) = 6p^2(1 - p)^2$ , que também é o produto do número de resultados com  $X = 2$  e a probabilidade qualquer de tais resultados. Em geral,

$$b(x; n, p) = \left\{ \begin{array}{l} \text{número de seqüências de} \\ \text{comprimento } n \text{ consistindo de } x \text{ } S's \end{array} \right\} \cdot \left\{ \begin{array}{l} \text{probabilidade de qualquer resultado} \\ \text{particular de tal seqüência} \end{array} \right\}$$

Como a ordenação de  $S$  e  $F$  não é importante, o segundo fator da equação anterior é  $p^x(1 - p)^{n-x}$  (por exemplo: as primeiras  $x$  tentativas resultarem em  $S$  e as últimas  $n - x$  resultarem em  $F$ ). O primeiro fator é o número de maneiras de escolher  $x$  das  $n$  tentativas serem  $S$ , isto é, o número de combinações de tamanho  $x$  que podem ser construídas com  $n$  objetos distintos (tentativas, no caso).

TEOREMA

$$b(x; n, p) = \begin{cases} \binom{n}{x} p^x (1 - p)^{n-x} & x = 0, 1, 2, \dots, n \\ 0 & \text{caso contrário} \end{cases}$$

**Exemplo 3.30**

Cada um de seis consumidores de refrigerante selecionados aleatoriamente recebe um copo com o refrigerante  $S$  e um com o refrigerante  $F$ . Os copos são idênticos, exceto por um código no fundo que identifica o refrigerante. Suponha que não haja uma tendência de preferência entre os consumidores. Então  $p = P(\text{um indivíduo selecionado prefere } S) = 0,5$ , de forma que  $X = \text{número de consumidores entre os seis que preferem } S, X \sim \text{Bin}(6, 0,5)$ .

Dessa forma,

$$P(X = 3) = b(3; 6, 0,5) = \binom{6}{3}(0,5)^3(0,5)^3 = 20(0,5)^6 = 0,313$$

A probabilidade de pelo menos três preferirem  $S$  é

$$P(3 \leq X) = \sum_{x=3}^6 b(x; 6, 0,5) = \sum_{x=3}^6 \binom{6}{x}(0,5)^x(0,5)^{6-x} = 0,656$$

e a probabilidade de no máximo um preferir  $S$  é

$$P(X \leq 1) = \sum_{x=0}^1 b(x; 6, 0,5) = 0,109 \quad \blacksquare$$

### Uso das tabelas binomiais

Mesmo para um valor relativamente pequeno de  $n$ , o cálculo da probabilidade binomial é tedioso. A Tabela do Apêndice A.1 exibe a FDA  $F(x) = P(X \leq x)$  para  $n = 5, 10, 15, 20, 25$  em combinação com determinados valores de  $p$ . Diversas outras probabilidades podem ser calculadas, usando-se a proposição de FDA da Seção 3.2.

#### NOTAÇÃO

Para  $X$  (distribuída como (o sinal é um Til, que não significa aproximadamente))  $\sim \text{Bin}(n, p)$ , a FDA será representada por

$$P(X \leq x) = B(x; n, p) = \sum_{y=0}^x b(y; n, p) \quad x = 0, 1, \dots, n$$

#### Exemplo 3.31

Suponha que 20% de todas as cópias de um livro-texto apresentem falha em um determinado teste de resistência de encadernação. Seja  $X$  o número de cópias que apresentam falhas entre 15 cópias selecionadas aleatoriamente. Então  $X$  tem distribuição binomial com  $n = 15$  e  $p = 0,2$ .

1. A probabilidade de no máximo 8 apresentarem falha é

$$P(X \leq 8) = \sum_{y=0}^8 b(y; 15, 0,2) = B(8; 15, 0,2)$$

que é a entrada na linha  $x = 8$  e coluna  $p = 0,2$  da tabela binomial  $n = 15$ . Na Tabela do Apêndice A.1, a probabilidade é  $B(8; 15, 0,2) = 0,999$ .

2. A probabilidade de exatamente 8 apresentarem falha é

$$P(X = 8) = P(X \leq 8) - P(X \leq 7) = B(8; 15, 0,2) - B(7; 15, 0,2)$$

que é a diferença entre duas entradas consecutivas na coluna  $p = 0,2$ . O resultado é  $0,999 - 0,996 = 0,003$ .

3. A probabilidade de no mínimo 8 apresentarem falha é

$$\begin{aligned} P(X \geq 8) &= 1 - P(X \leq 7) = 1 - B(7; 15, 0,2) \\ &= 1 - \left( \begin{array}{l} \text{entrada em } x = 7 \\ \text{linha da coluna } p = 0,2 \end{array} \right) \\ &= 1 - 0,996 = 0,004 \end{aligned}$$

4. Finalmente, a probabilidade de 4 a 7, inclusive, apresentarem falha é

$$\begin{aligned} P(4 \leq X \leq 7) &= P(X = 4, 5, 6 \text{ ou } 7) = P(X \leq 7) - P(X \leq 3) \\ &= B(7; 15, 0,2) - B(3; 15, 0,2) = 0,996 - 0,648 = 0,348 \end{aligned}$$

Observe que a última probabilidade é a diferença entre as entradas nas linhas  $x = 7$  e  $x = 3$  e não das linhas  $x = 7$  e  $x = 4$ . ■

### Exemplo 3.32

Um fabricante de eletrônicos alega que no máximo 10% de seus geradores precisam de reparo no período de garantia. Para investigar a declaração, técnicos de um laboratório de teste compraram 20 unidades e as submeteram a um teste acelerado para simular o uso durante o período de garantia. Seja  $p$  a probabilidade de um gerador precisar de reparo durante o período (proporção de unidades que necessitam de reparo). Os técnicos do laboratório devem decidir se os dados resultantes do experimento suportam a alegação de  $p \leq 0,10$ . Seja  $X$  o número entre os 20 da amostra que precisam de reparo, de forma que  $X \sim \text{Bin}(20, p)$ . Considere a regra de decisão.

Rejeitar a alegação de  $p \leq 0,10$  em favor da conclusão de que  $p > 0,10$  se  $x \geq 5$  (onde  $x$  é o valor observado de  $X$ ) e considerar a alegação aceitável se  $x \leq 4$ .

A probabilidade de a alegação ser rejeitada quando  $p = 0,10$  (conclusão incorreta) é

$$P(X \geq 5 \text{ quando } p = 0,10) = 1 - B(4; 20, 0,1) = 1 - 0,957 = 0,043$$

A probabilidade de a alegação não ser rejeitada quando  $p = 0,20$  (tipo diferente de conclusão incorreta) é

$$P(X \leq 4 \text{ quando } p = 0,2) = B(4; 20, 0,2) = 0,630$$

A primeira probabilidade é relativamente pequena, mas a segunda é intoleravelmente grande. Quando  $p = 0,20$ , o que significa que o fabricante errou grosseiramente para menos no percentual de unidades que necessitam de reparo, e a regra de decisão estabelecida for usada, 63% de todas as amostras serão incluídas na declaração do fabricante ser julgada aceitável!

Pode-se pensar que a probabilidade deste segundo tipo de conclusão errônea poderia ser menor se o valor de corte 5 fosse alterado para outro valor na regra de decisão. Entretanto, apesar de a substituição de 5 por um número menor resultar em uma probabilidade menor que 0,630, a outra probabilidade aumentaria. A única forma de diminuir as duas “probabilidades de erro” é basear a regra de decisão em um experimento que envolva muito mais unidades. ■

Observe que um valor tabelado igual a 0 significa apenas que a probabilidade é 0 com três dígitos de significativos, pois todas as entradas da tabela são positivas. Programas estatísticos como o MINITAB geram  $b(x; n, p)$  ou  $B(x; n, p)$  para quaisquer valores especificados de  $n$  e  $p$ . No Capítulo 4, apresentaremos um método rápido e preciso de aproximações de probabilidades binomiais quando  $n$  for grande.

## Média e variância de $X$

Para  $n = 1$ , a distribuição binomial se torna a distribuição de Bernoulli. Pelo Exemplo 3.17, o valor médio de uma variável Bernoulli é  $\mu = p$ , de forma que o número esperado de  $S$  em uma tentativa única é  $p$ . Como um experimento binomial consiste de  $n$  tentativas, a intuição sugere que, para  $X \sim \text{Bin}(n, p)$ ,  $E(X) = np$ , o produto do número de tentativas pela probabilidade de sucesso em uma única tentativa. A expressão de  $V(X)$  não é tão intuitiva.

### PROPOSIÇÃO

Se  $X \sim \text{Bin}(n, p)$ , então  $E(X) = np$ ,  $V(X) = np(1 - p) = npq$ , e  $\sigma_x = \sqrt{npq}$  (onde  $q = 1 - p$ ).

Dessa forma, o cálculo da média e da variância de uma va binomial não requer o cálculo de somatórias. A demonstração do resultado de  $E(X)$  está esquematizada no Exercício 60.

### Exemplo 3.33

Se 75% de todas as compras em uma determinada loja forem feitas com cartão de crédito e  $X$  for a quantidade de compras feitas com cartão de crédito entre 10 compras selecionadas aleatoriamente, então  $X \sim \text{Bin}(10, 0,75)$ . Portanto,  $E(X) = np = (10)(0,75) = 7,5$ ,  $V(X) = npq = 10(0,75)(0,25) = 1,875$ , e  $\sigma = \sqrt{1,875}$ . Novamente, apesar de  $X$  só poder assumir valores inteiros,  $E(X)$  não precisa ser um inteiro. Se executarmos um grande número de experimentos binomiais independentes, cada um com  $n = 10$  tentativas e  $p = 0,75$ , o número médio de  $S$  por experimento será próximo de 7,5. ■

## Exercícios | Seção 3.4 (44–63)

44. Calcule as seguintes probabilidades binomiais diretamente pela fórmula de  $b(x; n, p)$ :
- $b(3; 8, 0,6)$
  - $b(5; 8, 0,6)$
  - $P(3 \leq X \leq 5)$  quando  $n = 8$  e  $p = 0,6$
  - $P(1 \leq X)$  quando  $n = 12$  e  $p = 0,1$
45. Use a Tabela do Apêndice A.1 para obter as probabilidades a seguir:
- $B(4; 10, 0,3)$
  - $b(4; 10, 0,3)$
  - $b(6; 10, 0,7)$
  - $P(2 \leq X \leq 4)$  quando  $X \sim \text{Bin}(10, 0,3)$
  - $P(2 \leq X)$  quando  $X \sim \text{Bin}(10, 0,3)$
  - $P(X \leq 1)$  quando  $X \sim \text{Bin}(10, 0,7)$
  - $P(2 < X < 6)$  quando  $X \sim \text{Bin}(10, 0,3)$
46. Quando as placas de circuito integrado usadas na fabricação de CD-players são testadas, a porcentagem de placas com defeitos no longo prazo é igual a 5%. Seja  $X =$  número de placas com defeito em uma amostra aleatória de tamanho  $n = 25$ , de forma que  $X \sim \text{Bin}(25, 0,05)$ .
- Determine  $P(X \leq 2)$ .
  - Determine  $P(X \geq 5)$ .
  - Determine  $P(1 \leq X \leq 4)$ .
  - Qual é a probabilidade de que nenhuma das 25 placas apresente defeito?
  - Calcule o valor esperado e o desvio padrão de  $X$ .
47. Uma empresa de cristais finos sabe por experiência que 10% de suas taças possuem defeitos cosméticos e devem ser classificadas como “de segunda linha”.
- Entre seis taças selecionadas aleatoriamente, qual é a probabilidade de uma ser de segunda linha?
  - Entre seis taças selecionadas aleatoriamente, qual é a probabilidade de no mínimo duas serem de segunda linha?
  - Se as taças forem examinadas uma a uma, qual será a probabilidade de no máximo cinco terem de ser selecionadas para encontrar quatro que não sejam de segunda linha?
48. Suponha que apenas 25% de todos os motoristas parem completamente um cruzamento com semáforos vermelhos para todas as direções quando não há outros carros à vista. Qual é a probabilidade de que, entre 20 motoristas selecionados aleatoriamente chegando em um cruzamento nessas condições,
- no máximo seis parem totalmente?
  - exatamente seis parem completamente?
  - ao menos seis parem completamente?
  - quantos dos 20 motoristas você espera que parem completamente?
49. O Exercício 29 (Seção 3.3) forneceu a fmp de  $Y$ , o número de violações do código de trânsito de um indivíduo selecionado aleatoriamente com seguro de uma certa empresa. Qual é a probabilidade de que entre 15 de tais indivíduos selecionados aleatoriamente,
- pelo menos 10 não tenham violações?
  - menos da metade tenha ao menos uma violação?
  - o número das pessoas com ao menos uma violação esteja entre 5 e 10, inclusive?<sup>2</sup>
50. Um determinado tipo de raquete de tênis possui duas versões: média e grande. Sessenta por cento de todos os clientes de certa loja querem a versão grande.
- Entre 10 clientes selecionados aleatoriamente que querem esse tipo de raquete, qual é a probabilidade de ao menos seis quererem a versão grande?
  - Entre 10 clientes selecionados aleatoriamente, qual é a probabilidade de o número dos que desejam a versão grande estar dentro do intervalo de um desvio padrão da média?
  - A loja possui atualmente sete raquetes de cada versão. Qual é a probabilidade de os 10 próximos clientes que querem essa raquete conseguirem obter a versão desejada do estoque atual?
51. Vinte por cento de todos os telefones de um determinado tipo são enviados para reparo em garantia. Desses, 60% podem ser reparados, enquanto os outros 40%

<sup>2</sup>“Entre  $a$  e  $b$ , inclusive” é equivalente a  $(a \leq X \leq b)$ .

devem ser substituídos. Se uma empresa comprar 10 desses telefones, qual é a probabilidade de exatamente dois serem substituídos em garantia?

52. O College Board relata que 2% dos 2 milhões de alunos que fazem o SAT cada ano recebem acomodações especiais por causa de deficiência física documentada (*Los Angeles Times*, July 16, 2002). Considere uma amostra aleatória de 25 estudantes que fizeram o teste recentemente.
- Qual é a probabilidade de exatamente 1 ter recebido acomodação especial?
  - Qual é a probabilidade de ao menos 1 ter recebido acomodação especial?
  - Qual é a probabilidade de ao menos 2 terem recebido acomodação especial?
  - Qual é a probabilidade de o número entre os 25 que receberam acomodação especial estar dentro de 2 desvios padrão do número que você espera serem acomodados?
  - Suponha que um estudante que não receba acomodação especial tenha 3 horas para fazer o exame e um que receba tenha 4,5. Qual você espera que seja o tempo médio dos 25 estudantes selecionados?
53. Suponha que 90% de todas as pilhas de certo fabricante tenham voltagens aceitáveis. Um determinado tipo de lanterna necessita de duas pilhas tipo D, e ela só funciona se as duas pilhas tiverem voltagem aceitável. Entre 10 lanternas selecionadas aleatoriamente, qual é a probabilidade de pelo menos nove funcionarem? Que hipótese você fez no decorrer da resposta à pergunta proposta?
54. Um grande lote de componentes chegou em um distribuidor e só pode ser classificado como aceitável se a proporção de componentes com defeito for no máximo 0,10. O distribuidor decide selecionar 10 componentes aleatoriamente e aceitar o lote apenas se o número de componentes defeituosos da amostra for no máximo 2.
- Qual é a probabilidade de o lote ser aceito quando a proporção real de itens com defeito for 0,01? 0,05? 0,10? 0,20? 0,25?
  - Seja  $p$  a proporção real de itens com defeito no lote. Um gráfico de  $P(\text{lote é aceito})$  em função de  $p$ , com  $p$  no eixo horizontal e  $P(\text{lote é aceito})$  no eixo vertical é denominado *curva característica de operação* do plano de aceitação da amostragem. Use os resultados da parte (a) para fazer o gráfico da curva para  $0 \leq p \leq 1$ .
  - Repita as partes (a) e (b) com “1” substituindo “2” no plano de aceitação da amostragem.
  - Repita as partes (a) e (b) com “15” substituindo “10” no plano de aceitação da amostragem.
  - Qual dos três planos de amostragem é mais satisfatório e por quê?
55. Uma norma que exige a instalação de um detector de fumaça em todas as casas construídas anteriormente está em vigor há um ano em certa cidade. O corpo de bombeiros está preocupado porque muitas casas continuam sem detectores. Seja  $p$  = proporção real de casas que possuem detectores e suponha que uma amostra aleatória de 25 lares seja inspecionada. Se a amostra indica fortemente que pouco menos de 80% de todas as casas possuem detector, o corpo de bombeiros fará uma campanha por um programa de inspeção obrigatória. Por causa do custo do programa, o corpo de bombeiros prefere não pedir essas inspeções a menos que a amostra renda evidências fortes que comprovem sua necessidade. Seja  $X$  o número de casas com detectores entre as 25 da amostra. Considere rejeitar a hipótese de  $p \geq 0,8$  se  $x \leq 15$ .
- Qual é a probabilidade de a hipótese ser rejeitada se o valor real de  $p$  for 0,8?
  - Qual é a probabilidade de não rejeitar a hipótese quando  $p = 0,7$ ? Quando  $p = 0,6$ ?
  - Como as “probabilidades de erro” das partes (a) e (b) mudam se o valor 15 na regra de decisão for alterado para 14?
56. Uma ponte cobra um pedágio de US\$ 1,00 para carros de passeio e de US\$ 2,50 para outros veículos. Suponha que, durante o dia, 60% de todos os veículos sejam carros de passeio. Se 25 veículos cruzarem a ponte durante um determinado período do dia, qual será a receita esperada resultante? [Sugestão: Seja  $X$  = número de carros de passeio. A receita total  $h(X)$  é uma função linear de  $X$ .]
57. Um estudante que está tentando escrever um trabalho para um curso tem a escolha de dois tópicos: A e B. Se o aluno escolher o tópico A, solicitará dois livros por empréstimo da biblioteca, e, se escolher B, serão solicitados quatro livros. O estudante acredita que, para escrever um bom trabalho, precisa receber e usar ao menos metade dos livros selecionados para cada tópico escolhido. Se a probabilidade de um livro solicitado chegar em tempo for de 0,9 e os livros chegam independentemente um do outro, que tópico o aluno deve escolher para maximizar a probabilidade de escrever um bom artigo? E se a probabilidade for de apenas 0,5 em vez de 0,9?
58. a. Para  $n$  determinado, há valores de  $p$  ( $0 \leq p \leq 1$ ) para os quais  $V(X) = 0$ ? Explique.  
b. Para que valor de  $p$   $V(X)$  é maximizada? [Sugestão: faça o gráfico de  $V(X)$  em função de  $p$  ou calcule a derivada.]
59. a. Demonstre que  $b(x; n, 1 - p) = b(n - x; n, p)$ .  
b. Demonstre que  $B(x; n, 1 - p) = 1 - B(n - x - 1; n, p)$ . [Sugestão: no máximo  $x$  S's serão equivalentes a pelo menos  $(n - x)$  F's.]  
c. O que as partes (a) e (b) implicam sobre a necessidade de se incluir valores de  $p$  maiores do que 0,5 na Tabela A.1 do Apêndice?
60. Demonstre que  $E(X) = np$  quando  $X$  é uma variável aleatória binomial. [Sugestão: primeiro expresse  $E(X)$  como uma soma com limite inferior  $x = 1$ . Depois fature

$np$ , assumindo  $y = x - 1$  de forma que a soma varie de  $y = 0$  a  $y = n - 1$  e mostre que a soma é igual a 1.]

61. Os clientes de um posto de gasolina pagam com cartão de crédito (A), cartão de débito (B) ou dinheiro (C). Assuma que clientes sucessivos façam escolhas independentes, com  $P(A) = 0,5$ ,  $P(B) = 0,2$  e  $P(C) = 0,3$ .
- Entre os próximos 100 clientes, qual será a média e a variância do número dos que pagam com cartão de débito? Explique seu raciocínio.
  - Responda à parte (a) para o número entre os 100 que não pagam em dinheiro.
62. Uma limusine de aeroporto pode acomodar até quatro passageiros em qualquer corrida. A empresa aceitará um máximo de seis reservas e os passageiros devem ter reservas. Pelos registros anteriores, 20% de todos os que fazem reservas não aparecem para a corrida. Responda as seguintes perguntas, assumindo independência quando apropriado.

- Se forem feitas seis reservas, qual é a probabilidade de ao menos um indivíduo com reserva não poder ser acomodado na corrida?
- Se forem feitas seis reservas, qual é o número esperado de lugares disponíveis quando a limusine parte?
- Suponha que a distribuição de probabilidade do número de reservas feitas seja dada na tabela a seguir.

Número de reservas	3	4	5	6
Probabilidade	0,1	0,2	0,3	0,4

Seja  $X$  = número de passageiros de uma corrida selecionada aleatoriamente. Calcule a função de distribuição de probabilidade de  $X$ .

63. Refira-se à desigualdade de Chebyshev fornecida no Exercício 43. Calcule  $P(|X - \mu| \geq k\sigma)$  para  $k = 2$  e  $k = 3$  quando  $X \sim \text{Bin}(20, 0,5)$  e compare com o limite superior correspondente. Repita para  $X \sim \text{Bin}(20, 0,75)$ .

### 3.5 Distribuições binomiais hipergeométrica e negativa

As distribuições binomiais hipergeométrica e negativa são ambas intimamente relacionadas à distribuição binomial. Enquanto a distribuição binomial é o modelo aproximado de amostragem sem reposição de uma população ( $S$ - $F$ ) dicotômica finita, a distribuição hipergeométrica é o modelo de probabilidade para o número de  $S$ 's em uma amostra. A  $va$  binomial  $X$  é o número de  $S$  quando o número de tentativas  $n$  está estabelecido, enquanto a distribuição binomial negativa resulta da fixação do número de  $S$  desejado, deixando que o número de tentativas seja aleatório.

#### Distribuição hipergeométrica

As hipóteses que levam à distribuição hipergeométrica são as seguintes:

- A população ou o conjunto de onde é retirada a amostra consiste de  $N$  indivíduos, objetos ou elementos (população finita).
- Cada indivíduo é classificado como sucesso ( $S$ ) ou falha ( $F$ ) e há  $M$  sucessos na população.
- É selecionada uma amostra sem reposição de  $n$  indivíduos de forma que cada subconjunto de tamanho  $n$  seja igualmente provável de ser escolhido.

A variável aleatória de interesse é  $X$  = número de  $S$ 's na amostra. A distribuição de probabilidade de  $X$  depende dos parâmetros  $n$ ,  $M$  e  $N$ , de forma que queremos calcular  $P(X = x) = h(x; n, M, N)$ .

#### Exemplo 3.34

Durante determinado período, um escritório de tecnologia da informação de uma universidade recebeu 20 ordens de serviço de problemas com impressoras, das quais 8 de impressoras a laser e 12 a jato de tinta. Uma amostra de 5 dessas ordens de serviço será selecionada para inclusão em uma pesquisa de satisfação do cliente. Suponha que as 5 sejam selecionadas de forma completamente aleatória para que qualquer subconjunto de tamanho 5 tenha a mesma possibilidade de ser selecionado (imagine colocar os números 1, 2, ..., 20 em 20 tarjas de papel, misturá-las e escolher cinco delas). Qual será a probabilidade de exatamente  $x$  ( $x = 0, 1, 2, 3, 4$  ou  $5$ ) das ordens de serviço selecionadas serem de impressoras a jato de tinta?



Neste exemplo, o tamanho da população é  $N = 20$ , o tamanho da amostra é  $n = 5$  e o número de  $S$  (jato de tinta =  $S$ ) e  $F$  da população são  $M = 12$  e  $N - M = 8$ , respectivamente. Considere o valor  $x = 2$ . Como todos os resultados (cada um consistindo em 5 ordens) são igualmente prováveis,

$$P(X = 2) = h(2; 5, 12, 20) = \frac{\text{número de resultados com } X = 2}{\text{número de resultados possíveis}}$$

O número de resultados possíveis no experimento é o número de maneiras de selecionar 5 dos 20 objetos sem se importar com a ordem, isto é,  $\binom{20}{5}$ . Para contar o número de resultados tendo  $X = 2$ , observe que há  $\binom{12}{2}$  formas de selecionar 2 das ordens de serviço de impressoras a jato de tinta e para cada forma há  $\binom{8}{3}$  maneiras de selecionar as 3 ordens de serviço de impressoras a laser para completar a amostra. A regra do produto do Capítulo 2 fornece então  $\binom{12}{2}\binom{8}{3}$  como o número de resultados com  $X = 2$ , assim

$$h(2; 5, 12, 20) = \frac{\binom{12}{2}\binom{8}{3}}{\binom{20}{5}} = \frac{77}{323} = 0,238 \quad \blacksquare$$

Em geral, se o tamanho da amostra  $n$  for menor que o número de sucessos da população ( $M$ ), o maior valor possível de  $X$  será  $n$ . Entretanto, se  $M < n$  (por exemplo, uma amostra de tamanho 25 e apenas 15 sucessos na população), então  $X$  pode ser no máximo  $M$ . De forma similar, sempre que o número de falhas da população ( $N - M$ ) exceder o tamanho da amostra, o menor valor possível de  $X$  será 0 (porque todos os indivíduos da amostra podem ser falhas). Entretanto, se  $N - M < n$ , o menor valor possível de  $X$  será  $n - (N - M)$ . Resumindo, os valores possíveis de  $X$  satisfazem à restrição máx. $(0, n - (N - M)) \leq x \leq \text{mín.}(n, M)$ . Um argumento paralelo ao do exemplo anterior fornece a fmp de  $X$ .

#### PROPOSIÇÃO

Se  $X$  for o número de  $S$  de uma amostra completamente aleatória de tamanho  $n$  tirada de uma população constituída de  $M$   $S$ 's e  $(N - M)$   $F$ 's, então a distribuição de probabilidade de  $X$ , denominada **distribuição hipergeométrica**, será dada por

$$P(X = x) = h(x; n, M, N) = \frac{\binom{M}{x}\binom{N - M}{n - x}}{\binom{N}{n}} \quad (3.15)$$

para um inteiro  $x$  que satisfaça máx. $(0, n - N + M) \leq x \leq \text{mín.}(n, M)$ .

No Exemplo 3.34,  $n = 5$ ,  $M = 12$  e  $N = 20$ , de forma que  $h(x; 5, 12, 20)$  para  $x = 0, 1, 2, 3, 4, 5$  pode ser obtida, substituindo-se esses números na Equação (3.15).

#### Exemplo 3.35

Cinco indivíduos de uma população animal supostamente ameaçada de extinção em certa região foram capturados, marcados e liberados para se misturarem à população. Após terem uma oportunidade de cruzarem, foi selecionada uma amostra aleatória de 10 desses animais. Seja  $X$  = número de animais marcados na segunda amostra. Se, na verdade, houver 25 animais desse tipo na região, qual será a probabilidade de (a)  $X = 2$ ? (b)  $X \leq 2$ ?

Os valores dos parâmetros são  $n = 10$ ,  $M = 5$  (5 animais marcados na população) e  $N = 25$ , assim

$$h(x; 10, 5, 25) = \frac{\binom{5}{x}\binom{20}{10 - x}}{\binom{25}{10}} \quad x = 0, 1, 2, 3, 4, 5$$

Para a parte (a),

$$P(X = 2) = h(2; 10, 5, 25) = \frac{\binom{5}{2} \binom{20}{8}}{\binom{25}{10}} = 0,385$$

Para a parte (b),

$$\begin{aligned} P(X \leq 2) &= P(X = 0, 1, \text{ ou } 2) = \sum_{x=0}^2 h(x; 10, 5, 25) \\ &= 0,057 + 0,257 + 0,385 = 0,699 \end{aligned}$$

Estão disponíveis tabelas abrangentes da distribuição hipergeométrica, mas, como a distribuição possui três parâmetros, as tabelas requerem muito mais espaço do que as da distribuição binomial. O MINITAB e outros programas estatísticos geram probabilidades hipergeométricas facilmente.

Como no caso binomial, há expressões simples para  $E(X)$  e  $V(X)$  para vas hipergeométricas.

**PROPOSIÇÃO**

A média e a variância da va hipergeométrica  $X$  com fmp  $h(x; n, M, N)$  são

$$E(X) = n \cdot \frac{M}{N} \quad V(X) = \left( \frac{N - n}{N - 1} \right) \cdot n \cdot \frac{M}{N} \cdot \left( 1 - \frac{M}{N} \right)$$

A razão  $M/N$  é a proporção de  $S$  na população. Se substituirmos  $M/N$  por  $p$  em  $E(X)$  e  $V(X)$ , obtemos

$$\begin{aligned} E(X) &= np \\ V(X) &= \left( \frac{N - n}{N - 1} \right) \cdot np(1 - p) \end{aligned} \tag{3.16}$$

A Expressão (3.16) mostra que as médias das vas binomial e hipergeométrica são iguais, enquanto as variâncias das duas vas diferem pelo fator  $(N - n)/(N - 1)$ , freqüentemente denominado **fator de correção de população finita**. Esse fator é menor do que 1, de forma que a variável hipergeométrica tem variância menor do que a va binomial. O fator de correção pode ser escrito como  $(1 - n/N)/(1 - 1/N)$ , que é aproximadamente 1 quando  $n$  é pequeno em relação a  $N$ .

**Exemplo 3.36** (Continuação do Exemplo 3.35)

No exemplo de marcação dos animais,  $n = 10$ ,  $M = 5$ , e  $N = 25$ , de forma que  $p = \frac{5}{25} = 0,2$  e

$$\begin{aligned} E(X) &= 10(0,2) = 2 \\ V(X) &= \frac{15}{24} (10)(0,2)(0,8) = (0,625)(1,6) = 1 \end{aligned}$$

Se a amostragem fosse feita com reposição,  $V(X) = 1,6$ .

Suponha que o tamanho da população  $N$  não seja conhecido, de forma que o valor  $x$  é observado e queremos estimar  $N$ . É razoável igualar a proporção da amostra observada de  $S$ ,  $x/n$ , à proporção da população,  $M/N$ , fornecendo a estimativa

$$\hat{N} = \frac{M \cdot n}{x}$$

Se  $M = 100$ ,  $n = 40$  e  $x = 16$ , então  $\hat{N} = 250$ .

Nossa regra prática geral da Seção 3.4 afirmava que, se a amostragem foi feita sem reposição mas  $n/N$  era no máximo 0,05, a distribuição binomial podia ser usada para calcular probabilidades aproximadas, envolvendo o número de  $S$  na amostra. Segue uma definição mais precisa: considere que o tamanho da população  $N$  e o número  $S$ 's da população  $M$ , cresçam à medida que a relação  $M/N$  aproxima-se de  $p$ . Então  $h(x; n, M, N)$  se aproxima de  $b(x; n, p)$ . Para  $n/N$  pequeno, os dois são aproximadamente iguais, desde que  $p$  não esteja muito perto tanto de 0 como de 1. Essa é a base de nossa regra prática.

## Distribuição Binomial Negativa

A  $va$  e a distribuição binomial negativa são baseadas em experimentos que satisfaçam às condições a seguir:

1. O experimento consiste de uma seqüência de tentativas independentes.
2. Cada tentativa resulta em sucesso ( $S$ ) ou em falha ( $F$ ).
3. A probabilidade de sucesso é constante de uma tentativa para outra, então  $P(S \text{ na tentativa } i) = p$  para  $i = 1, 2, 3, \dots$
4. O experimento continua (as tentativas são executadas) até ser observado um total de  $r$  sucessos, sendo  $r$  um inteiro positivo.

A variável aleatória de interesse é  $X =$  número de falhas que precedem o  $r$ -ésimo sucesso.  $X$  é denominada **variável aleatória binomial negativa** porque, em contraste com a  $va$  binomial, o número de sucessos é fixo e o de tentativas é aleatório.

Valores possíveis de  $X$  são  $0, 1, 2, \dots$ . Represente por  $nb(x; r, p)$  a fmp de  $X$ . O evento  $\{X = x\}$  é equivalente a  $\{r - 1 S$ 's nas primeiras  $(x + r - 1)$  tentativas e um  $S$  na  $(x + r)$ ésima tentativa} (por exemplo: se  $r = 5$  e  $x = 10$ , deve haver quatro  $S$  nas primeiras 14 tentativas e a tentativa 15 deve ser um  $S$ ). Como as tentativas são independentes,

$$\begin{aligned} nb(x; r, p) &= P(X = x) \\ &= P(r - 1 S's \text{ nas primeiras } x + r - 1 \text{ tentativas}) \cdot P(S) \end{aligned} \quad (3.17)$$

A primeira probabilidade do membro direito da Expressão (3.17) é a probabilidade binomial

$$\binom{x + r - 1}{r - 1} p^{r-1} (1 - p)^x \quad \text{onde } P(S) = p$$

### PROPOSIÇÃO

A fmp da  $va$  binomial negativa  $X$  com parâmetros  $r =$  número de  $S$ 's e  $p = P(S)$  é

$$nb(x; r, p) = \binom{x + r - 1}{r - 1} p^r (1 - p)^x \quad x = 0, 1, 2, \dots$$

### Exemplo 3.37

Um pediatra deseja convocar cinco casais, cada um esperando seu primeiro filho, para participarem de um novo regime de parto natural. Seja  $p = P$  (um casal selecionado aleatoriamente concorda em participar). Se  $p = 0,2$ , qual é a probabilidade de 15 casais serem solicitados antes de serem encontrados cinco que concordem em participar? Isto é, com  $S = \{\text{concorda em participar}\}$ , qual é a probabilidade de ocorrerem 10  $F$  antes do quinto  $S$ ? Substituindo  $r = 5$ ,  $p = 0,2$  e  $x = 10$  em  $nb(x; r, p)$ , temos

$$nb(10; 5, 0,2) = \binom{14}{4} (0,2)^5 (0,8)^{10} = 0,034$$

A probabilidade de no máximo 10  $F$  serem observados (no máximo 15 casais serem entrevistados) é

$$P(X \leq 10) = \sum_{x=0}^{10} nb(x; 5, 0,2) = (0,2)^5 \sum_{x=0}^{10} \binom{x+4}{4} (0,8)^x = 0,164 \quad \blacksquare$$

Em algumas fontes de informações, a va binomial negativa é considerada como o número de tentativas  $X + r$  em vez de número de falhas.

No caso particular de  $r = 1$ , a fmp é

$$nb(x; 1, p) = (1 - p)^x p \quad x = 0, 1, 2, \dots \quad (3.18)$$

No Exemplo 3.10, deduzimos a fmp em relação ao número de tentativas necessárias para se obter o primeiro  $S$  e aquela fmp é semelhante à Expressão (3.18). Tanto  $X =$  número de  $F$  como  $Y =$  número de tentativas ( $= 1 + X$ ) são denominadas **variáveis aleatórias geométricas** na literatura especializada e a fmp da Expressão (3.18) é denominada **distribuição geométrica**.

No Exemplo 3.18, foi demonstrado que o valor esperado de tentativas até a obtenção do primeiro  $S$  é  $1/p$ , de forma que o valor esperado de  $F$  até a obtenção do primeiro  $S$  é  $(1/p) - 1 = (1 - p)/p$ . Intuitivamente, esperaríamos ver  $r \cdot (1 - p)/p$   $F$  antes do  $r$ -ésimo  $S$  e esta é, na verdade,  $E(X)$ . Também há uma fórmula simples para  $V(X)$ .

#### PROPOSIÇÃO

Se  $X$  é uma va binomial negativa com fmp  $nb(x; r, p)$ , então

$$E(X) = \frac{r(1 - p)}{p} \quad V(X) = \frac{r(1 - p)}{p^2}$$

Finalmente, desenvolvendo o coeficiente binomial de  $p^r(1 - p)^x$  e fazendo algumas simplificações, vê-se que  $nb(x; r, p)$  é bem definido, mesmo quando  $r$  não é inteiro. Verifica-se que a *distribuição binomial negativa geral* ajusta muito bem os dados observados em uma grande variedade de aplicações.

### Exercícios Seção 3.5 (64–74)

64. Certo tipo de câmera digital é oferecida em duas versões de três megapixel e quatro megapixel. Uma loja de câmeras recebeu uma encomenda de 15 dessas câmeras, das quais seis com resolução de três megapixel. Suponha que cinco delas sejam selecionadas aleatoriamente para serem estocadas atrás do balcão. As outras 10 são colocadas na área de armazenagem. Seja  $X =$  número de câmeras de três megapixel entre as cinco selecionadas para armazenagem atrás do balcão.
- Que tipo de distribuição tem  $X$  (nome e valores de todos os parâmetros)?
  - Calcule  $P(X = 2)$ ,  $P(X \leq 2)$  e  $P(X \geq 2)$ .
  - Calcule o valor médio e o desvio padrão de  $X$ .
65. Doze refrigeradores de um determinado tipo foram devolvidos ao distribuidor por causa de um ruído audível, oscilante e agudo que faziam quando em funcionamento. Suponha que sete desses refrigeradores possuam um compressor defeituoso e os outros cinco tenham problemas graves. Se os refrigeradores forem examinados em ordem aleatória, seja  $X$  o número de refrigeradores examinados entre os seis primeiros com compressores defeituosos. Calcule os seguintes valores:
- $P(X = 5)$
  - $P(X \leq 4)$
  - A probabilidade de  $X$  exceder o valor da média em mais de 1 desvio padrão.
  - Considere uma entrega de 400 refrigeradores, dos quais 40 possuem compressores com defeito. Se  $X$  for o número de compressores defeituosos entre 15 refrigeradores selecionados aleatoriamente, descreva uma forma menos cansativa de calcular (ao menos aproximadamente)  $P(X \leq 5)$  em vez de usar a fmp hipergeométrica.
66. Um instrutor que lecionou estatística para engenheiros para duas turmas no semestre passado, a primeira com 20 alunos e a segunda com 30, decidiu pedir aos alunos um projeto semestral. Após a entrega de todos os projetos, o instrutor os organizou aleatoriamente antes de corrigi-los. Considere os primeiros 15 projetos a serem corrigidos.
- Qual é a probabilidade de exatamente 10 projetos serem da segunda turma?

- b. Qual é a probabilidade de pelo menos 10 projetos serem da segunda turma?
- c. Qual é a probabilidade de ao menos 10 projetos serem da mesma turma?
- d. Qual é o valor médio e o desvio padrão do número de projetos entre esses 15 que pertencem à segunda turma?
- e. Qual é o valor médio e o desvio padrão do número de projetos que não estão entre os 15 primeiros que pertencem à segunda turma?
67. Um geólogo coletou 10 amostras de rocha basáltica e 10 de granito. Instruiu o geólogo assistente de laboratório para selecionar aleatoriamente 15 amostras para análise.
- a. Qual é o número de amostras de granito selecionadas para análise?
- b. Qual é a probabilidade de todas as amostras de um dos dois tipos de rocha serem selecionadas para análise?
- c. Qual é a probabilidade de o número de amostras de granito selecionadas para análise estarem dentro de 1 desvio padrão do valor médio?
68. Um diretor de pessoal que entrevistará 11 engenheiros seniores para quatro cargos marcou seis entrevistas para o primeiro dia e cinco para o segundo. Assuma que os candidatos serão entrevistados em ordem aleatória.
- a. Qual é a probabilidade de  $x$  dos quatro melhores candidatos serem entrevistados no primeiro dia?
- b. Dos quatro melhores candidatos, quantos podem ser esperados para entrevistas no primeiro dia?
69. Vinte pares de indivíduos inscritos em um torneio de *bridge* foram designados com os números 1, ..., 20. Na primeira parte do torneio, os 20 pares são divididos aleatoriamente em 10 pares leste-oeste e 10 pares norte-sul.
- a. Qual é a probabilidade de  $x$  dos 10 melhores pares jogarem leste-oeste?
- b. Qual é a probabilidade de todos os cinco melhores pares jogarem na mesma direção?
- c. Se há  $2n$  pares, qual é a fmp de  $X =$  número entre os  $n$  melhores pares que jogam leste-oeste? Quais são  $E(X)$  e  $V(X)$ ?
70. Um alerta de poluição de segundo estágio foi feito em uma determinada área de Los Angeles County, onde há 50 indústrias. Um inspetor visitará 10 indústrias selecionadas aleatoriamente para verificação de violações dos regulamentos.
- a. Se 15 empresas realmente estiverem violando ao menos um regulamento, qual é a fmp do número de empresas visitadas pelo inspetor que estão violando ao menos um regulamento?
- b. Se houver 500 empresas na área, das quais 150 com violação, faça a aproximação da fmp da parte (a) para uma fmp mais simples.
- c. Para  $X =$  número entre as 10 empresas visitadas que estão violando os regulamentos, calcule  $E(X)$  e  $V(X)$  para a fmp exata e aproximada na parte (b).
71. Suponha que  $p = P(\text{nascimento de menino}) = 0,5$ . Um casal quer ter exatamente duas meninas e terá filhos até essa condição ser satisfeita.
- a. Qual é a probabilidade de a família ter  $x$  filhos homens?
- b. Qual é a probabilidade de a família ter quatro filhos?
- c. Qual é a probabilidade de a família ter no máximo quatro filhos?
- d. Quantos filhos homens espera-se que essa família tenha? Quantos filhos espera-se que essa família tenha?
72. Uma família decide ter filhos até ter três do mesmo sexo. Assumindo  $P(H) = P(M) = 0,5$ , qual é a fmp de  $X =$  número de filhos na família?
73. Três irmãos e suas esposas decidem ter filhos até que cada família tenha duas meninas. Qual é a fmp de  $X =$  número total de filhos homens nascidos nas famílias? Qual é  $E(X)$  e como pode ser comparada ao número esperado de filhos nascidos em cada família?
74. O indivíduo A tem um dado vermelho e B tem um verde (ambos justos). Se cada um lançar o dado até obterem cinco “duplas” (1-1, ..., 6-6), qual será a fmp de  $X =$  número total de vezes que um dado é lançado? Quais são  $E(X)$  e  $V(X)$ ?

### 3.6 | Distribuição de probabilidade de Poisson

As distribuições binomial, hipergeométrica e binomial negativa foram deduzidas a partir de um experimento consistindo de tentativas ou retiradas e na aplicação das leis da probabilidade aos diversos resultados do experimento. Não há um experimento simples que sirva de base para a distribuição de Poisson, apesar de descrevermos simplificarmente como ela pode ser obtida por meio de certas operações limitantes.

#### DEFINIÇÃO

Uma variável aleatória  $X$  tem **distribuição de Poisson** com parâmetro  $\lambda$  ( $\lambda > 0$ ) se a fmp de  $X$  for

$$p(x; \lambda) = \frac{e^{-\lambda} \lambda^x}{x!} \quad x = 0, 1, 2, \dots$$

Geralmente, o valor de  $\lambda$  é uma taxa por unidade de tempo ou por unidade de área. A letra  $e$  em  $p(x; \lambda)$  é a base dos logaritmos naturais. Seu valor numérico é aproximadamente 2,71828. Como  $\lambda$  é positivo,  $p(x; \lambda) > 0$  para todos os valores possíveis de  $x$ . O fato de  $\sum_{x=0}^{\infty} p(x; \lambda) = 1$  é uma consequência da expansão de  $e^\lambda$  da série infinita de Maclaurin, que aparece na maioria dos livros-texto de cálculo:

$$e^\lambda = 1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \dots = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} \tag{3.19}$$

Se os dois termos extremos da Expressão (3.19) forem multiplicados por  $e^{-\lambda}$  e depois  $e^{-\lambda}$  for colocado dentro da somatória, o resultado será

$$1 = \sum_{x=0}^{\infty} e^{-\lambda} \frac{\lambda^x}{x!}$$

o que mostra que  $p(x; \lambda)$  satisfaz a segunda condição necessária para definir uma fmp.

**Exemplo 3.38**

Seja  $X$  o número de certo tipo de animais capturados em uma armadilha durante certo período de tempo. Suponha que  $X$  tenha uma distribuição de Poisson com  $\lambda = 4,5$ , de forma que, em média, cada armadilha contém 4,5 animais. [O artigo “Dispersal Dynamics of the Bivalve *Gemma Gemma* in a Patchy Environment (*Ecological Monographs*, 1995, p. 1-20) sugere esse modelo. O bivalvo *Gemma gemma* é um pequeno molusco]. A probabilidade de uma armadilha conter exatamente cinco animais é

$$P(X = 5) = \frac{e^{-4,5}(4,5)^5}{5!} = 0,1708$$

A probabilidade de uma armadilha conter no máximo cinco animais é

$$P(X \leq 5) = \sum_{x=0}^5 \frac{e^{-4,5}(4,5)^x}{x!} = e^{-4,5} \left[ 1 + 4,5 + \frac{(4,5)^2}{2!} + \dots + \frac{(4,5)^5}{5!} \right] = 0,7029 \quad \blacksquare$$

**A distribuição de Poisson como um limite**

A base lógica para o uso da distribuição de Poisson em muitas situações é fornecida pela proposição a seguir.

**PROPOSIÇÃO**

Suponha que na fmp binomial  $b(x; n, p)$ , tenhamos  $n \rightarrow \infty$  e  $p \rightarrow 0$  de tal forma que  $np$  se aproxime de um valor  $\lambda > 0$ . Então  $b(x; n, p) \rightarrow p(x; \lambda)$ .

De acordo com essa proposição, em qualquer experimento binomial em que  $n$  é grande e  $p$  é pequeno,  $b(x; n, p) \approx p(x; \lambda)$ , onde  $\lambda = np$ . Como regra prática, tal aproximação pode ser aplicada se  $n \geq 100$ ,  $p \leq 0,01$  e  $np \leq 20$ .

**Exemplo 3.39**

Se uma editora de livros não-técnicos se esforça para garantir que seus livros não possuem erros tipográficos, de forma que a probabilidade de uma página conter um erro desse tipo é de 0,005 e os erros são independentes de página para página, qual é a probabilidade de um de seus romances de 400 páginas conter exatamente uma página com erros? No máximo três páginas com erros?

Com  $S$  representando uma página com ao menos um erro e  $F$  uma página sem erros, o número  $X$  de páginas que contém ao menos um erro é uma va binomial com  $n = 400$  e  $p = 0,005$ , de forma que  $np = 2$ . Queremos

$$P(X = 1) = b(1; 400, 0,005) \approx p(1; 2) = \frac{e^{-2}(2)^1}{1!} = 0,271$$

De forma similar,

$$P(X \leq 3) \approx \sum_{x=0}^3 p(x; 2) = \sum_{x=0}^3 e^{-2} \frac{2^x}{x!} = 0,135 + 0,271 + 0,271 + 0,180 = 0,857 \quad \blacksquare$$

A Tabela A.2 do Apêndice mostra a FDA  $F(x; \lambda)$  para  $\lambda = 0,1; 0,2; \dots, 1, 2; \dots, 10; 15$  e  $20$ . Por exemplo: se  $\lambda = 2$ , então  $P(X \leq 3) = F(3; 2) = 0,857$ , como no Exemplo 3.39, enquanto  $P(X = 3) = F(3; 2) - F(2; 2) = 0,180$ . De forma alternativa, muitos programas estatísticos geraram  $p(x; \lambda)$  e  $F(x; \lambda)$  quando solicitado.

## Média e variância de $X$

Como  $b(x; n, p) \rightarrow p(x; \lambda)$  quando  $n \rightarrow \infty, p \rightarrow 0, np \rightarrow \lambda$ , a média e a variância de uma variável binomial tendem para as de uma variável de Poisson. Os limites são  $np \rightarrow \lambda$  e  $np(1-p) \rightarrow \lambda$ .

### PROPOSIÇÃO

Se  $X$  tiver distribuição de Poisson com parâmetro  $\lambda$ , então  $E(X) = V(X) = \lambda$ .

Esses resultados também podem ser deduzidos diretamente das definições de média e variância.

**Exemplo 3.40** (Continuação do Exemplo 3.38)

Tanto o número de animais capturados como a variância dessa quantidade são iguais a  $4,5$  e  $\sigma_x = \sqrt{\lambda} = \sqrt{4,5} = 2,12$ . ■

## Processo de Poisson

Uma aplicação muito importante da distribuição de Poisson surge juntamente com a ocorrência de eventos de um tipo particular no decorrer do tempo. Como exemplo, suponha que, iniciando em um instante denominado  $t = 0$ , estejamos interessados em contar o número de pulsos radioativos registrados em um contador Geiger. Façamos as seguintes hipóteses sobre a maneira como os pulsos ocorrem:

1. Existe um parâmetro  $\alpha > 0$  tal que, para qualquer intervalo de tempo curto  $\Delta t$ , a probabilidade de ser recebido exatamente um pulso é  $\alpha \cdot \Delta t + o(\Delta t)$ .<sup>3</sup>
2. A probabilidade de ser recebido mais de um pulso durante  $\Delta t$  é  $o(\Delta t)$  [que, juntamente com a Hipótese 1 implica que a probabilidade de nenhum pulso durante  $\Delta t$  é  $1 - \alpha \cdot \Delta t - o(\Delta t)$ ].
3. O número de pulsos recebidos durante o intervalo de tempo  $\Delta t$  é independente da quantidade recebida antes do intervalo de tempo.

Informalmente, a Hipótese 1 diz que, para um intervalo de tempo curto, a probabilidade de se receber um único pulso é aproximadamente proporcional à duração do intervalo de tempo, onde  $\alpha$  é a constante de proporcionalidade. Seja agora  $P_k(t)$  a probabilidade de serem recebidos  $k$  pulsos pelo contador durante um intervalo de tempo de duração  $t$ .

### PROPOSIÇÃO

$P_k(t) = e^{-\alpha t} \cdot (\alpha t)^k / k!$ , de forma que o número de pulsos durante um intervalo de tempo  $t$  é uma va de Poisson com parâmetro  $\lambda = \alpha t$ . O número esperado de pulsos durante qualquer um de tais intervalos de tempo será então  $\alpha t$ , de forma que o número esperado durante uma unidade de intervalo de tempo será  $\alpha$ .

<sup>3</sup> Uma quantidade é  $o(\Delta t)$  (leia "o minúsculo de delta t") se, à medida que  $\Delta t$  se aproxima de 0,  $o(\Delta t)/\Delta t$  também o faz. Isto é,  $o(\Delta t)$  é ainda mais desprezível que o próprio  $\Delta t$ . A quantidade  $(\Delta t)^2$  possui essa propriedade, mas  $\text{sen}(\Delta t)$  não.

**Exemplo 3.41**

Suponha que os pulsos cheguem no contador em uma taxa média de seis por minuto, de modo que  $\alpha = 6$ . Para determinar a probabilidade de pelo menos um pulso ser recebido em um intervalo de meio minuto, observe que o número de pulsos em tal intervalo tem distribuição de Poisson com parâmetro  $\alpha t = 6(0,5) = 3$  (0,5 min. é usado porque  $\alpha$  é expresso como taxa por minuto). Então, sendo  $X =$  número de pulsos recebidos no intervalo de 30 segundos,

$$P(1 \leq X) = 1 - P(X = 0) = 1 - \frac{e^{-3}(3)^0}{0!} = 0,950 \quad \blacksquare$$

Se nas Hipóteses 1-3 substituirmos “pulso” por “evento”, então o número de eventos que ocorrem durante um intervalo de tempo fixo  $t$  tem distribuição de Poisson com parâmetro  $\alpha t$ . Qualquer processo que tenha essa distribuição é denominado **processo de Poisson** e  $\alpha$  é denominada *taxa do processo*. Outros exemplos de situações que originam um processo de Poisson incluem monitoramento do *estado* de um sistema de computadores ao longo do tempo, com as quebras constituindo os eventos de interesse; registro do número de acidentes em uma instalação industrial ao longo do tempo, tempo de resposta em uma central telefônica; e observação da quantidade de chuviscos de raios cósmicos em um determinado observatório ao longo do tempo.

Em vez de observar eventos ao longo do tempo, considere apenas observar eventos de um determinado tipo que ocorrem em uma região bi ou tridimensional. Por exemplo: podemos selecionar uma determinada região  $R$  de uma floresta em um mapa ir para aquela região e contar o número de árvores. Cada árvore representa um evento que ocorre em um ponto certo no espaço. Com hipóteses similares a 1-3, podemos demonstrar que o número de eventos que ocorrem em uma região  $R$  tem distribuição de Poisson com parâmetro  $\alpha \cdot a(R)$ , onde  $a(R)$  é a área de  $R$ . A quantidade  $\alpha$  é o número esperado de eventos por unidade de área ou de volume.

---

**Exercícios** | Seção 3.6 (75–89)

75. Seja  $X$  o número de falhas na superfície de uma caldeira de um determinado tipo selecionada aleatoriamente, com distribuição de Poisson de parâmetro  $\lambda = 5$ . Use a Tabela A.2 do Apêndice para calcular as probabilidades a seguir:
- a.  $P(X \leq 8)$     b.  $P(X = 8)$     c.  $P(9 \leq X)$   
d.  $P(5 \leq X \leq 8)$     e.  $P(5 < X < 8)$
76. Suponha que o número  $X$  de tornados observados em uma determinada região durante o período de um ano tenha uma distribuição de Poisson com  $\lambda = 8$ .
- a. Calcule  $P(X \leq 5)$ .  
b. Calcule  $P(6 \leq X \leq 9)$ .  
c. Calcule  $P(10 \leq X)$ .  
d. Qual é a probabilidade de o número observado de tornados exceder o número esperado por mais de 1 desvio padrão?
77. Suponha que o número de motoristas que trafega entre certa origem e destino durante um período de tempo determinado tenha uma distribuição de Poisson com parâmetro  $\lambda = 20$  (sugerido no artigo “Dynamic Ride Sharing: Theory and Practice,” *J. of Transp. Engr.*, 1997, p. 308-312). Qual é a probabilidade de o número de motoristas ser:
- a. No máximo 10?  
b. Exceder 20?  
c. Estar entre 10 e 20, inclusive? Estar estritamente entre 10 e 20?  
d. Estar dentro de dois desvios padrão do valor médio?
78. Considere gravar algo em um disco rígido e enviá-lo para um certificador contar o número de pulsos ausentes. Suponha que esse número  $X$  tenha uma distribuição de Poisson com parâmetro  $\lambda = 0,2$ . (Sugerido em “Average Sample Number for SemiCurtailed Sampling Using the Poisson Distribution,” *J. Quality Technology*, 1983, p. 126-129).
- a. Qual é a probabilidade de um disco ter exatamente um pulso ausente?  
b. Qual é a probabilidade de um disco ter dois pulsos ausentes?  
c. Se dois discos forem selecionados de forma independente, qual é a probabilidade de nenhum conter pulsos ausentes?
79. Um artigo no *Los Angeles Times* (3 de dezembro de 1993) relata que 1 em 200 pessoas possui o gene recessivo que causa câncer de cólon hereditário. Em uma amostra de 1000 indivíduos, qual é a distribuição aproximada do número dos que possuem o gene? Use essa distribuição para calcular a probabilidade aproximada de:
- a. Entre 5 e 8 (inclusive) possuírem o gene.  
b. Ao menos 8 possuírem o gene.



80. Suponha que apenas 0,10% de todos os computadores de certo tipo apresentem falhas de CPU durante o período de garantia. Considere uma amostra de 10.000 computadores.
- Qual é o valor esperado e o desvio padrão do número de computadores da amostra que apresentam defeito?
  - Qual é a probabilidade (aproximada) de mais de 10 computadores da amostra apresentarem defeito?
  - Qual é a probabilidade (aproximada) de nenhum computador da amostra apresentar defeito?
81. Suponha que pequenas aeronaves pousem em um aeroporto, de acordo com um processo de Poisson, com a taxa  $\alpha = 8$  por hora, de forma que o número de pousos durante um período de tempo de  $t$  horas é uma va de Poisson com parâmetro  $\lambda = 8t$ .
- Quais são a probabilidade de exatamente seis aeronaves pequenas chegarem durante um período de uma hora? Ao menos seis? Ao menos 10?
  - Qual é o valor esperado e o desvio padrão do número de pequenas aeronaves que chegam durante um período de 90 minutos?
  - Qual é a probabilidade de ao menos 20 aeronaves pequenas chegarem durante um período de  $2\frac{1}{2}$  horas? De no máximo 10 chegarem nesse período?
82. O número de pessoas que chegam para tratamento em um pronto-socorro pode ser modelado por um processo de Poisson com taxa de cinco por hora.
- Qual é a probabilidade de exatamente quatro pessoas chegarem em certa hora?
  - Qual é a probabilidade de ao menos quatro pessoas chegarem em certa hora?
  - Quantas pessoas você espera que cheguem em um período de 45 minutos?
83. O número de solicitações de assistência recebido por um serviço de guincho é um processo de Poisson com taxa  $\alpha = 4$  por hora.
- Calcule a probabilidade de exatamente dez solicitações chegarem em um certo período de 2 horas.
  - Se os operadores do serviço de guincho tirarem 30 minutos para almoço, qual é a probabilidade de não perderem nenhum chamado de assistência?
  - Quantas ligações você espera que ocorram durante o almoço?
84. Em testes de placas de circuitos, a probabilidade de falha em um diodo é de 0,01. Suponha que uma placa de circuito contenha 200 diodos.
- Quantos diodos espera-se que apresentem falhas e qual é o desvio padrão desse valor?
  - Qual é a probabilidade (aproximada) de ao menos quatro diodos apresentarem falha em uma placa selecionada aleatoriamente?
  - Se cinco placas forem enviadas a um determinado cliente, qual é a probabilidade de ao menos quatro funcionarem corretamente? (Uma placa só funciona corretamente se todos os seus diodos funcionarem.)
85. O artigo “Reliability-Based Service-Life Assessment of Aging Concrete Structures” (*J. Structural Engr.*, 1993, p. 1600-1621) sugere que um processo de Poisson pode ser usado para representar a ocorrência de cargas estruturais no tempo. Suponha que o tempo médio entre as ocorrências de cargas seja 0,5 ano.
- Quantas cargas podem ser esperadas durante um período de dois anos?
  - Qual é a probabilidade de mais de cinco cargas ocorrerem durante um período de dois anos?
  - Quanto tempo deve ter um período para que a probabilidade de não ocorrerem cargas seja no máximo 0,1?
86. Seja  $X$  uma distribuição de Poisson com parâmetro  $\lambda$ . Demonstre que  $E(X) = \lambda$  diretamente da definição de valor esperado. (*Sugestão:* O primeiro termo da soma é igual a 0, então  $x$  pode ser cancelado. Fatore  $\lambda$  e mostre que o restante soma 1.)
87. Suponha que as árvores sejam distribuídas em uma floresta de acordo com um processo bidimensional de Poisson com parâmetro  $\alpha$ . O número esperado de árvores por acre é igual a 80.
- Qual é a probabilidade de que em um quarto de acre haja no máximo 16 árvores?
  - Se a floresta cobrir 85.000 acres, qual será o número esperado de árvores na floresta?
  - Suponha que você selecione um ponto na floresta e construa um círculo de raio 0,1 milha. Seja  $X =$  número de árvores na região circular. Qual é a fmp de  $X$ ? (*Sugestão:* 1 milha quadrada = 640 acres.)
88. Automóveis chegam em uma estação de inspeção de equipamentos veiculares de acordo com um processo de Poisson com taxa  $\alpha = 10$  por hora. Suponha que, com probabilidade 0,5, um veículo chegue sem violações de equipamento.
- Qual é a probabilidade de exatamente 10 veículos chegarem durante uma hora e nenhum ter violações?
  - Para qualquer  $y \geq 10$  fixo, qual é a probabilidade de  $y$  veículos chegarem em uma hora, todos sem violações?
  - Qual é a probabilidade de 10 carros sem violações chegarem na próxima hora? [*Sugestão:* Some as probabilidades na parte (b) entre  $y = 10$  e  $\infty$ .]
89. a. Em um processo de Poisson, o que deve acontecer no intervalo de tempo  $(0, t)$  e no intervalo  $(t, t + \Delta t)$  para que nenhum evento ocorra em todo o intervalo  $(0, t + \Delta t)$ ? Use esta Hipótese e as Hipóteses 1-3 para escrever a relação entre  $P_0(t + \Delta t)$  e  $P_0(t)$ .
- b. Use o resultado da parte (a) para escrever uma expressão para a diferença  $P_0(t + \Delta t) - P_0(t)$ . Divida então por  $\Delta t$  e considere que  $\Delta t \rightarrow 0$  para obter uma

equação envolvendo  $(d/dt)P_0(t)$ , derivada de  $P_0(t)$  em relação a  $t$ .

- c. Verifique se  $P_0(t) = e^{-\alpha t}$  satisfaz à equação da parte (b).
- d. Pode-se demonstrar de forma similar às partes (a) e (b) que as  $P_k(t)$  devem satisfazer ao sistema de equações diferenciais.

$$\frac{d}{dt} P_k(t) = \alpha P_{k-1}(t) - \alpha P_k(t)$$

$$k = 1, 2, 3, \dots$$

Verifique se  $P_k(t) = e^{-\alpha t} (\alpha t)^k / k!$  satisfaz ao sistema. (Na verdade, esta é a única solução.)

Exercícios suplementares (90–116)

- 90.** Considere um baralho formado por sete cartas, marcadas 1, 2, ..., 7. Três dessas cartas são selecionadas aleatoriamente. Defina uma  $W$  como  $W =$  soma dos números resultantes e calcule a fmp de  $W$ . Depois, calcule  $\mu$  e  $\sigma^2$ . [Sugestão: Considere os resultados como não-ordenados, de forma que (1, 3, 7) e (3, 1, 7) não sejam considerados resultados diferentes. Então haverá 35 resultados que podem ser relacionados. (Esse tipo de  $W$ , na verdade, surge juntamente com o teste de hipótese denominado teste de soma de postos de Wilcoxon, em que há uma amostra  $x$  e uma amostra  $y$  e  $W$  é a soma dos postos dos  $x$  na amostra combinada.)]

- 91.** Após embaralhar um baralho de 52 cartas, um crupiê dá cinco cartas. Seja  $X =$  número de naipes representados na mão de cinco cartas.

a. Mostre que a fmp de  $X$  é

$x$	1	2	3	4
$p(x)$	0,002	0,146	0,588	0,264

[Sugestão:  $p(1) = 4P(\text{todos de espadas})$ ,  $p(2) = 6P(\text{apenas espadas e copas com no mínimo um de cada})$  e  $p(4) = 4P(2 \text{ espadas} \cap \text{um de cada outro naipe})$ .]

b. Calcule  $\mu$ ,  $\sigma^2$ , e  $\sigma$ .

- 92.** A  $va$  binomial negativa  $X$  foi definida como o número de  $F$ 's antes do  $r$ -ésimo  $S$ . Seja  $Y =$  número de tentativas necessárias para obtenção do  $r$ -ésimo  $S$ . Da mesma forma que a fmp de  $X$  foi deduzida, deduza a fmp de  $Y$ .

- 93.** De todos os clientes que compram portões automáticos de garagem, 75% compram um modelo acionado por corrente. Seja  $X =$  número entre os próximos 15 compradores que selecionam o modelo acionado por corrente.

- a. Qual é a fmp de  $X$ ?
- b. Calcule  $P(X > 10)$ .
- c. Calcule  $P(6 \leq X \leq 10)$ .
- d. Calcule  $\mu$  e  $\sigma^2$ .

e. Se a loja atualmente tiver em estoque 10 modelos acionados por corrente e oito acionados por eixo, qual é a probabilidade de os pedidos desses 15 clientes serem todos atendidos com o estoque existente?

- 94.** Um amigo recentemente planejou uma viagem para um acampamento. Ele tinha duas lanternas, uma que exigia uma pilha de 6V e outra que usava duas pilhas tamanho

D. Ele já havia colocado na mala duas pilhas de 6V e quatro tamanho D. Suponha que a probabilidade de uma determinada pilha funcionar seja  $p$  e que as pilhas funcionam ou não, independentemente uma da outra. Nosso amigo quer levar apenas uma lanterna. Para que valores de  $p$  ele deve levar a lanterna de 6V?

- 95.** Um sistema  $k$ -de- $n$  funciona se, e somente se, pelo menos  $k$  dos  $n$  componentes individuais do sistema funcionarem. Se os componentes individuais funcionam independentemente um do outro, cada um com probabilidade 0,9, qual é a probabilidade de um sistema 3-de-5 funcionar?

- 96.** Um fabricante de pilhas de lanterna deseja controlar a qualidade de seus produtos rejeitando os lotes em que a proporção de pilhas com voltagem inaceitável seja muito alta. Assim, em cada lote grande (10.000 pilhas), 25 serão selecionadas e testadas. Se ao menos cinco delas gerarem voltagens inaceitáveis, todo o lote será rejeitado. Qual é a probabilidade de um lote ser rejeitado se

- a. 5% das pilhas do lote tiverem voltagens inaceitáveis?
- b. 10% das pilhas do lote tiverem voltagens inaceitáveis?
- c. 20% das pilhas do lote tiverem voltagens inaceitáveis?
- d. O que aconteceria às probabilidades nas partes (a)–(c) se o número crítico de rejeição fosse aumentado de 5 para 6?

- 97.** Das pessoas que passam por um detector de metal de um aeroporto, 0,5% o ativam. Seja  $X =$  número de pessoas que ativam o detector entre um grupo de 500 selecionado aleatoriamente.

- a. Qual é a fmp (aproximada) de  $X$ ?
- b. Calcule  $P(X = 5)$ .
- c. Calcule  $P(5 \leq X)$ .

- 98.** Uma empresa de consultoria educacional está tentando decidir se alunos de ensino médio que nunca usaram uma calculadora podem resolver um tipo de problema mais facilmente com uma calculadora que usa lógica polonesa reversa ou com uma que não a usa. Uma amostra de 25 alunos é selecionada e eles podem praticar em ambas as calculadoras. Então, cada aluno é solicitado a resolver um problema com a

calculadora polonesa reversa e um similar com a outra. Seja  $p = P(S)$ , onde  $S$  indica que o aluno resolveu o problema mais rapidamente com a lógica polonesa reversa do que com a outra e seja  $X =$  número de  $S$ 's.

- a. Se  $p = 0,5$ , qual é  $P(7 \leq X \leq 18)$ ?
  - b. Se  $p = 0,8$ , qual é  $P(7 \leq X \leq 18)$ ?
  - c. Se a Hipótese  $p = 0,5$  for rejeitada quando  $X \leq 7$  ou  $X \geq 18$ , qual é a probabilidade de rejeitar a hipótese quando ela estiver correta?
  - d. Se a decisão de rejeitar a hipótese  $p = 0,5$  for como na parte (c), qual é a probabilidade de ela não ser rejeitada quando  $p = 0,6$ ? E quando  $p = 0,8$ ?
  - e. Que regra de decisão você escolheria para rejeitar a hipótese  $p = 0,5$  se você quisesse que a probabilidade da parte (c) fosse no máximo 0,01?
- 99.** Considere uma doença cuja presença pode ser identificada por um teste sanguíneo. Seja  $p$  a probabilidade de um indivíduo selecionado aleatoriamente ter a doença. Suponha que  $n$  indivíduos sejam selecionados independentemente para teste. Uma forma de proceder é fazer um teste separado para cada uma das  $n$  amostras sanguíneas. Uma abordagem potencialmente mais econômica — teste em grupo — foi apresentada durante a Segunda Guerra Mundial para identificar homens com sífilis entre recrutas do exército. Primeiro, separe parte de cada amostra de sangue, misture-as e faça um único teste. Se ninguém tiver a doença, o resultado será negativo e apenas um teste terá sido necessário. Se ao menos um indivíduo tiver a doença, o teste da amostra misturada resultará positivo e nesse caso serão processados os  $n$  testes individuais. Se  $p = 0,1$  e  $n = 3$ , qual será o número esperado de testes usando esse procedimento? Qual será o número esperado quando  $n = 5$ ? [O artigo "Random Multiple Access Communication and Group Testing" (*IEEE Trans. on Commun.*, 1984, p. 769-774) aplicou essas idéias a um sistema de comunicação em que a dicotomia era usuário ativo/ocioso em vez de doente/não-doente.]
- 100.** Seja  $p_1$  a probabilidade de qualquer símbolo de certo código ser transmitido com erro em um sistema de comunicação. Assuma que os erros ocorram em símbolos diferentes, independentemente um do outro. Suponha também que com probabilidade  $p_2$ , o símbolo errado seja corrigido no recebimento. Seja  $X$  o número de símbolos corrigidos em um bloco de mensagem consistindo de  $n$  símbolos (após o final do processo de correção). Qual é a distribuição de probabilidade de  $X$ ?
- 101.** O comprador de um gerador exige  $c$  partidas consecutivas antes de acertar a unidade. Assuma que os resultados das partidas individuais sejam independentes um do outro. Seja  $p$  a probabilidade de certa partida ter sucesso. A variável aleatória de interesse é  $X =$  número de partidas que devem ser processadas antes da aceitação. Forneça a fmp de  $X$  para o caso  $c = 2$ . Se  $p = 0,9$ , qual será  $P(X \leq 8)$ ? [Sugestão: para  $x \geq 5$ , expresse  $p(x)$  "de forma recursiva" em termos da fmp calculada com os valores menores  $x - 3, x - 4, \dots, 2$ .] (Esse problema foi sugerido pelo artigo "Evaluation of a Start-Up Demonstration Test," *J. Quality Technology*, 1983, p. 103-106.)
- 102.** Um plano de fidelidade de viagens de executivos foi desenvolvido por uma companhia aérea sob a premissa de que 10% de seus clientes atuais se qualificariam para participar.
- a. Admitindo a validade dessa premissa, entre 25 clientes atuais selecionados aleatoriamente, qual é a probabilidade de entre 2 e 6 (inclusive) se qualificarem para o programa?
  - b. Admitindo novamente a validade da premissa, qual é o número esperado de clientes que se qualificam e o desvio padrão do número dos que se qualificam em uma amostra aleatória de 100 clientes?
  - c. Seja  $X$  o número em uma amostra de 25 clientes atuais que se qualificam para o programa. Considere rejeitar a premissa da empresa na hipótese de  $p > 0,10$  se  $x \geq 7$ . Qual é a probabilidade de a premissa da empresa ser rejeitada quando for válida?
  - d. Refira-se à regra de decisão apresentada na parte (c). Qual é a probabilidade de a premissa da empresa não ser rejeitada mesmo que  $p = 0,20$  (ou seja, 20% se qualificam)?
- 103.** Quarenta por cento das sementes de milho possuem espigas simples e as outras 60%, espigas duplas. Uma semente com espiga simples produzirá uma planta com espigas simples em 29% das vezes, enquanto uma semente com espiga dupla produzirá uma planta com espigas simples em 26% das vezes. Considere a seleção aleatória de 10 sementes.
- a. Qual é a probabilidade de exatamente cinco dessas sementes terem espiga única e produzirem uma planta com espigas únicas?
  - b. Qual é a probabilidade de exatamente cinco espigas possuírem espigas únicas? Qual é a probabilidade de no máximo cinco plantas terem espigas únicas?
- 104.** Um julgamento não resultou em veredicto, porque oito membros do júri eram favoráveis ao veredicto de culpado e os outros quatro eram favoráveis ao de absolvição. Se os jurados saírem da sala aleatoriamente e cada um dos quatro primeiros for interpelado por um repórter buscando uma entrevista, qual será a fmp de  $X =$  número de jurados favoráveis à absolvição, entre os entrevistados? Quantos dos favoráveis à absolvição, você espera que sejam entrevistados?
- 105.** Um serviço de reservas emprega cinco operadores de informações que recebem solicitações de informações independentemente um do outro de acordo com um processo de Poisson com taxa  $\alpha = 2$  por minuto.
- a. Qual é a probabilidade de que, durante certo período de 1 minuto, o primeiro operador não receba solicitações?

- b. Qual é a probabilidade de que, durante certo período de um minuto, quatro dos cinco operadores não recebam solicitações?
- c. Escreva uma expressão para a probabilidade de em um dado período de um minuto, todos os operadores recebam exatamente o mesmo número de solicitações.

**106.** Gafanhotos distribuem-se aleatoriamente em um vasto campo, de acordo com uma distribuição de Poisson, com parâmetro  $\alpha = 2$  por jarda quadrada. Qual é o tamanho do raio  $R$  de uma região amostral circular para que a probabilidade de se encontrar pelo menos um gafanhoto na região seja de 0,99?

**107.** Uma banca de jornal faz o pedido de cinco cópias de uma edição de certa revista de fotografia. Seja  $X =$  número de indivíduos que chegam para comprar a revista. Se  $X$  tiver distribuição de Poisson com parâmetro  $\lambda = 4$ , qual será o número esperado de cópias vendidas?

**108.** Os indivíduos A e B começam a jogar uma seqüência de partidas de xadrez. Seja  $S = \{A \text{ ganha a partida}\}$  e suponha que os resultados de partidas sucessivas sejam independentes com  $P(S) = p$  e  $P(F) = 1 - p$  (nunca há empate). Eles jogarão até que um ganhe 10 partidas. Seja  $X =$  número de partidas jogadas (com valores possíveis 10, 11, ..., 19).

- a. Para  $x = 10, 11, \dots, 19$ , obtenha a expressão de  $p(x) = P(X = x)$ .
- b. Se puder haver empates, com  $p = P(S)$ ,  $q = P(F)$ ,  $1 - p - q = P$  (empate), quais serão os valores possíveis de  $X$ ? Qual é  $P(20 \leq X)$ ? [Sugestão:  $P(20 \leq X) = 1 - P(X < 20)$ .]

**109.** Certo teste de presença de uma certa doença tem probabilidade 0,20 de fornecer um resultado positivo falso (indicando que o indivíduo tem a doença quando não for o caso) e probabilidade 0,10 de fornecer um resultado negativo falso. Suponha que 10 indivíduos sejam testados, cinco dos quais têm a doença e cinco não. Seja  $X =$  número de resultados positivos.

- a.  $X$  tem distribuição binomial? Explique seu raciocínio.
- b. Qual é a probabilidade de exatamente três dos 10 resultados serem positivos?

**110.** A fmp geral da binomial negativa é dada por

$$nb(x; r, p) = k(r, x) \cdot p^r(1 - p)^x$$

$$x = 5, 0, 1, 2, \dots$$

Seja  $X$  o número de certo tipo de planta encontrada em uma determinada região, calcule a distribuição para  $p = 0,3$  e  $r = 2,5$ . Qual é  $P(X = 4)$ ? Qual é a probabilidade de ao menos uma planta ser encontrada?

**111.** Defina uma função  $p(x; \lambda, \mu)$  por

$$p(x; \lambda, \mu)$$

$$= \begin{cases} \frac{1}{2} e^{-\lambda} \frac{\lambda^x}{x!} + \frac{1}{2} e^{-\mu} \frac{\mu^x}{x!} & x = 0, 1, 2, \dots \\ 0 & \text{caso contrário} \end{cases}$$

- a. Mostre que  $p(x; \lambda, \mu)$  satisfaz às duas condições necessárias para especificação de uma fmp. [Nota: Se uma empresa emprega dois digitadores, um dos quais comete erros tipográficos na taxa de  $\lambda$  por página e o outro na taxa de  $\mu$  por página e cada um deles faz metade do trabalho de digitação da empresa, então  $p(x; \lambda, \mu)$  é a fmp de  $X =$  número de erros em uma página selecionada aleatoriamente.]
- b. Se o primeiro digitador (taxa  $\lambda$ ) digita 60% de todas as páginas, qual é a fmp de  $X$  da parte (a)?
- c. Qual é  $E(X)$  de  $p(x; \lambda, \mu)$  dado por essa expressão?
- d. Qual é o valor de  $\sigma^2$  de  $p(x; \lambda, \mu)$  obtido por aquela expressão?

**112.** A moda de uma variável aleatória discreta  $X$  com fmp  $p(x)$  é o valor  $x^*$  para o qual  $p(x)$  é o maior (o valor mais provável de  $x$ ).

- a. Seja  $X \sim \text{Bin}(n, p)$ . Considerando a razão  $b(x + 1; n, p)/b(x; n, p)$ , demonstre que  $b(x; n, p)$  aumenta com  $x$ , desde que  $x < np - (1 - p)$ . Conclua que a moda  $x^*$  é o inteiro que satisfaz  $(n + 1)p - 1 \leq x^* \leq (n + 1)p$ .
- b. Demonstre que, se  $X$  tem distribuição de Poisson com parâmetro  $\lambda$ , a moda é o maior inteiro menor que  $\lambda$ . Se  $\lambda$  for um inteiro, mostre que  $\lambda - 1$  e  $\lambda$  são modas.

**113.** O dispositivo de armazenagem de um computador possui 10 faixas concêntricas numeradas 1, 2, ..., 10 da mais externa para a mais interna e um único braço de acesso. Seja  $p_i =$  a probabilidade de certa solicitação de dados levar o braço para a faixa  $i$  ( $i = 1, \dots, 10$ ). Assuma que as faixas acessadas em buscas sucessivas são independentes. Seja  $X =$  número de faixas sobre as quais o braço de acesso passa durante duas solicitações sucessivas (excluindo a faixa da qual o braço acaba de sair, de forma que os valores possíveis de  $X$  são  $x = 0, 1, \dots, 9$ ). Calcule a fmp de  $X$ . [Sugestão:  $P(\text{o braço está na faixa } i \text{ e } X = j) = P(X = j | \text{braço em } i) \cdot p_i$ . Após a probabilidade condicional ser expressa em termos de  $p_1, \dots, p_{10}$  pela lei da probabilidade total, a probabilidade desejada é obtida pela somatória dos valores  $i$ .]

**114.** Se  $X$  é uma va hipergeométrica, demonstre diretamente pela definição que  $E(X) = nM/N$  (considere apenas o caso  $n < M$ ). [Sugestão: Fatore  $nM/N$  na soma de  $E(X)$  e mostre que os termos da soma têm a forma  $h(y; n - 1, M - 1, N - 1)$ , onde  $y = x - 1$ .]

**115.** Use o fato de

$$\sum_{\text{todos } x} (x - \mu)^2 p(x) \geq \sum_{x: |x - \mu| \geq k\sigma} (x - \mu)^2 p(x)$$

para demonstrar a desigualdade de Chebyshev dada no Exercício 43.

**116.** O processo de Poisson simples da Seção 3.6 é caracterizado por uma taxa constante  $\alpha$  em que os eventos ocorrem por unidade de tempo. Uma generalização disso é supor a probabilidade de ocorrência de

exatamente um evento no intervalo  $[t, t + \Delta t]$  é  $\alpha(t) \cdot \Delta t + o(\Delta t)$ . Pode-se, então, demonstrar que o número de eventos que ocorrem durante um intervalo  $[t_1, t_2]$  tem uma distribuição de Poisson com parâmetro

$$\lambda = \int_{t_1}^{t_2} \alpha(t) dt$$

A ocorrência de eventos ao longo do tempo nessa situação é denominada *processo de Poisson não-homogêneo*. O artigo “Inference Based on Retrospec-

tive Ascertainment”, *J. Amer Stat. Assoc.*, 1989, p. 360-372, considera a função intensidade

$$\alpha(t) = e^{a+bt}$$

apropriada para eventos que envolvem a transmissão do HIV (vírus da AIDS) via transfusões de sangue. Suponha que  $a = 2$  e  $b = 0,6$  (próximos aos valores sugeridos no artigo), com o tempo em anos.

- a. Qual é o número esperado de eventos no intervalo  $[0, 4]$ ? Em  $[2, 6]$ ?
- b. Qual é a probabilidade de no máximo 15 eventos ocorrerem no intervalo  $[0, 0,9907]$ ?

---

## Bibliografia

JOHNSON, Norman, KOTZ, Samuel e KEMP, Adrienne, *Discrete Univariate Distributions*. Wiley, Nova York, 1992. Uma enciclopédia de informações sobre distribuições discretas.

OLKIN, Ingram, DERMAN, Cyrus e GLENER, Leon, *Probability Models and Applications* (2. ed.), Macmillan, Nova York, 1994. Contém uma profunda discussão

das propriedades gerais de distribuições discretas e contínuas e resultados para distribuições específicas.

ROSS, Sheldon, *Introduction to Probability Models* (7 ed.), Academic Press, Nova York, 2003. Boa fonte de material sobre o processo de Poisson e generalizações, além de ser uma boa introdução a outros tópicos de probabilidade aplicada.

# Variáveis Aleatórias Contínuas e Distribuições de Probabilidade

## Introdução

Como mencionado no início do Capítulo 3, nem todas as variáveis aleatórias são discretas. Neste capítulo, estudamos o segundo tipo geral de variável aleatória que aparece em diversos problemas práticos. As Seções 4.1 e 4.2 apresentam as definições básicas e as propriedades das variáveis aleatórias contínuas e suas distribuições de probabilidades. Na Seção 4.3, estudamos em detalhe a variável aleatória normal e sua distribuição, inquestionavelmente a mais importante e útil da probabilidade e estatística. As seções 4.4 e 4.5 discutem outras distribuições contínuas freqüentemente usadas no trabalho prático. Na Seção 4.6, apresentamos um método de avaliação da consistência dos dados da amostra com a distribuição especificada.

## 4.1 Variáveis Aleatórias Contínuas e Funções de Densidade de Probabilidade

Uma variável aleatória (va) é discreta se os seus valores possíveis constituírem tanto um conjunto finito como puderem ser relacionados em uma seqüência infinita (uma lista em que haja um primeiro elemento, um segundo etc.). Uma variável aleatória cujo conjunto de valores possíveis consiste de um intervalo completo de números não é discreta.

### Variáveis Aleatórias Contínuas

#### DEFINIÇÃO

Uma variável aleatória  $X$  é dita contínua se o seu conjunto de valores possíveis consistir do intervalo completo de todos os valores, isto é, se, para cada  $A < B$ , qualquer valor  $x$  entre  $A$  e  $B$  for possível.

**Exemplo 4.1**

Se no estudo de ecologia de um lago fizermos medidas de profundidade em locais selecionados aleatoriamente, então  $X =$  profundidade nesse local é uma va contínua. Neste caso,  $A$  é a profundidade mínima na região da amostragem e  $B$  é a profundidade máxima. ■

**Exemplo 4.2**

Se um composto químico for selecionado aleatoriamente e determinarmos seu pH  $X$ , então,  $X$  é uma va contínua porque qualquer valor de PH entre 0 e 14 é possível. Caso se saiba mais sobre o composto selecionado para análise, o conjunto de valores possíveis pode ser um subintervalo de  $[0, 14]$ , tal como  $5,5 \leq x \leq 6,5$ , mas  $X$  ainda seria contínua. ■

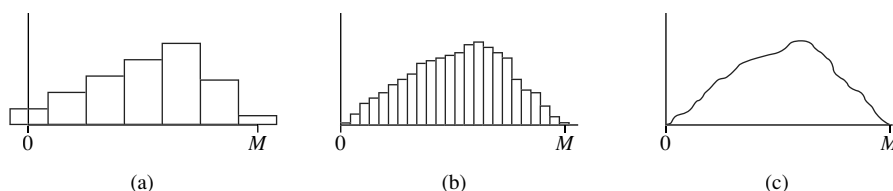
Se a escala de medida de  $X$  puder ser subdividida tanto quanto se desejar, a variável será contínua. Caso contrário, a variável será discreta. Por exemplo: se a variável for altura ou comprimento, poderá ser medida em quilômetros, metros, centímetros e assim por diante, de forma que a variável é contínua. Se, entretanto,  $X =$  valor de uma conta de gás mensal selecionada aleatoriamente, a menor unidade de medida será o centavo, fazendo com que qualquer valor de  $X$  seja um múltiplo de US\$ 0,01 e  $X$  seja uma variável discreta.

Alguém pode argumentar que variáveis como altura, peso e temperatura, apesar de serem, em princípio, contínuas, na prática, as limitações dos instrumentos de medida nos restringem a um mundo discreto (embora possa ser bastante subdividido). Entretanto, os modelos contínuos normalmente se aproximam muito bem de situações do mundo real e a matemática contínua (cálculo) freqüentemente se mostra mais fácil de trabalhar do que a matemática de variáveis e distribuições discretas.

**Distribuições de Probabilidade de Variáveis Contínuas**

Suponha que a variável de interesse  $X$  seja a profundidade de um lago em um ponto da superfície escolhido aleatoriamente. Seja  $M =$  profundidade máxima (em metros), de modo que qualquer número do intervalo  $[0, M]$  seja um valor possível de  $X$ . Se “considerarmos  $X$  discreta”, arredondando a profundidade para o valor mais próximo de metro, então os valores possíveis são inteiros não-negativos menores ou iguais a  $M$ . A distribuição discreta resultante das profundidades pode ser ilustrada usando-se um histograma de probabilidade. Se desenharmos o histograma de forma que a área do retângulo acima de qualquer inteiro possível  $k$  seja a proporção do lago cuja profundidade seja  $k$  (com arredondamento para o metro seguinte), então a área total de todos os retângulos será 1. Um histograma possível é exibido na Figura 4.1(a).

Se a profundidade for medida com arredondamento para o centímetro seguinte e for usado o mesmo eixo de medidas da Figura 4.1(a), cada retângulo no histograma de probabilidades será muito mais estreito, apesar da área total de todos os retângulos permanecer 1. Um histograma possível é ilustrado na Figura 4.1(b). Sua aparência é mais ajustada do que o histograma da Figura 4.1(a). Se continuarmos executando medidas de forma mais e mais precisa, a seqüência resultante terá uma curva cada vez mais ajustada, como a ilustrada na Figura 4.1(c). Como, para cada histograma, a área de todos os retângulos é igual a 1, a área total sob a curva ajustada também é 1. A probabilidade de a profundidade em um ponto selecionado aleatoriamente estar entre  $a$  e  $b$  é igual à área sob a curva ajustada entre os dois referidos pontos  $a$  e  $b$ . É exatamente uma curva ajustada do tipo ilustrado na Figura 4.1(c) que especifica uma distribuição de probabilidade contínua.



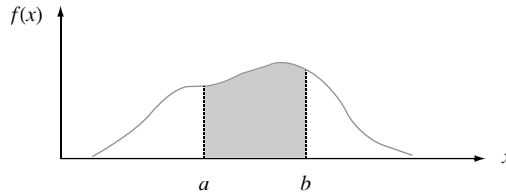
**Figura 4.1** (a) Histograma de probabilidade da profundidade medida com arredondamento em metros; (b) histograma de probabilidade da profundidade medida com arredondamento em centímetros; (c) limite de uma seqüência de histogramas discretos

**DEFINIÇÃO**

Seja  $X$  uma va contínua. A **distribuição de probabilidade** ou **função de densidade de probabilidade** (fdp) de  $X$  será, então, uma função  $f(x)$  tal que, para quaisquer dois números  $a$  e  $b$  com  $a \leq b$ ,

$$P(a \leq X \leq b) = \int_a^b f(x) dx$$

Isto é, a probabilidade de  $X$  ter um determinado valor no intervalo  $[a, b]$  é a área contida entre o intervalo e abaixo da curva da função de densidade, conforme ilustrado na Figura 4.2. O gráfico de  $f(x)$  normalmente é denominado *curva de densidade*.



**Figura 4.2**  $P(a \leq X \leq b)$  = a área abaixo da curva de densidade, entre  $a$  e  $b$

Para que  $f(x)$  seja uma fdp legítima, deve satisfazer às duas condições a seguir:

1.  $f(x) \geq 0$  para todos os  $x$
2.  $\int_{-\infty}^{\infty} f(x) dx = \text{área abaixo do gráfico de } f(x) = 1$

**Exemplo 4.3**

A direção de uma imperfeição em relação a uma linha de referência em um objeto circular como um pneu, um rotor de freio ou um volante normalmente apresenta alguma incerteza. Considere a linha de referência que conecta a válvula do pneu até o ponto central e seja  $X$  o ângulo medido no sentido horário até o local da imperfeição. Uma fdp possível de  $X$  é

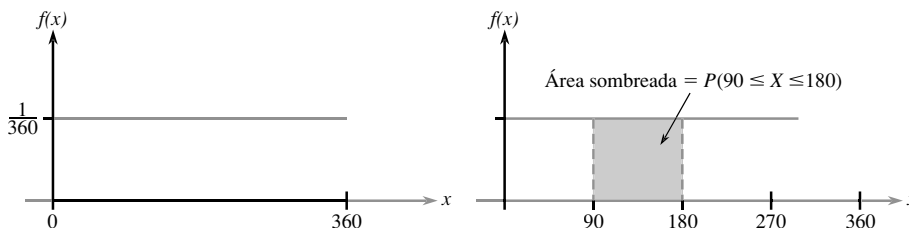
$$f(x) = \begin{cases} \frac{1}{360} & 0 \leq x < 360 \\ 0 & \text{caso contrário} \end{cases}$$

A fdp está ilustrada na Figura 4.3. Claramente,  $f(x) \geq 0$ . A área abaixo da curva de densidade é apenas a área de um retângulo: (altura)(base) =  $(\frac{1}{360})(360) = 1$ . A probabilidade de o ângulo estar entre  $90^\circ$  e  $180^\circ$  é

$$P(90 \leq X \leq 180) = \int_{90}^{180} \frac{1}{360} dx = \frac{x}{360} \Big|_{x=90}^{x=180} = \frac{1}{4} = 0,25$$

A probabilidade de o ângulo de ocorrência estar dentro de  $90^\circ$  da linha de referência é

$$P(0 \leq X \leq 90) + P(270 \leq X < 360) = 0,25 + 0,25 = 0,50$$



**Figura 4.3** A fdp e a probabilidade do Exemplo 4.3





Como sempre  $0 \leq a \leq b \leq 360$  no Exemplo 4.3,  $P(a \leq X \leq b)$  depende apenas da largura  $b - a$  do intervalo,  $X$  é dita ter distribuição uniforme.

## DEFINIÇÃO

Uma va contínua  $X$  é dita ter **distribuição uniforme** no intervalo  $[A, B]$  se a fdp de  $X$  for

$$f(x; A, B) = \begin{cases} \frac{1}{B - A} & A \leq x \leq B \\ 0 & \text{caso contrário} \end{cases}$$

O gráfico de qualquer fdp uniforme possui a aparência do gráfico da Figura 4.3, exceto pelo fato de que o intervalo da densidade positiva é  $[A, B]$  em vez de  $[0, 360]$ .

No caso discreto, uma função de massa de probabilidade (fmp) nos diz como as pequenas “falhas” daquele tipo de massa, de diversas magnitudes, estão distribuídas de modo contínuo ao longo do eixo das medidas. No caso contínuo, a densidade da probabilidade está “suja” de forma contínua ao longo do intervalo de valores possíveis. Quando a densidade está mantida uniformemente ao longo do intervalo, resulta uma fdp uniforme, como na Figura 4.3.

Quando  $X$  é uma va discreta, é designada uma probabilidade positiva a cada valor possível, mas este não é o caso quando  $X$  é contínua.

## PROPOSIÇÃO

Se  $X$  é uma va contínua, então para qualquer número  $c$ ,  $P(X = c) = 0$ . Além disso, para quaisquer dois números  $a$  e  $b$  com  $a < b$ ,

$$\begin{aligned} P(a \leq X \leq b) &= P(a < X \leq b) \\ &= P(a \leq X < b) = P(a < X < b) \end{aligned} \quad (4.1)$$

Em palavras, a probabilidade atribuída a qualquer valor específico é zero e a probabilidade de um intervalo não depende da inclusão ou não de seus pontos extremos. Essas propriedades decorrem dos fatos de que a área abaixo do gráfico de  $f(x)$  e acima do único valor  $c$  é nula, e que a área sob o gráfico acima de um intervalo não é afetada pela exclusão ou inclusão dos pontos extremos do intervalo. Se  $X$  for discreta e houver uma massa de probabilidade positiva tanto em  $X = a$  como em  $X = b$ , então as quatro probabilidades da Expressão (4.1) diferirão uma da outra.

O fato de uma distribuição contínua atribuir probabilidade zero para cada valor tem uma analogia física. Considere uma haste circular sólida com área da seção =  $1 \text{ pol}^2$ . Posicione a haste ao longo do eixo das medidas e suponha que sua densidade em qualquer ponto  $x$  seja dada pelo valor  $f(x)$  de uma função de densidade. Então, se a haste for seccionada nos pontos  $a$  e  $b$  e o segmento removido, a quantidade de massa removida é  $\int_a^b f(x) dx$ . Se a haste for cortada apenas no ponto  $c$ , nenhuma massa será removida. A massa é atribuída a segmentos do intervalo da haste, mas não a pontos individuais.

## Exemplo 4.4

“Tempo de avanço” no fluxo do tráfego é o tempo entre o instante em que um carro termina de passar por um ponto fixo e o instante em que o próximo carro começa a passar por esse ponto. Seja  $X$  = tempo de avanço para dois carros consecutivos escolhidos ao acaso, em uma estrada durante um período de tráfego intenso. A seguinte fdp de  $X$  é essencialmente a sugerida em “The Statistical Properties of Freeway Traffic” (*Transp. Research*, vol. 11, p. 221-228):

$$f(x) = \begin{cases} 0,15e^{-0,15(x-0,5)} & x \geq 0,5 \\ 0 & \text{caso contrário} \end{cases}$$

O gráfico de  $f(x)$  é mostrado na Figura 4.4. Não há densidade associada aos tempos de avanço inferiores a 0,5 e a densidade do avanço decresce exponencialmente rápida à medida que  $x$  aumenta a partir de 0,5. Claramente,  $f(x) \geq 0$ . Para mostrar que  $\int_{-\infty}^{\infty} f(x) dx = 1$ , usamos a expressão do cálculo  $\int_a^{\infty} e^{-kx} dx = (1/k)e^{-k \cdot a}$ . Então

$$\begin{aligned} \int_{-\infty}^{\infty} f(x) dx &= \int_{0,5}^{\infty} 0,15e^{-0,15(x-0,5)} dx = 0,15e^{0,075} \int_{0,5}^{\infty} e^{-0,15x} dx \\ &= 0,15e^{0,075} \cdot \frac{1}{0,15} e^{-(0,15)(0,5)} = 1 \end{aligned}$$

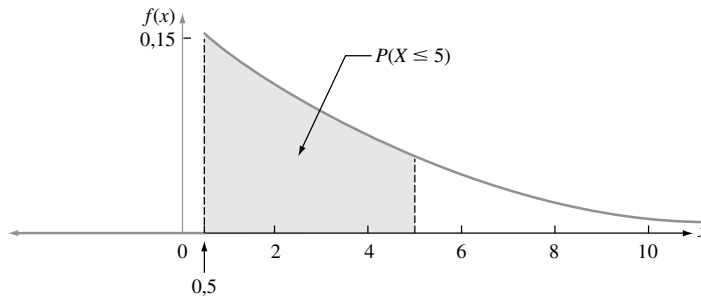


Figura 4.4 A curva de densidade do tempo de avanço do Exemplo 4.4

A probabilidade do tempo de avanço ser no máximo 5 segundos é

$$\begin{aligned} P(X \leq 5) &= \int_{-\infty}^5 f(x) dx = \int_{0,5}^5 0,15e^{-0,15(x-0,5)} dx \\ &= 0,15e^{0,075} \int_{0,5}^5 e^{-0,15x} dx = 0,15e^{0,075} \cdot \left. -\frac{1}{0,15} e^{-0,15x} \right|_{x=0,5}^{x=5} \\ &= e^{0,075}(-e^{-0,75} + e^{-0,075}) = 1,078(-0,472 + 0,928) = 0,491 \\ &= P(\text{menos de 5 segundos}) = P(X < 5) \end{aligned}$$

Diferentemente das distribuições discretas, tais como a binomial, a hipergeométrica e a binomial negativa, a distribuição de qualquer va contínua dada normalmente não pode ser deduzida por meio de argumentos probabilísticos simples. Pelo contrário, deve-se fazer uma escolha cuidadosa da fdp com base em conhecimentos anteriores e dados disponíveis. Felizmente, há algumas famílias gerais de fdps que funcionam bem em diversas situações experimentais, várias delas discutidas posteriormente, neste capítulo.

Assim como no caso discreto, normalmente é útil imaginar a população de interesse constituída de  $X$  valores em vez de indivíduos ou objetos. A fdp é, então, um modelo para distribuição dos valores da população numérica e, pelo modelo, pode-se calcular diversas características da população (como a média).

**Exercícios** | Seção 4.1 (1–10)

1. Seja  $X$  o tempo que um livro de uma reserva de duas horas, na biblioteca de uma faculdade, é examinado por um estudante selecionado aleatoriamente e suponha que  $X$  tenha função de densidade

$$f(x) = \begin{cases} 0,5x & 0 \leq x \leq 2 \\ 0 & \text{caso contrário} \end{cases}$$

Calcule as probabilidades a seguir:

- a.  $P(X \leq 1)$

- b.  $P(0,5 \leq X \leq 1,5)$   
 c.  $P(1,5 < X)$

2. Suponha que a temperatura de reação  $X$  (em °C), em um certo processo químico, tenha função de densidade uniforme para  $A = -5$  e  $B = 5$ .

- a. Calcule  $P(X < 0)$ .  
 b. Calcule  $P(-2,5 < X < 2,5)$ .  
 c. Calcule  $P(-2 \leq X \leq 3)$ .

- d. Para  $k$  satisfazendo  $-5 < k < k + 4 < 5$ , calcule  $P(k < X < k + 4)$ .
3. Suponha que o erro envolvido ao se fazer uma certa medida seja uma va contínua  $X$  com fdp

$$f(x) = \begin{cases} 0,09375(4 - x^2) & -2 \leq x \leq 2 \\ 0 & \text{caso contrário} \end{cases}$$

- a. Faça o gráfico de  $f(x)$ .
- b. Calcule  $P(X > 0)$ .
- c. Calcule  $P(-1 < X < 1)$ .
- d. Calcule  $P(X < -0,5 \text{ ou } X > 0,5)$ .
4. Seja  $X$  a tensão de vibração<sup>1</sup> (psi). Em uma palheta de turbina com certa velocidade de vento em um túnel aerodinâmico. O artigo "Blade Fatigue Life Assessment with Application to VAWTS" (*J. Solar Energy Engr.*, 1982, p. 107–111) propõe a distribuição de Rayleigh, com fdp

$$f(x; \theta) = \begin{cases} \frac{x}{\theta^2} \cdot e^{-x^2/(2\theta^2)} & x > 0 \\ 0 & \text{caso contrário} \end{cases}$$

como modelo da distribuição de  $X$ .

- a. Verifique se  $f(x; \theta)$  é uma fdp legítima.
- b. Suponha que  $\theta = 100$  (valor sugerido: em um gráfico do artigo). Qual é a probabilidade de  $X$  ser no máximo 200? Menos que 200? Pelo menos 200?
- c. Qual é a probabilidade de  $X$  estar entre 100 e 200 (novamente assumindo  $\theta = 100$ )?
- d. Forneça uma expressão para  $P(X \leq x)$ .
5. Um professor de faculdade nunca finaliza sua aula antes do final do horário e sempre termina dentro de dois minutos após o horário. Seja  $X =$  tempo entre o fim do horário e o fim da aula e suponha que a fdp de  $X$  seja

$$f(x) = \begin{cases} kx^2 & 0 \leq x \leq 2 \\ 0 & \text{caso contrário} \end{cases}$$

- a. Determine o valor de  $k$ . [Sugestão: a área total abaixo do gráfico de  $f(x)$  é 1].
- b. Qual é a probabilidade de a aula terminar dentro de 1 min. do final do horário?
- c. Qual é a probabilidade de a aula continuar além do horário por 60 a 90 seg?
- d. Qual é a probabilidade de a aula continuar por pelo menos 90 segundos após o final do horário?
6. O peso de tração real de um cartucho de som que está ajustado para rodas com 3g em um certo aparelho pode ser considerado uma va contínua  $X$  com fdp

$$f(x) = \begin{cases} k[1 - (x - 3)^2] & 2 \leq x \leq 4 \\ 0 & \text{caso contrário} \end{cases}$$

- a. Faça o gráfico de  $f(x)$ .
- b. Determine o valor de  $k$ .
- c. Qual é a probabilidade de o peso real de tração ser maior do que o nominal?
- d. Qual é a probabilidade de o peso real estar dentro de 0,25g de distância em relação ao nominal?

- e. Qual é a probabilidade de o peso real diferir do nominal em mais de 0,5g?

7. Acredita-se que o tempo  $X$  (min.) para um assistente de laboratório preparar o equipamento para certo experimento tenha distribuição uniforme com  $A = 25$  e  $B = 35$ .

- a. Escreva a fdp de  $X$  e faça o gráfico.
- b. Qual é a probabilidade de o tempo de preparação exceder 33 minutos?
- c. Qual é a probabilidade de o tempo de preparação estar dentro de 2 minutos em relação à média? [Sugestão: leia  $\mu$  no gráfico de  $f(x)$ .]
- d. Para qualquer  $a$  tal que  $25 < a < a + 2 < 35$ , qual é a probabilidade de o tempo de preparação estar entre  $a$  e  $a + 2$  min.?

8. Para ir trabalhar, primeiro preciso tomar um ônibus perto da minha casa e depois mudar de ônibus. Se o tempo de espera (em minutos) em cada parada tem uma distribuição uniforme com  $A = 0$  e  $B = 5$ , pode ser mostrado que meu tempo total de espera  $Y$  tem a fdp

$$f(y) = \begin{cases} \frac{1}{25}y & 0 \leq y < 5 \\ \frac{2}{5} - \frac{1}{25}y & 5 \leq y \leq 10 \\ 0 & y < 0 \text{ ou } y > 10 \end{cases}$$

- a. Desenhe o gráfico da fdp de  $Y$ .
- b. Demonstre que  $\int_{-\infty}^{\infty} f(y) dy = 1$ .
- c. Qual é a probabilidade de o tempo total de espera ser no máximo 3 min.?
- d. Qual é a probabilidade de o tempo total de espera ser no máximo 8 min.?
- e. Qual é a probabilidade de o tempo total de espera estar entre 3 e 8 min.?
- f. Qual é a probabilidade de o tempo total de espera ser menos de 2 minutos ou mais de 6 minutos?

9. Considere novamente a fdp  $X =$  tempo de avanço fornecida no Exemplo 4.4.

- a. Qual é a probabilidade de o tempo de avanço ser no máximo 6 segundos?
- b. Mais de 6 segundos? Ao menos 6 segundos?
- c. Estar entre 5 e 6 segundos?

10. Uma família de fdps que tem sido usada para aproximar a distribuição de renda, tamanho da população da cidade e tamanho das empresas é a família de curvas de Pareto. A família tem dois parâmetros,  $k$  e  $\theta$ , ambos  $> 0$  e a fdp é

$$f(x; k, \theta) = \begin{cases} \frac{k \cdot \theta^k}{x^{k+1}} & x \geq \theta \\ 0 & x < \theta \end{cases}$$

- a. Faça o gráfico de  $f(x; k, \theta)$ .
- b. Demonstre que a área total abaixo do gráfico é igual a 1.
- c. Se a va  $X$  tiver fdp  $f(x; k, \theta)$ , para qualquer  $b > \theta$ , fixe obtenha uma expressão para  $P(X \leq b)$ .
- d. Para  $\theta < a < b$ , obtenha uma expressão para a probabilidade  $P(a \leq X \leq b)$ .

<sup>1</sup> psi é a abreviatura do *pounds per square inch* (libras por polegada quadrada). (NT)

## 4.2 Funções de Distribuição Acumulada e Valores Esperados

Muitos dos conceitos mais importantes apresentados no estudo de distribuições discretas também desempenham papel importante nas distribuições contínuas. Definições análogas às do Capítulo 3 envolvem a substituição da somatória pelo símbolo de integral.

### Função Distribuição Acumulada

A função distribuição acumulada (fdc)  $F(x)$  de uma va discreta  $X$  fornece, para qualquer número específico  $x$ , a probabilidade  $P(X \leq x)$ . Ela é obtida somando-se fmp  $p(y)$  para todos os valores possíveis  $y$  que satisfaçam  $y \leq x$ . A fdc de uma va contínua fornece as mesmas probabilidades  $P(X \leq x)$  e é obtida pela integração da fdp  $f(y)$  entre os limites  $-\infty$  e  $x$ .

DEFINIÇÃO

A **função distribuição acumulada**  $F(x)$  de uma va contínua  $X$  é definida para cada número  $x$  por

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(y) dy$$

Para cada  $x$ ,  $F(x)$  é a área abaixo da curva de densidade à esquerda de  $x$ . Essa propriedade é ilustrada na Figura 4.5, onde  $F(x)$  aumenta com ajuste à medida que  $x$  aumenta.

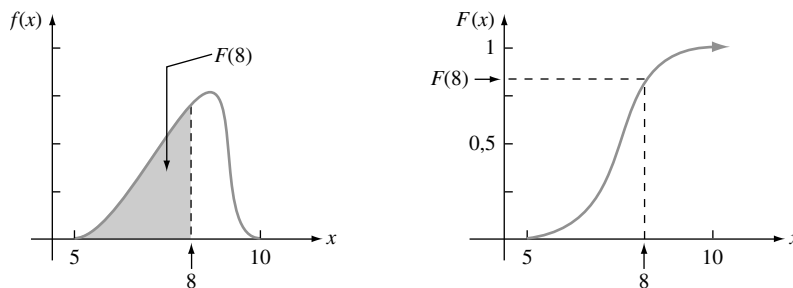


Figura 4.5 Uma fdp e a fdc associada

#### Exemplo 4.5

Seja  $X$ , a espessura de uma determinada chapa de metal, com distribuição uniforme no intervalo  $[A, B]$ . A função de densidade é exibida na Figura 4.6. Para  $x < A$ ,  $F(x) = 0$ , porque não há área sob o gráfico da função de densidade à esquerda de  $x$ . Para  $x \geq B$ ,  $F(x) = 1$ , já que toda a área está acumulada à esquerda de  $x$ . Finalmente, para  $A \leq x \leq B$ ,

$$F(x) = \int_{-\infty}^x f(y) dy = \int_A^x \frac{1}{B-A} dy = \frac{1}{B-A} \cdot y \Big|_{y=A}^{y=x} = \frac{x-A}{B-A}$$

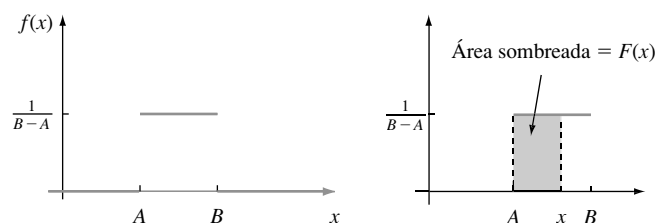


Figura 4.6 A fdp para distribuição uniforme

A fdc completa é

$$F(x) = \begin{cases} 0 & x < A \\ \frac{x - A}{B - A} & A \leq x < B \\ 1 & x \geq B \end{cases}$$

O gráfico desta fdc é exibido na Figura 4.7.

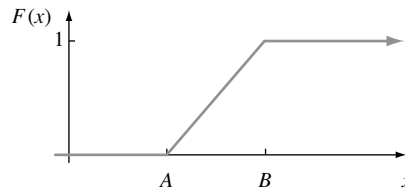


Figura 4.7 A fdc de distribuição uniforme

### Uso de $F(x)$ para Calcular as Probabilidades

A importância da fdc neste caso, assim como no de vas discretas, é que as probabilidades dos diversos intervalos podem ser calculadas por uma fórmula ou lidas de uma tabela de  $F(x)$ .

**PROPOSIÇÃO**

Seja  $X$  uma va contínua com fdp  $f(x)$  e fdc  $F(x)$ . Então, para qualquer número  $a$ ,

$$P(X > a) = 1 - F(a)$$

e, para quaisquer dois números  $a$  e  $b$  com  $a < b$ ,

$$P(a \leq X \leq b) = F(b) - F(a)$$

A Figura 4.8 ilustra a segunda parte dessa proposição. A probabilidade desejada é a área sombreada abaixo da curva de densidade entre  $a$  e  $b$ , e é igual à diferença entre as duas áreas acumuladas sombreadas.

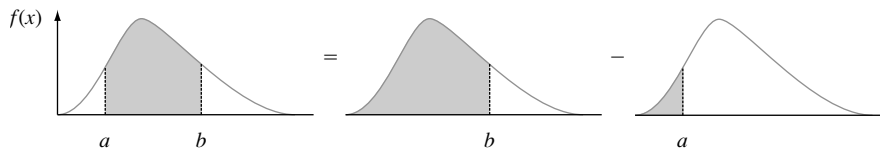


Figura 4.8 Cálculo de  $P(a \leq X \leq b)$  pelas de probabilidades acumuladas

**Exemplo 4.6**

Suponha que a fdp da grandeza  $X$  de uma carga dinâmica em uma ponte (em newtons) seja dada por

$$f(x) = \begin{cases} \frac{1}{8} + \frac{3}{8}x & 0 \leq x \leq 2 \\ 0 & \text{caso contrário} \end{cases}$$

Para qualquer número  $x$  entre 0 e 2,

$$F(x) = \int_{-\infty}^x f(y) dy = \int_0^x \left( \frac{1}{8} + \frac{3}{8}y \right) dy = \frac{x}{8} + \frac{3}{16}x^2$$

Dessa forma,

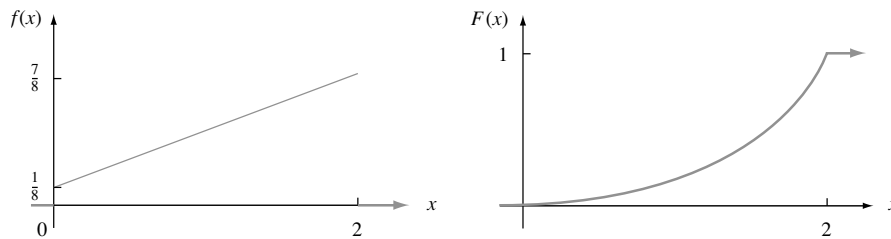
$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{8} + \frac{3}{16}x^2 & 0 \leq x \leq 2 \\ 1 & 2 < x \end{cases}$$

Os gráficos de  $f(x)$  e  $F(x)$  são exibidos na Figura 4.9. A probabilidade de a carga estar entre 1 e 1,5 é

$$\begin{aligned} P(1 \leq X \leq 1,5) &= F(1,5) - F(1) \\ &= \left[ \frac{1}{8}(1,5) + \frac{3}{16}(1,5)^2 \right] - \left[ \frac{1}{8}(1) + \frac{3}{16}(1)^2 \right] \\ &= \frac{19}{64} = 0,297 \end{aligned}$$

A probabilidade de a carga exceder 1 é

$$\begin{aligned} P(X > 1) &= 1 - P(X \leq 1) = 1 - F(1) = 1 - \left[ \frac{1}{8}(1) + \frac{3}{16}(1)^2 \right] \\ &= \frac{11}{16} = 0,688 \end{aligned}$$



**Figura 4.9** A fdp e a fdc do Exemplo 4.6

Uma vez obtida a fdc, qualquer probabilidade envolvendo  $X$  é facilmente calculada sem necessidade de integração.

### Obtendo $f(x)$ a partir de $F(x)$

Para  $X$  discreta, a fmp é obtida pela fdc calculando-se a diferença entre dois valores  $F(x)$ . A expressão contínua análoga de uma diferença é a derivada. O resultado a seguir é uma consequência do Teorema Fundamental do Cálculo.

**PROPOSIÇÃO**

Se  $X$  for uma va contínua com fdp  $f(x)$  e fdc  $F(x)$  então, para qualquer  $x$  em que a derivada  $F'(x)$  existir,  $F'(x) = f(x)$ .

**Exemplo 4.7** (Continuação do Exemplo 4.5)

Quando  $X$  tem distribuição uniforme,  $F(x)$  é diferenciável exceto nos pontos  $x = A$  e  $x = B$ , onde o gráfico de  $F(x)$  forma cantos agudos. Como  $F(x) = 0$  para  $x < A$  e  $F(x) = 1$  para  $x > B$ ,  $F'(x) = 0 = f(x)$  para tal  $x$ . Para  $A < x < B$ ,

$$F'(x) = \frac{d}{dx} \left( \frac{x - A}{B - A} \right) = \frac{1}{B - A} = f(x)$$

## Percentis da Distribuição Contínua

Quando dizemos que a pontuação do teste de um indivíduo estava no 85° percentil da população, queremos dizer que 85% de todas as pontuações da população estavam abaixo daquela pontuação e 15% acima dela. De forma similar, o 40° percentil é a pontuação que excede 40% de todas as pontuações e é excedido por 60%.

### DEFINIÇÃO

Seja  $p$  um número entre 0 e 1. O **100 $p$ -ésimo percentil** da distribuição de uma va contínua  $X$ , representada por  $\eta(p)$  é definido por

$$p = F(\eta(p)) = \int_{-\infty}^{\eta(p)} f(y) dy \quad (4.2)$$

De acordo com a Expressão (4.2),  $\eta(p)$  é o valor no eixo das medidas tal que 100 $p$ % da área abaixo do gráfico de  $f(x)$  está à esquerda de  $\eta(p)$  e 100(1 -  $p$ )% está à direita. Portanto,  $\eta(0,75)$ , o 75° percentil, é tal que a área abaixo do gráfico de  $f(x)$  à esquerda de  $\eta(0,75)$  é 0,75. A Figura 4.10 ilustra a definição.

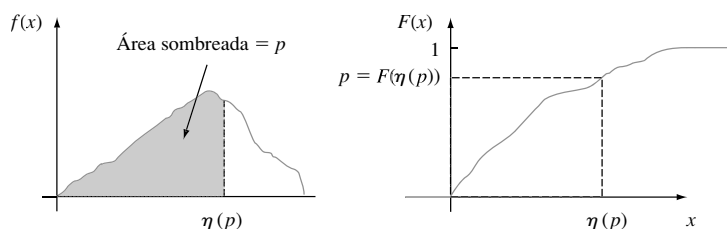


Figura 4.10 O (100 $p$ )-ésimo percentil de uma distribuição contínua

### Exemplo 4.8

A distribuição da quantidade de cascalho (em toneladas) vendida por uma empresa de materiais de construção em uma semana é uma va contínua  $X$  com fdp

$$f(x) = \begin{cases} \frac{3}{2} (1 - x^2) & 0 \leq x \leq 1 \\ 0 & \text{caso contrário} \end{cases}$$

A fdc de vendas para qualquer  $x$  entre 0 e 1 é

$$F(x) = \int_0^x \frac{3}{2} (1 - y^2) dy = \frac{3}{2} \left( y - \frac{y^3}{3} \right) \Big|_{y=0}^{y=x} = \frac{3}{2} \left( x - \frac{x^3}{3} \right)$$

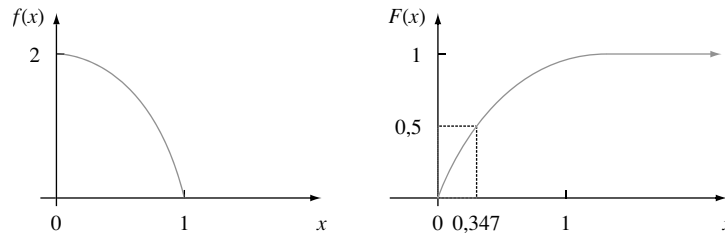
Os gráficos de  $f(x)$  e  $F(x)$  são exibidos na Figura 4.11. O (100 $p$ )-ésimo percentil dessa distribuição satisfaz a equação

$$p = F(\eta(p)) = \frac{3}{2} \left[ \eta(p) - \frac{(\eta(p))^3}{3} \right]$$

isto é,

$$(\eta(p))^3 - 3\eta(p) + 2p = 0$$

Para o 50° percentil,  $p = 0,5$  e a equação a ser resolvida é  $\eta^3 - 3\eta + 1 = 0$ ; a solução é  $\eta = \eta(0,5) = 0,347$ . Se a distribuição continuar igual de semana para semana, no longo prazo, 50% de todas as semanas resultarão em vendas de menos de 0,347 toneladas e 50% em mais de 0,347 toneladas.

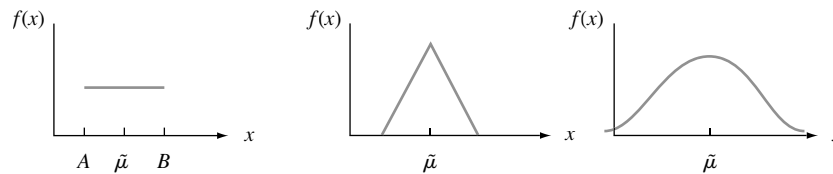


**Figura 4.11** A fdp e a fdc do Exemplo 4.8

**DEFINIÇÃO**

A **mediana** de uma distribuição contínua, representada por  $\tilde{\mu}$ , é o 50º percentil, de forma que  $\tilde{\mu}$  satisfaz  $0,5 = F(\tilde{\mu})$ . Isto é, metade da área abaixo da curva de densidade está à esquerda de  $\tilde{\mu}$  e metade à direita.

Uma distribuição contínua cuja fdp é **simétrica** — que significa que o gráfico da fdp à esquerda de um ponto é uma imagem especular do gráfico à direita desse ponto — tem mediana  $\tilde{\mu}$  igual ao ponto de simetria, já que metade da área abaixo da curva encontra-se em cada lado desse ponto. A Figura 4.12 fornece diversos exemplos. O valor do erro na medida de uma quantidade física é frequentemente suposto como tendo distribuição simétrica.



**Figura 4.12** Medianas de distribuições simétricas

**Valores Esperados de Variáveis Aleatórias Contínuas**

Para uma variável aleatória discreta  $X$ ,  $E(X)$  foi obtido pela soma de  $x \cdot p(x)$  para os valores possíveis de  $X$ . Neste caso, substituímos a soma pela integração e a fmp pela fdp para obtermos uma média contínua ponderada.

**DEFINIÇÃO**

O **valor médio** ou **esperado** de uma va contínua  $X$  com fdp  $f(x)$  é

$$\mu_x = E(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx$$

**Exemplo 4.9** (Continuação do Exemplo 4.8)

A fdp das vendas semanais de cascalho  $X$  era

$$f(x) = \begin{cases} \frac{3}{2}(1 - x^2) & 0 \leq x \leq 1 \\ 0 & \text{caso contrário} \end{cases}$$

então

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x \cdot f(x) dx = \int_0^1 x \cdot \frac{3}{2}(1 - x^2) dx \\ &= \frac{3}{2} \int_0^1 (x - x^3) dx = \frac{3}{2} \left( \frac{x^2}{2} - \frac{x^4}{4} \right) \Bigg|_{x=0}^{x=1} = \frac{3}{8} \end{aligned}$$



Quando a fdp  $f(x)$  especifica um modelo para a distribuição dos valores em uma população numérica, então  $\mu$  é a média da população, a medida de localização ou centro da população mais freqüentemente usada.

Freqüentemente queremos calcular o valor esperado de alguma função  $h(X)$  da va  $X$ . Se imaginarmos  $h(X)$  como uma nova va  $Y$ , as técnicas da estatística matemática podem ser usadas para deduzir a fdp de  $Y$ , e  $E(Y)$  pode ser calculada a partir da definição. Felizmente, como no caso discreto, há um modo mais fácil de calcularmos  $E[h(X)]$ .

## PROPOSIÇÃO

Se  $X$  for uma va contínua com fdp  $f(x)$  e  $h(X)$  for qualquer função de  $X$ , então

$$E[h(X)] = \mu_{h(X)} = \int_{-\infty}^{\infty} h(x) \cdot f(x) dx$$

## Exemplo 4.10

Dois espécies estão competindo em uma região pelo controle de uma quantidade limitada de um determinado recurso. Seja  $X$  = a proporção do recurso controlado pela espécie 1 e suponha que  $X$  tem fdp

$$f(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{caso contrário} \end{cases}$$

que é uma distribuição uniforme no intervalo  $[0, 1]$ . (Em seu livro *Ecological Diversity*, E. C. Pielou denomina o caso como o modelo da “vareta quebrada” para alocação de recursos, dada a sua analogia com a quebra de uma vareta em um ponto escolhido aleatoriamente.) A espécie que controla a maioria dos recursos controla a quantidade

$$h(X) = \max(X, 1 - X) = \begin{cases} 1 - X & \text{se } 0 \leq X < \frac{1}{2} \\ X & \text{se } \frac{1}{2} \leq X \leq 1 \end{cases}$$

O valor esperado controlado pela espécie que tem controle majoritário é portanto

$$\begin{aligned} E[h(X)] &= \int_{-\infty}^{\infty} \max(x, 1 - x) \cdot f(x) dx = \int_0^1 \max(x, 1 - x) \cdot 1 dx \\ &= \int_0^{1/2} (1 - x) \cdot 1 dx + \int_{1/2}^1 x \cdot 1 dx = \frac{3}{4} \end{aligned}$$

Para  $h(X)$ , função linear,  $E[h(X)] = E(aX + b) = aE(X) + b$ .

## Variância da Variável Aleatória Contínua

## DEFINIÇÃO

A **variância** de uma variável aleatória contínua  $X$  com fdp  $f(x)$  e média  $\mu$  é

$$\sigma_X^2 = V(X) = \int_{-\infty}^{\infty} (x - \mu)^2 \cdot f(x) dx = E[(X - \mu)^2]$$

O **desvio padrão** (DP) de  $X$  é  $\sigma_X = \sqrt{V(X)}$ .

Como no caso discreto,  $\sigma_X^2$  é o quadrado do desvio médio ou esperado em relação à média  $\mu$  e fornece uma medida da dispersão da distribuição ou da população dos valores de  $x$ . A forma mais fácil de calcular  $\sigma^2$  é usar novamente uma fórmula prática.

**PROPOSIÇÃO**

$$V(X) = E(X^2) - [E(X)]^2$$

**Exemplo 4.11** (Continuação do Exemplo 4.9)

Para  $X =$  venda semanal de cascalho, calculamos  $E(X) = \frac{3}{8}$ . Como

$$\begin{aligned} E(X^2) &= \int_{-\infty}^{\infty} x^2 \cdot f(x) dx = \int_0^1 x^2 \cdot \frac{3}{2} (1 - x^2) dx \\ &= \int_0^1 \frac{3}{2} (x^2 - x^4) dx = \frac{1}{5} \end{aligned}$$

$$V(X) = \frac{1}{5} - \left(\frac{3}{8}\right)^2 = \frac{19}{320} = 0,059 \quad \text{e} \quad \sigma_X = 0,244 \quad \blacksquare$$

Quando  $h(X)$  é uma função linear e  $V(X) = \sigma^2$ ,  $V[h(X)] = V(aX + b) = a^2 \cdot \sigma^2$  e  $\sigma_{aX+b} = |a| \cdot \sigma$ .

**Exercícios** | Seção 4.2 (11–25)

**11.** A fdc da duração da retirada  $X$  descrita no Exercício 1 é

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x^2}{4} & 0 \leq x < 2 \\ 1 & 2 \leq x \end{cases}$$

Use tais condições para calcular os itens a seguir

- a.  $P(X \leq 1)$
- b.  $P(0,5 \leq X \leq 1)$
- c.  $P(X > 0,5)$
- d. A mediana da duração da retirada  $\tilde{\mu}$  [resolva  $0,5 = F(\tilde{\mu})$ ]
- e.  $F'(x)$  para obter a função de densidade  $f(x)$
- f.  $E(X)$
- g.  $V(X)$  e  $\sigma_X$
- h. Se o aluno que retira o livro tem uma quantia a pagar  $h(X) = X^2$ , quando a duração da retirada é  $X$ , calcule a taxa esperada  $E[h(X)]$ .

**12.** A fdc de  $X$  (erro de medida) do Exercício 3 for

$$F(x) = \begin{cases} 0 & x < -2 \\ \frac{1}{2} + \frac{3}{32} \left(4x - \frac{x^3}{3}\right) & -2 \leq x < 2 \\ 1 & 2 \leq x \end{cases}$$

- a. Calcule  $P(X < 0)$ .
- b. Calcule  $P(-1 < X < 1)$ .
- c. Calcule  $P(0,5 < X)$ .
- d. Demonstre que  $f(x)$  está de acordo com o fornecido no Exercício 3, pela obtenção de  $F'(x)$ .
- e. Demonstre que  $\tilde{\mu} = 0$ .

**13.** O Exemplo 4.4 apresentou o conceito de tempo de avanço no fluxo de tráfego e propôs uma distribuição particular para  $X =$  o tempo de avanço entre dois carros

consecutivos selecionados aleatoriamente (em segundos). Suponha que, em um ambiente de tráfego diferente, a distribuição do tempo de avanço tenha a forma

$$f(x) = \begin{cases} \frac{k}{x^4} & x > 1 \\ 0 & x \leq 1 \end{cases}$$

- a. Determine o valor de  $k$  para o qual  $f(x)$  é uma fdp legítima.
  - b. Obtenha a função de distribuição acumulada.
  - c. Use a fdc de (b) para determinar a probabilidade de o tempo de avanço exceder 2 segundos e também a probabilidade de ele estar entre 2 e 3 segundos.
  - d. Obtenha o valor médio do tempo de avanço e seu desvio padrão.
  - e. Qual é a probabilidade de o tempo de avanço estar dentro de 1 desvio padrão em relação à média?
- 14.** O artigo “Modeling Sediment and Water Column Interactions for Hydrophobic Pollutants” (*Water Research*, 1984, p. 1169-1174) sugere a distribuição uniforme no intervalo (7,5, 20) como modelo para a profundidade (cm) da camada de bioturbação em sedimentos de uma determinada região.
- a. Qual é a média e a variância da profundidade?
  - b. Qual é a fdc da profundidade?
  - c. Qual é a probabilidade de a profundidade observada ser no máximo 10? Entre 10 e 15?
  - d. Qual é a probabilidade de a profundidade observada estar dentro de 1 desvio padrão em relação ao valor da média? E dentro de 2 desvios padrão?
- 15.** Seja  $X$  o espaço ocupado por um produto colocado em um recipiente de 1 pé cúbico. A fdp de  $X$  é

$$f(x) = \begin{cases} 90x^8(1 - x) & 0 < x < 1 \\ 0 & \text{caso contrário} \end{cases}$$

- a. Desenhe o gráfico da fdp. Determine então a fdc de  $X$  e desenhe seu gráfico.
- b. Qual é a  $P(X \leq 0,5)$  [isto é,  $F(0,5)$ ]?
- c. Usando o item (a), qual é  $P(0,25 < X \leq 0,5)$ ? Qual é  $P(0,25 \leq X \leq 0,5)$ ?
- d. Qual é o 75° percentil da distribuição?
- e. Calcule  $E(X)$  e  $\sigma_x$ .
- f. Qual é a probabilidade de  $X$  estar a mais de 1 desvio padrão em relação ao valor da média?

16. Responda aos itens (a)-(f) do Exercício 15 com  $X$  = tempo de palestra além da hora dada no Exercício 5.

17. Considere a fdp de  $X$  = peso de tração real fornecido no Exercício 6.

- a. Determine e desenhe o gráfico de  $X$ .
- b. Determine  $\tilde{\mu}$  pelo gráfico de  $f(x)$ .
- c. Calcule  $E(X)$  e  $V(X)$ .

18. Considere  $X$  tendo distribuição uniforme no intervalo  $[A, B]$ .

- a. Determine uma expressão para o  $(100p)$ -ésimo percentil.
- b. Calcule  $E(X)$ ,  $V(X)$ , e  $\sigma_x$ .
- c. Calcule  $E(X^n)$  para  $n$  inteiro positivo.

19. Seja  $X$  uma va contínua com fdc.

$$F(x) = \begin{cases} 0 & x \leq 0 \\ \frac{x}{4} \left[ 1 + \ln\left(\frac{4}{x}\right) \right] & 0 < x \leq 4 \\ 1 & x > 4 \end{cases}$$

[Esse tipo de fdc é sugerido no artigo “Variability in Measured Bedload-Transport Rates” (*Water Resources Bull.*, 1985, p. 39-48) como modelo para determinada variável hidrológica]. Qual é

- a.  $P(X \leq 1)$ ?
- b.  $P(1 \leq X \leq 3)$ ?
- c. A fdp de  $X$ ?

20. Considere a fdp do tempo total de espera  $Y$  para dois ônibus

$$f(y) = \begin{cases} \frac{1}{25} y & 0 \leq y < 5 \\ \frac{2}{5} - \frac{1}{25} y & 5 \leq y \leq 10 \\ 0 & \text{caso contrário} \end{cases}$$

apresentado no Exercício 8.

- a. Calcule e desenhe a fdc de  $Y$ . [*Sugestão*: considere separadamente  $0 \leq y < 5$  e  $5 \leq y \leq 10$  ao calcular  $F(y)$ . Um gráfico da fdp pode ser útil].
- b. Determine uma expressão para o  $(100p)$ -ésimo percentil. (*Sugestão*: considere separadamente  $0 < p < 0,5$  e  $0,5 < p < 1$ .)
- c. Calcule  $E(Y)$  e  $V(Y)$ . Como elas se comparam ao tempo de espera estimado e à variância de um único

ônibus quando o tempo tem distribuição uniforme no intervalo  $[0, 5]$ ?

21. Um ecologista deseja marcar uma região de amostragem circular com raio de 10 m. Entretanto, o raio da região resultante é, na verdade, uma variável aleatória  $R$  com fdp

$$f(r) = \begin{cases} \frac{3}{4} [1 - (10 - r)^2] & 9 \leq r \leq 11 \\ 0 & \text{caso contrário} \end{cases}$$

Qual é a área esperada da região circular resultante?

22. A demanda semanal de gás propano (em milhares de galões) de uma instalação é uma va  $X$  com fdp

$$f(x) = \begin{cases} 2 \left( 1 - \frac{1}{x^2} \right) & 1 \leq x \leq 2 \\ 0 & \text{caso contrário} \end{cases}$$

- a. Calcule a fdc de  $X$ .
- b. Determine uma expressão para o  $(100p)$ -ésimo percentil. Qual é o valor de  $\mu$ ?
- c. Calcule  $E(X)$  e  $V(X)$ .
- d. Se 1,5 mil galões estão em estoque no começo da semana e nenhuma entrega está programada para a semana, quanto desse volume se espera ter no fim da semana? *Sugestão*: seja  $h(x)$  = quantidade restante quando a demanda =  $x$ .

23. Se a temperatura de fusão de certo componente for uma variável aleatória com valor médio 120 °C e desvio padrão 2 °C, quais são a temperatura média e o desvio padrão em °F? (*Sugestão*: °F = 1,8 °C + 32.)

24. Seja  $X$  com fdp de Pareto

$$f(x; k, \theta) = \begin{cases} \frac{k \cdot \theta^k}{x^{k+1}} & x \geq \theta \\ 0 & x < \theta \end{cases}$$

apresentado no Exercício 10.

- a. Se  $k > 1$ , calcule  $E(X)$ .
- b. O que se pode dizer sobre  $E(X)$  se  $k = 1$ ?
- c. Se  $k > 2$ , mostre que  $V(X) = k\theta^2(k - 1)^{-2}(k - 2)^{-1}$ .
- d. Se  $k = 2$ , o que pode ser dito sobre  $V(X)$ ?
- e. Que condições em  $k$  são necessárias para assegurar que  $E(X^n)$  seja finita?

25. Seja  $X$  a temperatura em °C em que ocorre uma reação química e seja  $Y$  a temperatura em °F (então  $Y = 1,8X + 32$ ).

- a. Se a mediana da distribuição  $X$  é  $\tilde{\mu}$ , demonstre que  $1,8\tilde{\mu} + 32$  é a mediana da distribuição  $Y$ .
- b. Como o 90° percentil da distribuição  $Y$  se relaciona com o 90° percentil da distribuição  $X$ ? Demonstre suas conjecturas.
- c. De maneira geral, se  $Y = aX + b$ , como qualquer percentil da distribuição  $Y$  se relaciona com o percentil correspondente da distribuição  $X$ ?

## 4.3 | A Distribuição Normal

A distribuição normal é a mais importante de todas em probabilidade e estatística. Muitas populações numéricas possuem distribuições que podem ser ajustadas aproximadamente por uma curva normal apropriada. Os exemplos incluem alturas, pesos e outras características físicas (o famoso artigo da *Biometrika*, de 1903 “On the Laws of Inheritance in Man”, discutiu muitos exemplos desse tipo), erros de medida em experimentos científicos, medidas antropométricas em fósseis, tempos de reação em experimentos psicológicos, medidas de inteligência e aptidão, pontuações em testes variados e numerosas medidas e indicadores econômicos. Mesmo quando a distribuição em questão é discreta, a curva normal frequentemente fornece aproximação excelente. Além disso, ainda que as próprias variáveis individuais não sejam normalmente distribuídas, as somas e as médias das variáveis terão uma distribuição aproximadamente normal sob condições adequadas. Essa é a essência do Teorema do Limite Central discutido no próximo Capítulo.

### DEFINIÇÃO

Diz-se que uma va contínua  $X$  possui uma **distribuição normal** com parâmetros  $\mu$  e  $\sigma$  (ou  $\mu$  e  $\sigma^2$ ), onde  $-\infty < \mu < \infty$  e  $0 < \sigma$ , se a fdp de  $X$  for

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)} \quad -\infty < x < \infty \quad (4.3)$$

Novamente,  $e$  denota a base do sistema de logaritmos naturais e é aproximadamente igual a 2,71828 e  $\pi$  representa a constante matemática familiar com valor aproximado 3,14159. A afirmação de que  $X$  possui distribuição normal com parâmetros  $\mu$  e  $\sigma^2$  é normalmente abreviada por  $X \sim N(\mu, \sigma^2)$ .

Claramente  $f(x; \mu, \sigma) \geq 0$  para qualquer número  $x$ , mas as técnicas do cálculo de multivariados são usadas para demonstrar que  $\int_{-\infty}^{\infty} f(x; \mu, \sigma) dx = 1$ . Demonstra-se que  $E(X) = \mu$  e  $V(X) = \sigma^2$ , de forma que os parâmetros são a média e o desvio padrão de  $X$ . A Figura 4.13 apresenta gráficos de  $f(x; \mu, \sigma)$  para diversos pares  $(\mu, \sigma)$ . Cada gráfico é simétrico em torno de  $\mu$  e tem forma de sino, de modo que o centro do sino (ponto de simetria) é tanto a média como a mediana da distribuição. O valor de  $\sigma$  é a distância de  $\mu$  até os pontos de inflexão da curva (os pontos em que a curva muda de direção). Valores grandes de  $\sigma$  geram gráficos com grande dispersão em torno de  $\mu$ , enquanto valores pequenos de  $\sigma$  fornecem gráficos com picos altos acima de  $\mu$  e a maior parte da área do gráfico é muito próxima de  $\mu$ . Assim,  $\sigma$  grande implica boa possibilidade de se observar um valor de  $X$  afastado de  $\mu$ , enquanto essa possibilidade é muito improvável para  $\sigma$  pequeno.

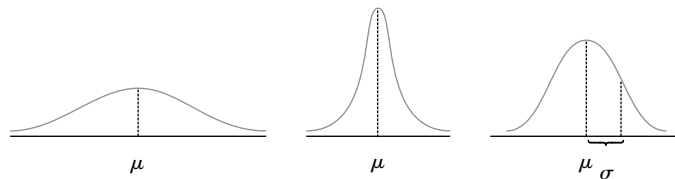


Figura 4.13 Gráficos de fdp's normais

### Distribuição Normal Padrão

Para calcular  $P(a \leq X \leq b)$  quando  $X$  é uma va normal com parâmetros  $\mu$  e  $\sigma$ , devemos calcular

$$\int_a^b \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)} dx \quad (4.4)$$

Nenhuma das técnicas de integração-padrão podem ser usadas para calcular a Expressão (4.4). Em vez disso, quando  $\mu = 0$  e  $\sigma = 1$ , a Expressão (4.4) foi calculada numericamente e tabulada para valores determinados de  $a$  e  $b$ . Essa tabela também é usada para calcular probabilidades de quaisquer outros valores de  $\mu$  e  $\sigma$  que estejam em consideração.

## DEFINIÇÃO

A distribuição normal com parâmetros  $\mu = 0$  e  $\sigma = 1$  é denominada **distribuição normal padrão**. Uma variável aleatória que tenha distribuição normal padrão é denominada **variável aleatória normal padrão** e é denotada por  $Z$ . A fdp de  $Z$  é

$$f(z; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \quad -\infty < z < \infty$$

A fdc de  $Z$  é  $P(Z \leq z) = \int_{-\infty}^z f(y; 0, 1)$ , que será denotada por  $\Phi(z)$ .

A distribuição normal padrão frequentemente não serve como modelo para uma população natural. Ao contrário, é uma distribuição de referência a partir da qual podem ser obtidas informações sobre outras distribuições normais. A Tabela A.3 do Apêndice fornece  $\Phi(z) = P(Z \leq z)$ , a área abaixo do gráfico da fdp normal padrão à esquerda de  $z$ , para  $z = -3,49, -3,48, \dots, 3,48, 3,49$ . A Figura 4.14 ilustra o tipo de área acumulada (probabilidade) tabulada na Tabela A.3. Pela tabela podem ser calculadas diversas outras probabilidades que envolvem  $Z$ .

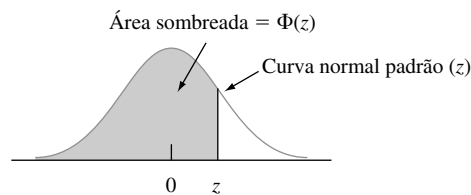


Figura 4.14 Áreas normais padrão acumuladas tabuladas na Tabela A.3 do Apêndice

## Exemplo 4.12

Calcule as seguintes probabilidades normais padrão: (a)  $P(Z \leq 1,25)$ , (b)  $P(Z > 1,25)$ , (c)  $P(Z \leq -1,25)$  e (d)  $P(-0,38 \leq Z \leq 1,25)$ .

- a.  $P(Z \leq 1,25) = \Phi(1,25)$ , probabilidade tabulada na Tabela A.3 do Apêndice na interseção da linha 1,2 com a coluna 0,05. O número mostrado é 0,8944, então  $P(Z \leq 1,25) = 0,8944$ . Consulte a Figura 4.15(a).

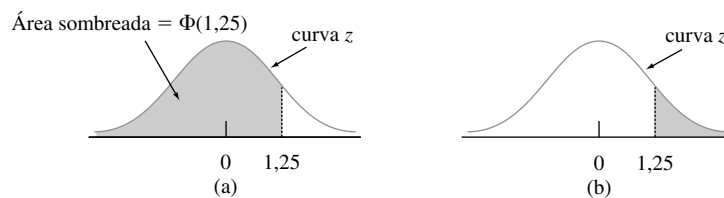


Figura 4.15 Áreas da curva normal (probabilidades) do Exemplo 4.12

- b.  $P(Z > 1,25) = 1 - P(Z \leq 1,25) = 1 - \Phi(1,25)$ , a área abaixo da curva normal padrão à direita de 1,25 (área sob a cauda superior). Como  $\Phi(1,25) = 0,8944$ , então  $P(Z > 1,25) = 0,1056$ . Como  $Z$  é uma variável contínua,  $P(Z \geq 1,25)$  também é igual a 0,1056. Consulte a Figura 4.15 (b).
- c.  $P(Z \leq -1,25) = \Phi(-1,25)$ , a área sob a cauda inferior. Diretamente da Tabela A.3 do Apêndice,  $\Phi(-1,25) = 0,1056$ . Por simetria da curva normal, esta é a mesma resposta do item (b).
- d.  $P(-0,38 \leq Z \leq 1,25)$  é a área abaixo da curva normal padrão acima do intervalo cujo ponto extremo esquerdo é  $-0,38$  e o direito é 1,25. Da Seção 4.2, se  $X$  é uma variável contínua com fdc  $F(x)$ , então  $P(a \leq X \leq b) = F(b) - F(a)$ . O que fornece  $P(-0,38 \leq Z \leq 1,25) = \Phi(1,25) - \Phi(-0,38) = 0,8944 - 0,3520 = 0,5424$ . (Consulte a Figura 4.16.)

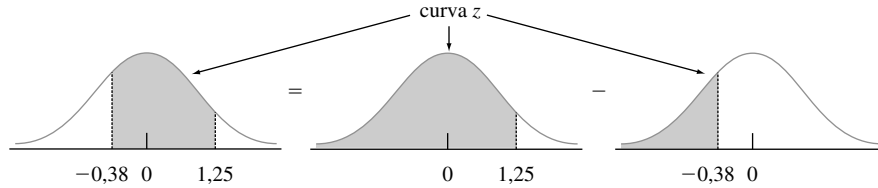


Figura 4.16  $P(-0,38 \leq Z \leq 1,25)$  como a diferença entre duas áreas acumuladas

## Percentis da Distribuição Normal Padrão

Para qualquer  $p$  entre 0 e 1, a Tabela A.3 do Apêndice é usada para obter o  $(100p)$ -ésimo percentil da distribuição normal padrão.

### Exemplo 4.13

O 99º percentil da distribuição normal padrão é o valor no eixo horizontal tal que a área sob a curva à esquerda do valor seja 0,9900. A Tabela A.3 do Apêndice fornece para  $z$  fixo a área abaixo da curva normal à esquerda de  $z$ , neste caso temos a área e queremos o valor de  $z$ . Este é o problema “inverso” de  $P(Z \leq z) = ?$  de modo que a tabela é usada de maneira inversa: encontre na tabela o valor 0,9900; a linha e a coluna onde se acha o valor identificam o 99º percentil de  $z$ . No caso, 0,9901 está na linha 2,3 e coluna 0,03, então o 99º percentil é (aproximadamente)  $z = 2,33$ . (Veja a Figura 4.17.) Por simetria, o primeiro percentil é o negativo do 99º percentil, então é igual a  $-2,33$  (1% fica abaixo do primeiro e acima do 99º). (Veja a Figura 4.18.)

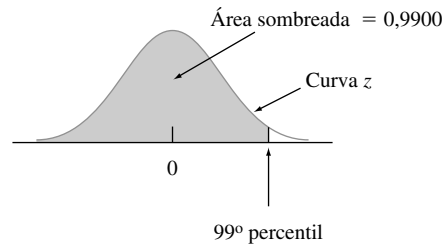


Figura 4.17 Determinação do 99º percentil

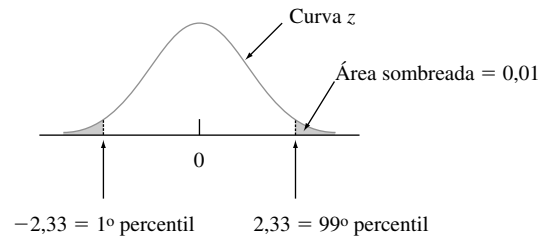


Figura 4.18 Relação entre o 1º e o 99º percentis

Em geral, o  $(100p)$ -ésimo percentil é identificado pela linha e coluna da Tabela A.3 do Apêndice onde se encontra a entrada  $p$  (por exemplo: o 67º percentil é obtido encontrando-se 0,6700 no corpo da tabela, que fornece  $z = 0,44$ ). Se  $p$  não estiver na tabela, normalmente é usado o número mais próximo a ele, apesar de que a interpolação linear fornece uma resposta mais precisa. Por exemplo: para determinar o 95º percentil, procuramos 0,9500 na tabela. Apesar de 0,9500 não aparecer, os valores 0,9495 e 0,9505 são mostrados e correspondem a  $z = 1,64$  e 1,65, respectivamente. Como 0,9500 está na metade da diferença entre as duas probabilidades mostradas, usaremos 1,645 como o 95º percentil e  $-1,645$  como o 5º percentil.

### Notação $z_\alpha$

Em inferência estatística, precisaremos dos valores do eixo das medidas que encerram pequenas áreas da cauda abaixo da curva normal padrão.

#### Notação $z_\alpha$

$z_\alpha$  representará o valor no eixo das medidas para o qual uma área abaixo da curva  $z$  fica à direita de  $z_\alpha$ . (Veja a Figura 4.19.)

Por exemplo:  $z_{0,10}$  contém a área sob a cauda superior 0,10 e  $z_{0,01}$  contém a área sob a cauda superior 0,01.

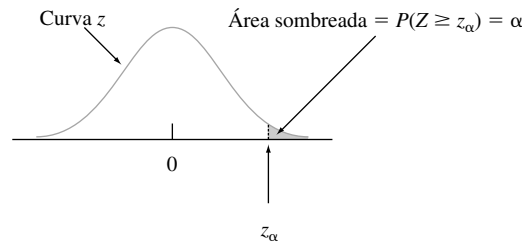


Figura 4.19 Notação  $z_\alpha$  ilustrada

Como  $\alpha$  da área abaixo da curva normal padrão está à direita de  $z_\alpha$ ,  $1 - \alpha$  da área está à esquerda de  $z_\alpha$ . Portanto,  $z_\alpha$  é o  $100(1 - \alpha)$ -ésimo percentil da distribuição normal padrão. Por simetria, a área abaixo da curva normal e à esquerda de  $-z_\alpha$  também é  $\alpha$ . Os  $z_\alpha$ s normalmente são denominados **valores críticos de  $z$** . A Tabela 4.1 relaciona os percentis normais padrão mais usados e os valores  $z_\alpha$ .

Tabela 4.1 Percentis normais padrão e valores críticos

Percentil	90	95	97,5	99	99,5	99,9	99,95
$\alpha$ (área sob a cauda)	0,1	0,05	0,025	0,01	0,005	0,001	0,0005
$z_\alpha = 100(1 - \alpha)$ -ésimo percentil	1,28	1,645	1,96	2,33	2,58	3,08	3,27

#### Exemplo 4.14

$z_{0,05}$  é o  $100(1 - 0,05)$ ésimo = 95º percentil da distribuição normal padrão, de modo que  $z_{0,05} = 1,645$ . A área abaixo da curva normal padrão à esquerda de  $-z_{0,05}$  também é 0,05. (Veja a Figura 4.20).

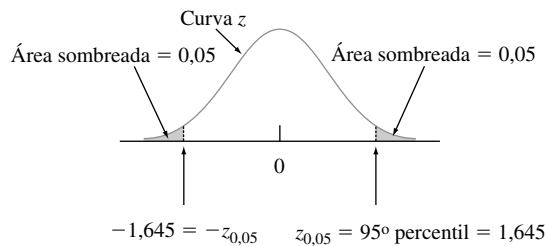


Figura 4.20 Determinação de  $z_{0,05}$

### Distribuições Normais Não-Padrão

Quando  $X \sim N(\mu, \sigma^2)$ , as probabilidades que envolvem  $X$  são calculadas por “padronização”. A **variável padrão** é  $(X - \mu)/\sigma$ . A subtração de  $\mu$  desloca a média de  $\mu$  para zero e depois a divisão por  $\sigma$  escalona a variável, de forma que o desvio padrão passa a ser 1 em vez de  $\sigma$ .

**PROPOSIÇÃO**

Se  $X$  tem distribuição normal com média  $\mu$  e desvio padrão  $\sigma$ , então

$$Z = \frac{X - \mu}{\sigma}$$

tem distribuição normal padrão. Assim,

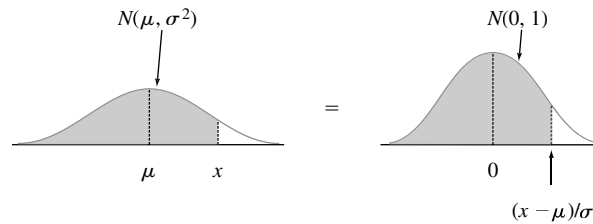
$$\begin{aligned} P(a \leq X \leq b) &= P\left(\frac{a - \mu}{\sigma} \leq Z \leq \frac{b - \mu}{\sigma}\right) \\ &= \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right) \end{aligned}$$

$$P(X \leq a) = \Phi\left(\frac{a - \mu}{\sigma}\right) \quad P(X \geq b) = 1 - \Phi\left(\frac{b - \mu}{\sigma}\right)$$

A idéia-chave da proposição é que, pela padronização, qualquer probabilidade que envolva  $X$  é expressa como uma probabilidade que envolve uma va normal padrão  $Z$ , de forma que a Tabela A.3 do Apêndice pode ser usada. A figura 4.21 ilustra o caso. A proposição é demonstrada escrevendo a fdc de  $Z = (X - \mu)/\sigma$  como

$$P(Z \leq z) = P(X \leq \sigma z + \mu) = \int_{-\infty}^{\sigma z + \mu} f(x; \mu, \sigma) dx$$

Usando um resultado do cálculo, integral pode ser diferenciada em relação à  $z$  para fornecer a fdp desejada  $f(z; 0, 1)$ .



**Figura 4.21** Igualdade das áreas de curvas normais não-padrão e padrão

**Exemplo 4.15**

O tempo que um motorista leva para reagir às luzes de freio em um veículo em desaceleração é crucial para evitar colisões traseiras. O artigo “Fast-Rise Brake Lamp as a Collision-Prevention Device” (*Ergonomics*, 1993, p. 391-395) sugere que o tempo de reação de uma resposta no trânsito a um sinal de frenagem com luzes de freio convencionais pode ser modelado com uma distribuição normal de média 1,25 segundo e desvio padrão 0,46 segundo. Qual é a probabilidade de que o tempo de reação esteja entre 1,00 e 1,75 segundo? Se representarmos por  $X$  o tempo de reação, a padronização fornece

$$1,00 \leq X \leq 1,75$$

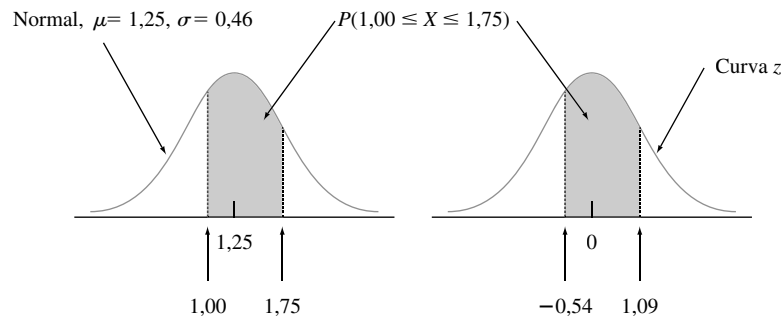
se, e somente se,

$$\frac{1,00 - 1,25}{0,46} \leq \frac{X - 1,25}{0,46} \leq \frac{1,75 - 1,25}{0,46}$$

Dessa forma,

$$\begin{aligned} p(1,00 \leq X \leq 1,75) &= P\left(\frac{1,00}{0,46} \leq Z \leq \frac{1,75 - 1,25}{0,46}\right) \\ &= P(-0,54 \leq Z \leq 1,09) = \Phi(1,09) - \Phi(-0,54) \\ &= 0,8621 - 0,2946 = 0,5675 \end{aligned}$$





**Figura 4.22** Curvas normais do Exemplo 4.15

A Figura 4.22 ilustra o problema. De forma similar, se considerarmos que 2 segundos é um tempo de resposta muito longo, a probabilidade de que o tempo real de resposta exceda esse valor será

$$P(X > 2) = P\left(Z > \frac{2 - 1,25}{0,46}\right) = P(Z > 1,63) = 1 - \Phi(1,63) = 0,0516$$

A padronização nada mais é que calcular uma distância em relação à média e então expressar essa distância novamente como certo número de desvios padrão. Por exemplo: se  $\mu = 100$  e  $\sigma = 15$ , então  $x = 130$  corresponde a  $z = (130 - 100)/15 = 30/15 = 2,00$ . Isto é, 130 está a 2 desvios padrão acima (à direita) do valor da média. De forma similar, padronizar 85 fornece  $(85 - 100)/15 = -1,00$ , de forma que 85 está 1 desvio padrão abaixo da média. A tabela  $z$  se aplica a qualquer distribuição normal, desde que pensemos em termos de número de desvios padrão em relação à média.

#### Exemplo 4.16

A voltagem de quebra de um diodo de um tipo específico tem distribuição normal. Qual é a probabilidade dessa voltagem estar a 1 desvio padrão de seu valor médio? Esta pergunta pode ser respondida sem conhecer  $\mu$  ou  $\sigma$ , desde que a distribuição seja normal. Em outras palavras, a resposta é a mesma para qualquer distribuição normal:

$$\begin{aligned} P(X \text{ está no intervalo de 1 desvio padrão da média}) &= P(\mu - \sigma \leq X \leq \mu + \sigma) \\ &= P\left(\frac{\mu - \sigma - \mu}{\sigma} \leq Z \leq \frac{\mu + \sigma - \mu}{\sigma}\right) \\ &= P(-1,00 \leq Z \leq 1,00) \\ &= \Phi(1,00) - \Phi(-1,00) = 0,6826 \end{aligned}$$

A probabilidade de  $X$  estar no intervalo de 2 desvios-padrão é  $P(-2,00 \leq Z \leq 2,00) = 0,9544$  e no intervalo de 3 desvios padrão é  $P(-3,00 \leq Z \leq 3,00) = 0,9974$ .

Os resultados do Exemplo 4.16 são frequentemente relatados sob a forma de porcentagem e são denominados *regra empírica* (porque evidências empíricas mostram que histogramas de dados reais frequentemente podem ser aproximados por curvas normais).

Se a distribuição de população de uma variável for (aproximadamente) normal, então

1. Cerca de 68% dos valores estão a 1 DP da média.
2. Cerca de 95% dos valores estão a 2 DPs da média.
3. Cerca de 99,7% dos valores estão a 3 DPs da média.

É incomum observar um valor de uma população normal muito mais distante de 2 desvios padrão de  $\mu$ . Esses resultados serão importantes no desenvolvimento dos procedimentos de teste de hipóteses nos capítulos posteriores.

## Percentis de uma Distribuição Normal Arbitrária

O  $(100p)$ -ésimo percentil de uma distribuição normal com média  $\mu$  e desvio padrão  $\sigma$  pode ser facilmente relacionado com o  $(100p)$ -ésimo percentil da distribuição normal padrão.

PROPOSIÇÃO

$$(100p)\text{-ésimo percentil para a normal } (\mu, \sigma) = \mu + \left[ (100p)\text{-ésimo para a normal padronizada} \right] \cdot \sigma$$

Outra forma de expressar a proposição: se  $z$  é o percentil desejado da distribuição normal padrão, então o percentil desejado da distribuição normal  $(\mu, \sigma)$  está a  $z$  desvios padrão de  $\mu$ .

### Exemplo 4.17

A quantidade de água destilada produzida por certa máquina tem distribuição normal com valor médio de 64 onças e desvio padrão de 0,78 onça. Qual o tamanho de recipiente  $c$  que assegurará que ocorra transbordamento em apenas 0,5% das vezes? Se  $X$  representa a quantidade usada, a condição desejada é que  $P(X > c) = 0,005$ , ou, de forma equivalente, que  $P(X \leq c) = 0,995$ . Assim,  $c$  é o 99,5º percentil da distribuição normal com  $\mu = 64$  e  $\sigma = 0,78$ . O 99,5º percentil da distribuição normal padrão é 2,58, então

$$c = \eta(0,995) = 64 + (2,58)(0,78) = 64 + 2,0 = 66 \text{ onças}$$

A questão está ilustrada na Figura 4.23.

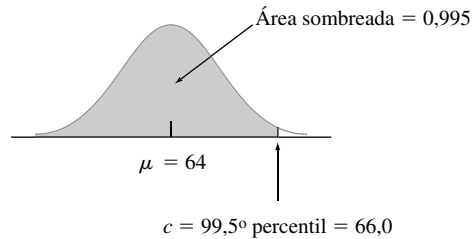


Figura 4.23 Distribuição da quantidade produzida do Exemplo 4.17

## A Distribuição Normal e as Populações Discretas

A distribuição normal é usada com freqüência como aproximação para a distribuição de valores de uma população discreta. Em tais situações, devemos tomar cuidado redobrado para nos assegurarmos de que as probabilidades sejam calculadas de forma precisa.

### Exemplo 4.18

O QI em uma determinada população (medido pelo teste-padrão) é sabido ter distribuição aproximadamente normal com  $\mu = 100$  e  $\sigma = 15$ . Qual é a probabilidade de um indivíduo selecionado aleatoriamente ter um QI de pelo menos 125? Sendo  $X = \text{QI}$  de uma pessoa selecionada aleatoriamente, queremos saber  $P(X \geq 125)$ . A tentação aqui é padronizar  $X \geq 125$  imediatamente, como foi feito nos exemplos anteriores. O problema, entretanto, é que a população de QI é discreta, visto que os QI's são valores inteiros, de forma que a curva normal é a aproximação de um histograma de probabilidade discreta, conforme ilustrado na Figura 4.24.

Os retângulos do histograma estão centrados em inteiros, de forma que o QI de pelo menos 125 corresponde aos retângulos começando em 124,5, conforme sombreado na Figura 4.24. Assim, queremos a área abaixo da curva normal de aproximação à direita de 124,5. A padronização desse valor fornece  $P(Z \geq 1,63) = 0,0516$ . Se tivéssemos padronizado  $X \geq 125$ , teríamos obtido  $P(Z \geq 1,67) = 0,0475$ . A diferença não é grande, mas a resposta 0,0516 é mais precisa. De forma similar,  $P(X = 125)$  seria aproximada pela área entre 124,5 e 125,5, já que a área abaixo da curva normal acima do valor 125 é zero.

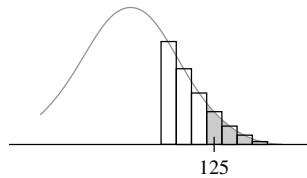


Figura 4.24 Aproximação normal de uma distribuição discreta

A correção da descontinuidade da distribuição subjacente no Exemplo 4.18 é frequentemente denominada **correção de continuidade**. Ela é útil na seguinte aplicação da distribuição normal, para cálculo de probabilidades da binomial.

### Aproximação Normal para a Distribuição Binomial

Lembre-se de que o valor médio e o desvio padrão de uma variável aleatória  $X$  são  $\mu_x = np$  e  $\sigma_x = \sqrt{npq}$ , respectivamente. A Figura 4.25 exibe um histograma de probabilidade binomial para a distribuição binomial com  $n = 20$ ,  $p = 0,6$  [de forma que  $\mu = 20(0,6) = 12$  e  $\sigma = \sqrt{20(0,6)(0,04)} = 2,19$ ]. Uma curva normal com valor médio e desvio padrão iguais aos valores correspondentes para a distribuição binomial foi sobreposta no histograma de probabilidade. Apesar de o histograma ter um pouco de inclinação (porque  $p \neq 0,5$ ), a curva normal fornece uma aproximação muito boa, especialmente na parte central da figura. A área de qualquer retângulo (probabilidade de qualquer valor  $X$  particular), exceto aqueles nas extremidades das caudas, pode ser aproximada com precisão pela área correspondente da curva normal. Por exemplo:  $P(X = 10) = B(10; 20, 0,6) - B(9; 20, 0,6) = 0,117$ , enquanto a área sob a curva normal entre 9,5 e 10,5 é  $P(-1,14 \leq Z \leq -0,68) = 0,1212$ .

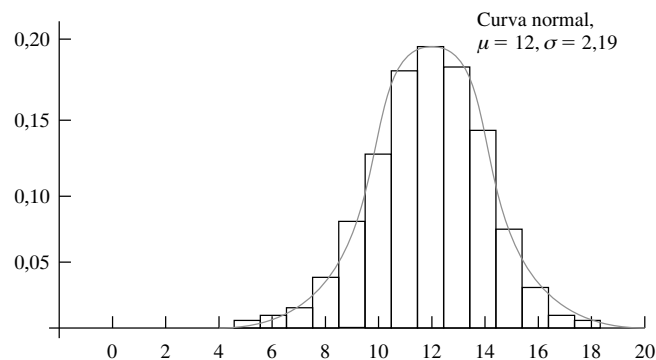


Figure 4.25 Histograma de probabilidade binomial para  $n = 20$ ,  $p = 0,6$  com curva de aproximação normal sobreposta

De forma mais geral, contanto que o histograma de probabilidade não apresente muita inclinação, as probabilidades binomiais podem ser bem aproximadas pelas áreas da curva normal. Habitualmente se diz que  $X$  tem uma distribuição aproximadamente normal.

#### PROPOSIÇÃO

Seja  $X$  uma va binomial com base em  $n$  tentativas com probabilidade de sucesso  $p$ . Então, se o histograma de probabilidade binomial não tiver muita inclinação,  $X$  terá uma distribuição aproximadamente normal com  $\mu = np$  e  $\sigma = \sqrt{npq}$ . Em particular, para  $x =$  valor possível de  $X$ ,

$$\begin{aligned}
 P(X \leq x) &= B(x; n, p) \approx \left( \begin{array}{l} \text{área abaixo da curva normal} \\ \text{à esquerda de } x + 0,5 \end{array} \right) \\
 &= \Phi\left(\frac{x + 0,5 - np}{\sqrt{npq}}\right)
 \end{aligned}$$

Na prática, a aproximação é adequada desde que  $np \geq 10$  e  $nq \geq 10$ .

Se  $np < 10$  ou  $nq < 10$ , a distribuição terá muita inclinação para que a curva normal (simétrica) forneça uma aproximação precisa.

#### Exemplo 4.19

Suponha que 25% de todos os motoristas habilitados de um determinado estado não tenham seguro. Represente por  $X$  o número de motoristas sem seguro em uma amostra aleatória de tamanho 50 (com um quê de perversidade, o sucesso é o motorista não ter seguro), de forma que  $p = 0,25$ . Então  $\mu = 12,5$  e  $\sigma = 3,06$ . Como  $np = 50(0,25) = 12,5 \geq 10$  e  $nq = 37,5 \geq 10$ , a aproximação pode ser aplicada com segurança:

$$\begin{aligned} P(X \leq 10) &= B(10; 50, 0,25) \approx \Phi\left(\frac{10 + 0,5 - 12,5}{3,06}\right) \\ &= \Phi(-0,65) = 0,2578 \end{aligned}$$

De forma similar, a probabilidade de inclusive entre 5 e 15 motoristas selecionados não terem seguro é

$$\begin{aligned} P(5 \leq X \leq 15) &= B(15; 50, 0,25) - B(4; 50, 0,25) \\ &\approx \Phi\left(\frac{15,5 - 12,5}{3,06}\right) - \Phi\left(\frac{4,5 - 12,5}{3,06}\right) = 0,8320 \end{aligned}$$

As probabilidades exatas são 0,2622 e 0,8348, respectivamente, de forma que as aproximações são muito boas. No último cálculo, a probabilidade  $P(5 \leq X \leq 15)$  é aproximada pela área sob a curva normal entre 4,5 e 15,5. A correção de continuidade é usada para os limites superior e inferior. ■

Quando o objetivo da investigação é fazer uma inferência sobre a proporção  $p$  de uma população, o interesse enfocará a proporção amostral de sucessos  $X/n$  em vez de  $X$ . Como essa proporção é  $X$  multiplicada pela constante  $1/n$ , também terá uma distribuição aproximadamente normal (com média  $\mu = p$  e desvio padrão  $\sigma = \sqrt{pq/n}$ ) desde que  $np \geq 10$  e  $nq \geq 10$ . Essa aproximação normal é a base de diversos procedimentos inferenciais a serem discutidos nos capítulos posteriores.

É muito difícil dar uma prova direta da validade da aproximação normal (a primeira é de Laplace, há mais de 150 anos). No próximo capítulo, veremos que ela é uma consequência de um importante resultado geral denominado Teorema do Limite Central.

---

## Exercícios | Seção 4.3 (26–54)

26. Seja  $Z$  uma variável aleatória normal e calcule as probabilidades a seguir, fazendo as ilustrações quando apropriado.
- $P(0 \leq Z \leq 2,17)$
  - $P(0 \leq Z \leq 1)$
  - $P(-2,50 \leq Z \leq 0)$
  - $P(-2,50 \leq Z \leq 2,50)$
  - $P(Z \leq 1,37)$
  - $P(-1,75 \leq Z)$
  - $P(-1,50 \leq Z \leq 2,00)$
  - $P(1,37 \leq Z \leq 2,50)$
  - $P(1,50 \leq Z)$
  - $P(|Z| \leq 2,50)$
27. Em cada caso, determine o valor da constante  $c$  que torna correta a declaração de probabilidade.
- $\Phi(c) = 0,9838$
  - $P(0 \leq Z \leq c) = 0,291$
  - $P(c \leq Z) = 0,121$
  - $P(-c \leq Z \leq c) = 0,668$
  - $P(c \leq |Z|) = 0,016$
28. Encontre os percentis a seguir para a distribuição normal. Faça a interpolação quando apropriado.
- 91º
  - 9º
  - 75º
  - 25º
  - 6º
29. Determine  $z_\alpha$  para os itens a seguir:
- $\alpha = 0,0055$
  - $\alpha = 0,09$
  - $\alpha = 0,663$
30. Se  $X$  é uma va normal com média 80 e desvio padrão 10, calcule as probabilidades a seguir, usando padronização:

- a.  $P(X \leq 100)$   
 b.  $P(X \leq 80)$   
 c.  $P(65 \leq X \leq 100)$   
 d.  $P(70 \leq X)$   
 e.  $P(85 \leq X \leq 95)$   
 f.  $P(|X - 80| \leq 10)$
31. Suponha que a força que age sobre uma coluna que ajuda a suportar um edifício tenha distribuição normal com média 15,0 kips e desvio padrão 1,25 kips. Qual é a probabilidade de a força:
- Ser no máximo 18 kips?
  - Estar entre 10 e 12 kips?
  - Diferir de 15,0 kips por no máximo 2 desvios padrão?
32. O artigo “Reliability of Domestic-Waste Biofilm Reactors” (*J. of Envir. Engr.*, 1995, p. 785-790) sugere que a concentração de substrato (mg/cm<sup>3</sup>) de influência para um reator tem distribuição normal com  $\mu = 0,30$  e  $\sigma = 0,06$ .
- Qual é a probabilidade de a concentração exceder 0,25?
  - Qual é a probabilidade de a concentração ser no máximo 0,10?
  - Como você caracterizaria os maiores 5% de todos os valores de concentração?
33. Suponha que o diâmetro de certo tipo de árvores na altura do tronco tenha distribuição normal com  $\mu = 8,8$  e  $\sigma = 2,8$ , conforme sugerido pelo artigo “Simulating a Harvester-Forwarder Softwood Thinning” (*Forest Products J.*, May 1997, p. 36-41).
- Qual é a probabilidade de uma árvore selecionada aleatoriamente ter um diâmetro de no mínimo 10 polegadas? Exceder 10 polegadas?
  - Qual é a probabilidade de o diâmetro de uma árvore selecionada aleatoriamente exceder 20 polegadas?
  - Qual é a probabilidade de o diâmetro de uma árvore selecionada aleatoriamente estar entre 5 e 10 polegadas?
  - Que valor  $c$  faz com que o intervalo  $(8,8 - c, 8,8 + c)$  inclua 98% de todos os valores de diâmetro?
  - Se quatro árvores forem selecionadas de forma independente, qual é a probabilidade de ao menos uma ter diâmetro maior que 10 polegadas?
34. Há duas máquinas disponíveis para corte de rolhas para garrafas de vinho. A primeira produz rolhas com diâmetros que possuem uma distribuição normal com média 3 cm e desvio padrão 0,1 cm. A segunda máquina produz rolhas com diâmetros que possuem uma distribuição normal com média 3,04 cm e desvio padrão 0,02 cm. As rolhas aceitáveis possuem diâmetros entre 2,9 cm e 3,1 cm. Que máquina tem maior probabilidade de produzir uma rolha aceitável?
35. a. Se uma distribuição normal tem  $\mu = 30$  e  $\sigma = 5$ , qual é o 91<sup>o</sup> percentil da distribuição?  
 b. Qual é o 6<sup>o</sup> percentil da distribuição?  
 c. A largura de uma linha gravada em um chipe de circuito integrado tem distribuição normal com média 3,000  $\mu\text{m}$  e desvio-padrão 0,140. Que valor separa os 10% mais largos de todas as linhas dos outros 90%?
36. O artigo “Monte Carlo Simulation-Tool for Better Understanding of LRFD” (*J. Structural Engr.*, 1993, p. 1586-1599) sugere que a resistência de rendimento (ksi) de aço de graduação A36 tem distribuição normal com  $\mu = 43$  e  $\sigma = 4,5$ .
- Qual é a probabilidade de a resistência ser no máximo 40? Maior que 60?
  - Que valor de resistência separa os 75% mais fortes dos outros?
37. O dispositivo de abertura automática de um pára-quadras de carga militar foi projetado para abrir quando estiver a 200 m do solo. Suponha que a altitude de abertura tenha uma distribuição normal com média 200 m e desvio padrão 30 m. Haverá dano no equipamento se o pára-quadras abrir a uma altitude inferior a 100 m. Qual é a probabilidade de haver dano ao equipamento em pelo menos um de cinco pára-quadras lançados independentemente?
38. A leitura de temperatura de um par termoeletrico colocado em um meio de temperatura constante tem distribuição normal com média  $\mu$  e desvio-padrão  $\sigma$ . Que valor  $\sigma$  precisaria existir para assegurar que 95% de todas as leituras estão dentro de  $0,1^\circ$  de  $\mu$ ?
39. A distribuição da resistência de resistores de um tipo específico é normal, 10% de todos os equipamentos apresentam resistência maior que 10,256 ohms e 5% com resistência menor que 9,671 ohms. Quais são os valores da média e do desvio padrão da distribuição das resistências?
40. Se o comprimento da rosca de um parafuso tem distribuição normal, qual é a probabilidade de o comprimento da rosca de um parafuso selecionado aleatoriamente:
- estar a 1,5 DPs do valor da média?
  - estar a mais de 2,5 DPs do valor da média?
  - estar entre 1 e 2 DPs do valor da média?
41. Uma máquina que produz rolamentos inicialmente foi configurada para que o diâmetro real médio dos rolamentos produzidos seja de 0,500 polegadas. Um rolamento é aceitável se o diâmetro está dentro de 0,004 polegadas desse valor-alvo. Suponha, entretanto, que a configuração tenha sido alterada durante o curso da produção, de forma que os rolamentos tenham diâmetros com distribuição normal com média 0,499 polegadas e desvio padrão 0,002 polegadas. Que porcentagem dos rolamentos produzidos não será aceitável?
42. A dureza Rockwell de um metal é determinada pela pressão de uma ponta rígida na superfície do metal e, em seguida, pela medição da profundidade de penetração das pontas. Suponha que a dureza Rockwell de uma determinada liga tenha distribuição normal com média 70 e desvio padrão 3. (A dureza Rockwell é medida em uma escala contínua.)

- a. Se uma amostra for aceitável apenas se a sua dureza estiver entre 67 e 75, qual é a probabilidade de uma amostra selecionada aleatoriamente ter dureza aceitável?
- b. Se o intervalo aceitável de dureza for  $(70 - c, 70 + c)$ , para que valor de  $c$  95% de todas as amostras teriam uma dureza aceitável?
- c. Se o intervalo aceitável for o da parte (a) e a dureza de cada uma de 10 amostras selecionadas aleatoriamente for determinada de forma independente, qual será o número esperado de amostras aceitáveis dentre os 10?
- d. Qual é a probabilidade de no máximo oito de 10 amostras selecionadas de forma independente terem dureza inferior a 73,84? (*Sugestão*:  $Y =$  número entre as 10 amostras com dureza inferior a 73,84 é uma variável binomial. Qual é  $p$ ?)
43. A distribuição do peso de pacotes enviados de uma determinada forma é normal com média 12 libras e desvio padrão 3,5 libras. O serviço de encomendas deseja determinar um valor  $c$  acima do qual haverá um acréscimo. Que valor de  $c$  faz com que 99% de todos os pacotes estejam no mínimo 1 libra abaixo do peso de acréscimo?
44. Suponha que a Tabela A.3 do Apêndice contenha  $\Phi(z)$  apenas para  $z \geq 0$ . Explique como você pode calcular:
- a.  $P(-1,72 \leq Z \leq -0,55)$
- b.  $P(-1,72 \leq Z \leq 0,55)$
- É necessário tabelar  $\Phi(z)$  para  $z$  negativo? Que propriedade da curva normal padrão justifica sua resposta?
45. Considere os bebês nascidos no intervalo “normal” de 37-43 semanas de gestação. Muitos dados apóiam a suposição de que, para os bebês nascidos nos Estados Unidos, o peso de nascimento possui distribuição normal com média 3432 g e desvio padrão 482 g. [O artigo “Are Babies Normal” (*The American Statistician*, 1999, p. 298-302) analisou dados de um ano específico. Para uma escolha significativa de intervalos de classe, um histograma não tinha aparência normal, mas, após mais investigações, determinou-se que isso acontecia devido a alguns hospitais medirem o peso em gramas e outros em onças (com aproximação para a próxima onça) e depois converterem para gramas. Uma escolha de intervalos de classe modificada resultou um histograma bem-descrito por uma distribuição normal.]
- a. Qual é a probabilidade de o peso de nascimento de um bebê selecionado aleatoriamente exceder 4000 gramas? Estar entre 3000 e 4000 gramas?
- b. Qual é a probabilidade de o peso de um bebê selecionado aleatoriamente ser inferior a 2000 gramas ou superior a 5000 gramas?
- c. Qual é a probabilidade de o peso de nascimento de um bebê selecionado aleatoriamente exceder 7 libras?
- d. Como você caracterizaria o 0,1% mais extremo de todos os pesos de nascimento?
- e. Se  $X$  for uma variável aleatória com distribuição normal e  $a$  for uma constante numérica ( $a \neq 0$ ), então  $Y = aX$  também terá uma distribuição normal. Use isso para determinar a distribuição do peso de nascimento expresso em libras (formato, média e desvio padrão) e então calcule novamente a probabilidade da parte (c). Como isso se compara à resposta anterior?
46. Em resposta às preocupações sobre o conteúdo nutricional dos *fast foods*, o McDonald’s anunciou que usará um novo óleo na fritura de suas batatas. O produto diminuirá substancialmente os níveis de ácido de gordura trans e aumentará a quantidade de gordura poliinsaturada mais benéfica. A empresa alega que 97 dentre 100 pessoas não conseguem detectar diferença no gosto entre o óleo antigo e o novo. Assumindo que esse valor esteja correto (como proporção de longo prazo), qual é a probabilidade aproximada de, em uma amostra de 1000 indivíduos que compraram fritas no McDonald’s,
- a. ao menos 40 sentirem a diferença de gosto entre os dois óleos?
- b. no máximo 5% sentirem a diferença de gosto entre os dois óleos?
47. A desigualdade de Chebyshev, apresentada no Exercício 43 (Capítulo 3), é válida para distribuições discretas e contínuas. Ela afirma que, para qualquer número  $k$  que satisfaça  $k \geq 1$ ,  $P(|X - \mu| \geq k\sigma) \leq 1/k^2$  (consulte o Exercício 43 no Capítulo 3 para obter uma interpretação). Obtenha essa probabilidade no caso de uma distribuição normal para  $k = 1, 2$  e  $3$  e a compare ao limite superior.
48. Seja por  $X$  o número de falhas em um rolo de 100 m de fita magnética (uma variável de valores inteiros). Suponha que  $X$  tenha uma distribuição aproximadamente normal com  $\mu = 25$  e  $\sigma = 5$ . Use a correção de continuidade para calcular a probabilidade de o número de falhas:
- a. estar entre 20 e 30, inclusive.
- b. ser no máximo 30. Inferior a 30.
49. Assuma  $X$  com uma distribuição binomial de parâmetros  $n = 25$  e  $p$ . Calcule cada uma das probabilidades a seguir, usando a aproximação normal (com correção de continuidade) para os casos  $p = 0,5, 0,6$  e  $0,8$  e compare com as probabilidades exatas calculadas pela Tabela A.1 do Apêndice.
- a.  $P(15 \leq X \leq 20)$
- b.  $P(X \leq 15)$
- c.  $P(20 \leq X)$
50. Suponha que 10% de todos os eixos de aço produzidos por um processo específico não estejam em conformidade mas possam ser retrabalhados (em vez de serem sucateados). Considere uma amostra aleatória de 200 eixos e represente por  $X$  o número dentre eles que não estejam conforme e possam ser retrabalhados. Qual é a probabilidade (aproximada) de  $X$  ser:
- a. No máximo 30?

- b. Menos que 30?  
c. Estar entre 15 e 25 (inclusive)?
51. Suponha que apenas 40% de todos os motoristas de um estado específico usem o cinto de segurança regularmente. É selecionada uma amostra de 500 motoristas. Qual é a probabilidade de:
- a. entre 180 e 230 (inclusive) motoristas usarem o cinto de segurança regularmente?  
b. menos de 175 dos motoristas da amostra usarem o cinto regularmente? Menos de 150?
52. Mostre que a relação entre um percentil normal geral e o percentil correspondente  $z$  atende ao que foi dito nesta seção.
53. a. Mostre que, se  $X$  tem distribuição normal com parâmetros  $\mu$  e  $\sigma$ , então  $Y = aX + b$  (função linear de  $X$ ) também tem distribuição normal. Quais são os parâmetros da distribuição de  $Y$  [ou seja,  $E(Y)$  e  $V(Y)$ ]? [Sugestão: escreva a FDA de  $Y$ ,  $P(Y \leq y)$  como uma integral que envolve a fdp de  $X$  e então faça a diferenciação em relação a  $y$  para obter a fdp de  $Y$ .]
- b. Se, quando medida em  $^{\circ}\text{C}$ , a temperatura tem distribuição normal com média 115 e desvio padrão 2, o que se pode dizer sobre a distribuição da temperatura medida em  $^{\circ}\text{F}$ ?
54. Não há uma boa fórmula para a fdc normal padrão  $\Phi(z)$ , mas foram publicadas diversas aproximações razoáveis. A aproximação a seguir foi obtida em “Approximations for Hand Calculators Using Small Integer Coefficients” (*Mathematics of Computation*, 1977, p. 214-222). Para  $0 \leq z \leq 5,5$ ,

$$P(Z \geq z) = 1 - \Phi(z) \approx 0,5 \exp \left\{ - \left[ \frac{(83z + 351)z + 562}{z} \right] \right\}$$

O erro relativo dessa aproximação é inferior a 0,042%. Use isso para calcular aproximações das probabilidades a seguir e as compare com as probabilidades obtidas na Tabela A.3 do Apêndice, sempre que possível.

- a.  $P(Z \geq 1)$   
b.  $P(Z < -3)$   
c.  $P(-4 < Z < 4)$   
d.  $P(Z > 5)$

## 4.4 | A Distribuição Gama e Seus Parentes

O gráfico de qualquer fdp normal tem formato de sino e é simétrico. Há várias situações práticas em que a variável de interesse do investigador pode ter distribuição com inclinação. Uma família de fdp que fornece diversos formatos de distribuições com inclinação é a família gama. Para defini-la, precisamos antes apresentar uma função que tem importante papel em vários ramos da matemática.

### DEFINIÇÃO

Para  $\alpha > 0$ , a **função gama**  $\Gamma(\alpha)$  é definida por

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx \quad (4.5)$$

As propriedades mais importantes da função gama são as seguintes:

1. Para qualquer  $\alpha > 1$ ,  $\Gamma(\alpha) = (\alpha - 1) \cdot \Gamma(\alpha - 1)$  [via integração por partes]
2. Para qualquer inteiro positivo,  $n$ ,  $\Gamma(n) = (n - 1)!$
3.  $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

Pela Expressão (4.5), se tivermos

$$f(x; \alpha) = \begin{cases} \frac{x^{\alpha-1} e^{-x}}{\Gamma(\alpha)} & x \geq 0 \\ 0 & \text{caso contrário} \end{cases} \quad (4.6)$$

então  $f(x; \alpha) \geq 0$  e  $\int_0^{\infty} f(x; \alpha) dx = \Gamma(\alpha)/\Gamma(\alpha) = 1$ , então  $f(x; \alpha)$  satisfaz às duas propriedades básicas de uma fdp.

## A Família de Distribuições Gama

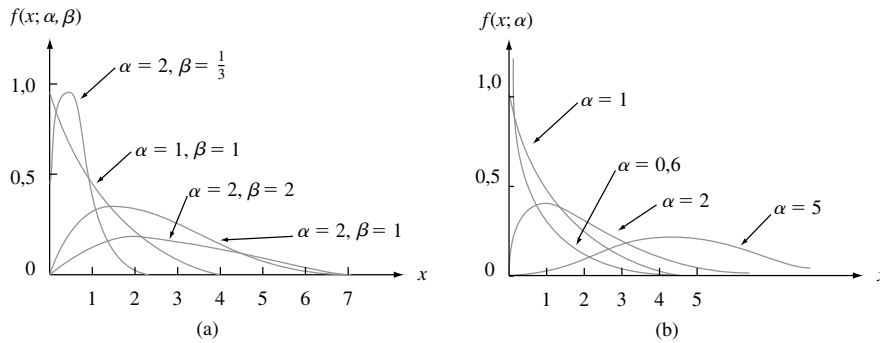
**DEFINIÇÃO**

Diz-se que uma variável aleatória contínua  $X$  tem uma **distribuição gama** se a fdp de  $X$  é

$$f(x; \alpha, \beta) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} & x \geq 0 \\ 0 & \text{caso contrário} \end{cases} \quad (4.7)$$

onde os parâmetros  $\alpha$  e  $\beta$  satisfazem  $\alpha > 0, \beta > 0$ . A **distribuição gama padrão** tem  $\beta = 1$ , de forma que a fdp de uma va gama padrão é dada por (4.6).

A Figura 4.26(a) ilustra os gráficos da fdp gama  $f(x; \alpha, \beta)$  para diversos pares  $(\alpha, \beta)$ , enquanto a Figura 4.26(b) apresenta os gráficos da fdp gama padrão. Para a fdp-padrão, quando  $\alpha \leq 1$ ,  $f(x; \alpha)$  for estritamente decrescente à medida que  $x$  aumenta a partir de 0. Quando  $\alpha > 1$ ,  $f(x; \alpha)$  aumenta a partir de 0 em  $x = 0$  até um máximo e depois decresce. O parâmetro  $\beta$  em (4.7) é denominado *parâmetro de escala* porque os valores diferentes de 1 esticam ou comprimem a fdp na direção de  $x$ .



**Figura 4.26** (a) Curvas de densidade gama; (b) curvas de densidade gama padrão

$E(X)$  e  $E(X^2)$  podem ser obtidos a partir de uma integração razoavelmente direta e então  $V(X) = E(X^2) - [E(X)]^2$ .

A média e a variância de uma variável aleatória  $X$  com distribuição gama  $f(x; \alpha, \beta)$  são

$$E(X) = \mu = \alpha\beta \quad V(X) = \sigma^2 = \alpha\beta^2$$

Quando  $X$  é uma va gama-padrão, a fdc de  $X$ ,

$$F(x; \alpha) = \int_0^x \frac{y^{\alpha-1} e^{-y}}{\Gamma(\alpha)} dy \quad x > 0 \quad (4.8)$$

é denominada **função gama incompleta** [algumas vezes a função gama incompleta se refere à Expressão (4.8) sem o denominador  $\Gamma(\alpha)$  no integrando]. Há extensas tabelas de  $F(x; \alpha)$  disponíveis. Na Tabela A.4 do Apêndice, apresentamos uma pequena tabulação para  $\alpha = 1, 2, \dots, 10$  e  $x = 1, 2, \dots, 15$ .

**Exemplo 4.20**

Suponha que o tempo de reação  $X$  de um indivíduo selecionado aleatoriamente a um certo estímulo possui uma distribuição gama padrão com  $\alpha = 2$  segundos. Como

$$P(a \leq X \leq b) = F(b) - F(a)$$



quando  $X$  é contínuo,

$$P(3 \leq X \leq 5) = F(5; 2) - F(3; 2) = 0,960 - 0,801 = 0,159$$

A probabilidade de o tempo de reação ser mais de 4 segundos é

$$P(X > 4) = 1 - P(X \leq 4) = 1 - F(4; 2) = 1 - 0,908 = 0,092 \quad \blacksquare$$

A função gama incompleta também pode ser usada para calcular probabilidades que envolvam distribuições gama não-padrão.

#### PROPOSIÇÃO

Seja  $X$  com uma distribuição gama de parâmetros  $\alpha$  e  $\beta$ . Então, para qualquer  $x > 0$ , a fdc de  $X$  é dada por

$$P(X \leq x) = F(x; \alpha, \beta) = F\left(\frac{x}{\beta}; \alpha\right)$$

onde  $F(\cdot; \alpha)$  é a função gama incompleta.<sup>2</sup>

#### Exemplo 4.21

Suponha que o tempo de sobrevivência  $X$  em semanas de um camundongo macho selecionado aleatoriamente exposto a 240 rads de radiação gama tenha distribuição gama com  $\alpha = 8$  e  $\beta = 15$ . (Dados em *Survival Distributions: Reliability Applications in the Biomedical Services*, de A. J. Gross e V. Clark, sugerem  $\alpha \approx 8,5$  e  $\beta \approx 13,3$ .) O tempo esperado de sobrevivência é  $E(X) = (8)(15) = 120$  semanas, enquanto  $V(X) = (8)(15)^2 = 1800$  e  $\sigma_X = \sqrt{1800} = 42,43$  semanas. A probabilidade de um camundongo sobreviver entre 60 e 120 semanas é

$$\begin{aligned} P(60 \leq X \leq 120) &= P(X \leq 120) - P(X \leq 60) \\ &= F(120/15; 8) - F(60/15; 8) \\ &= F(8; 8) - F(4; 8) = 0,547 - 0,051 = 0,496 \end{aligned}$$

A probabilidade de um camundongo sobreviver ao menos 30 semanas é

$$\begin{aligned} P(X \geq 30) &= 1 - P(X < 30) = 1 - P(X \leq 30) \\ &= 1 - F(30/15; 8) = 0,999 \end{aligned} \quad \blacksquare$$

## A Distribuição Exponencial

A família de distribuições exponenciais fornece modelos probabilísticos largamente usados em engenharia e em várias disciplinas da ciência.

#### DEFINIÇÃO

Diz-se que  $X$  tem uma **distribuição exponencial** com parâmetro  $\lambda$  ( $\lambda > 0$ ) se a fdp de  $X$  for

$$f(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{caso contrário} \end{cases} \quad (4.9)$$

A fdp exponencial é um caso especial da Expressão (4.7) da fdp gama geral em que  $\alpha = 1$  e  $\beta$  foi substituído por  $1/\lambda$  [alguns autores usam a forma  $(1/\beta)e^{-x/\beta}$ ]. A média e a variância de  $X$  são

$$\mu = \alpha\beta = \frac{1}{\lambda} \quad \sigma^2 = \alpha\beta^2 = \frac{1}{\lambda^2}$$

<sup>2</sup> MINITAB e outros programas estatísticos calculam  $F(x; \alpha, \beta)$ , uma vez que os valores de  $x$ ,  $\alpha$ , e  $\beta$  são especificados.

Tanto a média como o desvio padrão da distribuição exponencial são iguais a  $1/\lambda$ . São exibidos diversos gráficos das fdps exponenciais na Figura 4.27.

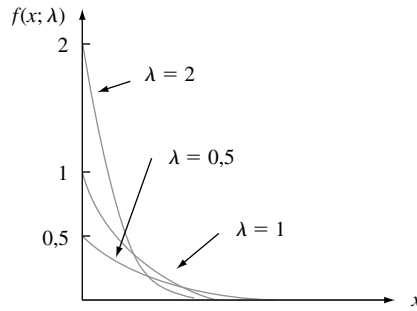


Figura 4.27 Curvas de densidade exponencial

Diferentemente da fdp gama geral, a fdp exponencial pode ser facilmente integrada. Em particular, a fdc de  $X$  é

$$F(x; \lambda) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \geq 0 \end{cases}$$

**Exemplo 4.22**

Suponha que o tempo de resposta  $X$  em um terminal de computador *on-line* específico (o tempo entre o final de uma consulta de um usuário e o começo da resposta do sistema para essa consulta) tenha distribuição exponencial com tempo de resposta esperado igual a 5 segundos. Então  $E(X) = 1/\lambda = 5$ , com  $\lambda = 0,2$ . A probabilidade de o tempo de resposta ser no máximo 10 segundos é

$$P(X \leq 10) = F(10; 0,2) = 1 - e^{-(0,2)(10)} = 1 - e^{-2} = 1 - 0,135 = 0,865$$

A probabilidade de o tempo de resposta estar entre 5 e 10 segundos é

$$\begin{aligned} P(5 \leq X \leq 10) &= F(10; 0,2) - F(5; 0,2) \\ &= (1 - e^{-2}) - (1 - e^{-1}) = 0,233 \end{aligned}$$

A distribuição exponencial é usada freqüentemente como modelo para distribuição dos tempos entre a ocorrência de eventos sucessivos, tais como clientes chegando em uma unidade de atendimento ou chamadas em uma central telefônica. O motivo disso é que a distribuição exponencial está fortemente relacionada ao processo de Poisson discutido no Capítulo 3.

**PROPOSIÇÃO**

Suponha que o número de eventos que ocorrem em um intervalo de tempo de duração  $t$  tenha distribuição de Poisson com parâmetro  $\alpha t$  (onde  $\alpha$ , a taxa do processo do evento, é o número esperado de eventos que ocorrem em uma unidade de tempo) e que os números das ocorrências em intervalos não-sobrepostos sejam independentes um do outro. Então a distribuição do tempo decorrido entre a ocorrência de dois eventos sucessivos é exponencial com parâmetro  $\lambda = \alpha$ .

Apesar de uma demonstração completa estar além do escopo do livro, o resultado é facilmente verificado para o tempo  $X_1$ , até que ocorra o primeiro evento:

$$\begin{aligned} P(X_1 \leq t) &= 1 - P(X_1 > t) = 1 - P[\text{nenhum evento em } (0, t)] \\ &= 1 - \frac{e^{-\alpha t} \cdot (\alpha t)^0}{0!} = 1 - e^{-\alpha t} \end{aligned}$$

que é exatamente a FDA da distribuição exponencial.

#### Exemplo 4.23

Suponha que sejam recebidas chamadas em um serviço telefônico de “auxílio a suicidas” 24 horas, de acordo com um processo de Poisson com taxa  $\alpha = 0,5$  chamada por dia. Então o número de dias  $X$  entre chamadas sucessivas tem uma distribuição exponencial com valor de parâmetro 0,5, de forma que a probabilidade de haver mais de 2 dias entre chamadas é

$$P(X > 2) = 1 - P(X \leq 2) = 1 - F(2; 0,5) = e^{-(0,5)(2)} = 0,368$$

O tempo esperado entre chamadas sucessivas é  $1/0,5 = 2$  dias. ■

Outra aplicação importante da distribuição exponencial é modelar a distribuição do tempo de vida de um componente. O motivo parcial da popularidade dessas aplicações é a propriedade de “falta de memória” da distribuição exponencial. Suponha que o tempo de vida de um componente seja distribuído exponencialmente com o parâmetro  $\lambda$ . Após o componente ser colocado em serviço, saímos por um período de  $t_0$  horas e então retornamos para encontrar o componente ainda funcionando. Nesse caso, qual é a probabilidade de ele durar ao menos  $t$  horas adicionais? Em símbolos, queremos  $P(X \geq t + t_0 | X \geq t_0)$ . Pela definição de probabilidade condicional,

$$P(X \geq t + t_0 | X \geq t_0) = \frac{P[(X \geq t + t_0) \cap (X \geq t_0)]}{P(X \geq t_0)}$$

Entretanto, o evento,  $X \geq t_0$  no numerador é redundante, já que os dois eventos podem ocorrer se, e somente se,  $X \geq t + t_0$ . Portanto,

$$P(X \geq t + t_0 | X \geq t_0) = \frac{P(X \geq t + t_0)}{P(X \geq t_0)} = \frac{1 - F(t + t_0; \lambda)}{1 - F(t_0; \lambda)} = e^{-\lambda t}$$

Tal probabilidade condicional é idêntica à probabilidade original  $P(X \geq t)$  de o componente ter durado  $t$  horas. Assim, a *distribuição do tempo de vida adicional é exatamente igual à distribuição original de tempo de vida*, de forma que em cada ponto no tempo o componente não dá sinais de desgaste. Em outras palavras, a distribuição do tempo de vida restante é independente da idade atual.

Embora a propriedade de ausência de memória possa ser justificada pelo menos, aproximadamente, em muitos problemas aplicados, em outras situações os componentes se deterioram ou ocasionalmente melhoram com o uso (ao menos até certo ponto). Modelos mais gerais de tempo de vida são fornecidos pelas distribuições gama, Weibull e lognormal (as duas últimas serão discutidas na próxima seção).

## A Distribuição Qui-Quadrado

### DEFINIÇÃO

Seja  $\nu$  um inteiro positivo. Diz-se que uma variável aleatória  $X$  possui uma **distribuição qui-quadrado** com parâmetro  $\nu$  se a fdp de  $X$  for a densidade gama com  $\alpha = \nu/2$  e  $\beta = 2$ . A fdp de uma va qui-quadrado será

$$f(x; \nu) = \begin{cases} \frac{1}{2^{\nu/2} \Gamma(\nu/2)} x^{(\nu/2)-1} e^{-x/2} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (4.10)$$

O parâmetro  $\nu$  é denominado **número de graus de liberdade** (gl) de  $X$ . O símbolo  $\chi^2$  frequentemente é usado no lugar de “qui-quadrado”.

A distribuição qui-quadrado é importante por ser a base de diversos procedimentos de inferência estatística. O motivo disso é que as distribuições qui-quadrado estão intimamente relacionadas a distribuições normal (consulte o Exercício 65). Discutiremos a distribuição qui-quadrado com mais detalhes nos capítulos sobre inferência.

## Exercícios | Seção 4.4 (55–65)

55. Calcule os dados a seguir:
- $\Gamma(6)$
  - $\Gamma(5/2)$
  - $F(4; 5)$  (a função gama incompleta)
  - $F(5; 4)$
  - $F(0; 4)$
56. Seja  $X$  uma distribuição gama-padrão com  $\alpha = 7$ . Calcule o seguinte:
- $P(X \leq 5)$
  - $P(X < 5)$
  - $P(X > 8)$
  - $P(3 \leq X \leq 8)$
  - $P(3 < X < 8)$
  - $P(X < 4 \text{ ou } X > 6)$
57. Suponha que o tempo gasto por um aluno selecionado aleatoriamente que usa um terminal conectado a uma instalação de computador com *time-sharing* tem uma distribuição gama com média de 20 minutos e variância de  $80 \text{ min}^2$ .
- Quais são os valores de  $\alpha$  e  $\beta$ ?
  - Qual é a probabilidade de um aluno usar o terminal por no máximo 24 minutos?
  - Qual é a probabilidade de um aluno passar entre 20 e 40 minutos usando o terminal?
58. Suponha que, quando um certo tipo de transistor é sujeito a um teste de vida acelerado, o tempo de vida  $X$  (em semanas) possui distribuição gama com média 24 semanas e desvio padrão 12 semanas.
- Qual é a probabilidade de um transistor durar entre 12 e 24 semanas?
  - Qual é a probabilidade de um transistor durar no máximo 24 semanas? A mediana da distribuição do tempo de vida é inferior a 24? Por que ou por que não?
  - Qual é o 99º percentil da distribuição do tempo de vida?
  - Suponha que o teste seja finalizado após  $t$  semanas. Que valor de  $t$  faz com que apenas 0,5% de todos os transistores estejam operando no fim do teste?
59. Seja  $X =$  tempo entre duas chegadas sucessivas no guichê de atendimento rápido de um banco local. Se  $X$  possui distribuição exponencial com  $\lambda = 1$  (que é idêntico a uma distribuição gama-padrão com  $\alpha = 1$ ), calcule os itens a seguir:
- O tempo esperado entre duas chegadas sucessivas
  - O desvio padrão do tempo entre chegadas sucessivas
  - $P(X \leq 4)$
  - $P(2 \leq X \leq 5)$
60. Seja  $X$  a distância (m) que um animal viaja desde seu local de nascimento até o primeiro local vago que encontra. Suponha que, para ratos-cangurus de rabo de bandeira,  $X$  possui uma distribuição exponencial com parâmetro  $\lambda = 0,01386$  (como sugerido no artigo “Competition and Dispersal from Multiple Nests,” *Ecology*, 1997, p. 873-883).
- Qual é a probabilidade de a distância ser no máximo 100 m? No máximo 200 m? Entre 100 e 200 m?
  - Qual é a probabilidade de a distância exceder a média por mais de 2 desvios padrão?
  - Qual é o valor da mediana da distância?
61. Diversas experiências com determinado tipo de ventilador, usados em motores a diesel, indicam que a distribuição exponencial sugere um bom modelo para cálculo do tempo até uma falha. Suponha que o tempo médio seja 25.000 horas. Qual é a probabilidade de:
- um ventilador selecionado aleatoriamente durar pelo menos 20.000 horas? No máximo 30.000 horas? Entre 20.000 e 30.000 horas?
  - o tempo de vida de um ventilador exceder o valor médio em mais de 2 desvios padrão? Mais de 3 desvios padrão?
62. O caso especial da distribuição gama em que  $\alpha$  é um inteiro positivo  $n$  é denominado distribuição de Erlang. Se substituirmos  $\beta$  por  $1/\lambda$  na Expressão (4.7), a fdp Erlang será
- $$f(x; \lambda, n) = \begin{cases} \frac{\lambda(\lambda x)^{n-1} e^{-\lambda x}}{(n-1)!} & x \geq 0 \\ 0 & x < 0 \end{cases}$$
- É possível demonstrar que, se os tempos entre os eventos sucessivos forem independentes, cada um com uma distribuição exponencial de parâmetro  $\lambda$ , o tempo total  $X$  antes de todos os  $n$  próximos eventos ocorrerem terá fdp  $f(x; \lambda, n)$ .
- Qual é o valor esperado de  $X$ ? Se o tempo (em minutos) entre as chegadas de clientes sucessivos for distribuído exponencialmente com  $\lambda = 0,5$ , quanto tempo podemos esperar que passe antes da chegada do décimo cliente?
  - Se o tempo entre as chegadas de clientes for distribuído exponencialmente com  $\lambda = 0,5$ , qual é a probabilidade de o décimo cliente (depois do que

acabou de chegar) chegue dentro dos próximos 30 minutos?

- c. O evento  $\{X \leq t\}$  ocorre se e somente se ao menos  $n$  eventos ocorrerem nas próximas  $t$  unidades de tempo. Use o fato de que o número de eventos que ocorrem em um intervalo de duração  $t$  possui uma distribuição de Poisson com parâmetro  $\lambda t$  para escrever uma expressão (envolvendo probabilidades de Poisson) para a FDA de Erlang  $F(t; \lambda, n) = P(X \leq t)$ .
63. Um sistema consiste de cinco componentes idênticos conectados em série, conforme ilustrado:



Se um componente falhar, todo o sistema falhará. Suponha que cada componente tenha um tempo de vida distribuído exponencialmente com  $\lambda = 0,01$  e que eles falhem de forma independente. Defina os eventos  $A_i = \{i\text{-ésimo componente que dura no mínimo } t \text{ horas}\}$ ,  $i = 1, \dots, 5$ , de forma que os  $A_i$  são eventos independentes. Seja  $X =$  momento da falha do sistema, isto é, o tempo de vida mais curto (mínimo) entre os cinco componentes.

- a. O evento  $\{X \geq t\}$  é equivalente a que evento que envolva  $A_1, \dots, A_5$ ?
  - b. Usando a independência dos  $A_i$ s, calcule  $P(X \geq t)$ . Obtenha então  $F(t) = P(X \leq t)$  e a fdp de  $X$ . Que tipo de distribuição tem  $X$ ?
  - c. Suponha que haja  $n$  componentes, cada um com tempo de vida exponencial com parâmetro  $\lambda$ . Que tipo de distribuição tem  $X$ ?
64. Se  $X$  tiver distribuição exponencial com parâmetro  $\lambda$ , deduza uma expressão geral para o percentil ( $100p$ ) da distribuição. Então faça a especialização para obter a mediana.
65. a. O evento  $\{X^2 \leq y\}$  é equivalente a que evento que envolve  $X$ ?
- b. Se  $X$  tiver distribuição normal padrão, use a parte (a) para escrever a integral que iguala  $P(X^2 \leq y)$ . Diferencie então essa expressão em relação a  $y$  para obter a fdp de  $X^2$  [o quadrado de uma variável  $N(0, 1)$ ]. Finalmente, mostre que  $X^2$  possui distribuição qui-quadrado com  $\nu = 1$  gl [consulte (4.10)]. (Sugestão: use a identidade a seguir).

$$\frac{d}{dy} \left\{ \int_{a(y)}^{b(y)} f(x) dx \right\} = f[b(y)] \cdot b'(y) - f[a(y)] \cdot a'(y)$$

## 4.5 | Outras Distribuições Contínuas

As famílias de distribuições normal, gama (incluindo exponencial) e uniforme fornecem ampla variedade de modelos probabilísticos para variáveis contínuas, mas há muitas situações práticas em que nenhum membro dessas famílias consegue se enquadrar bem em um conjunto de dados observados. Os estatísticos e outros investigadores desenvolveram outras famílias de distribuições que normalmente são apropriadas, na prática.

### A Distribuição de Weibull

A família de distribuições de Weibull foi apresentada pelo físico sueco Waloddi Weibull, em 1939. Seu artigo de 1951 “A Statistical Distribution Function of Wide Applicability” (*J. Applied Mechanics*, vol. 18, p. 293-297) apresenta diversas aplicações.

**DEFINIÇÃO**

Uma variável aleatória  $X$  tem distribuição de Weibull com parâmetros  $\alpha$  e  $\beta$  ( $\alpha > 0, \beta > 0$ ) se a fdp de  $X$  for

$$f(x; \alpha, \beta) = \begin{cases} \frac{\alpha}{\beta^\alpha} x^{\alpha-1} e^{-(x/\beta)^\alpha} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (4.11)$$

Em algumas situações, há justificativas teóricas para a utilização da distribuição de Weibull, mas em muitas aplicações  $f(x; \alpha, \beta)$  simplesmente fornece uma boa aproximação dos dados observados para valores específicos de  $\alpha$  e  $\beta$ . Quando  $\alpha = 1$ , a fdp é reduzida à distribuição exponencial (com  $\lambda = 1/\beta$ ), de forma que a distribuição

exponencial é um caso especial das distribuições gama e de Weibull. Entretanto, há distribuições gama que não são distribuições de Weibull e vice-versa, o que faz com que uma família não seja um subconjunto da outra. Tanto  $\alpha$  como  $\beta$  podem variar para se obter diversos formatos de distribuição diferentes, como ilustrado na Figura 4.28.  $\beta$  é um parâmetro de escala, de forma que valores diferentes esticam ou comprimem o gráfico na direção de  $x$ .

Integrar para obter  $E(X)$  e  $E(X^2)$  resulta em

$$\mu = \beta\Gamma\left(1 + \frac{1}{\alpha}\right) \quad \sigma^2 = \beta^2\left\{\Gamma\left(1 + \frac{2}{\alpha}\right) - \left[\Gamma\left(1 + \frac{1}{\alpha}\right)\right]^2\right\}$$

O cálculo de  $\mu$  e de  $\sigma^2$  necessita, portanto, da função gama.

A integração  $\int_0^x f(y; \alpha, \beta) dy$  pode ser feita facilmente para obter a fdc de  $X$ .

A fdc de uma va de Weibull com parâmetros  $\alpha$  e  $\beta$  é

$$F(x; \alpha, \beta) = \begin{cases} 0 & x < 0 \\ 1 - e^{-(x/\beta)^\alpha} & x \geq 0 \end{cases} \quad (4.12)$$

#### Exemplo 4.24

Nos últimos anos, a distribuição de Weibull tem sido usada para modelar emissões de poluentes de vários motores. Seja  $X$  o valor da emissão de  $\text{NO}_x$  (g/gal) a partir de certo tipo de motor de quatro tempos selecionado aleatoriamente e suponha que  $X$  possua uma distribuição de Weibull com  $\alpha = 2$  e  $\beta = 10$  (sugerido pelas informações do artigo “Quantification of Variability and Uncertainty in Lawn and Garden Equipment  $\text{NO}_x$  and Total Hydrocarbon Emission Factors,” *J. of the Air and Waste Management Assoc.*, 2002, p. 435-448). A curva de densidade correspondente tem a mesma aparência daquela da Figura 4.28 para  $\alpha = 2$ ,  $\beta = 1$ , exceto pelo fato de que agora os valores 50 e 100 substituem 5 e 10 no eixo horizontal (porque  $\beta$  é uma “parâmetro de escala”). Então

$$P(X \leq 10) = F(10; 2, 10) = 1 - e^{-(10/10)^2} = 1 - e^{-1} = 0,632$$

De forma similar,  $P(X \leq 25) = 0,998$ , de maneira que a distribuição está quase inteiramente concentrada nos valores entre 0 e 25. O valor  $c$ , que separa os 5% de todos os motores que emitem as maiores quantidades de  $\text{NO}_x$  dos 95% restantes, satisfaz

$$0,95 = 1 - e^{-(c/10)^2}$$

Isolando o termo exponencial em um lado, tirando os logaritmos e resolvendo a equação resultante, temos  $c \approx 17,3$  como o 95º percentil da distribuição de tempos de vida. ■

Freqüentemente, em situações práticas, o modelo de Weibull é razoável, exceto pelo fato de que o menor valor  $X$  possível pode ser um valor  $\gamma$  não assumido como zero (isso também se aplica a um modelo gama). A quantidade  $\gamma$  pode então ser vista como um terceiro parâmetro da distribuição, que foi o que Weibull fez em seu trabalho original. Para, digamos,  $\gamma = 3$ , todas as curvas na Figura 4.28 seriam deslocadas 3 unidades para a direita. Isso é equivalente a dizer que  $X - \gamma$  possui a fdp (4.11), de forma que a fdc de  $X$  é obtida pela substituição de  $x$  em (4.12) por  $x - \gamma$ .

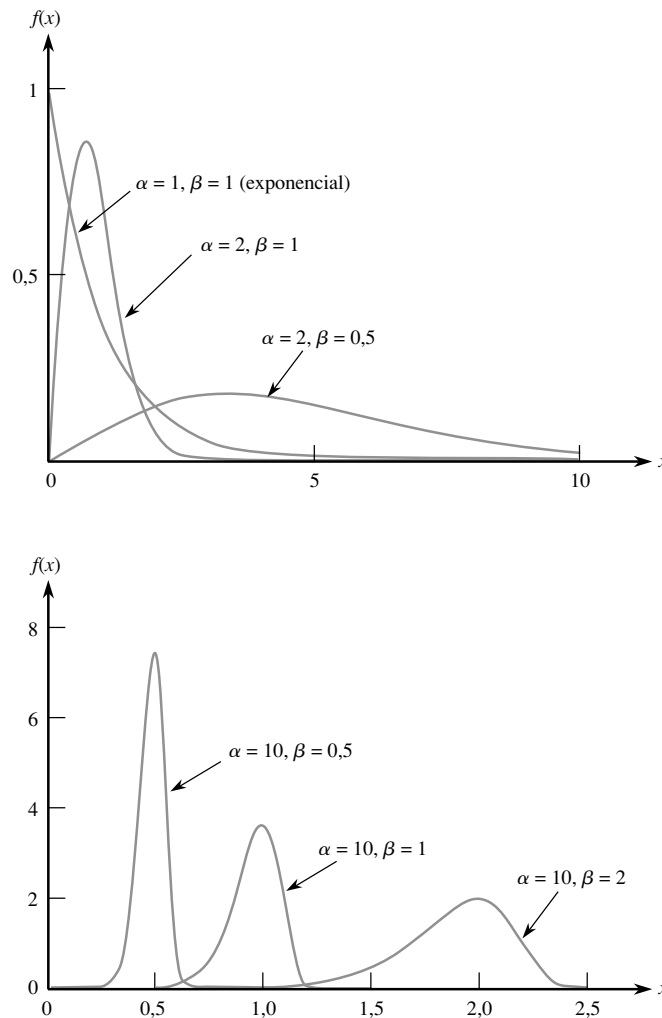


Figura 4.28 Curvas de densidade de Weibull

**Exemplo 4.25**

Seja  $X$  = perda de massa por corrosão de um pequeno prato quadrado de uma liga de magnésio imersa durante 7 dias em uma solução aquosa inibida de 20% de  $MgBr_2$ . Suponha que a menor perda de massa possível seja  $\gamma = 3$  e que o excesso  $X - 3$ , além desse mínimo, tenha distribuição de Weibull com  $\alpha = 2$  e  $\beta = 4$ . (Esse exemplo foi considerado em “Practical Applications of the Weibull Distribution,” *Industrial Quality Control*, Ago., 1964, p. 71-78. Os valores de  $\alpha$  e  $\beta$  foram escolhidos como 1,8 e 3,67, respectivamente, embora no artigo tenha sido usada uma escolha de parâmetros ligeiramente diferente.) A fdc de  $X$  será então

$$F(x; \alpha, \beta, \gamma) = F(x; 2, 4, 3) = \begin{cases} 0 & x < 3 \\ 1 - e^{-[(x-3)/4]^2} & x \geq 3 \end{cases}$$

Portanto,

$$P(X > 3,5) = 1 - F(3,5; 2, 4, 3) = e^{-0,0156} = 0,985$$

e

$$P(7 \leq X \leq 9) = 1 - e^{-2,25} - (1 - e^{-1}) = 0,895 - 0,632 = 0,263$$



## A Distribuição Lognormal

**DEFINIÇÃO**

Diz-se que uma va não-negativa  $X$  possui uma **distribuição lognormal** se a va  $Y = \ln(X)$  possui uma distribuição normal. A fdp resultante de uma va lognormal quando  $\ln(X)$  tiver distribuição normal com parâmetros  $\mu$  e  $\sigma$  é

$$f(x; \mu, \sigma) = \begin{cases} \frac{1}{\sqrt{2\pi} \sigma x} e^{-[\ln(x)-\mu]^2/(2\sigma^2)} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

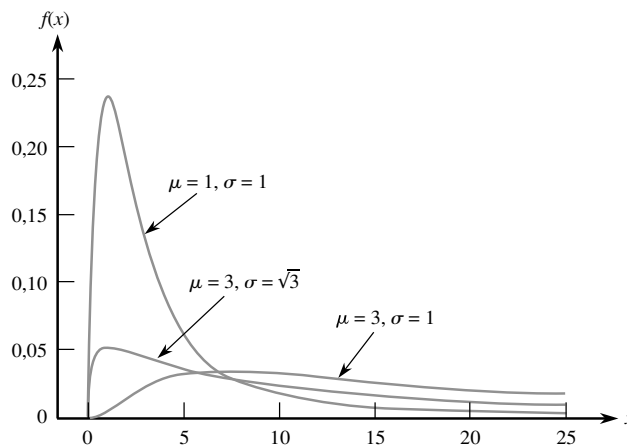
Tenha cuidado neste caso. Os parâmetros  $\mu$  e  $\sigma$  não são a média e o desvio padrão de  $X$  e sim de  $\ln(X)$ . A média e a variância de  $X$  são

$$E(X) = e^{\mu+\sigma^2/2} \quad V(X) = e^{2\mu+\sigma^2} \cdot (e^{\sigma^2} - 1)$$

No Capítulo 5, apresentaremos uma justificativa teórica para essa distribuição, relacionada ao Teorema do Limite Central, mas, assim como ocorre com outras distribuições, a lognormal pode ser usada como modelo, mesmo na ausência de tal justificativa. A Figura 4.29 ilustra os gráficos da fdp lognormal. Embora a curva normal seja simétrica, a lognormal possui inclinação positiva.

Como  $\ln(X)$  possui distribuição normal, a fdc de  $X$  pode ser expressa em termos da fdc  $\Phi(z)$  de uma va  $Z$  normal padrão. Para  $x \geq 0$ ,

$$\begin{aligned} F(x; \mu, \sigma) &= P(X \leq x) = P[\ln(X) \leq \ln(x)] \\ &= P\left(Z \leq \frac{\ln(x) - \mu}{\sigma}\right) = \Phi\left(\frac{\ln(x) - \mu}{\sigma}\right) \end{aligned} \tag{4.13}$$



**Figura 4.29** Curvas de densidade lognormal

**Exemplo 4.26**

A distribuição lognormal é usada frequentemente como modelo para diversas propriedades de materiais. O artigo “Reliability of Wood Joist Floor Systems with Creep” (*J. of Structural Engr.*, 1995, p. 946-954) sugere que a distribuição lognormal com  $\mu = 0,375$  e  $\sigma = 0,25$  é um modelo plausível para  $X =$  módulo da elasticidade



(MOE, em  $10^6$  psi) de sistemas de vigas de madeira construídos de hem-fir grau número 2. O valor da média e a variância de MOE são

$$E(X) = e^{0,375 + (0,25)^2/2} = e^{0,40625} = 1,50$$

$$V(X) = e^{0,8125}(e^{0,0625} - 1) = 0,1453$$

A probabilidade de MOE estar entre 1 e 2 é

$$\begin{aligned} P(1 \leq X \leq 2) &= P(\ln(1) \leq \ln(X) \leq \ln(2)) \\ &= P(0 \leq \ln(X) \leq 0,693) \\ &= P\left(\frac{0 - 0,375}{0,25} \leq Z \leq \frac{0,693 - 0,375}{0,25}\right) \\ &= \Phi(1,27) - \Phi(-1,50) = 0,8312 \end{aligned}$$

Que valor  $c$  faz com que apenas 1% dos sistemas tenha MOE que exceda  $c$ ? Desejamos o  $c$  para que

$$0,99 = P(X \leq c) = P\left(Z \leq \frac{\ln(c) - 0,375}{0,25}\right)$$

a partir de onde  $(\ln(c) - 0,375)/0,25 = 2,33$  e  $c = 2,605$ . Portanto, 2,605 é o 99º percentil da distribuição de MOE. ■

## Distribuição Beta

Todas as famílias de distribuições contínuas discutidas até agora, exceto a distribuição uniforme, têm densidade positiva em um intervalo infinito (apesar de normalmente a função densidade decrescer rapidamente para zero além de alguns desvios padrão em relação à média). A distribuição beta fornece densidade positiva apenas para  $X$  em um intervalo de comprimento finito.

### DEFINIÇÃO

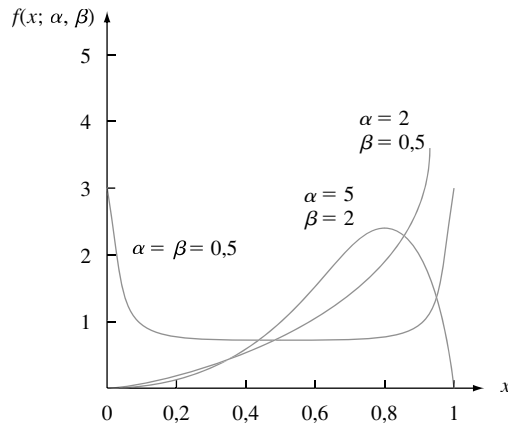
Diz-se que uma variável aleatória  $X$  possui **distribuição beta** com parâmetros  $\alpha$ ,  $\beta$  (ambos positivos),  $A$  e  $B$  se a fdp de  $X$  for

$$f(x; \alpha, \beta, A, B) = \begin{cases} \frac{1}{B - A} \cdot \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \cdot \Gamma(\beta)} \left(\frac{x - A}{B - A}\right)^{\alpha-1} \left(\frac{B - x}{B - A}\right)^{\beta-1} & A \leq x \leq B \\ 0 & \text{caso contrário} \end{cases}$$

O caso  $A = 0$ ,  $B = 1$  fornece a **distribuição beta-padrão**.

A Figura 4.30 ilustra diversas fdps beta-padrão. Os gráficos da fdp geral são similares, exceto pelo fato de que os valores são deslocados e então comprimidos ou estendidos para se adaptarem a  $[A, B]$ . A menos que  $\alpha$  e  $\beta$  sejam inteiros, a integração da fdp para calcular probabilidades é difícil, de forma que normalmente é usada uma tabela da função beta incompleta. A média e a variância de  $X$  são

$$\mu = A + (B - A) \cdot \frac{\alpha}{\alpha + \beta} \quad \sigma^2 = \frac{(B - A)^2 \alpha \beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)}$$



**Figura 4.30** Curvas de densidade beta-padrão

**Exemplo 4.27**

Os gerentes de projeto normalmente usam um método denominado PERT (técnica de revisão e avaliação de programa) para coordenar as diversas atividades de um projeto grande (uma aplicação de sucesso foi a construção da espaçonave Apollo.) Uma hipótese-padrão na análise PERT é que o tempo necessário para completar qualquer atividade específica depois de seu início tem uma distribuição beta com  $A =$  tempo otimista (se tudo correr bem) e  $B =$  tempo pessimista (se tudo correr mal). Suponha que, na construção de uma casa, o tempo  $X$  (em dias) necessário para a construção das fundações possui distribuição beta com  $A = 2, B = 5, \alpha = 2,$  e  $\beta = 3$ . Então  $\alpha/(\alpha + \beta) = 0,4$ , de forma que  $E(X) = 2 + (3)(0,4) = 3,2$ . Para esses valores de  $a$  e  $\beta$ , a fdp de  $X$  é uma função polinomial simples. A probabilidade de levar no máximo 3 dias para construir os alicerces é

$$\begin{aligned}
 P(X \leq 3) &= \int_2^3 \frac{1}{3} \cdot \frac{4!}{1!2!} \left(\frac{x-2}{3}\right) \left(\frac{5-x}{3}\right)^2 dx \\
 &= \frac{4}{27} \int_2^3 (x-2)(5-x)^2 dx = \frac{4}{27} \cdot \frac{11}{4} = \frac{11}{27} = 0,407
 \end{aligned}$$

A distribuição beta-padrão normalmente é usada para modelar a variação na proporção ou porcentagem de uma quantidade que ocorre em diferentes amostras, tal como a proporção de um dia de 24 horas em que um indivíduo está dormindo ou a proporção de determinado elemento em um composto químico.

**Exercícios** Seção 4.5 (66–80)

- 66.** O tempo de vida  $X$  (em centenas de horas) de certo tipo de válvula eletrônica possui uma distribuição Weibull com parâmetros  $\alpha = 2$  e  $\beta = 3$ . Calcule os itens a seguir:
- $E(X)$  e  $V(X)$
  - $P(X \leq 6)$
  - $P(1,5 \leq X \leq 6)$
- (Essa distribuição de Weibull é sugerida como modelo de tempo em serviço em “On the Assessment of Equipment Reliability: Trading Data Collection Costs for Precision,” *J. Engr Manuf.*, 1991, p. 105-109.)
- 67.** Os autores do artigo “A Probabilistic Insulation Life Model for Combined Thermal-Electrical Stresses”

(*IEEE Trans. on Elect. Insulation*, 1985, p. 519-522) declararam que “a distribuição de Weibull é largamente usada em problemas estatísticos relativos ao desgaste de materiais isolantes sólidos sujeitos ao desgaste e à tensão”. Eles propõem o uso da distribuição como um modelo para tempo (em horas) de falha de amostras de isolantes sólidos sujeitos a voltagem de AC. Os valores dos parâmetros dependem da voltagem e da temperatura. Suponha que  $\alpha = 2,5$  e  $\beta = 200$  (valores sugeridos pelos dados no artigo)

- Qual é a probabilidade de o tempo de vida de um espécime ser no máximo 250? Inferior a 250? Maior de 300?

- b. Qual é a probabilidade de o tempo de vida de uma amostra estar entre 100 e 250?
- c. Que valor faz com que 50% de todas as amostras tenham tempos de vida que excedam o valor?
68. Seja  $X =$  tempo (em  $10^{-1}$  semanas) desde a remessa de um produto com defeito até que o cliente o devolva. Suponha que o tempo mínimo de retorno seja  $\gamma = 3,5$  e que o excesso  $X - 3,5$  acima do mínimo possua uma distribuição de Weibull com parâmetros  $\alpha = 2$  e  $\beta = 1,5$  (consulte o artigo *Industrial Quality Control*, mencionado no Exemplo 4.25).
- Qual é a FDA de  $X$ ?
  - Quais são o tempo de retorno esperado e a variância do tempo de retorno? [*Sugestão*: Primeiro obtenha  $E(X - 3,5)$  e  $V(X - 3,5)$ .]
  - Calcule  $P(X > 5)$ .
  - Calcule  $P(5 \leq X \leq 8)$ .
69. Seja  $X$  com distribuição de Weibull e a fdp da Expressão (4.11). Demonstre que  $\mu = \beta\Gamma(1 + 1/\alpha)$ . (*Sugestão*: na integral de  $E(X)$ , faça a mudança de variável  $y = (x/\beta)^\alpha$ , de forma que  $x = \beta y^{1/\alpha}$ .)
70. a. No Exercício 66, qual a mediana do tempo de vida das válvulas? [*Sugestão*: use a Expressão (4.12).]  
 b. No Exercício 68, qual é a mediana do prazo de retorno?  
 c. Se  $X$  tiver uma distribuição de Weibull com a fdc da Expressão (4.12), obtenha uma expressão geral para o  $(100p)$ ésimo percentil da distribuição.  
 d. No Exercício 68, a empresa deseja recusar devoluções após  $t$  semanas. Para que valores de  $t$  apenas 10% de todas as devoluções serão recusadas?
71. Seja  $X$  o limite de resistência à tração (ksi) a  $-200^\circ$  de um corpo de prova de aço selecionado aleatoriamente que apresenta “fragilidade ao frio” em baixas temperaturas. Suponha que  $X$  tenha distribuição de Weibull com  $\alpha = 20$  e  $\beta = 100$ .
- Qual é a probabilidade de  $X$  ser no máximo 105?
  - Se forem selecionados vários corpos de prova um após o outro, qual é a probabilidade no longo prazo de obter valores de resistência entre 100 e 105 ksi?
  - Qual é o valor da mediana da distribuição de resistência?
72. Os autores do artigo em que foram obtidos os dados do Exercício 1.27 sugerem que um modelo de probabilidade razoável para a vida útil de uma broca é a distribuição lognormal com  $\mu = 4,5$  e  $\sigma = 0,8$ .
- Qual é o valor da média e o desvio padrão da vida útil?
  - Qual é a probabilidade da vida útil ser no máximo 100?
  - Qual é a probabilidade da vida útil ser no mínimo 200? Maior que 200?
73. Seja  $X =$  potência média horária (em decibéis) de sinais de rádio recebidos transmitidos entre duas cidades. Os autores do artigo “Families of Distributions for Hourly Median Power and Instantaneous Power of Received Radio Signals” (*J. Research National Bureau of Standards*, vol. 67D, 1963, p. 753-762) argumentam que a distribuição lognormal fornece um modelo probabilístico razoável para  $X$ . Se os valores dos parâmetros forem  $\mu = 3,5$  e  $\sigma = 1,2$ , calcule os itens a seguir:
- O valor médio e o desvio-padrão da potência recebida.
  - A probabilidade de a potência recebida estar entre 50 e 250 dB.
  - A probabilidade de  $X$  ser menor que o valor da média. Por que a probabilidade não é 0,5?
74. a. Use a Equação (4.13) para escrever uma fórmula para a mediana  $\tilde{\mu}$  da distribuição lognormal. Qual é a mediana da distribuição do Exercício 73?  
 b. Lembrando que  $z_\alpha$  é a notação do percentil  $100(1 - \alpha)$  da distribuição padrão normal, escreva uma expressão para o percentil  $100(1 - \alpha)$  da distribuição lognormal. No Exercício 73, que valor a potência recebida excederá apenas 5% do tempo?
75. Uma justificativa teórica baseada no mecanismo de falha de certo material admite a hipótese de que a resistência dúctil  $X$  de um material possui distribuição lognormal. Suponha que os parâmetros sejam  $\mu = 5$  e  $\sigma = 0,1$ .
- Calcule  $E(X)$  e  $V(X)$ .
  - Calcule  $P(X > 125)$ .
  - Calcule  $P(110 \leq X \leq 125)$ .
  - Qual é o valor da mediana da resistência dúctil?
  - Se 10 diferentes amostras de uma liga de aço desse tipo forem submetidas a um teste de resistência, quantas você espera que tenham ao menos 125?
  - Se 5% dos menores valores não forem aceitáveis, qual seria a resistência mínima aceitável?
76. O artigo “The Statistics of Phytotoxic Air Pollutants” (*J. Royal Stat. Soc.*, 1989, p. 183-198) sugere a distribuição lognormal com um modelo de concentração de  $\text{SO}_2$  sobre uma floresta específica. Suponha que os valores dos parâmetros sejam  $\mu = 1,9$  e  $\sigma = 0,9$ .
- Qual é o valor médio e o desvio padrão da concentração?
  - Qual é a probabilidade de a concentração ser no máximo 10? Entre 5 e 10?
77. Que condição de  $\alpha$  e  $\beta$  é necessária para que a fdp beta-padrão seja simétrica?
78. Suponha que a proporção  $X$  da área coberta por uma determinada planta em um quadrante selecionado aleatoriamente possui distribuição beta-padrão com  $\alpha = 5$  e  $\beta = 2$ .
- Calcule  $E(X)$  e  $V(X)$ .
  - Calcule  $P(X \leq 0,2)$ .
  - Calcule  $P(0,2 \leq X \leq 0,4)$ .
  - Qual é a proporção esperada da região de amostragem não coberta pela planta?
79. Assuma que  $X$  tenha densidade beta-padrão com parâmetros  $\alpha$  e  $\beta$ .

- a. Demonstre a fórmula de  $E(X)$  fornecida na seção.
- b. Calcule  $E[(1 - X)^m]$ . Se  $X$  representa a proporção de uma substância formada por um determinado ingrediente, qual é a proporção esperada que não é formada por ele?
80. É aplicada uma força em uma barra de aço de 20 pol. presa nas duas extremidades. Seja  $Y =$  distância da quebra da barra em relação à extremidade esquerda. Suponha que  $Y/20$  possui uma distribuição beta com  $E(Y) = 10$  e  $V(Y) = \frac{100}{7}$ .
- a. Quais são os parâmetros da distribuição beta-padrão correspondente?
- b. Calcule  $P(8 \leq Y \leq 12)$ .
- c. Calcule a probabilidade de a barra quebrar a mais de 2 pol. de onde você espera.

## 4.6 | Gráficos de Probabilidade

Um investigador normalmente obtém uma amostra numérica  $x_1, x_2, \dots, x_n$  e deseja saber se é plausível ela ter vindo de uma distribuição de população de um tipo específico (por exemplo, de uma distribuição normal). Muitos procedimentos formais de inferência estatística se baseiam na hipótese de que a distribuição da população é de um tipo especificado. O uso de tal procedimento não é apropriado se a distribuição de probabilidades subjacente for muito diferente do tipo assumido. Além disso, a compreensão da distribuição subjacente algumas vezes pode fornecer percepções dos mecanismos físicos envolvidos na geração dos dados. Uma forma eficiente de verificar uma hipótese de distribuição é construir o que chamamos de **gráfico de probabilidade**. A essência desse gráfico é que, se a distribuição na qual ele se baseia estiver correta, seus pontos estarão próximos a uma linha reta. Se a distribuição real for muito diferente da usada para criar o gráfico, os pontos diferem substancialmente do padrão linear.

### Percentis da Amostra

Os detalhes envolvidos na construção de gráficos de probabilidade são um pouco diferentes de fonte para fonte. A base da nossa construção é a comparação entre os percentis dos dados da amostra e os percentis correspondentes da distribuição consideradas. Lembre-se que o  $(100p)$ ésimo percentil de uma distribuição contínua com fdc  $F(\cdot)$  é o número  $\eta(p)$  que satisfaz  $F(\eta(p)) = p$ . Isto é,  $\eta(p)$  é o número na escala de medida tal que a área sob a curva de densidade à esquerda de  $\eta(p)$  é  $p$ . Assim, o 50º percentil  $\eta(0,5)$  satisfaz  $F(\eta(0,5)) = 0,5$ , e o 90º percentil satisfaz  $F(\eta(0,9)) = 0,9$ . Considere como exemplo a distribuição-padrão normal, para a qual representamos a fdc por  $\Phi(\cdot)$ . Na Tabela A.3 do Apêndice, encontramos o 20º percentil localizando a linha e a coluna em que 0,2000 (ou um número o mais próximo dele) aparece na tabela. Como 0,2005 aparece na interseção da linha -0,8 e da coluna 0,04, o 20º percentil é aproximadamente -0,84. De forma similar, o 25º percentil da distribuição normal-padrão é (usando interpolação linear) aproximadamente -0,675.

*Grosso modo*, os percentis da amostra são definidos da mesma forma que os da distribuição da população. O 50º percentil da amostra deve separar os 50% menores valores dos 50% maiores da amostra, o 90º percentil deve ser o limite entre 90% abaixo do valor e 10% acima, e assim por diante. Infelizmente encontramos problemas quando tentamos calcular os percentis da amostra para uma amostra determinada de  $n$  observações. Se, por exemplo,  $n = 10$ , podemos dividir 20% desses valores ou 30% dos dados, mas nenhum valor dividirá exatamente 23% das 10 observações. Para nos aprofundarmos, precisamos de uma definição operacional de percentis da amostra (local em que diferentes pessoas fazem coisas ligeiramente diferentes). Lembre-se de que, quando  $n$  é ímpar, a mediana da amostra, ou o 50º percentil, é o valor do meio na lista ordenada, por exemplo, o sexto maior valor quando  $n = 11$ . Isto significa considerar a observação do centro localizando-se metade na parte inferior dos dados e metade na superior. De forma similar, suponha  $n = 10$ . Então, se denominarmos o terceiro menor valor como 25º percentil, estaremos considerando metade desse valor o grupo inferior (formado pelas duas menores observações) e metade no grupo superior (as sete maiores observações). Tal consideração leva à seguinte definição geral de percentis da amostra:

## DEFINIÇÃO

Ordene as  $n$  observações da amostra da menor para a maior. A menor observação da lista é então definida como o  $[100(i - 0,5)/n]^{\circ}$  percentil da amostra.

Uma vez calculados os valores percentuais de  $100(i - 0,5)/n$  ( $i = 1, 2, \dots, n$ ), os percentis da amostra correspondentes aos percentuais intermediários são obtidos por interpolação linear. Por exemplo: se  $n = 10$ , os percentuais correspondentes às observações ordenadas serão  $100(1 - 0,5)/10 = 5\%$ ,  $100(2 - 0,5)/10 = 15\%$ ,  $25\%$ , ..., e  $100(10 - 0,5)/10 = 95\%$ . O 10º percentil estará a meio caminho entre o 5º (menor observação da amostra) e o 15º percentil (segunda menor observação da amostra). Para os nossos propósitos, tal interpolação não será necessária porque o gráfico de probabilidade se baseará somente nos percentuais  $100(i - 0,5)/n$  correspondentes às  $n$  observações da amostra.

### Um Gráfico de Probabilidade

Suponha agora que para os percentuais  $100(i - 0,5)/n$  ( $i = 1, \dots, n$ ) os percentis sejam determinados para uma distribuição de população específica cuja plausibilidade está sendo investigada. Se a amostra foi realmente selecionada na distribuição especificada, os percentis da amostra (observações ordenadas da amostra) devem estar relativamente perto dos percentis correspondentes da distribuição da população. Isso é, para  $i = 1, 2, \dots, n$  deve haver concordância razoável entre a  $i$ -ésima menor observação da amostra e o  $[100(i - 0,5)/n]$ ésimo percentil da distribuição especificada. Considere os pares (percentil da população, percentil da amostra), isso é, os pares

$$\left( \begin{array}{l} [100(i - 0,5)/n]\text{ésimo percentil} \\ \text{da distribuição} \end{array} , \begin{array}{l} i\text{-ésima menor observação} \\ \text{da amostra} \end{array} \right)$$

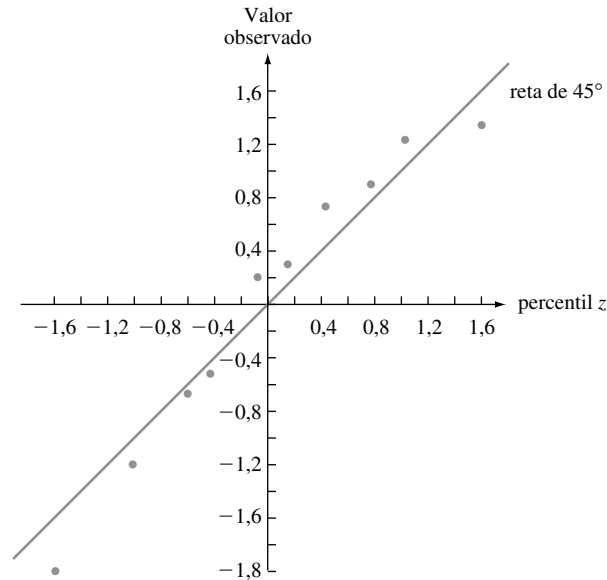
para  $i = 1, \dots, n$ . Cada par desse tipo é locado como um ponto em um sistema de coordenadas bidimensional. Se os percentis da amostra estiverem próximos aos percentis correspondentes da distribuição da população, o primeiro número de cada par será aproximadamente igual ao segundo. Os pontos do gráfico estão próximos a uma reta de  $45^\circ$ . Desvios substanciais dos pontos plotados em relação à reta de  $45^\circ$  geram dúvidas sobre a hipótese de a distribuição considerada ser a correta.

#### Exemplo 4.28

O valor de uma constante física é conhecido de um experimentador. Ele executa  $n = 10$  medidas independentes desse valor, usando certo dispositivo de medida, e registra os erros de medida resultantes (erro = valor observado - valor real). Essas observações são exibidas na tabela a seguir.

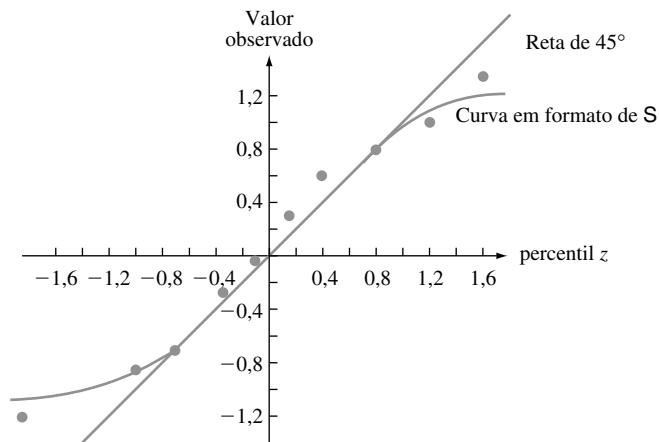
Percentagem	5	15	25	35	45
percentil $z$	-1,645	-1,037	-0,675	-0,385	-0,126
Observação da amostra	-1,91	-1,25	-0,75	-0,53	0,20
Percentagem	55	65	75	85	95
percentil $z$	0,126	0,385	0,675	1,037	1,645
Observação da amostra	0,35	0,72	0,87	1,40	1,56

É plausível que a variável aleatória *erro de medida* tenha uma distribuição normal padrão? Os percentis ( $z$ ) normais padronizados necessários também são exibidos na tabela. Portanto, os pontos no gráfico de probabilidade são  $(-1,645; -1,91)$ ,  $(-1,037; -1,25)$ , ... e  $(1,645; 1,56)$ . A Figura 4.31 (página 191) mostra o gráfico resultante. Apesar de os pontos se desviarem um pouco da reta de  $45^\circ$ , a impressão predominante é que a reta ajusta muito bem os pontos. O gráfico sugere que a distribuição normal padrão seja um modelo probabilístico razoável para o erro de medida.



**Figura 4.31** Gráfico de pares (percentil z, valor observado) dos dados do Exemplo 4.28: primeira amostra

A Figura 4.32 mostra um gráfico de pares (percentil z, observação) de uma segunda amostra de 10 observações. A reta de 45° fornece bom ajuste para a parte do meio da amostra, mas não para os extremos. O gráfico tem uma aparência bem-definida de formato S.



**Figura 4.32** Gráfico de pares (percentil z, valor observado) dos dados do Exemplo 4.28: segunda amostra

As duas menores observações da amostra são consideravelmente maiores que os percentis z correspondentes (os pontos na extremidade esquerda do gráfico estão muito acima da reta de 45°). De forma similar, as duas maiores observações são muito menores do que os percentis z associados. Esse gráfico indica que a distribuição normal padronizada não seria uma escolha plausível para o modelo de probabilidade que propiciou o aparecimento desses erros de medida observados. ■

Normalmente um investigador não está interessado em saber apenas se uma distribuição probabilística especificada, tal como a distribuição normal padrão (normal com  $\mu = 0$  e  $\sigma = 1$ ) ou a distribuição exponencial com  $\lambda = 0,1$ , é um modelo plausível para a distribuição da população da qual a amostra foi selecionada. Em vez disso, o investigador vai querer saber se algum membro de uma família de distribuições de probabilidade especifica um modelo plausível — a família de distribuições normais, a família de distribuições exponenciais, a família de distribuições de Weibull e assim por diante. Os valores dos parâmetros de uma distribuição

normalmente não são especificados de início. Se a família de distribuições de Weibull for considerada como modelo para os dados de vida útil, a questão será se há *quaisquer* valores dos parâmetros  $\alpha$  e  $\beta$  para os quais a distribuição de Weibull correspondentes fornece bom ajuste dos dados. Felizmente, quase sempre ocorre de apenas um gráfico de probabilidade ser suficiente para avaliar a plausibilidade de uma família. Se o gráfico se desviar demais de uma reta, nenhum membro da família pode ser usado. Quando o gráfico é bastante reto, é exigido mais trabalho para estimar os valores dos parâmetros (por exemplo, encontrar valores para  $\mu$  e  $\sigma$ ) que forneçam a melhor distribuição para o tipo especificado.

Enfoquemos a verificação de normalidade de um gráfico. Esse gráfico pode ser bastante útil no trabalho aplicado porque diversos processos estatísticos formais são apropriados (fornecem inferências precisas) apenas quando a distribuição da população é pelo menos aproximadamente normal. Tais procedimentos normalmente não devem ser usados se o gráfico de probabilidade normal mostrar um desvio muito pronunciado da linearidade. A chave para se construir um gráfico de probabilidade normal abrangente é a relação entre os percentis ( $z$ ) normais padronizados e os de outra distribuição:

$$\text{percentil de uma distribuição } (\mu, \sigma) \text{ normal} = \mu + \sigma \cdot (\text{percentil } z \text{ correspondente})$$

Considere primeiro o caso  $\mu = 0$ . Então, se cada observação for exatamente igual ao percentil normal correspondente para um valor particular de  $\sigma$ , os pares ( $\sigma \cdot [\text{percentil } z]$ , observação) estarão sobre uma reta de  $45^\circ$ , que tem coeficiente angular 1. Tal fato implica que os pares (percentil  $z$ , observação) localizam-se em uma reta que passa por (0, 0) (isto é, com intercepto  $y$  em 0) mas com coeficiente angular  $\sigma$  em vez de 1. O efeito de um valor de  $\mu$  diferente de zero é simplesmente mudar o intercepto  $y$  de 0 para  $\mu$ .

Um gráfico dos  $n$  pares

( $[100(i - 0,5)/n]$ ésimo percentil  $z$ ,  $i$ -ésima menor observação)

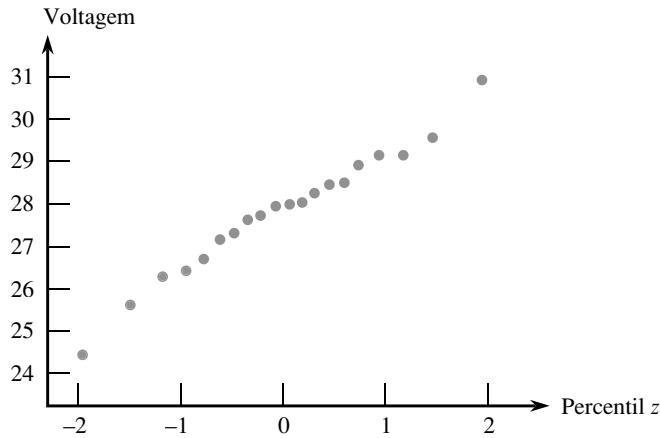
em um sistema de coordenadas bidimensional é denominado **gráfico de probabilidade normal**. Se as observações da amostra forem de fato obtidas a partir de uma distribuição normal com valor da média  $\mu$  e desvio padrão  $\sigma$ , os pontos devem estar próximos de uma linha reta com coeficiente angular  $\sigma$  e cruzando o eixo em  $\mu$ . Portanto, um gráfico em que os pontos estão próximos a uma linha reta sugere que a suposição de uma distribuição de população normal é plausível.

#### Exemplo 4.29

A amostra a seguir, formada por  $n = 20$  observações sobre a voltagem de corte dielétrica de uma resina epóxi foi exibida no artigo “Maximum Likelihood Estimation in the 3-Parameter Weibull Distribution (*IEEE Trans. on Dielectrics and Elec. Insul.*, 1996, p. 43-55). Os valores de  $(i - 0,5)/n$  para os quais os percentis  $z$  são obrigatórios são  $(1 - 0,5)/20 = 0,025$ ,  $(2 - 0,5)/20 = 0,075$ ,..., e  $0,975$ .

Observação	24,46	25,61	26,25	26,42	26,66	27,15	27,31	27,54	27,74	27,94
Percentil $z$	-1,96	-1,44	-1,15	-0,93	-0,76	-0,60	-0,45	-0,32	-0,19	-0,06
Observação	27,98	28,04	28,28	28,49	28,50	28,87	29,11	29,13	29,50	30,88
Percentil $z$	0,06	0,19	0,32	0,45	0,60	0,76	0,93	1,15	1,44	1,96

A Figura 4.33 mostra o gráfico de probabilidade normal resultante. O padrão no gráfico é bastante reto, indicando que é plausível a distribuição da população de voltagens de corte dielétrica ser normal.



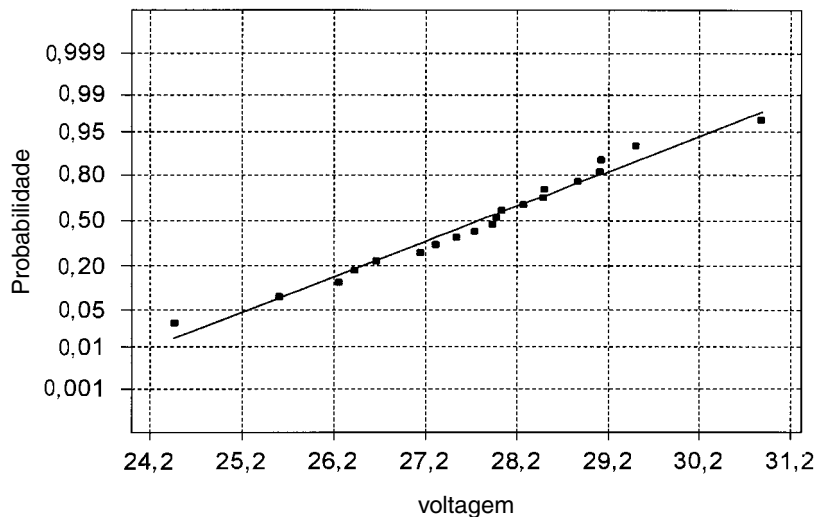
**Figura 4.33** Gráfico de probabilidade normal da amostra de voltagem de dielétrica de corte

Há uma versão alternativa do gráfico de probabilidade normal em que o eixo dos percentis  $z$  é substituído por um eixo de probabilidade não-linear. A escala dos eixos é construída de forma que os pontos devam novamente estar próximos a uma reta quando a distribuição da amostra for normal. A Figura 4.34 a seguir mostra um gráfico do tipo em MINITAB, para os dados de voltagem de corte do Exemplo 4.29.

Uma distribuição de população não-normal freqüentemente pode se enquadrar em uma das três categorias a seguir:

1. É simétrica e tem “caudas mais curtas” do que uma distribuição normal, isto é, a curva de densidade decresce mais rapidamente nas caudas do que a curva normal.
2. É simétrica e com cauda longa, em comparação à da distribuição normal.
3. É inclinada.

Uma distribuição uniforme tem cauda curta, visto que sua função de densidade decresce a zero fora de um intervalo finito. A função de densidade  $f(x) = 1/[\pi(1 + x^2)]$  para  $-\infty < x < \infty$  é um exemplo de distribuição de cauda longa, pois  $1/(1 + x^2)$  decresce muito menos rapidamente que  $e^{-x^2/2}$ . As distribuições lognormal e de Weibull estão entre as que possuem inclinação. Quando os pontos em um gráfico de probabilidade normal não tendem para a reta, o padrão freqüentemente sugere que a distribuição da população esteja em uma dessas três categorias.



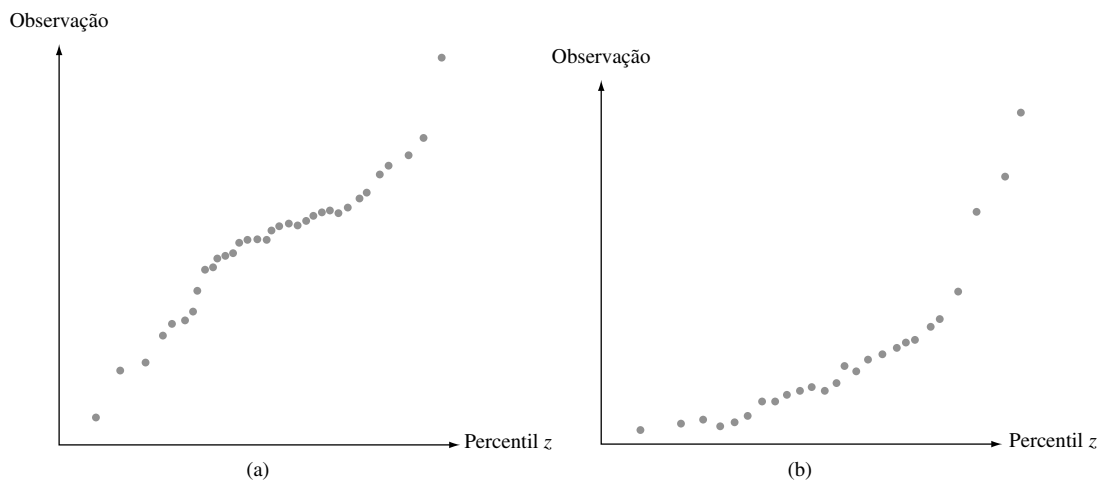
**Figura 4.34** Gráfico de probabilidade normal dos dados de voltagem de corte em MINITAB



Quando a distribuição da qual a amostra foi selecionada for de cauda curta, a menor e a maior observações normalmente não serão tão extremas quanto seria esperado de uma amostra aleatória normal. Visualize uma reta desenhada no centro do gráfico: os pontos na extrema direita tendem a estar abaixo da reta (valor observado < percentil  $z$ ), enquanto os pontos na extrema esquerda do gráfico tendem a estar acima da reta (valor observado > percentil  $z$ ). O resultado é um padrão em forma de S, como na Figura 4.32.

Uma amostra de distribuição de cauda longa também tende a produzir um gráfico com formato de S. Entretanto, em contraste com o caso da distribuição de cauda curta, a extremidade esquerda do gráfico se curva para baixo (observado < percentil  $z$ ), conforme mostrado na Figura 4.35(a) a seguir. Se a distribuição subjacente tiver inclinação positiva (uma cauda curta à esquerda e uma cauda longa à direita), as menores observações amostrais serão maiores que as esperadas de uma amostra normal, acontecendo o mesmo com as maiores observações. Nesse caso, os pontos em ambas as extremidades do gráfico estarão acima de uma reta que passa pela parte central, resultando um padrão curvo, conforme ilustrado na Figura 4.35(b). Uma amostra de distribuição lognormal geralmente produz tal padrão. Um gráfico de pares (percentil  $z$ ,  $\ln(x)$ ) deve então ter a aparência de uma reta.

Mesmo quando a distribuição da população é normal, os percentis da amostra não coincidirão exatamente com os percentis teóricos por causa da variabilidade da amostragem. Quanto os pontos do gráfico de probabilidade podem se desviar do padrão da reta antes que a hipótese de normalidade da população não seja mais plausível? A resposta a essa questão não é fácil. *Grosso modo*, uma pequena amostra de distribuição normal tem maior probabilidade de gerar um gráfico com padrão não-linear do que uma amostra grande. O livro *Fitting Equations to Data* (veja a bibliografia do Capítulo 13) apresenta os resultados de um estudo de simulação em que diversas amostras de diferentes tamanhos foram selecionadas de distribuições normais.



**Figura 4.35** Gráficos de probabilidade que sugerem distribuição não-normal: (a) um gráfico consistente com distribuição de cauda longa; (b) um gráfico consistente com distribuição de inclinação positiva

Os autores concluíram que normalmente há maior variação na aparência do gráfico de probabilidade para tamanhos de amostra inferiores a 30 e o padrão linear geralmente predomina apenas para tamanhos de amostras muito maiores. Quando um gráfico baseia-se em uma amostra de tamanho pequeno, apenas um desvio muito substancial pode ser usado como evidência conclusiva de não-normalidade. Comentário similar se aplica a gráficos de probabilidade para verificação de plausibilidade de outros tipos de distribuições.

## Além da Normalidade

Considere uma família de distribuições de probabilidade envolvendo dois parâmetros,  $\theta_1$  e  $\theta_2$ , e represente por  $F(x; \theta_1, \theta_2)$  as fdc correspondentes. A família de distribuições normais é uma delas, com  $\theta_1 = \mu$ ,  $\theta_2 = \sigma$ , e  $F(x; \mu, \sigma) = \Phi[(x - \mu)/\sigma]$ . Outro exemplo é a família de Weibull, com  $\theta_1 = \alpha$ ,  $\theta_2 = \beta$ , e

$$F(x; \alpha, \beta) = 1 - e^{-(x/\beta)^\alpha}$$

Outra família desse tipo é a família gama, para a qual a fdc é uma integral que envolve a função gama incompleta que não pode ser expressa de nenhuma forma mais simples.

Os parâmetros  $\theta_1$  e  $\theta_2$  são denominados parâmetros de **localização** e parâmetros de **escala**, respectivamente, se  $F(x; \theta_1, \theta_2)$  for uma função de  $(x - \theta_1)/\theta_2$ . Os parâmetros  $\mu$  e  $\sigma$  da família normal são parâmetros de localização e parâmetros de escala, respectivamente. A alteração de  $\mu$  desloca a localização da curva de densidade em formato de sino para a direita ou para a esquerda, e a alteração de  $\sigma$  amplia ou comprime a escala de medida (a escala do eixo horizontal quando a função densidade for plotada). Outro exemplo é dado pela fdc

$$F(x; \theta_1, \theta_2) = 1 - e^{-e^{(x-\theta_1)/\theta_2}} \quad -\infty < x < \infty$$

Diz-se que uma variável aleatória com essa fdc possui *distribuição de valor extremo*. É usada em aplicações que envolvem a vida útil de componente e a resistência do material.

Apesar de a fdc do valor extremo sugerir à primeira vista que  $\theta_1$  é o ponto de simetria da função de densidade e, portanto, ser a média e a mediana, esse não é o caso. Pelo contrário,  $P(X \leq \theta_1) = F(\theta_1; \theta_1, \theta_2) = 1 - e^{-1} = 0,632$ , e a função de densidade  $f(x; \theta_1, \theta_2) = F'(x; \theta_1, \theta_2)$  tem inclinação negativa (cauda longa mais baixa). De forma similar, o parâmetro da escala  $\theta_2$  não é o desvio padrão ( $\mu = \theta_1 - 0,5772\theta_2$  e  $\sigma = 1,283\theta_2$ ). Entretanto, a mudança do valor de  $\theta_1$  altera a localização da curva de densidade, enquanto a alteração de  $\theta_2$  redefine o eixo de medidas.

O parâmetro  $\beta$  da distribuição de Weibull é um parâmetro de escala, mas  $\alpha$  não é de localização. O parâmetro  $\alpha$  normalmente é mencionado como **parâmetro de formato**. Um comentário similar se aplica aos parâmetros  $\alpha$  e  $\beta$  da distribuição gama. Na forma usual, a função de densidade de qualquer membro da distribuição gama ou de Weibull é positiva para  $x \Rightarrow 0$  e zero em caso contrário. Um parâmetro de localização pode ser introduzido como um terceiro parâmetro  $\gamma$  (fizemos isso para a distribuição de Weibull) para deslocar a função de densidade de forma que seja positiva se  $x \Rightarrow \gamma$  e zero em caso contrário.

Quando a família em consideração possui apenas parâmetros de escala e de localização, a questão de algum membro da família ser ou não uma distribuição de população plausível pode ser resolvida por meio de um único gráfico de probabilidade de fácil construção. Primeiro são obtidos os percentis da distribuição-padrão, aquela com  $\theta_1 = 0$  e  $\theta_2 = 1$ , para porcentagens  $100(i - 0,5)/n$  ( $i = 1, \dots, n$ ). Os  $n$  (percentil padronizado, observação) pares fornecem os pontos do gráfico. Isso foi exatamente o que fizemos para obter um gráfico de probabilidade normal abrangente. Surpreendentemente, essa metodologia pode ser aplicada para obtenção de um gráfico de probabilidade de Weibull. O resultado-chave é que, se  $X$  tiver distribuição de Weibull com parâmetro de formato  $\alpha$  e parâmetro de escala  $\beta$ , a variável transformada  $\ln(X)$  tem uma distribuição de valor extremo com parâmetro de localização  $\theta_1 = \ln(\beta)$  e parâmetro de escala  $\alpha$ . Assim, um gráfico dos pares (percentil do valor extremo padronizado,  $\ln(x)$ ) que mostre padrão linear forte fornece suporte para a escolha da distribuição Weibull como modelo de população.

**Exemplo 4.30**

As observações a seguir são relativas à vida útil (em horas) do isolamento de aparelho elétrico quando a aceleração da tensão térmica e elétrica são mantidas com certos valores (“On the Estimation of Life of Power Apparatus Insulation Under Combined Electrical and Thermal Stress,” *IEEE Trans. on Electrical Insulation*, 1985, p. 70-78). O gráfico de probabilidade de Weibull exige que se calcule primeiro o 5º, 15º, ésimio ... e 95º percentis da distribuição de valor extremo padrão. O  $(100p)^{\alpha}$  percentil  $\eta(p)$  satisfaz

$$p = F(\eta(p)) = 1 - e^{-e^{\eta(p)}}$$

de onde  $\eta(p) = \ln[-\ln(1 - p)]$ .

<i>Percentil</i>	–2,97	–1,82	–1,25	–0,84	–0,51
<i>x</i>	282	501	741	851	1072
<i>ln(x)</i>	5,64	6,22	6,61	6,75	6,98
<i>Percentil</i>	–0,23	0,05	0,33	0,64	1,10
<i>x</i>	1122	1202	1585	1905	2138
<i>ln(x)</i>	7,02	7,09	7,37	7,55	7,67

Os pares  $(-2,97, 5,64)$ ,  $(-1,82, 6,22)$ , ...,  $(1,10, 7,67)$  estão plotados como pontos na Figura 4.36. A disposição dos pontos em reta é uma forte evidência de que se deve usar a distribuição de Weibull como modelo da vida útil do isolamento, conclusão a que o autor do artigo citado também chegou.

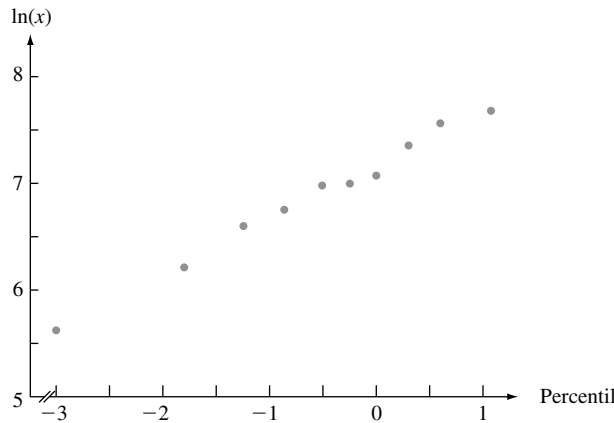
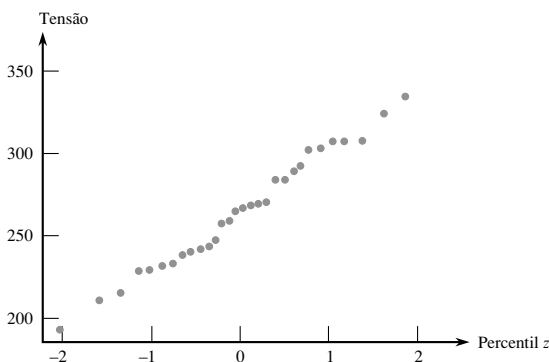


Figura 4.36 Gráfico de probabilidade de Weibull dos dados de vida útil do isolamento

A distribuição gama é um exemplo de uma família que envolve um parâmetro de formato para o qual não há transformação  $h(\cdot)$ , tal que  $h(X)$  tenha distribuição que depende apenas dos parâmetros de localização e escala. A construção de um gráfico de probabilidade requer primeiro a estimativa do parâmetro de formato pelos dados da amostra (alguns métodos de processamento são descritos no Capítulo 6). Às vezes, um investigador quer saber se a variável transformada  $X^\theta$  tem distribuição normal para algum valor de  $\theta$  (por convenção,  $\theta = 0$  é identificado com a transformação logarítmica, caso em que  $X$  possui distribuição lognormal). O livro *Graphical Methods for Data Analysis*, relacionado na bibliografia do Capítulo 1, aborda esse tipo de problema, bem como outras discussões dos gráficos de probabilidade.

**Exercícios** | Seção 4.6 (81–91)

81. O gráfico de probabilidade normal a seguir foi construído a partir de uma amostra de 30 leituras de tensão na malha de pontos de tubos de vídeo usados em monitores de computadores. Parece plausível que a distribuição da tensão seja normal?



82. Considere as 10 observações a seguir de vida útil de mancais (em horas):

152,7	172,0	172,5	173,3	193,0
204,7	216,5	234,9	262,6	422,6

Construa o gráfico de probabilidade normal e comente sobre a plausibilidade da distribuição normal como modelo de estimativa da vida útil de mancais (dados obtidos em “Modified Moment Estimation for the Three-Parameter Lognormal Distribution,” *J. Quality Technology*, 1985, p. 92 -99).

83. Construa um gráfico de probabilidade normal para a seguinte amostra de observações relativas à espessura da cobertura de tintas de baixa viscosidade (“Achieving a Target Value for a Manufacturing Process: A Case Study”, *J. of Quality Technology*, 1992, p. 22-26). Você se sentiria confortável ao estimar a média da espessura da população usando um método que assume uma distribuição de população normal?

0,83	0,88	0,88	1,04	1,09	1,12	1,29	1,31
1,48	1,49	1,59	1,62	1,65	1,71	1,76	1,83

84. O artigo “A Probabilistic Model of Fracture in Concrete and Size Effects on Fracture Toughness” (*Magazine of Concrete Res.*, 1996, p. 311-320) fornece argumentos que explicam por que a distribuição de dureza de fratura em corpos de prova de concreto tem uma distribuição de

Weibull, além de apresentar diversos histogramas de dados que parecem ajustados com curvas de Weibull superpostas. Considere a seguinte amostra de tamanho  $n = 18$  observações de dureza de concreto de alta resistência (consistente com um dos histogramas). Os valores de  $p_i = (i - 0,5)/18$  também são fornecidos.

Observação	0,47	0,58	0,65	0,69	0,72	0,74
$p_i$	0,0278	0,0833	0,1389	0,1944	0,2500	0,3056
Observação	0,77	0,79	0,80	0,81	0,82	0,84
$p_i$	0,3611	0,4167	0,4722	0,5278	0,5833	,6389
Observação	0,86	0,89	0,91	0,95	1,01	1,04
$p_i$	0,6944	0,7500	0,8056	0,8611	,9167	0,9722

Construa um gráfico de probabilidade de Weibull e comente.

85. Construa um gráfico de probabilidade normal para os dados de propagação de trincas por fadiga do Exercício 39 (Capítulo 1). Parece plausível que a vida da propagação tenha distribuição normal? Explique.

86. O artigo “The Load-Life Relationship for M50 Bearings with Silicon Nitride Ceramic Balls” (*Lubrication Engr.*, 1984, p. 153-159) relata os dados a seguir sobre a vida de carga de mancais (milhões revs.) para mancais testados com cargas de 6,45 kN.

47,1	68,1	68,1	90,8	103,6	106,0	115,0
126,0	146,6	229,0	240,0	240,0	278,0	278,0
289,0	289,0	367,0	385,9	392,0	505,0	

- a. Construa um gráfico de probabilidade normal. A normalidade é plausível?
- b. Construa um gráfico de probabilidade de Weibull. A família de distribuições Weibull é plausível?

87. Construa um gráfico de probabilidade que permita avaliar a normalidade da distribuição lognormal como um modelo para os dados de precipitação do Exercício 81 (Capítulo 1).

88. As observações a seguir são valores de precipitação no mês de março de um período de 30 anos em Minneapolis-St. Paul.

0,77	1,20	3,00	1,62	2,81	2,48
1,74	0,47	3,09	1,31	1,87	0,96
0,81	1,43	1,51	0,32	1,18	1,89
1,20	3,37	2,10	0,59	1,35	0,90
1,95	2,20	0,52	0,81	4,75	2,05

- a. Construa e interprete um gráfico de distribuição normal para o conjunto de dados.
- b. Calcule a raiz quadrada de cada valor e então construa um gráfico de probabilidade normal com base nos dados transformados. Parece plausível que a raiz quadrada da precipitação tenha distribuição normal?
- c. Repita a parte (b) após transformar por raízes cúbicas.

89. Use um software estatístico para construir um gráfico de probabilidade normal dos dados do limite de resistência à tração fornecidos no Exercício 13 do Capítulo 1 e comente.

90. Sejam as observações ordenadas da amostra  $y_1, y_2, \dots, y_n$  ( $y_1$  sendo o menor e  $y_n$  o maior). Sugerimos, para verificação de normalidade, que os pares  $(\Phi^{-1}((i - 0,5)/n), y_i)$  sejam plotados. Suponha que acreditemos que as observações são provenientes de uma distribuição com média 0 e assumamos  $w_1, \dots, w_n$  como os valores absolutos ordenados dos  $x_i$ s. Um gráfico **seminormal** é um gráfico de probabilidade dos  $w_i$ s. Mais especificamente, como  $P(|Z| \leq w) = P(-w \leq Z \leq w) = 2\Phi(w) - 1$ , um gráfico seminormal é um gráfico dos pares  $(\Phi^{-1}((i - 0,5)/n) + 1)/2, w_i)$ . A virtude desse gráfico é que os pequenos ou grandes *outliers* na amostra original serão exibidos apenas na extremidade superior e não nas duas extremidades. Construa um gráfico seminormal para a seguinte amostra de erros de medida e comente: -3,78, -1,27, 1,44, -0,39, 12,38, -43,40, 1,15, -3,96, -2,34, 30,84.

91. As seguintes observações de tempos de falha (milhares de horas) são resultantes de testes de vida aceleradas de 16 chips de certo tipo de circuito integrado:

82,8	11,6	359,5	502,5	307,8	179,7
242,0	26,5	244,8	304,3	379,1	212,6
229,9	558,9	366,7	204,6		

Use os percentis correspondentes da distribuição exponencial com  $\lambda = 1$  para construir um gráfico de probabilidade. Explique, então, por que o gráfico avalia a plausibilidade da amostra gerada a partir de qualquer distribuição exponencial.

Exercícios Suplementares (92–120)

92. Seja  $X =$  tempo que leva para uma cabeça de leitura/gravação encontrar um registro específico em uma memória em disco de um computador, após ter sido posicionada sobre a faixa correta. Se os discos executarem uma rotação em cada 25 milissegundos, uma hipótese razoável é que  $X$  seja distribuído uniformemente no intervalo  $[0, 25]$ .

- a. Calcule  $P(10 \leq X \leq 20)$ .
- b. Calcule  $P(X \geq 10)$ .
- c. Deduza a fdc  $F(X)$ .
- d. Calcule  $E(X)$  e  $\sigma_X$ .

93. Uma barra de 12 pol. fixa nas duas extremidades será submetida a uma tensão crescente até sua ruptura.

Seja  $Y$  = distância da ruptura da barra em relação à extremidade esquerda. Suponha que  $Y$  tenha fdp

$$f(y) = \begin{cases} \left(\frac{1}{24}\right)y\left(1 - \frac{y}{12}\right) & 0 \leq y \leq 12 \\ 0 & \text{caso contrário} \end{cases}$$

Calcule os dados a seguir:

- A fdc de  $Y$ , e faça o gráfico.
  - $P(Y \leq 4)$ ,  $P(Y > 6)$ , e  $P(4 \leq Y \leq 6)$
  - $E(Y)$ ,  $E(Y^2)$  e  $V(Y)$
  - A probabilidade de o ponto de ruptura ocorrer a mais de 2 pol. do ponto de ruptura esperado.
  - O comprimento esperado do menor segmento após a ruptura.
94. Seja  $X$  o tempo até a falha (em anos) de certo componente hidráulico. Suponha que a fdp de  $X$  seja  $f(x) = 32/(x + 4)^3$  para  $x > 0$ .
- Demonstre que se  $f(x)$  é uma fdp legítima.
  - Determine a fdc.
  - Use o resultado da parte (b) para calcular a probabilidade de o tempo até a falha estar entre 2 e 5 anos.
  - Qual é o tempo esperado até a falha?
  - Se o componente tiver valor residual de  $100/(4 + x)$ , quando o tempo até a falha for  $x$ , qual será o valor residual esperado?

95. O tempo de conclusão  $X$  de uma determinada tarefa possui fdc  $F(x)$  dada por

$$\begin{cases} 0 & x < 0 \\ \frac{x^3}{3} & 0 \leq x < 1 \\ 1 - \frac{1}{2}\left(\frac{7}{3} - x\right)\left(\frac{7}{4} - \frac{3}{4}x\right) & 1 \leq x \leq \frac{7}{3} \\ 1 & x \geq \frac{7}{3} \end{cases}$$

- Obtenha a fdp  $f(x)$  e faça o gráfico.
  - Calcule  $P(0,5 \leq X \leq 2)$ .
  - Calcule  $E(X)$ .
96. A tensão de ruptura de um determinado tipo de diodo escolhido aleatoriamente possui distribuição normal com valor de média 40 V e desvio padrão 1,5 V.
- Qual é a probabilidade de a voltagem de um diodo estar entre 39 e 42?
  - Que valor faz com que apenas 15% dos diodos tenham voltagens que excedam aquele valor?
  - Se quatro diodos forem selecionados de forma independente, qual será a probabilidade de ao menos um ter uma voltagem maior que 42?
97. O artigo "Computer Assisted Net Weight Control" (*Quality Progress*, 1983, p. 22-25) sugere uma distribuição normal com média 137,2 onças e desvio padrão 1,6 onça para o conteúdo real de um tipo específico de jarras. O conteúdo nominal é de 135 onças.
- Qual é a probabilidade de uma única jarra conter mais do que o conteúdo nominal?

- Entre 10 jarras selecionadas aleatoriamente, qual é a probabilidade de ao menos oito conterem mais do que o conteúdo nominal?
- Assumindo que a média permaneça em 137,2, para que valor o desvio padrão teria de ser alterado para que 95% de todas as jarras contenham mais do que o conteúdo nominal?

98. Quando as placas de circuito usadas na fabricação de CD-players são testadas, a porcentagem de peças com defeito no longo prazo é 5%. Suponha que um lote de 250 placas tenha sido recebido e que a condição de uma placa é independente das outras.
- Qual é a probabilidade aproximada de que ao menos 10% das placas do lote tenham defeito?
  - Qual é a probabilidade aproximada de que exatamente 10 peças tenham defeito no lote?
99. O artigo "Characterization of Room Temperature Damping in Aluminum-Indium Alloys" (*Metallurgical Trans.*, 1993, p. 1611-1619) sugere que o tamanho de grão de matriz Al ( $\mu\text{m}$ ) de uma liga formada por 2% de índio pode ser modelado com uma distribuição normal de valor médio 96 e desvio padrão 14.
- Qual é a probabilidade de o tamanho do grão exceder 100?
  - Qual é a probabilidade de o tamanho do grão estar entre 50 e 80?
  - Que intervalo  $(a, b)$  inclui os 90% centrais de todos os tamanhos de grãos (de forma que 5% estejam abaixo de  $a$  e 5% acima de  $b$ )?

100. O tempo de reação (em segundos) a um determinado estímulo é uma variável aleatória contínua com fdp

$$f(x) = \begin{cases} \frac{3}{2} \cdot \frac{1}{x^2} & 1 \leq x \leq 3 \\ 0 & \text{caso contrário} \end{cases}$$

- Deduza a fdc.
  - Qual é a probabilidade de o tempo de reação ser no máximo 2,5 segundos? Entre 1,5 e 2,5 segundos?
  - Calcule o tempo de reação esperado.
  - Calcule o desvio padrão do tempo de reação.
  - Se um indivíduo levar mais de 1,5 segundo para reagir, uma luz se acende e fica acesa até que mais um segundo passe ou até a pessoa reagir (o que ocorrer primeiro). Determine o tempo esperado de luz acesa. [Sugestão: Seja  $h(X)$  = tempo em que a luz fica acesa como função do tempo de reação  $X$ .]
101. Seja  $X$  a temperatura em que uma reação química acontece. Suponha que a fdp de  $X$  seja

$$f(x) = \begin{cases} \frac{1}{9}(4 - x^2) & -1 \leq x \leq 2 \\ 0 & \text{caso contrário} \end{cases}$$

- Faça o gráfico de  $f(x)$ .
- Determine a fdc e faça o gráfico.

- c. 0 é a temperatura mediana em que ocorre a reação? Caso não seja, essa temperatura mediana é maior ou menor que 0?
- d. Suponha que a reação ocorra de forma independente, uma vez em cada um de 10 diferentes laboratórios e que a fdp do tempo de reação em cada laboratório seja fornecida. Seja  $Y$  = número entre os 10 laboratórios em que a temperatura excede 1. Qual é o tipo de distribuição de  $Y$ ? (Forneça o nome e os valores dos parâmetros.)

- 102.** O artigo “Determination of the MTF of Positive Photoresists Using the Monte Carlo Method” (*Photographic Sci. and Engr.*, 1983, p. 254-260) propõe uma distribuição exponencial com parâmetro  $\lambda = 0,93$  como modelo de distribuição do comprimento da trajetória de um fóton ( $\mu\text{m}$ ) sob determinadas circunstâncias. Suponha que esse seja o modelo correto.
- a. Qual é o comprimento esperado da trajetória e qual é o desvio padrão?
  - b. Qual é a probabilidade de o comprimento da trajetória exceder 3,0? Qual é a probabilidade de o comprimento da trajetória estar entre 1,0 e 3,0?
  - c. Que valor é excedido por apenas 10% de todos os comprimentos de trajetórias?

- 103.** O artigo “The Prediction of Corrosion by Statistical Analysis of Corrosion Profiles” (*Corrosion Science*, 1985, p. 305-315) sugere a fdc a seguir para a profundidade  $X$  da cavidade mais profunda do experimento que envolve a exposição de aço carbono manganês à água do mar acidificada.

$$F(x; \alpha, \beta) = e^{-e^{-(x-\alpha)/\beta}} \quad -\infty < x < \infty$$

Os autores propõem os valores  $\alpha = 150$  e  $\beta = 90$ . Assuma esse modelo como o correto.

- a. Qual é a probabilidade de a profundidade da cavidade mais profunda ser no máximo 150? No máximo 300? Entre 150 e 300?
  - b. Abaixo de que valor a profundidade máxima do poço pode ser observada em 90% de todos os experimentos?
  - c. Qual é a função densidade de  $X$ ?
  - d. A função densidade pode ser mostrada como sendo unimodal (um único pico). Acima de que valor do eixo das medidas ocorre o pico? (Esse valor é a moda.)
  - e. Podemos demonstrar que  $E(X) \approx 0,5772\beta + \alpha$ . Qual é a média dos valores fornecidos de  $\alpha$  e  $\beta$ , e como ele se compara à mediana e à moda? Desenhe o gráfico da função densidade. (Nota: Ela é denominada distribuição de maior valor extremo.)
- 104.** Um componente tem uma vida útil  $X$  com distribuição exponencial de parâmetro  $\lambda$ .
- a. Se o custo operacional por unidade de tempo for  $c$ , qual será o custo operacional esperado desse componente durante sua vida útil?
  - b. Em vez de uma taxa de custo constante  $c$  como na parte (a), suponha que a taxa de custo seja  $c(1 - 0,5e^{ax})$  com  $a < 0$ , de forma que o custo por unidade de

tempo é menor do que  $c$  quando o componente é novo e fica mais caro à medida que o componente envelhece. Calcule agora o custo operacional esperado durante a vida útil do componente.

- 105.** A moda de uma distribuição contínua é o valor  $x^*$  que maximiza  $f(x)$ .
- a. Qual é a moda de uma distribuição normal com parâmetros  $\mu$  e  $\sigma$ ?
  - b. A distribuição uniforme com parâmetros  $A$  e  $B$  tem uma única moda? Por quê?
  - c. Qual é a moda de uma distribuição exponencial com parâmetro  $\lambda$ ? (Desenhe uma ilustração.)
  - d. Se  $X$  possui uma distribuição gama com parâmetros  $\beta$ , e  $\alpha > 1$ , determine a moda. [Sugestão:  $\ln[f(x)]$  será maximizada se, e somente se,  $f(x)$  o for, e é mais simples obter a derivada de  $\ln[f(x)]$ .]
  - e. Qual é a moda de uma distribuição qui quadrado com  $\nu$  graus de liberdade?
- 106.** O artigo “Error Distribution in Navigation” (*J. Institute of Navigation*, 1971, p. 429-442) sugere que a distribuição de frequência de erros positivos (magnitudes dos erros) é bem aproximada por uma distribuição exponencial. Seja  $X$  = o erro de posição lateral (milhas náuticas), que pode ser negativo ou positivo. Suponha que a fdp de  $X$  seja

$$f(x) = (0,1)e^{-0,2|x|} \quad -\infty < x < \infty$$

- a. Desenhe um gráfico de  $f(x)$  e demonstre que  $f(x)$  é uma fdp legítima (mostre que a integral é 1).
  - b. Deduza a fdc de  $X$  e faça o gráfico.
  - c. Calcule  $P(X \leq 0)$ ,  $P(X \leq 2)$ ,  $P(-1 \leq X \leq 2)$  e a probabilidade de se cometer um erro de mais de 2 milhas.
- 107.** Em alguns sistemas, um cliente é alocado a uma de duas instalações. Se o tempo de serviço de um cliente servido por uma instalação  $i$  tiver uma distribuição com parâmetros  $\lambda_i$  ( $i = 1, 2$ ) e  $p$  for a proporção de todos os clientes servidos pela instalação 1, a fdp de  $X$  = tempo de serviço de um cliente selecionado aleatoriamente será

$$f(x; \lambda_1, \lambda_2, p) = \begin{cases} p\lambda_1 e^{-\lambda_1 x} + (1-p)\lambda_2 e^{-\lambda_2 x} & x \geq 0 \\ 0 & \text{caso contrário} \end{cases}$$

Essa distribuição é freqüentemente denominada distribuição exponencial mista ou hiperexponencial. Ela também é proposta como modelo para quantidade de chuva em “Modeling Monsoon Affected Rainfall of Pakistan by Point Processes” (*J. Water Resources Planning and Mgmt.*, 1992, p. 671-688).

- a. Demonstre que  $f(x; \lambda_1, \lambda_2, p)$  realmente é uma fdp.
- b. Qual é a fdc  $F(x; \lambda_1, \lambda_2, p)$ ?
- c. Se  $X$  tiver  $f(x; \lambda_1, \lambda_2, p)$  como sua fdp, qual será  $E(X)$ ?
- d. Usando o fato de que  $E(X^2) = 2/\lambda^2$  quando  $X$  possui uma distribuição exponencial com parâmetro  $\lambda$ , calcule  $E(X^2)$  quando  $X$  tiver fdp  $f(x; \lambda_1, \lambda_2, p)$ . Depois, calcule  $V(X)$ .

- e. O coeficiente da variação de uma variável aleatória (ou distribuição) é  $CV = \sigma/\mu$ . Qual é o  $CV$  para uma  $va$  exponencial? O que você pode dizer sobre o valor  $CV$  quando  $X$  tem uma distribuição hiperexponencial?
- f. Qual é o  $CV$  de uma distribuição de Erlang com parâmetros  $\lambda$  e  $n$  como definido no Exercício 62? (Nota: No trabalho aplicado, o  $CV$  amostral é usado para decidir qual das três distribuições pode ser apropriada.)
108. Suponha que um estado específico permita que os indivíduos preencham restituições de imposto somente se o total das deduções em itens ultrapassar valores acima de US\$ 5000. Seja  $X$  (em milhares de dólares) o total das deduções em itens de um formulário escolhido aleatoriamente. Assuma que  $X$  tenha fdp

$$f(x; \alpha) = \begin{cases} k/x^\alpha & x \geq 5 \\ 0 & \text{caso contrário} \end{cases}$$

- a. Determine o valor de  $k$ . Que restrição é necessária para  $\alpha$ ?
- b. Qual é a fdc de  $X$ ?
- c. Qual é a dedução total esperada de um formulário escolhido aleatoriamente? Que restrição é necessária para  $\alpha$  tal que  $E(X)$  seja finito?
- d. Mostre que  $\ln(X/5)$  possui uma distribuição exponencial com parâmetro  $\alpha - 1$ .
109. Seja por  $I_i$  a corrente de entrada de um transistor e  $I_o$  a corrente de saída. O ganho de corrente então será proporcional a  $\ln(I_o/I_i)$ . Suponha que a constante de proporcionalidade seja 1 (que implica a escolha de uma unidade de medida específica), de forma que o ganho da corrente =  $X = \ln(I_o/I_i)$ . Assuma  $X$  como distribuição normal com  $\mu = 1$  e  $\sigma = 0,05$ .
- a. Que tipo de distribuição tem a razão  $I_o/I_i$ ?
- b. Qual é a probabilidade de a corrente de saída ser maior que o dobro da corrente de entrada?
- c. Qual é o valor esperado e a variância da razão da corrente de saída partir a de entrada?
110. O artigo "Response of SiC<sub>f</sub>/Si<sub>3</sub>N<sub>4</sub> Composites Under Static and Cyclic Loading — An Experimental and Statistical Analysis" (*J. of Engr. Materials and Technology*, 1997, p. 186-193) sugere que a resistência à tração (MPa) de compostos em condições específicas pode ser modelada por uma distribuição de Weibull com  $\alpha = 9$  e  $\beta = 180$ .
- a. Desenhe o gráfico da função de densidade.
- b. Qual é a probabilidade de a resistência de um espécime selecionado aleatoriamente exceder 175? Estar entre 150 e 175?
- c. Se dois espécimes selecionados aleatoriamente forem escolhidos de forma independente, qual será a probabilidade de ao menos um ter uma resistência entre 150 e 175?
- d. Que valor de resistência separa os 10% mais fracos de todos os espécimes dos 90% restantes?

111. Assuma  $Z$  com uma distribuição normal padrão e defina uma nova  $va$   $Y$  como  $Y = \sigma Z + \mu$ . Mostre que  $Y$  tem distribuição normal com parâmetros  $\mu$  e  $\sigma$ . (Sugestão:  $Y \leq y$  se, e somente se,  $Z \leq ?$  Use isso para determinar a fdc de  $Y$  e depois diferencie em relação a  $y$ .)
112. a. Suponha que a vida útil  $X$  de um componente, quando medido em horas, tenha uma distribuição gama com parâmetros  $\alpha$  e  $\beta$ . Seja  $Y =$  vida útil medida em minutos. Deduza a fdp de  $Y$ . (Sugestão:  $Y \leq y$  se e somente  $X \leq y/60$ . Use isso para obter a fdc de  $Y$  e depois diferencie para obter a fdp.)
- b. Se  $X$  tem distribuição gama com parâmetros  $\alpha$  e  $\beta$ , qual é a distribuição probabilidade  $Y = cX$ ?
113. Nos Exercícios 111 e 112, bem como em muitas outras situações, temos a fdp  $f(x)$  de  $X$  e queremos saber a fdp de  $Y = h(X)$ . Assuma que  $h(\cdot)$  é uma função inversível, de forma que  $y = h(x)$  pode ser resolvida para  $x$  para se obter  $x = k(y)$ . Podemos então mostrar que a fdp de  $Y$  é

$$g(y) = f[k(y)] \cdot |k'(y)|$$

- a. Se  $X$  tiver distribuição uniforme com  $A = 0$  e  $B = 1$ , encontre a fdp de  $Y = -\ln(X)$ .
- b. Resolva o Exercício 111 usando esse resultado.
- c. Resolva o Exercício 112 (b) usando esse resultado.
114. Com base nos dados de um experimento de lançamento de dardos, o artigo "Shooting Darts" (*Chance*, Summer, 1997, p. 16-19) propôs que os erros horizontais e verticais em relação a um alvo pontual devem ser independentes um do outro, cada um com distribuição normal de média 0 e variância  $\sigma^2$ . Podemos mostrar então que a fdp da distância  $V$  do alvo até o ponto de toque é
- $$f(v) = \frac{v}{\sigma^2} \cdot e^{-v^2/2\sigma^2} \quad v > 0$$
- a. Essa fdp é membro de que família apresentada neste capítulo?
- b. Se  $\sigma = 20$ mm (próximo ao valor sugerido no artigo), qual é a probabilidade de um dardo ser acertado dentro de 25mm (cerca de 1 pol.) do alvo?
115. O artigo "Three Sisters Give Birth on the Same Day" (*Chance*, Spring 2001, p. 23-25) usou o fato de três irmãs de Utah darem à luz em 11 de março de 1998 como base de proposta de algumas questões interessantes relativas a coincidências de nascimento.
- a. Desconsiderando anos bissextos e assumindo que os outros 365 dias são igualmente prováveis, qual é a probabilidade de que três nascimentos selecionados aleatoriamente ocorram todos em 11 de março? Indique quaisquer hipóteses existentes, se houver.
- b. Com as hipóteses usadas na parte (a), qual é a probabilidade de três nascimentos selecionados aleatoriamente ocorrerem no mesmo dia?
- c. O autor sugeriu que, com base em extensos dados, a duração da gestação (tempo entre a concepção e o nascimento) pode ser modelada como tendo

distribuição normal com valor médio 280 dias e desvio padrão 19,88 dias. As datas-limite para as três irmãs de Utah eram 15 de março, 1º de abril e 4 de abril, respectivamente. Supondo que o prazo esteja na média da distribuição, qual é a probabilidade de todos os nascimentos ocorrerem em 11 de março? (*Sugestão*: O desvio em relação à data de nascimento tem distribuição normal com média 0.)

- d. Explique como você usaria as informações da parte (c) para calcular a probabilidade de uma data de nascimento comum.

- 116.** Seja  $X$  a vida útil de um componente, com  $f(x)$  e  $F(x)$  sendo a fdp e a fdc de  $X$ . A probabilidade de o componente apresentar falha no intervalo  $(x, x + \Delta x)$  é aproximadamente  $f(x) \cdot \Delta x$ . A probabilidade condicional de o componente apresentar falha em  $(x, x + \Delta x)$ , dado que durou ao menos  $x$ , é  $f(x) \cdot \Delta x / [1 - F(x)]$ . Dividindo a expressão por  $\Delta x$  produz **o função de taxa de falha**:

$$r(x) = \frac{f(x)}{1 - F(x)}$$

Um aumento da função de taxa de falha indica que os componentes mais antigos têm mais probabilidade de apresentar desgaste, enquanto uma taxa de falha decrescente é evidência de aumento de confiabilidade com a idade. Na prática, uma falha “com formato de banheira” é assumida com frequência.

- a. Se  $X$  tiver distribuição exponencial, qual será  $r(x)$ ?
- b. Se  $X$  tiver distribuição de Weibull com parâmetros  $\alpha$  e  $\beta$ , qual será  $r(x)$ ? Para que valores de parâmetros  $r(x)$  aumentará? Para que valores de parâmetros  $r(x)$  diminuirá com  $x$ ?
- c. Como  $r(x) = -(d/dx)\ln[1 - F(x)]$ ,  $\ln[1 - F(x)] = -\int r(x) dx$ . Suponha

$$r(x) = \begin{cases} \alpha \left(1 - \frac{x}{\beta}\right) & 0 \leq x \leq \beta \\ 0 & \text{caso contrário} \end{cases}$$

de forma que, se um componente durar  $\beta$  horas, durará para sempre (apesar de não parecer razoável, o modelo é usado para estudar apenas o “desgaste inicial”). Quais são a fdc e a fdp de  $X$ ?

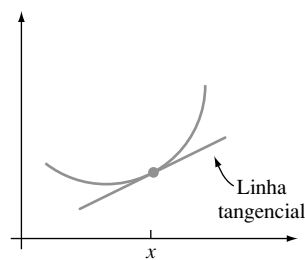
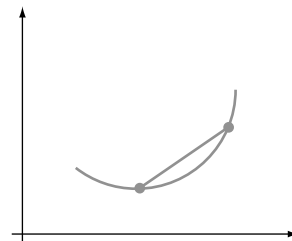
- 117.** Assuma  $U$  como tendo uma distribuição uniforme no intervalo  $[0, 1]$ . Os valores observados com essa distribuição podem ser obtidos com um gerador de números aleatórios. Seja  $X = -(1/\lambda)\ln(1 - U)$ .
- a. Mostre que  $X$  possui uma distribuição exponencial com parâmetro  $\lambda$ . [*Sugestão*: a fdc de  $X$  é  $F(x) = P(X \leq x)$ ;  $X \leq x$  é equivalente a  $U \leq ?$ ]
  - b. Como você usaria a parte (a) e um gerador de números aleatórios para obter valores observados de uma distribuição exponencial com parâmetro  $\lambda = 10$ ?
- 118.** Considere uma va  $X$  com média  $\mu$  e desvio padrão  $\sigma$  e seja  $g(X)$  uma função especificada de  $X$ . A série de

aproximação de Taylor de primeira ordem em relação a  $g(X)$  na vizinhança de  $\mu$  é

$$g(X) \approx g(\mu) + g'(\mu) \cdot (X - \mu)$$

O lado direito dessa equação é uma função linear de  $X$ . Se a distribuição de  $X$  estiver concentrada em um intervalo em que  $g(\cdot)$  é aproximadamente linear [por exemplo,  $\sqrt{x}$  é aproximadamente linear em  $(1, 2)$ ], a equação fornece aproximações para  $E(g(X))$  e  $V(g(X))$ .

- a. Forneça expressões para tais aproximações. (*Sugestão*: use regras de valor esperado e variância para uma função linear  $aX + b$ .)
  - b. Se a voltagem  $v$  em um meio for fixa mas a corrente  $I$  for aleatória, então a resistência também será uma variável aleatória relacionada a  $I$  por  $R = v/I$ . Se  $\mu_I = 20$  e  $\sigma_I = 0,5$ , calcule aproximações para  $\mu_R$  e  $\sigma_R$ .
- 119.** Uma função  $g(x)$  é convexa se a corda que conecta quaisquer dois pontos no gráfico da função ficar acima do gráfico. Quando  $g(x)$  é diferenciável, uma condição equivalente é que, para cada  $x$ , a reta tangente em  $x$  esteja totalmente acima ou abaixo do gráfico. (Veja as figuras abaixo.) Como  $g(\mu) = g(E(X))$  se compara a  $E(g(X))$ ? [*Sugestão*: a equação da tangente em  $x = \mu$  é  $y = g(\mu) + g'(\mu) \cdot (x - \mu)$ . Use a condição de convexidade, substitua  $X$  por  $x$  e calcule os valores esperados. *Nota*: a menos que  $g(x)$  seja linear, a desigualdade resultante (normalmente denominada desigualdade de Jensen) é estrita ( $<$  em vez de  $\leq$ ). Isso é válido para vas contínuas e discretas.]



- 120.** Assuma  $X$  com distribuição de Weibull de parâmetros  $\alpha = 2$  e  $\beta$ . Mostre que  $Y = 2X^2/\beta^2$  possui uma distribuição qui quadrado com  $\nu = 2$ . [*Sugestão*: a fdc de  $Y$  é  $P(Y \leq y)$ . Expresse essa probabilidade sob a forma  $P(X \leq g(y))$ , usando o fato de que  $X$  possui fdc da forma da Expressão (4.12) e diferencie em relação à  $y$  para obter a fdp de  $Y$ .]



---

## Bibliografia

BURY, Karl, *Statistical Distributions in Engineering*, Cambridge: Cambridge Univ. Press, 1999. Uma pesquisa informativa e de fácil leitura sobre distribuições e suas propriedades.

JOHNSON, Norman; KOTZ, Samuel e BALAKRISHNAN N. *Continuous Univariate Distributions*, Wiley: Nova York, 1994, vols. 1–2. Esses dois volumes apresentam uma pesquisa abrangente sobre as diversas distribuições contínuas.

NELSON, Wayne. *Applied Life Data Analysis*, Wiley. Nova York, 1982. Fornece uma discussão abrangente das distribuições e dos métodos usados na análise de dados de vida útil.

OLKIN, Ingram, DERMAN, Cyrus e GLESER, Leon, *Probability Models and Applications*. Macmillan: Nova York, 1994, (2 ed.). Boa cobertura de propriedades gerais e de distribuições específicas.

# Distribuições de Probabilidade Conjunta e Amostras Aleatórias

## Introdução

Nos capítulos 3 e 4, estudamos modelos probabilísticos com uma única variável aleatória. Muitos problemas de probabilidade e estatística nos conduzem a modelos que envolvem diversas variáveis aleatórias simultaneamente. Neste capítulo, discutiremos inicialmente os modelos probabilísticos para o comportamento conjunto de diversas variáveis aleatórias, enfatizando o caso em que as variáveis são independentes umas das outras. Depois estudaremos os valores esperados das funções de diversas variáveis aleatórias, incluindo covariância e correlação como medida do grau de relação entre duas variáveis.

As três últimas seções do capítulo consideram funções de  $n$  variáveis aleatórias  $X_1, X_2, \dots, X_n$ , enfocando especialmente as médias  $(X_1 + \dots + X_n)/n$ . Denominamos as funções desse tipo, que são elas mesmas variáveis aleatórias, uma estatística. Resultados da probabilidade são usados para obtenção de informações sobre a distribuição de uma estatística. O resultado principal desse tipo é o Teorema do Limite Central (CLT), a base de vários procedimentos inferenciais que envolvem tamanhos de amostras grandes.

---

## 5.1 Variáveis Aleatórias de Distribuição Conjunta

Há várias situações experimentais em que mais de uma variável aleatória será de interesse do investigador. Primeiro consideraremos as distribuições de probabilidade conjunta para duas vas discretas, depois para duas vas contínuas e finalmente para mais de duas variáveis.

## A Função de Massa de Probabilidade Conjunta Para Duas Variáveis Aleatórias Discretas

A função de massa de probabilidade (fmp) de uma única va discreta  $X$  especifica quanta massa de probabilidade é colocada em cada valor  $X$  possível. A fmp conjunta de duas vas discretas  $X$  e  $Y$  descreve quanta massa de probabilidade é colocada em cada par de valores possível  $(x, y)$ .

### DEFINIÇÃO

Sejam  $X$  e  $Y$  duas vas discretas definidas no espaço amostral  $\mathcal{S}$  de um experimento. A **função de massa de probabilidade conjunta**  $p(x, y)$  é definida para cada par de números  $(x, y)$  por

$$p(x, y) = P(X = x \text{ e } Y = y)$$

Seja  $A$  qualquer conjunto formado por pares de valores  $(x, y)$ . A probabilidade  $P[(X, Y) \in A]$  é obtida pela soma da fmp conjunta com os pares de  $A$ :

$$P[(X, Y) \in A] = \sum_{(x, y) \in A} p(x, y)$$

### Exemplo 5.1

Uma grande agência de seguros presta serviços a diversos clientes que compraram uma apólice residencial e outra de automóvel da mesma seguradora. Para cada tipo, deve ser especificado um valor dedutível. Para uma apólice de automóvel as opções são US\$ 100 e US\$ 250, enquanto, para uma apólice residencial, as opções são 0, US\$ 100 e US\$ 200. Suponha que um indivíduo com os dois referidos tipos seja selecionado aleatoriamente nos arquivos da seguradora. Sejam  $X$  = valor dedutível na apólice de automóvel e  $Y$  = valor dedutível na apólice residencial. Os pares  $(X, Y)$  possíveis são  $(100, 0)$ ,  $(100, 100)$ ,  $(100, 200)$ ,  $(250, 0)$ ,  $(250, 100)$  e  $(250, 200)$ ; a fmp conjunta especifica a probabilidade associada a cada um desses pares, com qualquer outro par de probabilidade zero. Suponha que a fmp conjunta seja dada na **tabela de probabilidade conjunta** a seguir:

$p(x, y)$		$y$		
		0	100	200
$x$	100	0,20	0,10	0,20
	250	0,05	0,15	0,30

Então,  $p(100, 100) = P(X = 100 \text{ e } Y = 100) = P(\text{US\$100 dedutível nas duas apólices}) = 0,10$ . A probabilidade  $P(Y \geq 100)$  é calculada pela soma das probabilidades de todos os pares  $(x, y)$  para os quais  $y \geq 100$ :

$$\begin{aligned} P(Y \geq 100) &= p(100, 100) + p(250, 100) + p(100, 200) + p(250, 200) \\ &= 0,75 \end{aligned}$$

Uma função  $p(x, y)$  pode ser usada como fmp conjunta desde que  $p(x, y) \geq 0$  para todos os  $x$  e  $y$  e  $\sum_x \sum_y p(x, y) = 1$ .

A fmp de apenas uma das variáveis é obtida pela soma de  $p(x, y)$  em relação aos valores da outra variável. O resultado é denominado *fmp marginal* porque, quando os valores  $p(x, y)$  são exibidos em uma tabela retangular, as somas são apenas totais marginais (linha ou coluna).

### DEFINIÇÃO

As **funções de massa de probabilidade marginais** de  $X$  e de  $Y$ , representadas respectivamente por  $p_x(x)$  e  $p_y(y)$ , são dadas por

$$p_x(x) = \sum_y p(x, y) \quad p_y(y) = \sum_x p(x, y)$$

Assim, para obter a fmp marginal de  $X$  calculado, digamos, para  $x = 100$ , as probabilidades  $p(100, y)$  são somadas para todos os valores possíveis  $y$ . Cálculo de cada  $X$  possível fornece a fmp marginal de  $X$  apenas (sem referência a  $Y$ ). Pelas fmps marginais, podem ser calculadas as probabilidades de eventos que envolvem apenas  $X$  ou apenas  $Y$ .

**Exemplo 5.2** (continuação do Exemplo 5.1)

Os valores de  $X$  possíveis são  $x = 100$  e  $x = 250$ , assim o cálculo dos totais de linhas da tabela de probabilidade conjunta fornece

$$p_x(100) = p(100, 0) + p(100, 100) + p(100, 200) = 0,50$$

e

$$p_x(250) = p(250, 0) + p(250, 100) + p(250, 200) = 0,50$$

A fmp marginal de  $X$  será então

$$p_x(x) = \begin{cases} 0,5 & x = 100, 250 \\ 0 & \text{caso contrário} \end{cases}$$

De forma similar, a fmp marginal de  $Y$  será obtida pelos totais da coluna como a seguir

$$p_y(y) = \begin{cases} 0,25 & y = 0, 100 \\ 0,50 & y = 200 \\ 0 & \text{caso contrário} \end{cases}$$

então  $P(Y \geq 100) = p_y(100) + p_y(200) = 0,75$  como antes. ■

## A Função de Densidade de Probabilidade Conjunta para Duas Variáveis Aleatórias Contínuas

A probabilidade de o valor observado de uma va contínua  $X$  estar em um conjunto unidimensional  $A$  (como um intervalo) é obtida integrando-se a fdp  $f(x)$  em relação ao conjunto  $A$ . De forma similar, a probabilidade de o par  $(X, Y)$  de vas contínuas estar em um conjunto bidimensional  $A$  (como um retângulo) é obtida pela integração de uma função denominada *função de densidade conjunta*.

### DEFINIÇÃO

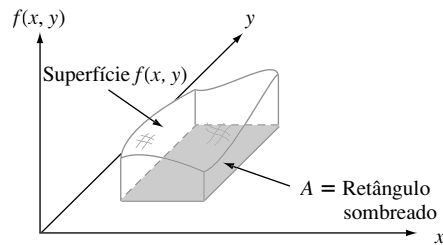
Seja  $X$  e  $Y$  vas contínuas. Então,  $f(x, y)$  é a **função de densidade de probabilidade conjunta** de  $X$  e  $Y$  se, para qualquer conjunto bidimensional,  $A$

$$P[(X, Y) \in A] = \int_A \int f(x, y) \, dx \, dy$$

Em particular, se  $A$  for o retângulo bidimensional  $\{(x, y): a \leq x \leq b, c \leq y \leq d\}$ , então

$$P[(X, Y) \in A] = P(a \leq X \leq b, c \leq Y \leq d) = \int_a^b \int_c^d f(x, y) \, dy \, dx$$

Para  $f(x, y)$  ser candidata a uma fdp conjunta, deve satisfazer  $f(x, y) \geq 0$  e  $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \, dx \, dy = 1$ . Podemos imaginar  $f(x, y)$  como especificação de uma superfície de altura  $f(x, y)$  acima do ponto  $(x, y)$  em um sistema tridimensional. Então  $P[(X, Y) \in A]$  é o volume abaixo dessa superfície e acima da região  $A$ , análogo à área abaixo de uma curva no caso unidimensional. Como ilustrado na Figura 5.1.



**Figura 5.1**  $P[(X, Y) \in A] = \text{volume sob a superfície de densidade acima de } A$

### Exemplo 5.3

Um banco opera tanto uma instalação de *drive-through* como em guichê de atendimento. Em um dia selecionado aleatoriamente, assuma  $X =$  a proporção de tempo em que a instalação de *drive-through* está em uso (ao menos um cliente está sendo atendido ou esperando para ser atendido) e  $Y =$  a proporção de tempo em que o guichê de atendimento está em uso. O conjunto de valores possíveis de  $(X, Y)$  é, então, o retângulo  $D = \{(x, y): 0 \leq x \leq 1, 0 \leq y \leq 1\}$ . Suponha que a fdp conjunta de  $(X, Y)$  seja dada por

$$f(x, y) = \begin{cases} \frac{6}{5}(x + y^2) & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{caso contrário} \end{cases}$$

Para provar que essa fdp é verdadeira, observe que  $f(x, y) \geq 0$  e

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \, dx \, dy &= \int_0^1 \int_0^1 \frac{6}{5} (x + y^2) \, dx \, dy \\ &= \int_0^1 \int_0^1 \frac{6}{5} x \, dx \, dy + \int_0^1 \int_0^1 \frac{6}{5} y^2 \, dx \, dy \\ &= \int_0^1 \frac{6}{5} x \, dx + \int_0^1 \frac{6}{5} y^2 \, dy = \frac{6}{10} + \frac{6}{15} = 1 \end{aligned}$$

A probabilidade de nenhuma das instalações estar ocupada em mais de um quarto do tempo é

$$\begin{aligned} P\left(0 \leq X \leq \frac{1}{4}, 0 \leq Y \leq \frac{1}{4}\right) &= \int_0^{1/4} \int_0^{1/4} \frac{6}{5} (x + y^2) \, dx \, dy \\ &= \frac{6}{5} \int_0^{1/4} \int_0^{1/4} x \, dx \, dy + \frac{6}{5} \int_0^{1/4} \int_0^{1/4} y^2 \, dx \, dy \\ &= \frac{6}{20} \cdot \frac{x^2}{2} \Big|_{x=0}^{x=1/4} + \frac{6}{20} \cdot \frac{y^3}{3} \Big|_{y=0}^{y=1/4} = \frac{7}{640} \\ &= 0,0109 \end{aligned}$$

Como acontece com as fmps conjuntas, cada uma das duas funções densidade marginais pode ser calculada pela fdp de  $X$  e  $Y$ .

#### DEFINIÇÃO

As **funções de densidade de probabilidade marginais** de  $X$  e  $Y$ , representadas respectivamente por  $f_X(x)$  e  $f_Y(y)$ , são dadas por

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) \, dy \quad \text{para } -\infty < x < \infty$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) \, dx \quad \text{para } -\infty < y < \infty$$

**Exemplo 5.4** (continuação do Exemplo 5.3)

A fdp marginal de  $X$ , que fornece a distribuição de probabilidades do tempo ocupado para a instalação de *drive-through* sem referência à janela de atendimento, é

$$f_x(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_0^1 \frac{6}{5} (x + y^2) dy = \frac{6}{5} x + \frac{2}{5}$$

para  $0 \leq x \leq 1$  e 0 em caso contrário. A fdp marginal de  $Y$  é

$$f_y(y) = \begin{cases} \frac{6}{5} y^2 + \frac{3}{5} & 0 \leq y \leq 1 \\ 0 & \text{caso contrário} \end{cases}$$

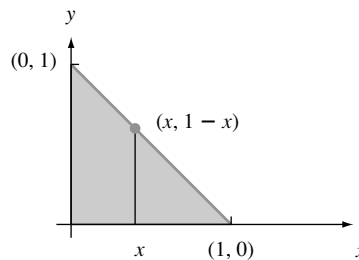
Então

$$P\left(\frac{1}{4} \leq Y \leq \frac{3}{4}\right) = \int_{1/4}^{3/4} f_y(y) dy = \frac{37}{80} = 0,4625 \quad \blacksquare$$

No Exemplo 5.3, a região da densidade conjunta positiva era um retângulo, o que tornou o cálculo das fdp marginais relativamente fácil. Considere agora um exemplo em que a região de densidade positiva é uma figura mais complicada.

**Exemplo 5.5**

Uma empresa de nozes comercializa latas luxuosas de nozes mistas com amêndoas, castanhas de caju e amendoins. Suponha que o peso líquido de cada lata seja exatamente 1 libra, mas que a contribuição do peso de cada tipo de noz seja aleatória. Como os três pesos devem somar 1, um modelo de probabilidade conjunta para quaisquer dois fornece todas as informações necessárias sobre o peso do terceiro tipo. Sejam  $X$  = peso das amêndoas em uma lata selecionada e  $Y$  = peso das castanhas de caju. Então a região de densidade positiva é  $D = \{(x, y): 0 \leq x \leq 1, 0 \leq y \leq 1, x + y \leq 1\}$ , região sombreada ilustrada na Figura 5.2.



**Figura 5.2** Região de densidade positiva do Exemplo 5.5

Agora considere que a fdp conjunta de  $(X, Y)$  seja

$$f(x, y) = \begin{cases} 24xy & 0 \leq x \leq 1, 0 \leq y \leq 1, x + y \leq 1 \\ 0 & \text{caso contrário} \end{cases}$$

Para qualquer  $x$  fixo,  $f(x, y)$  aumenta com  $y$ ; para  $y$  fixo,  $f(x, y)$  aumenta com  $x$ . A hipótese é apropriada porque a palavra *luxo* implica que a maior parte da lata deve ser formada por amêndoas e castanhas de caju em vez de amendoins, de forma que a função de densidade deve ser grande próximo ao limite superior e pequena perto da origem. A superfície determinada por  $f(x, y)$  tem declive positivo a partir de zero, conforme  $(x, y)$  se distancia dos eixos.

Claramente,  $f(x, y) \geq 0$ . Para demonstrar a segunda condição de uma fdp conjunta, lembre-se de que a integral dupla é calculada como integral iterativa, mantendo uma variável fixa (como  $x$  na Figura 5.2),

integrando em relação aos valores da outra variável que estão na reta que passa pelo valor da variável fixa e, finalmente, integrando em relação a todos os valores possíveis da variável fixa. Dessa forma,

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dy dx &= \int_D \int f(x, y) dy dx = \int_0^1 \left\{ \int_0^{1-x} 24xy dy \right\} dx \\ &= \int_0^1 24x \left\{ \frac{y^2}{2} \Big|_{y=0}^{y=1-x} \right\} dx = \int_0^1 12x(1-x)^2 dx = 1 \end{aligned}$$

Para calcular a probabilidade de dois tipos de nozes formarem juntos no máximo 50% da lata, assuma  $A = \{(x, y): 0 \leq x \leq 1, 0 \leq y \leq 1, \text{ e } x + y \leq 0,5\}$ , conforme mostrado na Figura 5.3. Então

$$P((X, Y) \in A) = \int_A \int f(x, y) dx dy = \int_0^{0,5} \int_0^{0,5-x} 24xy dy dx = 0,0625$$

A fdp marginal das amêndoas é obtida fixando-se  $X$  em  $x$  e integrando  $f(x, y)$  ao longo da linha vertical que passa por  $x$ :

$$f_x(x) = \int_{-\infty}^{\infty} f(x, y) dy = \begin{cases} \int_0^{1-x} 24xy dy = 12x(1-x)^2 & 0 \leq x \leq 1 \\ 0 & \text{caso contrário} \end{cases}$$

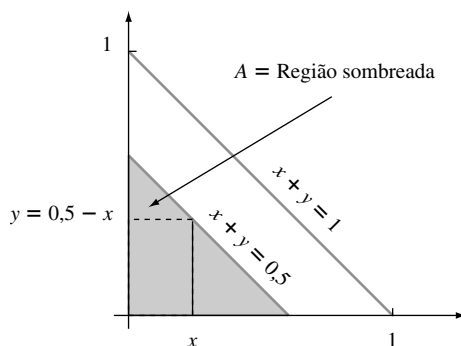


Figura 5.3 Cálculo de  $P[(X, Y) \in A]$  do Exemplo 5.5

Pela simetria de  $f(x, y)$  e a região  $D$ , a fdp marginal de  $Y$  é obtida pela substituição de  $x$  e  $X$  em  $f_x(x)$  por  $y$  e  $Y$ , respectivamente.

### Variáveis Aleatórias Independentes

Em muitas situações, as informações sobre o valor observado de uma das duas variáveis  $X$  e  $Y$  fornecem dados sobre o valor da outra variável. No Exemplo 5.1, a probabilidade marginal de  $X$  em  $x = 250$  era 0,5, como a probabilidade de  $X = 100$ . Se, entretanto, dissessem-nos que a pessoa selecionada tinha  $Y = 0$ , então,  $X = 100$  é quatro vezes mais provável que  $X = 250$ . Portanto, há uma dependência entre as duas variáveis.

No Capítulo 2, apontamos que uma forma de definir a independência de dois eventos é dizer que  $A$  e  $B$  são independentes, se  $P(A \cap B) = P(A) \cdot P(B)$ . Eis uma definição análoga da independência de duas vas.

**DEFINIÇÃO**

Duas variáveis aleatórias  $X$  e  $Y$  são **independentes** se, para cada par de valores  $x$  e  $y$ ,

$$p(x, y) = p_x(x) \cdot p_y(y) \quad \text{quando } X \text{ e } Y \text{ são discretas} \tag{5.1}$$

ou

$$f(x, y) = f_x(x) \cdot f_y(y) \quad \text{quando } X \text{ e } Y \text{ são contínuas}$$

Se (5.1) não for satisfeita para todos os  $(x, y)$ , então,  $X$  e  $Y$  são denominadas **dependentes**.

A definição diz que duas variáveis são independentes se sua fmp conjunta ou sua fdp for o produto das duas fmps ou fdps marginais.

### Exemplo 5.6

No caso do seguro dos Exemplos 5.1 e 5.2,

$$p(100, 100) = 0,10 \neq (0,5)(0,25) = p_X(100) \cdot p_Y(100)$$

de forma que  $X$  e  $Y$  não são independentes. A independência de  $X$  e  $Y$  exige que *cada* entrada na tabela de probabilidade conjunta seja o produto das probabilidades marginais da linha e da coluna correspondentes. ■

### Exemplo 5.7 (continuação do Exemplo 5.5)

Como  $f(x, y)$  tem a forma de um produto,  $X$  e  $Y$  podem parecer independentes. Entretanto, apesar de  $f_X(\frac{3}{4}) = f_Y(\frac{3}{4}) = \frac{9}{16}$ ,  $f(\frac{3}{4}, \frac{3}{4}) = 0 \neq \frac{9}{16} \cdot \frac{9}{16}$ , então as variáveis não são de fato independentes. Para serem independentes,  $f(x, y)$  deve ter a forma  $g(x) \cdot h(y)$  e a região de densidade positiva deve ser um retângulo cujos lados são paralelos aos eixos de coordenadas. ■

A independência de duas variáveis aleatórias é mais útil quando a descrição do experimento em estudo nos diz que  $X$  e  $Y$  não afetam uma à outra. Então, uma vez especificadas as fdps ou as fmps marginais, a fmp ou a fdp conjunta é simplesmente o produto das duas funções marginais. Segue que

$$P(a \leq X \leq b, c \leq Y \leq d) = P(a \leq X \leq b) \cdot P(c \leq Y \leq d)$$

### Exemplo 5.8

Suponha que as vidas úteis de dois componentes sejam independentes uma da outra e que a primeira vida útil,  $X_1$ , tenha distribuição exponencial com parâmetro  $\lambda_1$ , enquanto a segunda,  $X_2$ , tenha distribuição exponencial com parâmetro  $\lambda_2$ . A fdp conjunta será então

$$f(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) = \begin{cases} \lambda_1 e^{-\lambda_1 x_1} \cdot \lambda_2 e^{-\lambda_2 x_2} = \lambda_1 \lambda_2 e^{-\lambda_1 x_1 - \lambda_2 x_2} & x_1 > 0, x_2 > 0 \\ 0 & \text{caso contrário} \end{cases}$$

Sejam  $\lambda_1 = 1/1000$  e  $\lambda_2 = 1/1200$ , de forma que as vidas úteis esperadas sejam de 1000 e 1200 horas, respectivamente. A probabilidade de que ambas as vidas úteis sejam de no mínimo 1500 horas é

$$\begin{aligned} P(1500 \leq X_1, 1500 \leq X_2) &= P(1500 \leq X_1) \cdot P(1500 \leq X_2) \\ &= e^{-\lambda_1(1500)} \cdot e^{-\lambda_2(1500)} \\ &= (0,2231)(0,2865) = 0,0639 \end{aligned} \quad \blacksquare$$

## Mais de Duas Variáveis Aleatórias

Para modelar o comportamento conjunto de mais de duas variáveis aleatórias, estendemos o conceito de distribuição conjunta de duas variáveis.

### DEFINIÇÃO

Se  $X_1, X_2, \dots, X_n$  são variáveis aleatórias discretas, a fmp conjunta das variáveis é a função

$$p(x_1, x_2, \dots, x_n) = P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$$

Se as variáveis forem contínuas, a fdp conjunta de  $X_1, \dots, X_n$  será a função  $f(x_1, x_2, \dots, x_n)$  de forma que, para qualquer dos  $n$  intervalos  $[a_1, b_1], \dots, [a_n, b_n]$ ,

$$P(a_1 \leq X_1 \leq b_1, \dots, a_n \leq X_n \leq b_n) = \int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} f(x_1, \dots, x_n) dx_n \dots dx_1$$



Em um experimento binomial, cada tentativa pode resultar em um de somente dois resultados possíveis. Considere agora um experimento formado por  $n$  tentativas idênticas e independentes, em que cada tentativa pode resultar em um dos  $r$  resultados possíveis. Seja  $p_i = P(\text{resultado } i \text{ em qualquer tentativa específica})$  e defina as variáveis aleatórias por  $X_i = \text{o número de tentativas que resultam em } i \text{ (} i = 1, \dots, r \text{)}$ . Tal experimento é denominado **experimento multinomial**, e a fmp conjunta de  $X_1, \dots, X_r$ , é denominada **distribuição multinomial**. Usando um raciocínio de contagem análogo ao usado na demonstração da distribuição binomial, a fmp conjunta de  $X_1, \dots, X_r$  pode ser demonstrada

$$p(x_1, \dots, x_r) = \begin{cases} \frac{n!}{(x_1!)(x_2!) \cdots (x_r!)} p_1^{x_1} \cdots p_r^{x_r} & x_i = 0, 1, 2, \dots, \text{ com } x_1 + \cdots + x_r = n \\ 0 & \text{caso contrário} \end{cases}$$

O caso  $r = 2$  fornece a distribuição binomial, com  $X_1 = \text{número de sucessos}$  e  $X_2 = n - X_1 = \text{número de falhas}$ .

**Exemplo 5.9**

Se o alelo de cada uma das 10 seções da ervilha obtidas de forma independente é determinado e  $p_1 = P(AA)$ ,  $p_2 = P(Aa)$ ,  $p_3 = P(aa)$ ,  $X_1 = \text{número de AAs}$ ,  $X_2 = \text{número de Aas}$  e  $X_3 = \text{número de aas}$ , então

$$p(x_1, x_2, x_3) = \frac{10!}{(x_1!)(x_2!)(x_3!)} p_1^{x_1} p_2^{x_2} p_3^{x_3} \quad x_i = 0, 1, \dots \text{ e } x_1 + x_2 + x_3 = 10$$

Se  $p_1 = p_3 = 0,25$ ,  $p_2 = 0,5$ , então

$$\begin{aligned} P(X_1 = 2, X_2 = 5, X_3 = 3) &= p(2, 5, 3) \\ &= \frac{10!}{2! 5! 3!} (0,25)^2 (0,5)^5 (0,25)^3 = 0,0769 \end{aligned}$$

**Exemplo 5.10**

Quando um determinado método é usado para coletar um volume fixo de amostras de rochas de uma região, há quatro tipos de rochas resultantes. Represente por  $X_1, X_2$ , e  $X_3$  a proporção por volume dos tipos de rochas 1, 2 e 3 em uma amostra selecionada aleatoriamente (a proporção da rocha do tipo 4 é  $1 - X_1 - X_2 - X_3$ , de forma que uma variável  $X_4$  seria redundante). Se a fdp conjunta de  $X_1, X_2, X_3$  for

$$f(x_1, x_2, x_3) = \begin{cases} kx_1x_2(1 - x_3) & 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1, 0 \leq x_3 \leq 1, x_1 + x_2 + x_3 \leq 1 \\ 0 & \text{caso contrário} \end{cases}$$

então  $k$  será determinado por

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2, x_3) dx_3 dx_2 dx_1 \\ &= \int_0^1 \left\{ \int_0^{1-x_1} \left[ \int_0^{1-x_1-x_2} kx_1x_2(1 - x_3) dx_3 \right] dx_2 \right\} dx_1 \end{aligned}$$

Essa integral iterativa tem valor  $k/144$ , portanto,  $k = 144$ . A probabilidade de que as rochas dos tipos 1 e 2 juntas sejam responsáveis por no máximo 50% da amostra é

$$\begin{aligned} P(X_1 + X_2 \leq 0,5) &= \iiint_{\substack{0 \leq x_i \leq 1 \text{ para } i=1, 2, 3 \\ x_1 + x_2 + x_3 \leq 1, x_1 + x_2 \leq 0,5}} f(x_1, x_2, x_3) dx_3 dx_2 dx_1 \\ &= \int_0^{0,5} \left\{ \int_0^{0,5-x_1} \left[ \int_0^{1-x_1-x_2} 144x_1x_2(1 - x_3) dx_3 \right] dx_2 \right\} dx_1 \\ &= 0,6066 \end{aligned}$$

A noção de independência de mais de duas variáveis aleatórias é semelhante à noção de independência de mais de dois eventos.

**DEFINIÇÃO**

As variáveis aleatórias  $X_1, X_2, \dots, X_n$  são **independentes** se para *cada* subconjunto  $X_{i_1}, X_{i_2}, \dots, X_{i_n}$ , das variáveis (cada par, cada trio e assim por diante), a fmp ou fdp conjuntas do subconjunto for igual ao produto das fdps ou fmps marginais.

Assim, se as variáveis forem independentes com  $n = 4$ , a fmp ou fdp conjunta de quaisquer duas variáveis será o produto das duas marginais e de forma semelhante para quaisquer três e todas as quatro variáveis. O mais importante é, depois que soubermos que  $n$  variáveis são independentes, saber que a fmp ou fdp conjunta será o produto das  $n$  marginais.

**Exemplo 5.11**

Se  $X_1, \dots, X_n$  representam as vidas úteis de  $n$  componentes que operam de forma independente um do outro e cada vida útil é distribuída exponencialmente com parâmetro  $\lambda$ , então

$$\begin{aligned} f(x_1, x_2, \dots, x_n) &= (\lambda e^{-\lambda x_1}) \cdot (\lambda e^{-\lambda x_2}) \cdot \dots \cdot (\lambda e^{-\lambda x_n}) \\ &= \begin{cases} \lambda^n e^{-\lambda \sum x_i} & x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0 \\ 0 & \text{caso contrário} \end{cases} \end{aligned}$$

Se esses  $n$  componentes formam um sistema que apresentará falha assim que um único componente tiver uma falha, a probabilidade de o sistema durar mais do que o tempo  $t$  é

$$\begin{aligned} P(X_1 > t, \dots, X_n > t) &= \int_t^\infty \dots \int_t^\infty f(x_1, \dots, x_n) dx_1 \dots dx_n \\ &= \left( \int_t^\infty \lambda e^{-\lambda x_1} dx_1 \right) \cdot \dots \cdot \left( \int_t^\infty \lambda e^{-\lambda x_n} dx_n \right) \\ &= (e^{-\lambda t})^n = e^{-n\lambda t} \end{aligned}$$

Portanto,

$$P(\text{tempo de vida do sistema} \leq t) = 1 - e^{-n\lambda t} \quad \text{para } t \geq 0$$

que mostra que a vida útil do *sistema* tem uma distribuição exponencial com parâmetro  $n\lambda$ ; e o valor esperado do tempo de vida do sistema é  $1/n\lambda$ . ■

Em muitas situações experimentais a serem consideradas neste livro, a independência é uma suposição razoável, de forma que especificar a distribuição conjunta se reduz a decidir quais são as distribuições marginais apropriadas.

**Distribuições Condicionais**

Suponha que  $X$  = número de defeitos graves em um novo automóvel selecionado aleatoriamente e  $Y$  = número de defeitos menores no mesmo carro. Sabendo que o número de defeitos graves no automóvel selecionado é um, qual é a probabilidade de o carro ter no máximo três defeitos menores, isto é, qual é  $P(Y \leq 3 \mid X = 1)$ ? De forma semelhante, se  $X$  e  $Y$  representam as vidas úteis de dois componentes em um sistema e acontece de  $X = 100$ , qual é a probabilidade de  $Y \geq 200$  e qual é a vida útil esperada do segundo componente “condicional em relação” a esse valor de  $X$ ?

Questões desse tipo são respondidas pelo estudo das distribuições de probabilidades condicionais.

## DEFINIÇÃO

Sejam  $X$  e  $Y$  duas v.a.s contínuas com fdp conjunta  $f(x, y)$  e fdp marginal de  $X$ ,  $f_X(x)$ . Então, para qualquer valor  $x$  de  $X$  para o qual  $f_X(x) > 0$ , a **função de densidade de probabilidade condicional de  $Y$  dado que  $X = x$**  é

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)} \quad -\infty < y < \infty$$

Se  $X$  e  $Y$  forem discretas, substituir as fdps por fmps nesta definição fornecerá a **função de massa de probabilidade condicional de  $Y$  quando  $X = x$** .

Observe que a definição de  $f_{Y|X}(y|x)$  corresponde à de  $P(B|A)$ , a probabilidade condicional da ocorrência de  $B$ , dado que  $A$  tenha ocorrido. Uma vez que a fdp ou fmp condicional tiver sido determinada, questões do tipo propostas no início desta subseção podem ser respondidas pela integração ou pela soma de um conjunto apropriado de valores  $Y$ .

## Exemplo 5.12

Reconsidere a situação dos Exemplos 5.3 e 5.4 que envolvem  $X =$  proporção de tempo em que o guichê de automóveis de um banco está ocupado e  $Y =$  proporção análoga para o guichê de atendimento a pessoas. A fdp condicional de  $Y$  dado que  $X = 0,8$  é

$$f_{Y|X}(y|0,8) = \frac{f(0,8, y)}{f_X(0,8)} = \frac{1,2(0,8 + y^2)}{1,2(0,8) + 0,4} = \frac{1}{34}(24 + 30y^2) \quad 0 < y < 1$$

A probabilidade de o guichê de pessoas estar ocupado no máximo metade do tempo, dado que  $X = 0,8$ , é

$$P(Y \leq 0,5 | X = 0,8) = \int_{-\infty}^{0,5} f_{Y|X}(y|0,8) dy = \int_0^{0,5} \frac{1}{34}(24 + 30y^2) dy = 0,390$$

Usar a fdp marginal de  $Y$  fornece  $P(Y \leq 0,5) = 0,350$ . Além disso,  $E(Y) = 0,6$ , enquanto a proporção esperada de tempo em que o guichê de pessoas está ocupado, dado que  $X = 0,8$  (uma expectativa *condicional*) é

$$E(Y|X = 0,8) = \int_{-\infty}^{\infty} y \cdot f_{Y|X}(y|0,8) dy = \frac{1}{34} \int_0^1 y(24 + 30y^2) dy = 0,574 \quad \blacksquare$$

## Exercícios | Seção 5.1 (1–21)

- Um posto de gasolina tem ilhas de auto-serviço e de serviço completo. Em cada ilha, há uma única bomba de auto-serviço de gasolina comum com duas mangueiras. Sejam  $X =$  número de mangueiras em uso na ilha de auto-serviço em um momento específico e  $Y =$  número de mangueiras na ilha de serviço completo em uso naquele mesmo momento. A fmp de  $X$  e  $Y$  é mostrada na tabela a seguir:
 

$p(x, y)$		$y$		
		0	1	2
$x$	0	0,10	0,04	0,02
	1	0,08	0,20	0,06
	2	0,06	0,14	0,30

  - Qual é  $P(X = 1 \text{ e } Y = 1)$ ?
  - Calcule  $P(X \leq 1 \text{ e } Y \leq 1)$ .
  - Descreva o evento  $\{X \neq 0 \text{ e } Y \neq 0\}$  e calcule sua probabilidade.
  - Calcule a fmp marginal de  $X$  e de  $Y$ . Usando  $p_X(x)$ , qual é  $P(X \leq 1)$ ?
  - $X$  e  $Y$  são v.a.s independentes? Explique.
- Quando um automóvel é parado por uma patrulha de segurança, os pneus são verificados quanto ao desgaste, assim como o alinhamento de cada farol. Seja  $X$  o número de faróis que precisam de ajuste e  $Y$  o número de pneus com defeito.
  - Se  $X$  e  $Y$  são independentes com  $p_X(0) = 0,5$ ,  $p_X(1) = 0,3$ ,  $p_X(2) = 0,2$ , e  $p_Y(0) = 0,6$ ,  $p_Y(1) = 0,1$ ,  $p_Y(2) =$

$p_Y(3) = 0,05$ ,  $p_Y(4) = 0,2$ , mostre a fmp conjunta de  $(X, Y)$  em uma tabela de probabilidade conjunta.

- b. Calcule  $P(X \leq 1 \text{ e } Y \leq 1)$  pela tabela de probabilidade conjunta e verifique se ela é igual ao produto  $P(X \leq 1) \cdot P(Y \leq 1)$ .
  - c. Qual é  $P(X + Y = 0)$  (a probabilidade de nenhuma violação)?
  - d. Calcule  $P(X + Y \leq 1)$ .
3. Um determinado mercado tem uma fila de caixa expressa e uma fila de caixa superexpressa. Represente por  $X_1$  o número de clientes na fila da caixa expressa em um determinado horário do dia e por  $X_2$  o número de clientes na fila da caixa superexpressa no mesmo horário. Suponha que a fmp conjunta de  $X_1$  e  $X_2$  seja dada na tabela a seguir.

		$x_2$			
		0	1	2	3
$x_1$	0	0,08	0,07	0,04	0,00
	1	0,06	0,15	0,05	0,04
	2	0,05	0,04	0,10	0,06
	3	0,00	0,03	0,04	0,07
	4	0,00	0,01	0,05	0,06

- a. Qual é  $P(X_1 = 1, X_2 = 1)$ , isto é, a probabilidade de haver exatamente um cliente em cada fila?
  - b. Qual é  $P(X_1 = X_2)$ , isto é, a probabilidade de os números de clientes nas duas filas serem iguais?
  - c. Represente por  $A$  o evento de haver no mínimo dois clientes a mais em uma fila do que na outra. Expresse  $A$  em termos de  $X_1$  e  $X_2$ , e calcule a probabilidade desse evento.
  - d. Qual é a probabilidade de o número total de clientes nas duas filas ser exatamente quatro? E no mínimo quatro?
4. Retorne à situação descrita no Exercício 3.
- a. Determine a fmp marginal de  $X_1$  e, depois, calcule o número esperado de clientes na fila da caixa expressa.
  - b. Determine a fmp marginal de  $X_2$ .
  - c. Pela inspeção das probabilidades  $P(X_1 = 4)$ ,  $P(X_2 = 0)$ , e  $P(X_1 = 4, X_2 = 0)$ ,  $X_1$  e  $X_2$  são variáveis aleatórias independentes? Explique.
5. O número de clientes esperando no serviço de embalagens para presente em uma loja é uma va  $X$  com valores possíveis 0, 1, 2, 3, 4 e probabilidades correspondentes 0,1, 0,2, 0,3, 0,25, 0,15. Um cliente selecionado aleatoriamente terá 1, 2 ou 3 pacotes para embrulhar com probabilidades 0,6, 0,3 e 0,1, respectivamente. Seja  $Y =$  o número total de pacotes a serem embrulhados para os clientes que estão na fila (assuma que o número de pacotes de um cliente seja independente do número de pacotes de qualquer outro cliente).
- a. Determine  $P(X = 3, Y = 3)$ , isto é,  $p(3, 3)$ .
  - b. Determine  $p(4, 11)$ .
6. Seja  $X$  número de câmeras digitais Canon vendidas durante uma determinada semana por uma loja. A fmp de  $X$  é

$x$	0	1	2	3	4
$p_X(x)$	0,1	0,2	0,3	0,25	0,15

Sessenta por cento de todos os clientes que compram essas câmeras também compram uma garantia estendida. Seja  $Y$  o número de compradores que durante essa semana adquiriram garantia estendida.

- a. Qual é  $P(X = 4, Y = 2)$ ? [Sugestão: essa probabilidade é igual a  $P(Y = 2 | X = 4) \cdot P(X = 4)$ ; imagine agora que as quatro compras sejam quatro tentativas de um experimento binomial, com sucesso em uma tentativa correspondente à aquisição de uma garantia estendida].
  - b. Calcule  $P(X = Y)$ .
  - c. Determine a fmp conjunta de  $X$  e  $Y$  e, depois, a fmp marginal de  $Y$ .
7. A distribuição de probabilidades conjuntas do número  $X$  de carros e do número  $Y$  de ônibus por ciclo de semáforo em uma conversão à esquerda proposta é exibida na tabela de probabilidade conjunta a seguir.

		$y$		
		0	1	2
$x$	0	0,025	0,015	0,010
	1	0,050	0,030	0,020
	2	0,125	0,075	0,050
	3	0,150	0,090	0,060
	4	0,100	0,060	0,040
	5	0,050	0,030	0,020

- a. Qual é a probabilidade de haver exatamente um carro e exatamente um ônibus em um ciclo?
  - b. Qual é a probabilidade de haver no máximo um carro e no máximo um ônibus em um ciclo?
  - c. Qual é a probabilidade de haver exatamente um carro em um ciclo? E exatamente um ônibus?
  - d. Suponha que uma pista de conversão à esquerda tenha capacidade para cinco carros e um ônibus, equivalente a três carros. Qual é a probabilidade de excesso de fluxo em um ciclo?
  - e.  $X$  e  $Y$  são vas independentes? Explique.
8. Um armazém tem atualmente 30 componentes de certo tipo, dos quais 8 são fornecidos pelo fornecedor 1, 10 pelo fornecedor 2 e 12 pelo fornecedor 3. Seis deles serão aleatoriamente selecionados para uma montagem. Seja  $X =$  número de componentes selecionados do fornecedor 1,  $Y =$  número de componentes selecionados do fornecedor 2 e  $p(x, y)$  a fmp conjunta de  $X$  e  $Y$ .
- a. Qual é  $p(3, 2)$ ? [Sugestão: cada amostra de tamanho 6 tem a mesma probabilidade de ser selecionada. Portanto,  $p(3, 2) =$  (número de resultados com  $X = 3$  e  $Y = 2$ )/(número total de resultados). Use a regra do produto para contagem a fim de obter o numerador e o denominador].
  - b. Usando a lógica do item (a), obtenha  $p(x, y)$ . (Isso pode ser visto como uma distribuição hipergeométrica multivariada – amostragem sem reposição de

uma população finita consistindo de mais de duas categorias).

9. Admite-se que cada pneu dianteiro de um determinado tipo de veículo deve ter a pressão de 26 psi. Suponha que a pressão real em cada pneu seja uma variável aleatória  $X$  para o pneu direito e  $Y$  para o esquerdo, com fdp conjunta

$$f(x, y) = \begin{cases} K(x^2 + y^2) & 20 \leq x \leq 30, 20 \leq y \leq 30 \\ 0 & \text{caso contrário} \end{cases}$$

- Qual é o valor de  $K$ ?
  - Qual é a probabilidade de os dois pneus estarem com pressão inferior à ideal?
  - Qual é a probabilidade de a diferença de pressão de ar entre os dois pneus ser no máximo 2 psi?
  - Determine a distribuição (marginal) da pressão de ar só do pneu direito.
  - $X$  e  $Y$  são independentes?
10. Annie e Alvie combinaram de se encontrar entre 17 e 18 horas para jantar em um restaurante local de comida saudável. Seja  $X$  = a hora de chegada de Annie e  $Y$  = a de Alvie. Suponha que  $X$  e  $Y$  sejam independentes com distribuição uniforme no intervalo  $[5, 6]$ .
- Qual é a fdp conjunta de  $X$  e  $Y$ ?
  - Qual é a probabilidade de ambos chegarem entre 17h15 e 17h45?
  - Se o primeiro a chegar esperar apenas 10 minutos antes de ir comer em outro lugar, qual é a probabilidade de eles comerem no restaurante de comida saudável? [Sugestão: o evento de interesse é  $A = \{(x, y) : |x - y| \leq \frac{1}{6}\}$ .]

11. Dois professores diferentes enviaram suas provas finais para tiragem de cópias. Seja  $X$  o número de erros tipográficos na prova do primeiro professor e  $Y$  o número de tais erros para o segundo professor. Suponha que  $X$  tenha distribuição de Poisson com parâmetro  $\lambda$ ,  $Y$  tenha distribuição de Poisson com parâmetro  $\theta$  e que  $X$  e  $Y$  sejam independentes.
- Qual é a fmp conjunta de  $X$  e  $Y$ ?
  - Qual é a probabilidade de no máximo um erro ser cometido em ambos os exames combinados?
  - Obtenha a expressão geral para a probabilidade de o número total de erros nos dois exames ser  $m$  (onde  $m$  é um inteiro não-negativo). [Sugestão:  $A = \{(x, y) : x + y = m\} = \{(m, 0), (m - 1, 1), \dots, (1, m - 1), (0, m)\}$ . Agora some a fmp conjunta em  $(x, y) \in A$  e use o teorema binomial, que diz

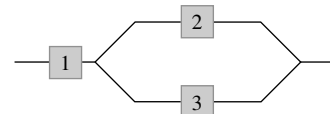
$$\sum_{k=0}^m \binom{m}{k} a^k b^{m-k} = (a + b)^m$$

para qualquer  $a, b$ .]

12. Dois componentes de um minicomputador possuem a fdp conjunta a seguir para seus tempos de vida útil  $X$  e  $Y$ .

$$f(x, y) = \begin{cases} xe^{-x(1+y)} & x \geq 0 \text{ e } y \geq 0 \\ 0 & \text{caso contrário} \end{cases}$$

- Qual é a probabilidade da vida útil  $X$  do primeiro componente exceder 3?
  - Quais são as fdps marginais de  $X$  e  $Y$ ? Os dois períodos de vida útil são independentes? Explique.
  - Qual é a probabilidade de o tempo de vida de pelo menos um componente exceder 3?
13. Você tem duas lâmpadas para uma determinada luminária. Seja  $X$  = vida útil da primeira lâmpada e  $Y$  = vida útil da segunda lâmpada (ambas de mil horas). Suponha que  $X$  e  $Y$  sejam independentes e que cada uma tenha distribuição exponencial com parâmetro  $\lambda = 1$ .
- Qual é a fdp conjunta de  $X$  e  $Y$ ?
  - Qual é a probabilidade de cada lâmpada durar no máximo 1000 horas (isso é,  $X \leq 1$  e  $Y \leq 1$ )?
  - Qual é a probabilidade de a vida útil total das duas lâmpadas ser no máximo 2? [Sugestão: desenhe a região  $A = \{(x, y) : x \geq 0, y \geq 0, x + y \leq 2\}$  antes da integração.]
  - Qual é a probabilidade de a vida útil total estar entre 1 e 2?
14. Suponha que você tenha 10 lâmpadas, que a vida útil de cada uma seja independente das demais e que cada vida útil tenha distribuição exponencial com parâmetro  $\lambda$ .
- Qual é a probabilidade de as 10 lâmpadas falharem antes do tempo  $t$ ?
  - Qual é a probabilidade de exatamente  $k$  das 10 lâmpadas falharem antes do tempo  $t$ ?
  - Suponha que nove lâmpadas tenham vida útil exponencialmente distribuída com parâmetro  $\lambda$  e que a lâmpada que sobrou tenha vida útil exponencialmente distribuída com parâmetro  $\theta$  (é feita por outro fabricante). Qual é a probabilidade de exatamente cinco das 10 lâmpadas falharem antes do tempo  $t$ ?
15. Considere um sistema formado por três componentes, conforme ilustração. O sistema continuará funcionando desde que o primeiro componente funcione e que o componente 2 ou o 3 funcionem também. Represente por  $X_1, X_2$  e  $X_3$  os tempos de vida dos componentes 1, 2 e 3, respectivamente. Suponha que os  $X_i$ s sejam independentes entre si e que cada  $X_i$  possua uma distribuição exponencial com parâmetro  $\lambda$ .



- Seja  $Y$  = tempo de vida do sistema. Obtenha a função de distribuição acumulada de  $Y$  e calcule a diferença para obter a fdp. [Sugestão:  $F(y) = P(Y \leq y)$ ; expresse o evento  $\{Y \leq y\}$  em termos de uniões e/ou interseções dos três eventos  $\{X_1 \leq y\}$ ,  $\{X_2 \leq y\}$ , e  $\{X_3 \leq y\}$ .]
  - Calcule a vida útil esperada do sistema.
16. a. Para  $f(x_1, x_2, x_3)$ , como dado no Exemplo 5.10, calcule a **função de densidade marginal conjunta** de  $X_1$  e somente de  $X_3$  (integrando em  $x_2$ ).

- b. Qual é a probabilidade de que as rochas dos tipos 1 e 3 juntas formem no máximo 50% da amostra? [Sugestão: use o resultado do item (a).]
- c. Calcule a fdp marginal apenas de  $X_1$ . [Sugestão: use o resultado do item (a).]
17. Um ecologista deseja selecionar um ponto dentro de uma região circular de amostragem, de acordo com uma distribuição uniforme (na prática, isso poderia ser feito selecionando-se primeiro uma direção e, depois, uma distância do centro naquela direção). Seja  $X =$  coordenada  $x$  do ponto selecionado e  $Y =$  coordenada  $y$  desse ponto. Se o círculo tiver centro em  $(0, 0)$  e raio  $R$ , então a fdp conjunta de  $X$  e  $Y$  é
- $$f(x, y) = \begin{cases} \frac{1}{\pi R^2} & x^2 + y^2 \leq R^2 \\ 0 & \text{caso contrário} \end{cases}$$
- a. Qual é a probabilidade de o ponto selecionado estar dentro de  $R/2$  do centro da região circular? [Sugestão: desenhe a região de densidade positiva  $D$ . Já que  $f(x, y)$  é constante em  $D$ , calcular a probabilidade resume-se em calcular uma área.]
- b. Qual é a probabilidade de  $X$  e  $Y$  serem diferentes de 0 no máximo  $R/2$ ?
- c. Responda ao item (b) com  $R/\sqrt{2}$  substituindo  $R/2$ .
- d. Qual é a fdp marginal de  $X$ ? De  $Y$ ?  $X$  e  $Y$  são independentes?
18. Refira-se ao Exercício 1 e responda as seguintes perguntas:
- a. Dado que  $X = 1$ , determine a fmp condicional de  $Y$ , isto é,  $p_{Y|X}(0 | 1)$ ,  $p_{Y|X}(1 | 1)$ , e  $p_{Y|X}(2 | 1)$ .
- b. Dado que duas mangueiras estão sendo utilizadas na ilha de auto-serviço, qual é a fmp condicional do número de mangueiras em uso na ilha de serviço completo?
- c. Use o resultado do item (b) para calcular a probabilidade condicional  $P(Y \leq 1 | X = 2)$ .
- d. Dado que duas mangueiras estão sendo utilizadas na ilha de serviço completo, qual é a fmp condicional do número em uso na ilha de auto-serviço?
19. A fdp conjunta de pressões dos pneus dianteiros é dada no Exercício 9.
- a. Determine a fdp condicional de  $Y$ , dado que  $X = x$ , e a fdp condicional de  $X$ , dado que  $Y = y$ .
- b. Se a pressão no pneu direito for 22 psi, qual é a probabilidade de o pneu esquerdo ter uma pressão de no mínimo 25 psi? Compare com  $P(Y \geq 25)$ .
- c. Se a pressão no pneu direito for 22 psi, qual é a pressão esperada no pneu esquerdo e qual é o desvio padrão da pressão desse pneu?
20. Sejam  $X_1, X_2,$  e  $X_3$  as vidas úteis dos componentes 1, 2 e 3 em um sistema de três componentes.
- a. Como você definiria a fdp condicional de  $X_3$ , dado que  $X_1 = x_1$  e  $X_2 = x_2$ ?
- b. Como você definiria a fdp condicional conjunta de  $X_2$  e  $X_3$  dado que  $X_1 = x_1$ ?
21. Que condição de  $f_{Y|X}(y|x)$  equivale à independência de  $X$  e  $Y$ ?

## 5.2 | Valores Esperados, Covariância e Correlação

Vimos anteriormente que qualquer função  $h(X)$  de uma única va  $X$  é por si mesma uma variável aleatória. Entretanto, para calcular  $E[h(X)]$ , não foi necessário obter a distribuição de probabilidades de  $h(X)$ ; ao contrário,  $E[h(X)]$  foi calculado como uma média ponderada dos valores de  $h(x)$ , em que a função do peso foi a fmp  $p(x)$  ou a fdp  $f(x)$  de  $X$ . Um resultado semelhante é obtido para a função  $h(X, Y)$  de duas variáveis aleatórias de distribuição conjunta.

### PROPOSIÇÃO

Sejam  $X$  e  $Y$  as vas de distribuição conjunta com fmp  $p(x, y)$  ou fdp  $f(x, y)$ , conforme as variáveis sejam discretas ou contínuas. Então, o valor esperado de uma função  $h(X, Y)$ , representada por  $E[h(X, Y)]$  ou  $\mu_{h(X, Y)}$ , é dado por

$$E[h(X, Y)] = \begin{cases} \sum_x \sum_y h(x, y) \cdot p(x, y) & \text{se } X \text{ e } Y \text{ forem discretos} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) \cdot f(x, y) dx dy & \text{se } X \text{ e } Y \text{ forem contínuos} \end{cases}$$

**Exemplo 5.13**

Cinco amigos compraram ingressos para um determinado show. Se os ingressos forem dos lugares de 1-5 em certa fileira, e os ingressos forem distribuídos aleatoriamente entre os cinco, qual é o número esperado de assentos que separam quaisquer dois dos cinco? Sejam  $X$  e  $Y$  os números dos lugares da primeira e da segunda pessoas, respectivamente. Possíveis pares de  $(X, Y)$  são  $\{(1, 2), (1, 3), \dots, (5, 4)\}$ , e a fmp conjunta de  $(X, Y)$  é

$$p(x, y) = \begin{cases} \frac{1}{20} & x = 1, \dots, 5; y = 1, \dots, 5; x \neq y \\ 0 & \text{caso contrário} \end{cases}$$

O número de lugares que separam as duas pessoas é  $h(X, Y) = |X - Y| - 1$ . A tabela a seguir fornece  $h(x, y)$  para cada par  $(x, y)$  possível.

$h(x, y)$		$x$				
		1	2	3	4	5
$y$	1	—	0	1	2	3
	2	0	—	0	1	2
	3	1	0	—	0	1
	4	2	1	0	—	0
	5	3	2	1	0	—

Dessa forma,

$$E[h(X, Y)] = \sum_{(x, y)} h(x, y) \cdot p(x, y) = \sum_{\substack{x=1 \\ x \neq y}}^5 \sum_{y=1}^5 (|x - y| - 1) \cdot \frac{1}{20} = 1 \quad \blacksquare$$

**Exemplo 5.14**

No Exemplo 5.5, a fdp conjunta da quantidade  $X$  de amêndoas e da quantidade  $Y$  de castanhas de caju em uma lata de castanhas de 1 libra foi

$$f(x, y) = \begin{cases} 24xy & 0 \leq x \leq 1, 0 \leq y \leq 1, x + y \leq 1 \\ 0 & \text{caso contrário} \end{cases}$$

Se 1 libra de amêndoas custar à empresa US\$ 1,00, 1 libra de castanhas de caju custar US\$ 1,50 e 1 libra de amendoins custar US\$ 0,50, então o custo total do conteúdo de uma lata é

$$h(X, Y) = (1)X + (1,5)Y + (0,5)(1 - X - Y) = 0,5 + 0,5X + Y$$

(uma vez que  $1 - X - Y$  do peso consiste em amendoins). O custo total esperado é

$$\begin{aligned} E[h(X, Y)] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) \cdot f(x, y) \, dx \, dy \\ &= \int_0^1 \int_0^{1-x} (0,5 + 0,5x + y) \cdot 24xy \, dy \, dx = \text{US\$ } 1,10 \quad \blacksquare \end{aligned}$$

O método para calcular o valor esperado de uma função  $h(X_1, \dots, X_n)$  de  $n$  variáveis aleatórias é semelhante ao método usado para duas variáveis aleatórias. Se os  $X_i$ s forem discretos,  $E[h(X_1, \dots, X_n)]$  é uma soma  $n$ -dimensional; se os  $X_i$ s forem contínuos, é uma integral  $n$ -dimensional.

**Covariância**

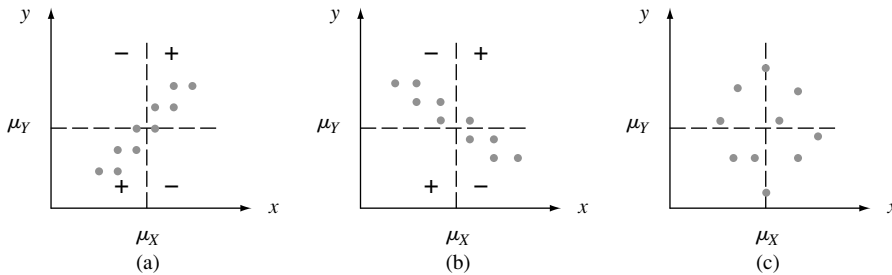
Quando duas variáveis aleatórias  $X$  e  $Y$  não são independentes, geralmente é de interesse avaliar quão fortemente estão relacionadas uma com a outra.

DEFINIÇÃO

A **covariância** entre duas vas  $X$  e  $Y$  é

$$\begin{aligned} \text{Cov}(X, Y) &= E[(X - \mu_X)(Y - \mu_Y)] \\ &= \begin{cases} \sum_x \sum_y (x - \mu_X)(y - \mu_Y)p(x, y) & X, Y \text{ discretas} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y)f(x, y) dx dy & X, Y \text{ contínuas} \end{cases} \end{aligned}$$

O fundamento lógico da definição é o seguinte. Suponha que  $X$  e  $Y$  tenham uma relação positiva forte entre si, pela qual queremos dizer que valores grandes de  $X$  tendem a ocorrer com valores grandes de  $Y$ , e valores pequenos de  $X$ , com valores pequenos de  $Y$ . Então, a maior parte da massa ou densidade de probabilidade estará associada a  $(x - \mu_X)$  e  $(y - \mu_Y)$ , ambos positivos ( $X$  e  $Y$  acima de suas respectivas médias) ou ambos negativos, de modo que o produto  $(x - \mu_X)(y - \mu_Y)$  tende a ser positivo. Dessa forma, para uma relação positiva forte,  $\text{Cov}(X, Y)$  deve ser positiva. Para uma relação negativa forte, os sinais de  $(x - \mu_X)$  e  $(y - \mu_Y)$  tenderão a ser opostos, levando a um produto negativo. Assim, para uma relação negativa forte,  $\text{Cov}(X, Y)$  deve ser negativa. Se  $X$  e  $Y$  não estão fortemente relacionadas, os produtos positivo e negativo tenderão a cancelar um ao outro, produzindo uma covariância próxima de 0. A Figura 5.4 ilustra as diferentes possibilidades. A covariância depende tanto do conjunto de pares possíveis como das probabilidades. Na Figura 5.4, as probabilidades poderiam ser trocadas sem alterar o conjunto de pares possíveis, o que pode mudar drasticamente o valor da  $\text{Cov}(X, Y)$ .



**Figura 5.4**  $p(x, y) = 1/10$  para cada um dos 10 pares correspondentes aos pontos indicados, (a) covariância positiva; (b) covariância negativa; (c) covariância próxima de zero

**Exemplo 5.15**

As fmps marginal e conjunta de  $X =$  quantidade dedutível das apólices de seguro de automóvel e  $Y =$  quantidade dedutível das apólices de seguro residencial no Exemplo 5.1 eram

		$y$				$x$			$y$		
	$p(x, y)$	0	100	200		100	250		0	100	200
$x$	100	0,20	0,10	0,20	$p_X(x)$	0,5	0,5	$p_Y(y)$	0,25	0,25	0,5
	250	0,05	0,15	0,30							

onde  $\mu_X = \sum xp_X(x) = 175$  e  $\mu_Y = 125$ . Portanto,

$$\begin{aligned} \text{Cov}(X, Y) &= \sum_{(x, y)} (x - 175)(y - 125)p(x, y) \\ &= (100 - 175)(0 - 125)(0,20) + \dots \\ &\quad + (250 - 175)(200 - 125)(0,30) \\ &= 1875 \end{aligned}$$





A fórmula simplificada a seguir para  $\text{Cov}(X, Y)$  facilita os cálculos.

PROPOSIÇÃO

$$\text{Cov}(X, Y) = E(XY) - \mu_X \cdot \mu_Y$$

De acordo com essa fórmula, não é necessária nenhuma subtração intermediária; somente no final do cálculo  $\mu_X \cdot \mu_Y$  é subtraído de  $E(XY)$ . A demonstração envolve expandir  $(X - \mu_X)(Y - \mu_Y)$  e depois calcular o valor esperado de cada termo separadamente. Observe que  $\text{Cov}(X, X) = E(X^2) - \mu_X^2 = V(X)$ .

**Exemplo 5.16** (continuação do Exemplo 5.5)

As fdps marginal e conjunta de  $X =$  quantidade de amêndoas e de  $Y =$  quantidade de castanhas de caju eram

$$f(x, y) = \begin{cases} 24xy & 0 \leq x \leq 1, 0 \leq y \leq 1, x + y \leq 1 \\ 0 & \text{caso contrário} \end{cases}$$

$$f_X(x) = \begin{cases} 12x(1-x)^2 & 0 \leq x \leq 1 \\ 0 & \text{caso contrário} \end{cases}$$

$f_Y(y)$  é obtido, substituindo  $x$  por  $y$  em  $f_X(x)$ . É fácil demonstrar que  $\mu_X = \mu_Y = \frac{2}{5}$ , e

$$\begin{aligned} E(XY) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x, y) dx dy = \int_0^1 \int_0^{1-x} xy \cdot 24xy dy dx \\ &= 8 \int_0^1 x^2(1-x)^3 dx = \frac{2}{15} \end{aligned}$$

Assim,  $\text{Cov}(X, Y) = \frac{2}{15} - \left(\frac{2}{5}\right)\left(\frac{2}{5}\right) = \frac{2}{15} - \frac{4}{25} = -\frac{2}{75}$ . A covariância negativa, neste caso, é razoável, porque uma quantidade maior de amêndoas na lata implica uma quantidade menor de castanhas de caju. ■

Pode parecer que a relação no exemplo do seguro seja bastante forte, uma vez que  $\text{Cov}(X, Y) = 1875$ , enquanto  $\text{Cov}(X, Y) = -\frac{2}{75}$ , no exemplo das castanhas, parece uma relação fraca. Infelizmente, a covariância possui uma deficiência grave que impossibilita a interpretação de um valor calculado. No exemplo do seguro, suponha que tivéssemos expressado a quantia dedutível em centavos e não em dólares. Então,  $100X$  substituiria  $X$ ,  $100Y$  substituiria  $Y$ , e a covariância resultante seria  $\text{Cov}(100X, 100Y) = (100)(100)\text{Cov}(X, Y) = 18.750.000$ . Se, por outro lado, a quantia dedutível tivesse sido expressa em centenas de dólares, a covariância calculada teria sido  $(0,01)(0,01)(1875) = 0,1875$ . A *deficiência da covariância é que seu valor calculado depende diretamente das unidades de medida*. Teoricamente, a escolha das unidades não deve afetar a medida da intensidade da relação. O que é obtido definindo-se a escala da covariância.

## Correlação

DEFINIÇÃO

O **coeficiente de correlação** de  $X$  e  $Y$ , representado por  $\text{Corr}(X, Y)$ ,  $\rho_{X,Y}$ , ou simplesmente  $\rho$ , é definido por

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sigma_X \cdot \sigma_Y}$$

**Exemplo 5.17**

É fácil verificar que no problema do seguro, do Exemplo 5.15,  $E(X^2) = 36.250$ ,  $\sigma_x^2 = 36.250 - (175)^2 = 5625$ ,  $\sigma_x = 75$ ,  $E(Y^2) = 22.500$ ,  $\sigma_y^2 = 6875$ , e  $\sigma_y = 82,92$ . O que dá

$$\rho = \frac{1875}{(75)(82,92)} = 0,301 \quad \blacksquare$$

A proposição a seguir mostra que  $\rho$  corrige a deficiência da  $\text{Cov}(X, Y)$ , além de sugerir como reconhecer a existência de uma relação (linear) forte.

**PROPOSIÇÃO**

1. Se  $a$  e  $c$  são ambos positivos ou negativos,

$$\text{Corr}(aX + b, cY + d) = \text{Corr}(X, Y)$$

2. Para quaisquer duas vas  $X$  e  $Y$ ,  $-1 \leq \text{Corr}(X, Y) \leq 1$ .

A proposição 1 diz precisamente que o coeficiente de correlação não é afetado por mudança linear das unidades de medida (se, por exemplo,  $X =$  temperatura em  $^{\circ}\text{C}$ , então,  $9X/5 + 32 =$  temperatura em  $^{\circ}\text{F}$ ). De acordo com a proposição 2, a relação positiva mais forte possível é evidenciada por  $\rho = +1$ , enquanto a relação negativa mais forte possível corresponde a  $\rho = -1$ . A demonstração da primeira proposição está esquematizada no Exercício 35, e a da segunda aparece no Exercício Suplementar 87, no final do capítulo. Para fins descritivos, a relação será descrita como forte se  $|\rho| \geq 0,8$ , moderada se  $0,5 < |\rho| < 0,8$ , e fraca se  $|\rho| \leq 0,5$ .

Se pensarmos que  $p(x, y)$  ou  $f(x, y)$  determinam um modelo matemático para a forma como as duas variáveis numéricas  $X$  e  $Y$  são distribuídas em determinada população (altura e peso, pontuação SAT verbal e pontuação SAT quantitativa etc.), então,  $\rho$  é uma característica ou parâmetro da população que mede o quão fortemente  $X$  e  $Y$  estão relacionados na população. No Capítulo 12, consideraremos uma amostra de pares  $(x_1, y_1), \dots, (x_n, y_n)$  da população. O coeficiente de correlação da amostra  $r$  será, então, definido e usado para fazer inferências sobre  $\rho$ .

O coeficiente de correlação  $\rho$  não é, na verdade, uma medida geral da força de uma relação.

**PROPOSIÇÃO**

1. Se  $X$  e  $Y$  são independentes, então  $\rho = 0$ , porém  $\rho = 0$  não implica independência.
2.  $\rho = 1$  ou  $-1$  se e somente se  $Y = aX + b$  para quaisquer números  $a$  e  $b$  com  $a \neq 0$ .

Essa proposição diz que  $\rho$  é uma medida do grau da relação **linear** entre  $X$  e  $Y$ , e somente quando as duas variáveis estiverem perfeitamente relacionadas de forma linear é que  $\rho$  assumirá os valores extremos positivo ou negativo. Um  $\rho$  menor que 1 em valor absoluto indica somente que a relação não é completamente linear, mas que ainda pode haver uma relação não-linear bastante forte. Além disso,  $\rho = 0$  não implica que  $X$  e  $Y$  sejam independentes, mas apenas que há ausência completa de relação linear. Quando  $\rho = 0$ ,  $X$  e  $Y$  são ditos **não-correlacionados**. Duas variáveis podem ser não-correlacionadas, porém altamente dependentes, pois pode existir uma relação não-linear forte; assim, cuidado para não tirar muitas conclusões a partir de  $\rho = 0$ .

**Exemplo 5.18**

Sejam  $X$  e  $Y$  vas discretas com fmp conjunta

$$p(x, y) = \begin{cases} \frac{1}{4} & (x, y) = (-4, 1), (4, -1), (2, 2), (-2, -2) \\ 0 & \text{caso contrário} \end{cases}$$

Os pontos que recebem massa de probabilidade positiva estão identificados no sistema de coordenadas  $(x, y)$  na Figura 5.5. É evidente, pela figura, que o valor de  $X$  é determinado completamente pelo valor de  $Y$  e vice-versa, de modo que as duas variáveis são totalmente dependentes. Entretanto, pela simetria  $\mu_X = \mu_Y = 0$  e  $E(XY) = (-4)\frac{1}{4} + (-4)\frac{1}{4} + (4)\frac{1}{4} + (4)\frac{1}{4} = 0$ , assim,  $\text{Cov}(X, Y) = E(XY) - \mu_X \cdot \mu_Y = 0$  e dessa forma  $\rho_{X,Y} = 0$ . Embora haja perfeita dependência, também há ausência completa de qualquer relação linear!

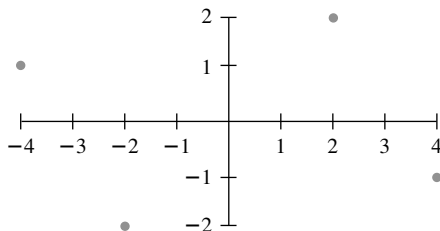


Figura 5.5 A população de pares do Exemplo 5.18

Um valor de  $\rho$  próximo de 1 não implica necessariamente que aumentar o valor de  $X$  cause um aumento em  $Y$ . Implica somente que valores grandes de  $X$  estão associados a valores grandes de  $Y$ . Por exemplo: na população de crianças, o tamanho do vocabulário e o número de cáries são correlacionados de forma bastante positiva, mas certamente não é verdade que as cáries façam o vocabulário aumentar. Ao contrário, os valores de ambas as variáveis tendem a aumentar com a idade das crianças, uma terceira variável. Para as crianças de certa idade, provavelmente haja correlação muito baixa entre o número de cáries e o tamanho do vocabulário. Resumindo, associação (uma alta correlação) não é a mesma coisa que causa.

**Exercícios** | Seção 5.2 (22–36)

22. Um instrutor fez perguntas curtas que consistem de duas partes. Para um estudante selecionado aleatoriamente, seja  $X =$  o número de pontos ganhos na primeira parte e  $Y =$  o número de pontos ganhos na segunda. Suponha que a fmp conjunta de  $X$  e  $Y$  seja dada na tabela a seguir.

$(x, y)$		$y$			
		0	5	10	15
$x$	0	0,02	0,06	0,02	0,10
	5	0,04	0,15	0,20	0,10
	10	0,01	0,15	0,14	0,01

- a. Se a pontuação constante do boletim for o número total de pontos ganhos nas duas partes, qual é a pontuação registrada esperada  $E(X + Y)$ ?
  - b. Se o máximo das duas pontuações for registrado, qual é a pontuação registrada esperada?
23. A diferença entre o número de clientes na fila da caixa expressa e o número de clientes na caixa superexpressa no Exercício 3 é  $X_1 - X_2$ . Calcule a diferença esperada.
24. Seis indivíduos, incluindo A e B, sentam-se ao redor de uma mesa redonda de forma totalmente aleatória. Suponha que os lugares sejam numerados de 1, ..., 6. Seja  $X =$  o número do assento de A e  $Y =$  o número do assento de B. Se A mandar um bilhete, ao redor da mesa,

para B, na direção em que estão mais próximos, quantos indivíduos (incluindo A e B) você espera que peguem o bilhete?

25. Um agrimensor deseja traçar uma região quadrada cujos lados têm comprimento  $L$ . Entretanto, devido a erro de medição, ele esboça um retângulo, em que os lados norte e sul têm comprimento  $X$ , e os lados leste e oeste, comprimento  $Y$ . Suponha que  $X$  e  $Y$  sejam independentes e que cada um seja distribuído uniformemente no intervalo  $[L - A, L + A]$  (onde  $0 < A < L$ ). Qual é a área esperada do retângulo resultante?
26. Considere uma pequena balsa que possa acomodar carros e ônibus. O pedágio para carros são US\$ 3, e para ônibus, US\$ 10. Sejam  $X$  e  $Y$  o número de carros e ônibus, respectivamente, levado em uma única viagem. Suponha que a distribuição conjunta de  $X$  e  $Y$  seja aquela dada na tabela do Exercício 7. Calcule a receita esperada de uma única viagem.
27. Annie e Alvie combinaram encontrar-se para almoçar juntos entre meio-dia e 13h. Represente por  $X$  a hora de chegada de Annie, por  $Y$  a de Alvie, e suponha que  $X$  e  $Y$  sejam independentes com fdps

$$f_X(x) = \begin{cases} 3x^2 & 0 \leq x \leq 1 \\ 0 & \text{caso contrário} \end{cases}$$

$$f_Y(y) = \begin{cases} 2y & 0 \leq y \leq 1 \\ 0 & \text{caso contrário} \end{cases}$$

Qual é o tempo esperado que a pessoa que chega primeiro deve aguardar pela outra? [*Sugestão*:  $h(X, Y) = |X - Y|$ .]

28. Mostre que, se  $X$  e  $Y$  são vas independentes, então  $E(XY) = E(X) \cdot E(Y)$ . Aplique essa expressão no Exercício 25. [*Sugestão*: considere o caso contínuo com  $f(x, y) = f_X(x) \cdot f_Y(y)$ .]
29. Calcule o coeficiente de correlação  $\rho$  de  $X$  e  $Y$  do Exemplo 5.16 (a covariância já foi calculada).
30. a. Calcule a covariância de  $X$  e  $Y$  no Exercício 22.  
b. Calcule  $\rho$  de  $X$  e  $Y$  no mesmo exercício.
31. a. Calcule a covariância entre  $X$  e  $Y$  no Exercício 9.  
b. Calcule o coeficiente de correlação  $\rho$  para  $X$  e  $Y$ .
32. Reconsidere as vidas úteis  $X$  e  $Y$  do componente do mini-computador, como descrito no Exercício 12. Determine  $E(XY)$ . O que se pode dizer sobre  $\text{Cov}(X, Y)$  e  $\rho$ ?
33. Use o resultado do Exercício 28 para mostrar que, quando  $X$  e  $Y$  são independentes,  $\text{Cov}(X, Y) = \text{Corr}(X, Y) = 0$ .
34. a. Lembrando-se da definição de  $\sigma^2$  para uma única va  $X$ , escreva uma fórmula apropriada para calcular a variância da função  $h(X, Y)$  de duas variáveis aleatórias. [*Sugestão*: lembre-se de que a variância é apenas um valor esperado especial.]  
b. Use essa fórmula para calcular a variância da pontuação registrada  $h(X, Y) [= \max(X, Y)]$  na parte (b) do Exercício 22.
35. a. Use as regras de valor esperado para demonstrar que  $\text{Cov}(aX + b, cY + d) = ac \text{Cov}(X, Y)$ .  
b. Use a parte (a) junto com as regras de variância e desvio padrão para demonstrar que  $\text{Corr}(aX + b, cY + d) = \text{Corr}(X, Y)$  quando  $a$  e  $c$  têm o mesmo sinal.  
c. O que acontece se  $a$  e  $c$  tiverem sinais contrários?
36. Mostre que, se  $Y = aX + b$  ( $a \neq 0$ ), então  $\text{Corr}(X, Y) = +1$  ou  $-1$ . Sob quais condições  $\rho = +1$ ?

## 5.3 | Estatísticas e Suas Distribuições

As observações de uma única amostra foram denotadas no Capítulo 1 por  $x_1, x_2, \dots, x_n$ . Considere a seleção de duas amostras diferentes de tamanho  $n$  oriundas da mesma distribuição da população. Os  $x_i$ s da segunda amostra serão sempre virtualmente diferentes, pelo menos um pouco, dos  $x_i$ s da primeira amostra. Por exemplo: uma primeira amostra de  $n = 3$  carros de um tipo específico pode resultar em consumos de combustível  $x_1 = 30,7$ ,  $x_2 = 29,4$ ,  $x_3 = 31,1$ , enquanto uma segunda amostra pode fornecer  $x_1 = 28,8$ ,  $x_2 = 30,0$  e  $x_3 = 31,1$ . Antes de obtermos os dados, há incerteza acerca do valor de cada  $x_i$ . Devido a essa incerteza, *antes* de os dados estarem disponíveis, consideramos cada observação como uma variável aleatória e denotamos a amostra por  $X_1, X_2, \dots, X_n$  (letras maiúsculas para variáveis aleatórias).

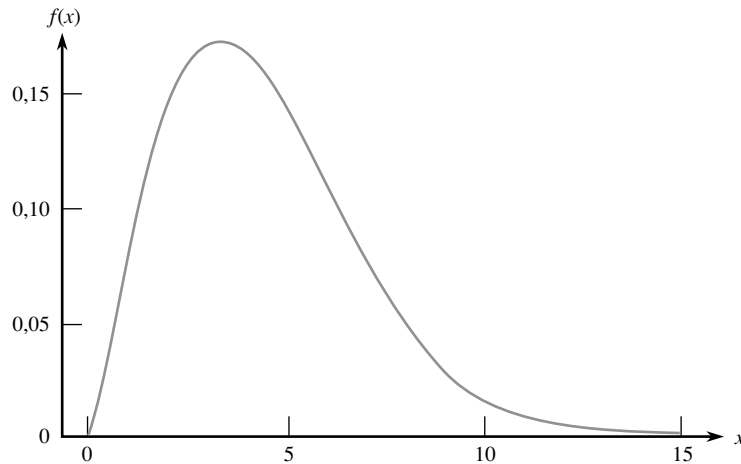
Essa variação nos valores observados implica, por sua vez, que o valor de qualquer função das observações da amostra – como a sua média, o seu desvio padrão ou a sua quarta dispersão – também varia de amostra para amostra. Isto é, antes de obter  $x_1, \dots, x_n$ , há incerteza quanto ao valor de  $\bar{x}$ , o valor de  $s$ , e assim por diante.

### Exemplo 5.19

Suponha que a resistência do material de um tipo específico de espécime selecionado aleatoriamente possui uma distribuição de Weibull com valores de parâmetro  $\alpha = 2$  (formato) e  $\beta = 5$  (escala). A curva de densidade correspondente é exibida na Figura 5.6. As fórmulas da Seção 4.5 fornecem

$$\mu = E(X) = 4,4311 \quad \tilde{\mu} = 4,1628 \quad \sigma^2 = V(X) = 5,365 \quad \sigma = 2,316$$

A média excede a mediana devido à inclinação positiva da distribuição.



**Figura 5.6** A curva de densidade de Weibull do Exemplo 5.19

Usamos MINITAB para gerar seis amostras diferentes, cada uma com  $n = 10$  a partir dessa distribuição (resistências do material para seis grupos diferentes de 10 espécimes cada). Os resultados são exibidos na Tabela 5.1, seguidos pelos valores da média, mediana e desvio padrão de cada uma das amostras. Repare primeiro que as 10 observações em qualquer amostra específica são diferentes das observações em qualquer outra amostra. Em segundo lugar, os seis valores da média amostral são diferentes uns dos outros, como os seis valores da mediana e do desvio padrão da amostra. O mesmo acontece com as médias trincadas de 10% da amostra, com suas quartas dispersões da amostra e assim por diante.

**Tabela 5.1** Amostras da distribuição de Weibull do Exemplo 5.19

Amostra	1	2	3	4	5	6
1	6,1171	5,07611	3,46710	1,55601	3,12372	8,93795
2	4,1600	6,79279	2,71938	4,56941	6,09685	3,92487
3	3,1950	4,43259	5,88129	4,79870	3,41181	8,76202
4	0,6694	8,55752	5,14915	2,49759	1,65409	7,05569
5	1,8552	6,82487	4,99635	2,33267	2,29512	2,30932
6	5,2316	7,39958	5,86887	4,01295	2,12583	5,94195
7	2,7609	2,14755	6,05918	9,08845	3,20938	6,74166
8	10,2185	8,50628	1,80119	3,25728	3,23209	1,75468
9	5,2438	5,49510	4,21994	3,70132	6,84426	4,91827
10	4,5590	4,04525	2,12934	5,50134	4,20694	7,26081
$\bar{x}$	4,401	5,928	4,229	4,132	3,620	5,761
$\tilde{x}$	4,360	6,144	4,608	3,857	3,221	6,342
$s$	2,642	2,062	1,611	2,124	1,678	2,496

Além disso, o valor da média amostral de qualquer amostra específica pode ser considerado uma *estimativa por pontos* (“ponto” porque é um único número, correspondente a um único ponto na linha dos números) da média da população  $\mu$ , cujo valor conhecido é 4,4311. Nenhuma das estimativas dessas seis amostras é idêntica à que está sendo estimada. As estimativas da segunda e da sexta amostras são muito grandes, enquanto a quinta amostra fornece uma subestimativa substancial. De forma semelhante, o desvio padrão da amostra fornece uma estimativa por ponto do desvio padrão da população. As seis estimativas resultantes estão erradas em pelo menos uma pequena quantidade.

Resumindo, os valores das observações individuais variam de amostra para amostra. Dessa forma, geralmente, o valor de qualquer quantidade calculada a partir dos dados amostrais e o valor de uma amostra característica usada como estimativa da característica da população correspondente, virtualmente, jamais coincidirão com o que está sendo estimado. ■

## DEFINIÇÃO

**Estatística** é qualquer quantidade cujo valor possa ser calculado com base nos dados da amostra. Antes de obter os dados, existe a incerteza quanto ao valor que resultará de qualquer estatística específica. Portanto, estatística é uma variável aleatória e será representada por uma letra maiúscula; uma letra minúscula será usada para representar o valor calculado ou observado da estatística.

Assim, a média amostral, considerada como estatística (antes de a amostra ser selecionada ou de o experimento ter sido feito), é representada por  $\bar{X}$ ; o valor calculado dessa estatística é  $\bar{x}$ . De forma semelhante,  $S$  representa o desvio padrão da amostra visto como estatística e seu valor calculado é  $s$ . Se amostras de dois tipos diferentes de tijolos forem selecionadas e as resistências à compressão individuais forem denotadas por  $X_1, \dots, X_m$  e  $Y_1, \dots, Y_n$ , respectivamente, então a estatística  $\bar{X} - \bar{Y}$ , diferença entre as duas resistências à compressão média amostral, geralmente é de grande interesse.

Qualquer estatística, sendo uma variável aleatória, possui uma distribuição de probabilidades. Em particular, a média amostral  $\bar{X}$  possui uma distribuição de probabilidades. Suponha, por exemplo, que  $n = 2$  componentes são selecionados aleatoriamente, e o número de quebras durante o período de garantia é determinado para cada um. Valores possíveis para o número médio de quebras  $\bar{X}$  da amostra são 0 (se  $X_1 = X_2 = 0$ ), 0,5 (se  $X_1 = 0$  e  $X_2 = 1$  ou  $X_1 = 1$  e  $X_2 = 0$ ), 1, 1,5, ...). A distribuição de probabilidade de  $\bar{X}$  específica  $P(\bar{X} = 0)$ ,  $P(\bar{X} = 0,5)$ , e assim por diante, a partir das quais outras probabilidades, como  $P(1 \leq \bar{X} \leq 3)$  e  $P(\bar{X} \geq 2,5)$  podem ser calculadas. De forma semelhante, se, para uma amostra de tamanho  $n = 2$ , os únicos valores possíveis da variância da amostra forem 0, 12,5 e 50 (que será o caso se  $X_1$  e  $X_2$  puderem adotar somente os valores 40, 45 e 50), então a distribuição de probabilidades de  $S^2$  fornece  $P(S^2 = 0)$ ,  $P(S^2 = 12,5)$  e  $P(S^2 = 50)$ . A distribuição de probabilidade de uma estatística às vezes é denominada **distribuição de amostragem** para enfatizar que ela descreve como a estatística varia de valor em todas as amostras que possam ser selecionadas.

## Amostras Aleatórias

A distribuição de probabilidade de qualquer estatística específica depende não somente da distribuição da população (normal, uniforme etc.) e do tamanho da amostra  $n$ , como também do método de amostragem. Considere a seleção de uma amostra de tamanho  $n = 2$  de uma população, formada apenas dos três valores 1, 5 e 10, e suponha que a estatística de interesse seja a variância da amostra. Se a amostragem é feita “com reposição”, então,  $S^2 = 0$  será o resultado se  $X_1 = X_2$ . Entretanto,  $S^2$  não pode ser igual a 0 se a amostragem for “sem reposição”. Assim,  $P(S^2 = 0) = 0$  para um método de amostragem, e essa probabilidade é positiva para o outro método. Nossa próxima definição descreve um método de amostragem geralmente encontrado (pelo menos aproximadamente) na prática.

## DEFINIÇÃO

Diz-se que as v.a.s  $X_1, X_2, \dots, X_n$  formam uma **amostra aleatória** (simples) de tamanho  $n$  se

1. Os  $X_i$ s são v.a.s independentes.
2. Todo  $X_i$  possui a mesma distribuição de probabilidades.

As Condições 1 e 2 podem ser parafraseadas, dizendo-se que os  $X_i$ s são *independente e identicamente distribuídos* (iid). Se a amostragem for com reposição ou de uma população (hipotética) infinita, as Condições 1 e 2 são satisfeitas exatamente. Tais condições serão satisfeitas aproximadamente se a amostragem for sem reposição, ainda que o tamanho da amostra  $n$  seja muito menor que o tamanho da população  $N$ . Na prática, se  $n/N \leq 0,05$  (no máximo 5% da população é coletada), podemos prosseguir como se os  $X_i$ s formassem uma amostra aleatória. A virtude desse método de amostragem é que a distribuição de probabilidade de qualquer estatística pode ser obtida com mais facilidade do que qualquer outro método de amostragem.

Existem dois métodos gerais para se obter informações sobre a distribuição da amostragem da estatística. Um método envolve cálculos com base nas regras de probabilidade; o outro envolve a realização de um experimento de simulação.

### Dedução da Distribuição da Amostragem de uma Estatística

As regras de probabilidade são usadas para obter a distribuição de uma estatística, contanto que seja uma função “bem simples” dos  $X_i$ s e que haja relativamente poucos valores de  $X$  diferentes na população; ou então, que a distribuição da população possua uma forma “apropriada”. Nossos dois próximos exemplos ilustram tais situações.

#### Exemplo 5.20

Um grande centro de manutenção de automóveis cobra US\$ 40, US\$ 45 e US\$ 50 para regulagem de carros de quatro, seis e oito cilindros, respectivamente. Se 20% das regulagens são feitas nos carros de quatro cilindros, 30% nos carros de seis cilindros e 50% nos carros de oito cilindros, então a distribuição de probabilidade da receita de uma única regulagem selecionada aleatoriamente é dada por

$x$	40	45	50	com $\mu = 46,5$ , $\sigma^2 = 15,25$
$p(x)$	0,2	0,3	0,5	

(5.2)

Suponha que, em certo dia, somente dois trabalhos de manutenção envolvam regulagem. Seja  $X_1 =$  receita da primeira regulagem e  $X_2 =$  a receita da segunda. Suponha que  $X_1$  e  $X_2$  sejam independentes, cada uma com a distribuição de probabilidades mostrada em (5.2) [de modo que  $X_1$  e  $X_2$  constituam uma amostra aleatória da distribuição (5.2)]. A Tabela 5.2 lista os pares possíveis  $(x_1, x_2)$ , a probabilidade de cada um [calculada usando-se (5.2) e a hipótese de independência], e os valores resultantes de  $\bar{x}$  e  $s^2$ . Para obter a distribuição de probabilidade de  $\bar{X}$ , a receita média amostral por regulagem, devemos considerar cada valor possível de  $\bar{x}$  e calcular sua probabilidade. Por exemplo:  $\bar{x} = 45$  ocorre três vezes na tabela com probabilidades 0,10, 0,09 e 0,10, assim

$$p_{\bar{x}}(45) = P(\bar{X} = 45) = 0,10 + 0,09 + 0,10 = 0,29$$

**Tabela 5.2** Resultados, probabilidades e valores de  $\bar{x}$  e  $s^2$  do Exemplo 5.20

$x_1$	$x_2$	$p(x_1, x_2)$	$\bar{x}$	$s^2$
40	40	0,04	40	0
40	45	0,06	42,5	12,5
40	50	0,10	45	50
45	40	0,06	42,5	12,5
45	45	0,09	45	0
45	50	0,15	47,5	12,5
50	40	0,10	45	50
50	45	0,15	47,5	12,5
50	50	0,25	50	0

De forma similar,

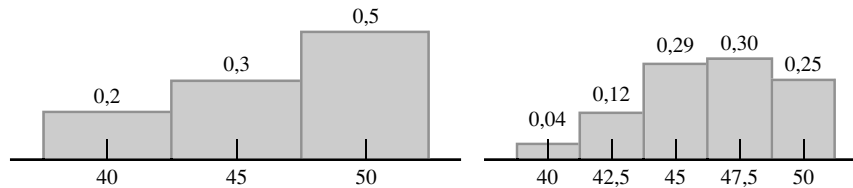
$$\begin{aligned} p_{s^2}(50) &= P(S^2 = 50) = P(X_1 = 40, X_2 = 50 \text{ ou } X_1 = 50, X_2 = 40) \\ &= 0,10 + 0,10 = 0,20 \end{aligned}$$

As distribuições de amostragem completas de  $\bar{X}$  e  $S^2$  são exibidas em (5.3) e (5.4).

$\bar{x}$	40	42,5	45	47,5	50	(5.3)
$p_{\bar{x}}(\bar{x})$	0,04	0,12	0,29	0,30	0,25	

$s^2$	0	12,5	50	(5.4)
$p_{s^2}(s^2)$	0,38	0,42	0,20	

A Figura 5.7 ilustra o histograma probabilístico da distribuição original (5.2) e da distribuição  $\bar{X}$  (5.3). A figura sugere, primeiro, que a média (valor esperado) da distribuição  $\bar{X}$  seja igual à média 46,5 da distribuição original, uma vez que ambos os histogramas parecem estar centralizados no mesmo lugar.



**Figura 5.7** Histogramas probabilísticos da distribuição subjacente e da distribuição  $\bar{X}$  do Exemplo 5.20

Pela (5.3),

$$\mu_{\bar{x}} = E(\bar{X}) = \sum \bar{x} p_{\bar{x}}(\bar{x}) = (40)(0,04) + \dots + (50)(0,25) = 46,5 = \mu$$

Em segundo lugar, parece que a distribuição  $\bar{X}$  possui dispersão (variabilidade) menor do que a distribuição original, uma vez que a massa de probabilidade moveu-se em direção à média. Novamente pela (5.3),

$$\begin{aligned} \sigma_{\bar{x}}^2 &= V(\bar{X}) = \sum \bar{x}^2 \cdot p_{\bar{x}}(\bar{x}) - \mu_{\bar{x}}^2 \\ &= (40)^2(0,04) + \dots + (50)^2(0,25) - (46,5)^2 \\ &= 7,625 = \frac{15,25}{2} = \frac{\sigma^2}{2} \end{aligned}$$

A variância de  $\bar{X}$  é precisamente metade da variância original (pois,  $n = 2$ ). O valor médio de  $S^2$  é

$$\begin{aligned} \mu_{s^2} &= E(S^2) = \sum s^2 \cdot p_{s^2}(s^2) \\ &= (0)(0,38) + (12,5)(0,42) + (50)(0,20) = 15,25 = \sigma^2 \end{aligned}$$

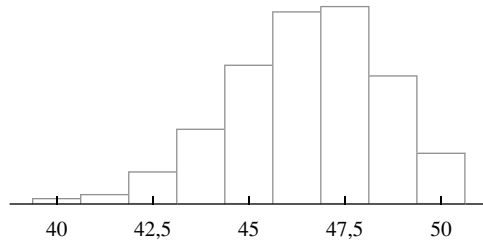
Isto é, a distribuição da amostragem  $\bar{X}$  está centrada na média da população  $\mu$  e a distribuição da amostragem  $S^2$ , está centrada na variância da população  $\sigma^2$ .

Se quatro regulagens foram feitas no dia de interesse, a receita média amostral  $\bar{X}$  seria baseada em uma amostra aleatória de quatro  $X_i$ s, cada um com distribuição (5.2). Eventualmente, mais cálculos produzem a fmp de  $\bar{X}$  para  $n = 4$  como

$\bar{x}$	40	41,25	42,5	43,75	45	46,25	47,5	48,75	50
$p_{\bar{x}}(\bar{x})$	0,0016	0,0096	0,0376	0,0936	0,1761	0,2340	0,2350	0,1500	0,0625

Pelos dados anteriores, para  $n = 4$ ,  $\mu_{\bar{x}} = 46,50 = \mu$  e  $\sigma_{\bar{x}}^2 = 3,8125 = \sigma^2/4$ . A Figura 5.8 é um histograma de probabilidade dessa fmp.





**Figura 5.8** Histograma probabilístico para  $\bar{X}$  com base em  $n = 4$  do Exemplo 5.20

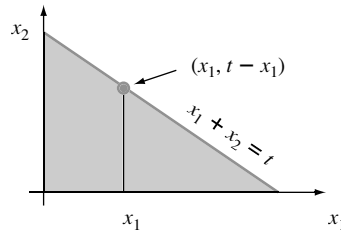
O Exemplo 5.20 sugere, antes de qualquer coisa, que o cálculo de  $p_{\bar{X}}(\bar{x})$  e  $p_{S^2}(s^2)$  é cansativo. Se a distribuição original (5.2) permitisse mais de três valores possíveis 40, 45 e 50, então, mesmo para  $n = 2$ , os cálculos teriam sido mais complexos. O exemplo sugere também, no entanto, que existem algumas relações gerais entre  $E(\bar{X})$ ,  $V(\bar{X})$ ,  $E(S^2)$ , e a média  $\mu$  e a variância  $\sigma^2$  da distribuição original, as quais são confirmadas na próxima seção. Agora, considere um exemplo em que a amostra aleatória é tirada de uma distribuição contínua.

**Exemplo 5.21**

O tempo gasto para atender a um cliente na caixa de um minimercado é uma variável aleatória que tem distribuição exponencial com parâmetro  $\lambda$ . Suponha que  $X_1$  e  $X_2$  sejam os tempos de atendimento para dois clientes diferentes, assumidos como independentes. Considere o tempo total de atendimento  $T_o = X_1 + X_2$  dos dois clientes também uma estatística. A fdc de  $T_o$  é para  $t \geq 0$ ,

$$\begin{aligned} F_{T_o}(t) &= P(X_1 + X_2 \leq t) = \iint_{\{(x_1, x_2): x_1 + x_2 \leq t\}} f(x_1, x_2) dx_1 dx_2 \\ &= \int_0^t \int_0^{t-x_1} \lambda e^{-\lambda x_1} \cdot \lambda e^{-\lambda x_2} dx_2 dx_1 = \int_0^t [\lambda e^{-\lambda x_1} - \lambda e^{-\lambda t}] dx_1 \\ &= 1 - e^{-\lambda t} - \lambda t e^{-\lambda t} \end{aligned}$$

A região de integração está ilustrada na Figura 5.9.



**Figura 5.9** Região de integração para obter a fdc de  $T_o$  do Exemplo 5.21

A fdp de  $T_o$  é obtida diferenciando  $F_{T_o}(t)$ :

$$f_{T_o}(t) = \begin{cases} \lambda^2 t e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{5.5}$$

Essa é uma fdp gama ( $\alpha = 2$  e  $\beta = 1/\lambda$ ). A fdp de  $\bar{X} = T_o/2$  é obtida da relação  $\{\bar{X} \leq \bar{x}\}$  se e somente se  $\{T_o \leq 2\bar{x}\}$  conforme

$$f_{\bar{X}}(\bar{x}) = \begin{cases} 4\lambda^2 \bar{x} e^{-2\lambda \bar{x}} & \bar{x} \geq 0 \\ 0 & \bar{x} < 0 \end{cases} \tag{5.6}$$

A média e a variância da distribuição exponencial subjacente são  $\mu = 1/\lambda$  e  $\sigma^2 = 1/\lambda^2$ . Pelas Expressões (5.5) e (5.6), pode-se demonstrar que  $E(\bar{X}) = 1/\lambda$ ,  $V(\bar{X}) = 1/(2\lambda^2)$ ,  $E(T_o) = 2/\lambda$  e  $V(T_o) = 2/\lambda^2$ . Tais resultados novamente sugerem algumas relações gerais entre médias e variâncias de  $\bar{X}$ ,  $T_o$  e a distribuição subjacente. ■

## Experimentos de Simulação

O segundo método para se obter informações sobre a distribuição de amostragem de uma estatística é fazer um experimento de simulação. Esse método geralmente é usado quando é muito difícil ou complicado calcular a derivada por meio das regras de probabilidade. Tal experimento sempre é feito virtualmente, com o auxílio de um computador. As seguintes características do experimento devem ser especificadas:

1. A estatística de interesse ( $\bar{X}$ ,  $S$ , uma média aparada específica etc.)
2. A distribuição da população (normal com  $\mu = 100$  e  $\sigma = 15$ , uniforme com limite inferior  $A = 5$  e limite superior  $B = 10$  etc.)
3. O tamanho da amostra  $n$  (por exemplo,  $n = 10$  ou  $n = 50$ )
4. O número de repetições  $k$  (por exemplo,  $k = 500$ )

Então, use um computador para obter  $k$  amostras aleatórias diferentes, cada uma de tamanho  $n$ , da distribuição da população designada. Para cada amostra, calcule o valor da estatística e construa um histograma dos  $k$  valores calculados. O histograma fornece a distribuição de amostragem *aproximada* da estatística. Quanto maior o valor de  $k$ , melhor é a tendência da aproximação (a distribuição de amostragem real surge quando  $k \rightarrow \infty$ ). Na prática,  $k = 500$  ou  $1000$  geralmente é suficiente, se a estatística for “bem simples”.

### Exemplo 5.22

A distribuição da população para nosso primeiro estudo de simulação é normal com  $\mu = 8,25$  e  $\sigma = 0,75$ , conforme ilustrado na Figura 5.10. [O artigo “Platelet Size in Myocardial Infarction” [Tamanho da Plaqueta no Infarto do Miocárdio] (*British Med. J.*, 1983, p. 449-451) sugere essa distribuição para o volume de plaquetas nos indivíduos sem histórico de problemas cardíacos graves.]

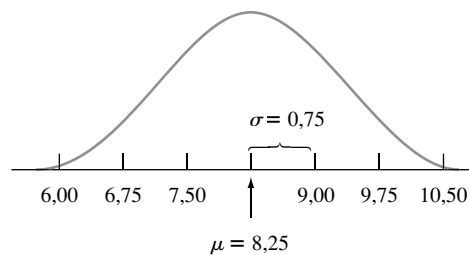


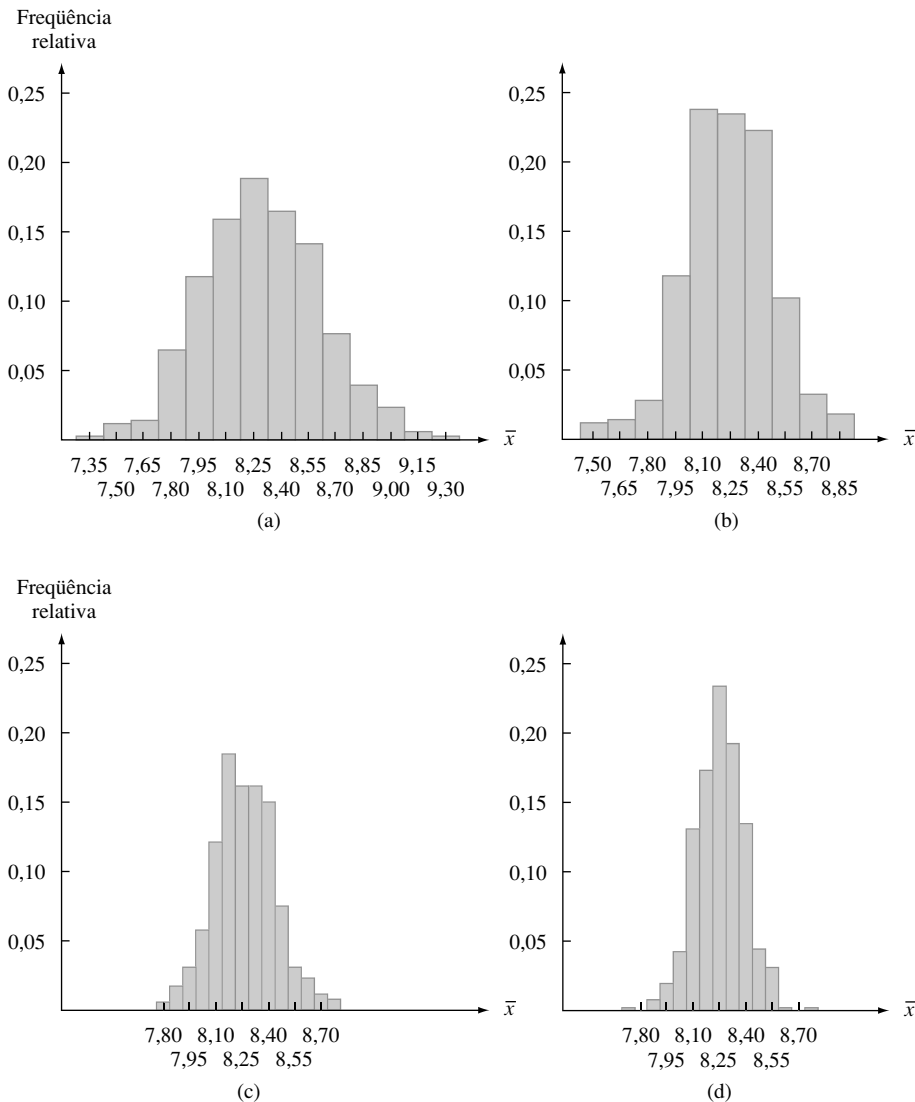
Figura 5.10 Distribuição normal, com  $\mu = 8,25$  e  $\sigma = 0,75$

Realmente fizemos quatro experimentos diferentes, com 500 repetições para cada um. No primeiro experimento, as 500 amostras de  $n = 5$  observações cada foram geradas usando-se o MINITAB, e os tamanhos da amostra para os outros três foram  $n = 10$ ,  $n = 20$  e  $n = 30$ , respectivamente. A média amostral foi calculada para cada amostra e os histogramas resultantes dos valores de  $\bar{x}$  são exibidos na Figura 5.11.

A primeira coisa a se observar sobre os histogramas é seu formato. Para uma aproximação razoável, cada quatro aproximações parecem com uma curva normal. A semelhança seria ainda mais impressionante se todo histograma fosse feito com base em mais de 500 valores de  $\bar{x}$ . Em segundo lugar, cada histograma está centrado aproximadamente em 8,25, a média da população que está servindo como amostra. Se os histogramas se baseassem em uma seqüência sem fim de valores de  $\bar{x}$ , seus centros seriam exatamente a média da população, 8,25.

O aspecto final a ser observado nos histogramas é sua dispersão relativa entre eles. Quanto menor o valor de  $n$ , maior a extensão com que a distribuição da amostragem se dispersa em relação ao valor médio. É por esse motivo que os histogramas de  $n = 20$  e  $n = 30$  baseiam-se em intervalos de classe menores que os dos dois

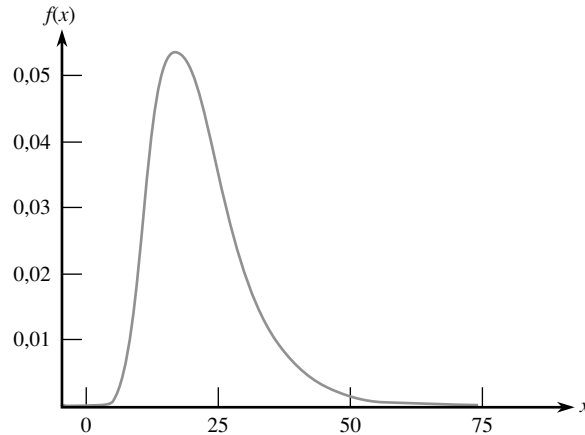
tamanhos menores da amostra. Para os tamanhos de amostra maiores, a maior parte dos valores de  $\bar{x}$  está bem próxima de 8,25. Esse é o efeito do cálculo da média. Quando  $n$  é pequeno, um único valor de  $x$  incomum pode resultar em um valor de  $\bar{x}$  longe do centro. Com um tamanho amostral maior, quaisquer valores de  $x$  incomuns quando incluídos no cálculo da média com os demais valores da amostra, ainda tendem a produzir um valor de  $\bar{x}$  próximo a  $\mu$ . Combinar essas percepções produz um resultado que desperta sua intuição:  **$\bar{X}$  determinado com base em  $n$  grande tende a estar mais próximo de  $\mu$  do que se fosse determinado com base em  $n$  pequeno.**



**Figura 5.11** Histogramas da amostra de  $\bar{x}$  com base em 500 amostras, cada uma consistindo de  $n$  observações: (a)  $n = 5$ ; (b)  $n = 10$ ; (c)  $n = 20$ ; (d)  $n = 30$

### Exemplo 5.23

Considere um experimento de simulação em que a distribuição da população seja bastante inclinada. A Figura 5.12 mais adiante mostra a curva de densidade da vida útil de determinado tipo de controle eletrônico [esta curva é realmente uma distribuição lognormal com  $E(\ln(X)) = 3$  e  $V(\ln(X)) = 0,4$ ]. Novamente, a estatística de interesse é a média amostral  $\bar{X}$ . O experimento utilizou 500 repetições e considerou os mesmos quatro tamanhos da amostra, como no Exemplo 5.22. Os histogramas resultantes, e um gráfico de probabilidade normal pelo MINITAB para os 500 valores de  $x$  com base em  $n = 30$ , são exibidos na Figura 5.13.



**Figura 5.12** Curva de densidade para o experimento de simulação do Exemplo 5.23  
 $[E(X) = \mu = 21,7584, V(X) = \sigma^2 = 82,1449]$

Ao contrário do caso normal, esses histogramas diferem no formato. Em particular, ficam progressivamente menos inclinados à medida que o tamanho da amostra  $n$  aumenta. A média dos 500 valores de  $\bar{x}$  para os quatro tamanhos diferentes da amostra está bem próxima do valor médio da distribuição da população. Se cada histograma fosse baseado em uma seqüência sem fim de valores de  $\bar{x}$ , não apenas 500, os quatro seriam centrados exatamente em 21,7584. Dessa forma, valores diferentes de  $n$  alteram o formato, mas não o centro da distribuição de  $\bar{X}$ . A comparação dos quatro histogramas na Figura 5.13 também mostra que, à medida que  $n$  aumenta, a dispersão dos histogramas diminui. Aumentar  $n$  resulta em um grau maior de concentração, em torno do valor médio da população, e faz o histograma se parecer mais com uma curva normal. O histograma da Figura 5.13(d) e o gráfico de probabilidade normal na Figura 5.13(e) fornecem evidência convincente de que o tamanho de uma amostra de  $n = 30$  é suficiente para superar a inclinação da distribuição da população e dar uma distribuição de amostragem de  $\bar{X}$  aproximadamente normal. ■

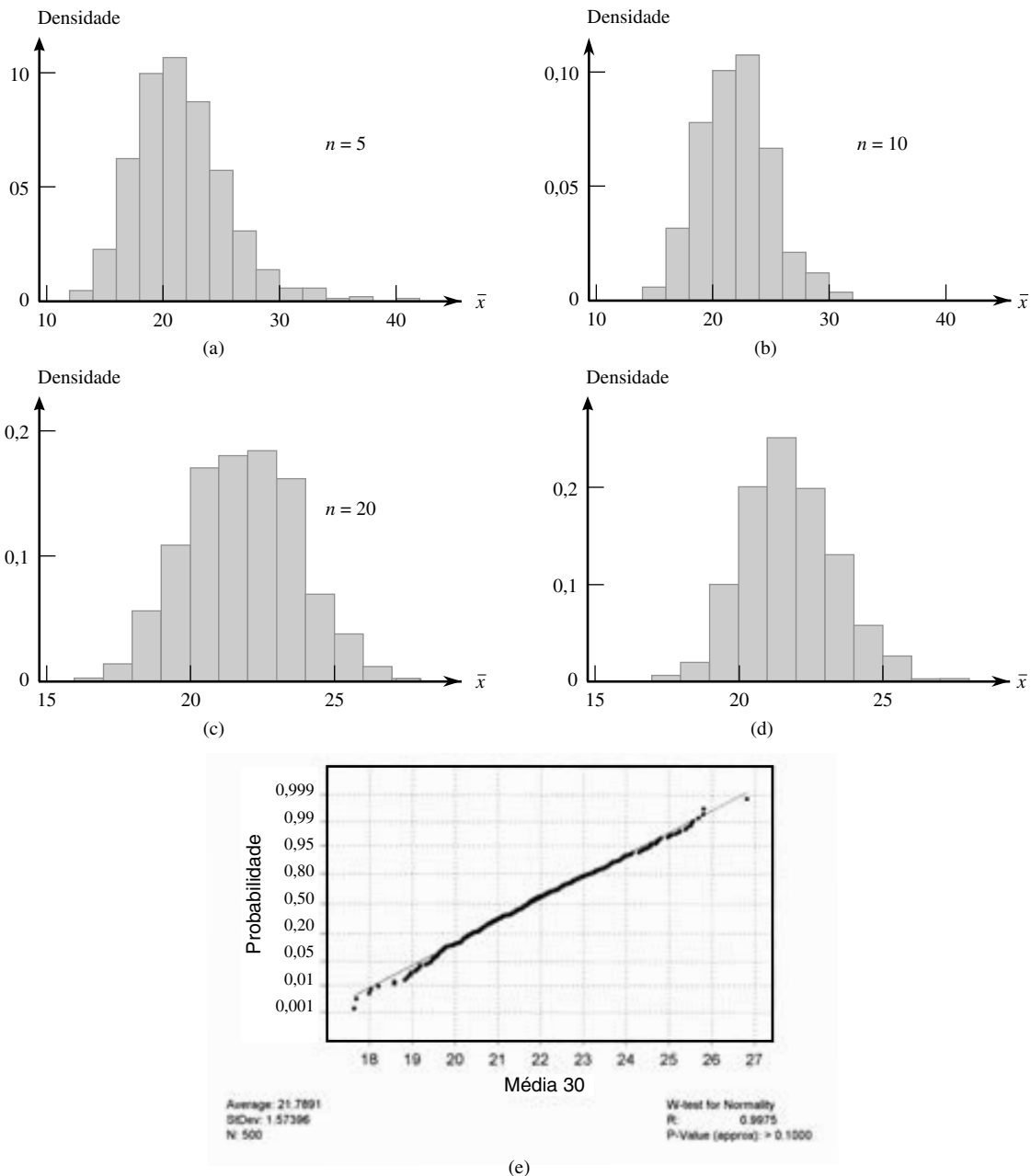
**Exercícios** | Seção 5.3 (37–45)

- 37.** Uma determinada marca de sabão para máquina de lavar louça é vendida em três tamanhos: 25 oz, 40 oz e 65 oz. Vinte por cento de todos os compradores escolhem a caixa de 25 oz, 50% escolhem a caixa de 40 oz e os 30% restantes escolhem a caixa de 65 oz. Sejam  $X_1$  e  $X_2$  os tamanhos dos pacotes escolhidos por dois compradores selecionados independentemente.
- Determine a distribuição de  $\bar{X}$ , calcule  $E(\bar{X})$  e compare a  $\mu$ .
  - Determine a distribuição da amostragem da variância da amostra  $S^2$ , calcule  $E(S^2)$  e compare com  $\sigma^2$ .
- 38.** Existem dois semáforos a caminho do meu trabalho. Seja  $X_1$  o número de sinais de trânsito em que eu tenho de parar, e suponha que a distribuição de  $X_1$  seja a seguinte:

$x_1$	0	1	2	$\mu = 1,1, \sigma^2 = 0,49$
$p(x_1)$	0,2	0,5	0,3	

Seja  $X_2$  o número de faróis de trânsito em que eu tenho de parar a caminho de casa;  $X_2$  é independente de  $X_1$ . Assuma que  $X_2$  possui a mesma distribuição que  $X_1$ , de modo que  $X_1, X_2$  é uma amostra aleatória  $n = 2$ .

- Seja  $T_o = X_1 + X_2$  e determine a distribuição de probabilidades de  $T_o$ .
  - Calcule  $\mu_{T_o}$ . Como ele está relacionado a  $\mu$ , a média da população?
  - Calcule  $\sigma_{T_o}^2$ . Como ele está relacionado a  $\sigma^2$ , variância da população?
- 39.** Sabe-se que 80% de todos os *zip drives* de marca A funcionam de forma satisfatória durante o período de garantia (são “sucessos”). Suponha que  $n = 10$  *drives* sejam selecionados aleatoriamente. Seja  $X$  = número de sucessos na amostra. A estatística  $X/n$  é a proporção da amostra (fração) de sucessos. Obtenha a distribuição de amostragem dessa estatística. [Sugestão: um valor possível de  $X/n$  é 0,3, correspondente a  $X = 3$ . Qual é a probabilidade desse valor (que tipo de variável aleatória é  $X$ )?]



**Figura 5.13** Resultados do experimento de simulação do Exemplo 5.23: (a) histograma de  $\bar{x}$  para  $n = 5$ ; (b) histograma de  $\bar{x}$  para  $n = 10$ ; (c) histograma de  $\bar{x}$  para  $n = 20$ ; (d) histograma de  $\bar{x}$  para  $n = 30$ ; (e) gráfico de probabilidade normal para  $n = 30$  (do MINITAB)

**40.** Uma caixa contém 10 envelopes fechados numerados de 1, ..., 10. Os cinco primeiros não têm dinheiro, os três seguintes têm US\$ 5 cada e, em cada um dos dois últimos, há uma nota de US\$ 10. Uma amostra de tamanho 3 é selecionada *com* reposição para obtermos uma amostra aleatória, e você ganha a quantia maior em qualquer um dos envelopes escolhidos. Se  $X_1$ ,  $X_2$  e  $X_3$  representam as quantias nos envelopes selecionados, a estatística de interesse é  $M = \text{o máximo de } X_1, X_2 \text{ e } X_3$ .

**a.** Obtenha a distribuição de probabilidades dessa estatística.

**b.** Descreva como você conduziria um experimento de simulação para comparar as distribuições de  $M$  dos diversos tamanhos da amostra. Como você faria para adivinhar a mudança da distribuição à medida que  $n$  aumenta?

**41.** Seja  $X$  o número de pacotes que estão sendo enviados por um cliente selecionado aleatoriamente em uma determinada empresa de remessas. Suponha que a distribuição de  $X$  seja como segue:

$x$	1	2	3	4
$p(x)$	0,4	0,3	0,2	0,1

- a. Considere uma amostra aleatória de tamanho  $n = 2$  (dois clientes) e seja  $\bar{X}$  o número médio da amostra de pacotes enviados. Obtenha a distribuição de probabilidades de  $\bar{X}$ .
  - b. Refira-se à parte (a) e calcule  $P(\bar{X} \leq 2,5)$ .
  - c. Considere novamente uma amostra aleatória de tamanho  $n = 2$ , mas agora enfoque a estatística  $R =$  o intervalo da amostra (diferença entre os valores maior e menor na amostra). Obtenha a distribuição de  $R$ . [Sugestão: calcule o valor de  $R$  para cada resultado e use as probabilidades da parte (a).]
  - d. Se uma amostra aleatória de tamanho  $n = 4$  for selecionada, qual é  $P(\bar{X} \leq 1,5)$ ? (Sugestão: você não deve listar todos os resultados possíveis, somente aqueles para os quais  $\bar{x} \leq 1,5$ .)
42. Uma empresa mantém três escritórios em certa região, cada um com dois funcionários. As informações sobre os salários anuais (milhares de dólares) seguem abaixo:

Escritório	1	1	2	2	3	3
nº de funcionários	1	2	3	4	5	6
salário	29,7	33,6	30,2	33,6	25,8	29,7

- a. Suponha que dois desses funcionários sejam selecionados aleatoriamente, dentre os seis (sem reposição). Determine a distribuição da amostragem  $\bar{X}$  do salário médio da amostra.

- b. Suponha que um dos três escritórios seja selecionado aleatoriamente. Sejam  $X_1$  e  $X_2$  os salários dos dois funcionários. Determine a distribuição de  $\bar{X}$ .
- c. Como  $E(\bar{X})$  das partes (a) e (b) é comparado ao salário médio da população  $\mu$ ?

- 43. Suponha que a quantidade de líquido despejada por certo equipamento seja uniformemente distribuída com limite inferior  $A = 8$  oz e limite superior  $B = 10$  oz. Descreva como você conduziria experimentos de simulação para comparar a distribuição de amostragem da quarta dispersão (da amostra) para tamanhos amostrais  $n = 5, 10, 20$  e  $30$ .
- 44. Faça um experimento de simulação usando um pacote estatístico ou outro software para estudar a distribuição de  $\bar{X}$  quando a distribuição da população é Weibull com  $\alpha = 2$  e  $\beta = 5$ , como no Exemplo 5.19. Considere os quatro tamanhos da amostra  $n = 5, 10, 20$  e  $30$  e, em cada caso, utilize 500 repetições. Para qual desses tamanhos de amostra a distribuição de  $\bar{X}$  parece ser aproximadamente normal?
- 45. Faça um experimento de simulação usando um pacote estatístico ou outro software para estudar a distribuição de  $\bar{X}$  quando a distribuição da população é lognormal com  $E(\ln(X)) = 3$  e  $V(\ln(X)) = 1$ . Considere os quatro tamanhos de amostra  $n = 10, 20, 30$  e  $50$  e, em cada caso, utilize 500 repetições. Para qual desses tamanhos de amostra a distribuição de  $\bar{X}$  parece ser aproximadamente normal?

## 5.4 A Distribuição da Média Amostral

A importância da média amostral  $\bar{X}$  surge de seu uso para tirar conclusões sobre a média da população  $\mu$ . Alguns dos procedimentos de inferência, usados com mais frequência, baseiam-se nas propriedades da distribuição de  $\bar{X}$ . A apresentação inicial dessas propriedades foi feita nos cálculos e experimentos de simulação da seção anterior, onde salientamos as relações entre  $E(\bar{X})$  e  $\mu$  e também entre  $V(\bar{X})$ ,  $\sigma^2$  e  $n$ .

### PROPOSIÇÃO

Sejam  $X_1, X_2, \dots, X_n$  elementos da amostra aleatória de uma distribuição com valor médio  $\mu$  e desvio padrão  $\sigma$ . Então,

- 1.  $E(\bar{X}) = \mu_{\bar{X}} = \mu$
- 2.  $V(\bar{X}) = \sigma_{\bar{X}}^2 = \sigma^2/n$  e  $\sigma_{\bar{X}} = \sigma/\sqrt{n}$

Além disso, com  $T_o = X_1 + \dots + X_n$  (o total da amostra),  $E(T_o) = n\mu$ ,  $V(T_o) = n\sigma^2$  e  $\sigma_{T_o} = \sqrt{n}\sigma$ .

As demonstrações desses resultados foram transferidas para a próxima seção. De acordo com o Resultado 1, a distribuição (isto é, probabilidade) da amostragem de  $\bar{X}$  é centrada precisamente na média da população da qual a amostra foi selecionada. O Resultado 2 mostra que a distribuição de  $\bar{X}$  se torna mais concentrada em torno de  $\mu$  à medida que o tamanho da amostra  $n$  aumenta. Com diferença marcante, a distribuição de  $T_o$  se dispersa mais à medida que  $n$  aumenta. A média desloca a probabilidade para o meio, enquanto o total dispersa a probabilidade em um intervalo de valores mais amplo.

### Exemplo 5.24

Em um teste de fadiga à tração de um espécime de titânio, o número esperado de ciclos para a primeira emissão acústica (usada para indicar o início da trinca) é  $\mu = 28.000$ , e o desvio padrão do número de ciclos é  $\sigma = 5000$ . Sejam  $X_1, X_2, \dots, X_{25}$  os elementos de uma amostra aleatória de tamanho 25, em que cada  $X_i$  é o número de ciclos em um espécime diferente, selecionado aleatoriamente. Então, o valor esperado do número de ciclos médio da amostra até a primeira emissão é  $E(\bar{X}) = \mu = 28.000$ , e o número total esperado de ciclos para os 25 espécimes é  $E(T_o) = n\mu = 25(28.000) = 700.000$ . Os desvios padrão de  $\bar{X}$  e  $T_o$  são

$$\sigma_{\bar{X}} = \sigma/\sqrt{n} = \frac{5000}{\sqrt{25}} = 1000$$

$$\sigma_{T_o} = \sqrt{n}\sigma = \sqrt{25}(5000) = 25.000$$

Se o tamanho da amostra aumenta para  $n = 100$ ,  $E(\bar{X})$  fica inalterado, mas  $\sigma_{\bar{X}} = 500$ , metade de seu valor anterior (o tamanho da amostra deve ser quadruplicado para reduzir à metade o desvio padrão de  $\bar{X}$ ). ■

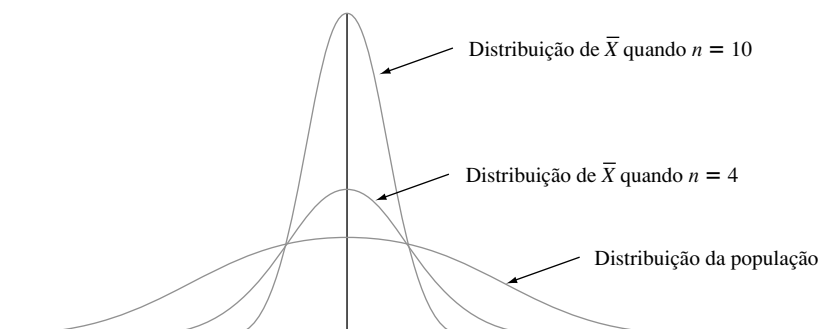
## Caso de Distribuição da População Normal

Ao reexaminar o experimento de simulação do Exemplo 5.22, observamos que quando a distribuição da população é normal, cada histograma de valores de  $\bar{x}$  é bem aproximado pela curva normal. Segue o resultado exato.

### PROPOSIÇÃO

Sejam  $X_1, X_2, \dots, X_n$  os elementos da amostra aleatória de uma distribuição *normal* com média  $\mu$  e desvio padrão  $\sigma$ . Então, para *qualquer*  $n$ ,  $\bar{X}$  é normalmente distribuído (com média  $\mu$  e desvio padrão  $\sigma/\sqrt{n}$ ), como é  $T_o$  (com média  $n\mu$  e desvio padrão  $\sqrt{n}\sigma$ ).<sup>1</sup>

Sabemos tudo sobre as distribuições de  $\bar{X}$  e  $T_o$  quando a distribuição da população é normal. Em particular, as probabilidades, como  $P(a \leq \bar{X} \leq b)$  e  $P(c \leq T_o \leq d)$ , podem ser obtidas simplesmente pela padronização. A Figura 5.14 ilustra a proposição.



**Figura 5.14** Distribuição da população normal e distribuições de amostragem de  $\bar{X}$

<sup>1</sup> É possível demonstrar o resultado de  $T_o$  quando  $n = 2$ , usando o método no Exemplo 5.21, mas os detalhes são confusos. A fórmula geral normalmente é demonstrada usando-se uma ferramenta teórica denominada *função de geração de momento*. Uma das referências do capítulo pode ser consultada para obter-se mais informações.

**Exemplo 5.25**

O tempo que um rato de determinada subespécie, selecionado aleatoriamente leva para encontrar o caminho em um labirinto é uma va distribuída normalmente com  $\mu = 1,5$  min e  $\sigma = 0,35$  min. Suponha que cinco ratos sejam selecionados. Sejam  $X_1, \dots, X_5$  seu tempo no labirinto. Assumindo que  $X_i$  é uma amostra aleatória dessa distribuição normal, qual é a probabilidade de o tempo total  $T_o = X_1 + \dots + X_5$  dos cinco estar entre 6 e 8 min? Pela proposição,  $T_o$  possui distribuição normal com  $\mu_{T_o} = n\mu = 5(1,5) = 7,5$  e variância  $\sigma_{T_o}^2 = n\sigma^2 = 5(0,1225) = 0,6125$ , assim,  $\sigma_{T_o} = 0,783$ . Para padronizar  $T_o$ , subtraia  $\mu_{T_o}$  e divida por  $\sigma_{T_o}$ :

$$P(6 \leq T_o \leq 8) = P\left(\frac{6 - 7,5}{0,783} \leq Z \leq \frac{8 - 7,5}{0,783}\right) \\ = P(-1,92 \leq Z \leq 0,64) = \Phi(0,64) - \Phi(-1,92) = 0,7115$$

A determinação da probabilidade de o tempo médio da amostra  $\bar{X}$  (uma variável distribuída normalmente) ser no máximo 2,0 min requer  $\mu_{\bar{X}} = \mu = 1,5$  e  $\sigma_{\bar{X}} = \sigma/\sqrt{n} = 0,35/\sqrt{5} = 0,1565$ . Então,

$$P(\bar{X} \leq 2,0) = P\left(Z \leq \frac{2,0 - 1,5}{0,1565}\right) = P(Z \leq 3,19) = \Phi(3,19) = 0,9993 \quad \blacksquare$$

**Teorema do Limite Central**

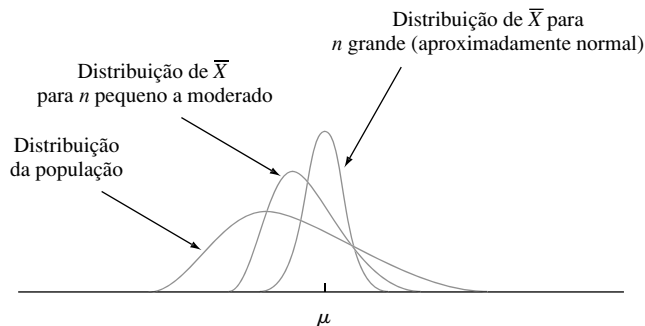
Quando os  $X_i$ s são distribuídos normalmente,  $\bar{X}$  também é para cada tamanho de amostra  $n$ . O experimento de simulação do Exemplo 5.23 sugere que, mesmo quando a distribuição da população é altamente não-normal, o cálculo da média produz uma distribuição em forma de sino mais acentuada do que a que está servindo como amostra. Uma hipótese razoável é que, se  $n$  for grande, uma curva normal adequada aproximará a distribuição real de  $\bar{X}$ . O enunciado formal de tal resultado é o teorema mais importante da probabilidade.

**TEOREMA**

**Teorema do Limite Central (TLC)**

$X_1, X_2, \dots, X_n$  formam a amostra aleatória de uma distribuição com média  $\mu$  e variância  $\sigma^2$ . Então, se  $n$  é suficientemente grande,  $\bar{X}$  tem aproximadamente uma distribuição normal com  $\mu_{\bar{X}} = \mu$  e  $\sigma_{\bar{X}}^2 = \sigma^2/n$ , e  $T_o$  também tem aproximadamente uma distribuição normal com  $\mu_{T_o} = n\mu$ ,  $\sigma_{T_o}^2 = n\sigma^2$ . Quanto maior o valor de  $n$ , melhor a aproximação.

A Figura 5.15 ilustra o Teorema do Limite Central. De acordo com o TLC, quando  $n$  é grande e queremos calcular a probabilidade, tal como  $P(a \leq \bar{X} \leq b)$ , precisamos somente “fingir” que  $\bar{X}$  é normal, padronizá-lo e usar a tabela normal. A resposta resultante será aproximadamente correta. A resposta exata pode ser obtida somente encontrando, primeiro, a distribuição de  $\bar{X}$ , assim, o TLC fornece um atalho de fato impressionante. A prova do teorema envolve matemática muito avançada.



**Figura 5.15** O Teorema do Limite Central ilustrado



**Exemplo 5.26**

Quando um lote de certo produto químico é preparado, a quantidade de uma impureza específica no lote é uma variável aleatória com valor médio de 4,0 g e desvio padrão de 1,5 g. Se 50 lotes forem preparados independentemente, qual é a probabilidade (aproximada) de a quantidade média de impureza  $\bar{X}$  da amostra estar entre 3,5 e 3,8 g? De acordo com a regra prática dita em poucas palavras,  $n = 50$  é grande o suficiente para que o TLC seja aplicável.  $\bar{X}$  então possui aproximadamente uma distribuição normal com valor médio  $\mu_{\bar{X}} = 4,0$  e  $\sigma_{\bar{X}} = 1,5/\sqrt{50} = 0,2121$ , assim,

$$\begin{aligned} P(3,5 \leq \bar{X} \leq 3,8) &\approx P\left(\frac{3,5 - 4,0}{0,2121} \leq Z \leq \frac{3,8 - 4,0}{0,2121}\right) \\ &= \Phi(-0,94) - \Phi(-2,36) = 0,1645 \end{aligned} \quad \blacksquare$$

**Exemplo 5.27**

Uma determinada organização de consumidores normalmente reporta o número de defeitos graves de cada carro novo examinado. Suponha que o número de tais defeitos de certo modelo seja uma variável aleatória com valor médio 3,2 e desvio padrão 2,4. Dentre 100 carros selecionados aleatoriamente desse modelo, qual é a probabilidade de o número médio da amostra de defeitos graves exceder 4? Seja  $X_i$  o número de defeitos graves do  $i$ -ésimo carro na amostra aleatória. Observe que  $X_i$  é uma va discreta, mas que o TLC é aplicável, se a variável de interesse for discreta ou contínua. Além disso, apesar de o desvio padrão dessa variável não-negativa ser grande em relação ao valor médio sugerir que sua distribuição é positivamente inclinada, o tamanho grande de amostra implica que  $\bar{X}$  tem a distribuição aproximadamente normal. Usando  $\mu_{\bar{X}} = 3,2$  e  $\sigma_{\bar{X}} = 0,24$ ,

$$P(\bar{X} > 4) \approx P\left(Z > \frac{4 - 3,2}{0,24}\right) = 1 - \Phi(3,33) = 0,0004 \quad \blacksquare$$

O TLC fornece compreensão clara do motivo pelo qual muitas variáveis aleatórias têm distribuições de probabilidade aproximadamente normais. Por exemplo: o erro de medição em um experimento científico pode ser visto como a soma de um número de irregularidades e erros subjacentes de pouca importância.

Embora a utilidade do TLC para inferência seja, em breve, aparente, o teor intuitivo do resultado faz muitos estudantes iniciantes sentirem dificuldade. Voltando à Figura 5.7, o histograma probabilístico à esquerda é um retrato da distribuição que está servindo como amostra. É discreta e bem inclinada, de modo que não parece ser uma distribuição normal. A distribuição de  $\bar{X}$  para  $n = 2$  começa a exibir certa simetria, e isso é até mais pronunciado para  $n = 4$  na Figura 5.8. A Figura 5.16 contém a distribuição de probabilidades de  $\bar{X}$  para  $n = 8$ , bem como um histograma probabilístico para essa distribuição. Com  $\mu_{\bar{X}} = \mu = 46,5$  e  $\sigma_{\bar{X}} = \sigma/\sqrt{n} = 3,905/\sqrt{8} = 1,38$ , se ajustarmos uma curva normal com tal média e desvio padrão por meio do histograma de  $\bar{X}$ , as áreas de retângulos do histograma probabilístico estarão razoavelmente bem aproximadas pelas áreas da curva normal, pelo menos na parte central da distribuição. O desenho de  $T_o$  é semelhante, exceto pelo fato de a escala horizontal ser muito mais dispersa, com  $T_o$  variando de 320 ( $\bar{x} = 40$ ) a 400 ( $\bar{x} = 50$ ).

Uma dificuldade prática na aplicação do TLC é saber quando  $n$  é suficientemente grande. O problema é que a precisão da aproximação de um  $n$  específico depende do formato da distribuição subjacente original que está servindo como amostra. Se a distribuição subjacente está próxima do formato de sino, então a aproximação será boa, mesmo para um  $n$  pequeno, ao passo que, se estiver longe desse formato, será necessário um  $n$  grande. Usaremos a regra prática a seguir, que geralmente é um pouco conservadora.

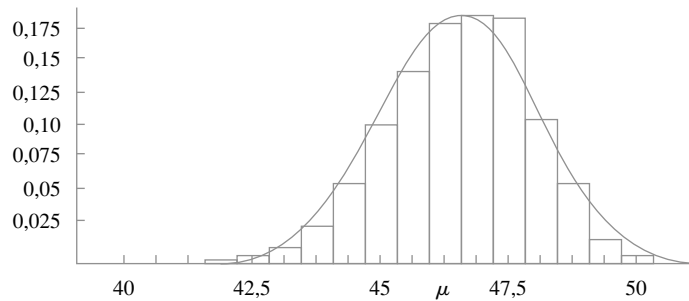
**Regra prática**

Se  $n > 30$ , o Teorema do Limite Central pode ser utilizado.

Existem distribuições de população para as quais mesmo com  $n = 40$  ou  $50$  não é suficiente, mas tais distribuições raramente são encontradas na prática.

$\bar{x}$	40	40,625	41,25	41,875	42,5	43,125
$p(\bar{x})$	0,0000	0,0000	0,0003	0,0012	0,0038	0,0112
$\bar{x}$	43,75	44,375	45	45,625	46,25	46,875
$p(\bar{x})$	0,0274	0,0556	0,0954	0,1378	0,1704	0,1746
$\bar{x}$	47,5	48,125	48,75	49,375	50	
$p(\bar{x})$	0,1474	0,0998	0,0519	0,0188	0,0039	

(a)



(b)

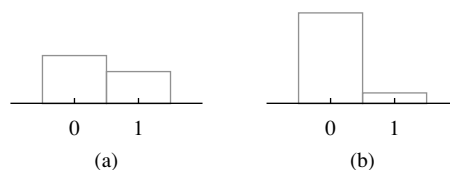
**Figura 5.16** (a) Distribuição de probabilidades de  $\bar{X}$  para  $n = 8$ ; (b) Histograma probabilístico e aproximação normal à distribuição de  $\bar{X}$  quando a distribuição original é como no Exemplo 5.20

### Outras Aplicações do Teorema do Limite Central

O TLC pode ser usado para justificar a aproximação normal da distribuição binomial discutida no Capítulo 4. Lembre-se de que uma variável binomial  $X$  é o número de sucessos em um experimento binomial que consiste de  $n$  tentativas independentes de sucesso/falha com  $p = P(S)$  para qualquer tentativa específica. Defina novas variáveis  $X_1, X_2, \dots, X_n$  por

$$X_i = \begin{cases} 1 & \text{se a } i\text{-ésima tentativa for um sucesso} \\ 0 & \text{se a } i\text{-ésima tentativa for um fracasso} \end{cases} \quad (i = 1, \dots, n)$$

Em virtude de as tentativas serem independentes e  $P(S)$  ser constante de tentativa para tentativa, os  $X_i$ s são iid (uma amostra aleatória de uma distribuição de Bernoulli). O TLC, então, implica que, se  $n$  for suficientemente grande, tanto a soma como a média dos  $X_i$ s têm aproximadamente distribuições normais. Quando os  $X_i$ s são somados, adiciona-se 1 a cada  $S$  que ocorre e 0 a cada  $F$ , assim,  $X_1 + \dots + X_n = X$ . A média amostral dos  $X_i$ s é  $X/n$ , a proporção da amostra de sucessos. Isso é,  $X$  e  $X/n$  são aproximadamente normais quando  $n$  é grande. O tamanho necessário da amostra para essa aproximação depende do valor de  $p$ : quando  $p$  está próximo de 0,5, a distribuição de cada  $X_i$  é razoavelmente simétrica (veja a Figura 5.17 na próxima página), ao passo que a distribuição é relativamente inclinada quando  $p$  está próximo de 0 ou 1. Usar a aproximação somente se  $np \geq 10$  e  $n(1 - p) \geq 10$  garante que  $n$  é grande o suficiente para superar qualquer inclinação na distribuição de Bernoulli subjacente.



**Figura 5.17** Duas distribuições de Bernoulli: (a)  $p = 0,4$  (razoavelmente simétrica); (b)  $p = 0,1$  (muito inclinada)

Lembre-se da Seção 4.5, em que  $X$  possui uma distribuição lognormal se  $\ln(X)$  tiver distribuição normal.

## PROPOSIÇÃO

Seja  $X_1, X_2, \dots, X_n$  uma amostra aleatória de uma distribuição para a qual somente valores positivos são possíveis [ $P(X_i > 0) = 1$ ]. Então, se  $n$  for suficientemente grande, o produto  $Y = X_1 X_2 \cdots X_n$  terá aproximadamente uma distribuição lognormal.

Para demonstrar a proposição, observe que

$$\ln(Y) = \ln(X_1) + \ln(X_2) + \cdots + \ln(X_n)$$

Uma vez que  $\ln(Y)$  é a soma de  $n$  [as  $\ln(X_i)$ s] independentes e identicamente distribuídas, é aproximadamente normal quando  $n$  é grande, pois, o próprio  $Y$  tem aproximadamente uma distribuição lognormal. Como exemplo da aplicabilidade desse resultado, Bury (*Statistical Models in Applied Science*, Wiley, p. 590) discute que o processo de danos no fluxo do plástico e propagação de trincas é multiplicativo, de modo que as variáveis, como porcentagem de alongamento e resistência à ruptura, têm aproximadamente distribuições lognormais.

---

**Exercícios** | Seção 5.4 (46–57)

- 46.** O diâmetro interno de um anel de pistão selecionado casualmente é uma variável aleatória com valor médio de 12 cm e desvio padrão de 0,04 cm.
- Se  $\bar{X}$  é o diâmetro médio para uma amostra aleatória de  $n = 16$  anéis, onde a distribuição de amostragem de  $\bar{X}$  está centrada, qual é o desvio padrão da distribuição de  $\bar{X}$ ?
  - Responda as questões propostas na parte (a) para um tamanho de amostra de  $n = 64$  anéis.
  - Para qual das duas amostras aleatórias, a da parte (a) ou a da parte (b),  $\bar{X}$  é mais provável de estar dentro de 0,01 cm de 12 cm? Explique seu raciocínio.
- 47.** Consulte o Exercício 46. Suponha que a distribuição do diâmetro seja normal.
- Calcule  $P(11,99 \leq \bar{X} \leq 12,01)$  quando  $n = 16$ .
  - Qual é a probabilidade de o diâmetro médio da amostra exceder 12,01 quando  $n = 25$ ?
- 48.** Sejam  $X_1, X_2, \dots, X_{100}$  os pesos líquidos reais de 100 sacos de fertilizantes de 50 lb selecionados aleatoriamente.
- Se o peso esperado de cada saco for 50 e a variância 1, calcule  $P(49,75 \leq \bar{X} \leq 50,25)$  (aproximadamente), usando o TLC.
  - Se o peso esperado for 49,8 lb, e não 50 lb, de modo que na média os sacos não estejam muito cheios, calcule  $P(49,75 \leq \bar{X} \leq 50,25)$ .
- 49.** Há 40 alunos em uma classe de estatística básica. Com base nos anos de experiência, o instrutor sabe que o tempo necessário para corrigir o primeiro questionário, escolhido aleatoriamente, é uma variável aleatória com valor esperado de 6 min e desvio padrão de 6 min.
- Se os tempos de correção forem independentes e o instrutor começar a corrigir às 18h50 sem parar, qual é a probabilidade (aproximada) de terminar as correções antes de o noticiário das 23h, na televisão, começar?
  - Se as notícias de esportes começam às 23h10, qual é a probabilidade de ele perder parte das notícias, se esperar até que a correção seja concluída antes de ligar a televisão?
- 50.** A resistência à ruptura de um rebite possui valor médio de 10.000 psi e desvio padrão de 500 psi.
- Qual é a probabilidade de a resistência à ruptura média de uma amostra aleatória de 40 rebites estar entre 9.900 e 10.200?
  - Se o tamanho da amostra fosse 15, e não 40, a probabilidade pedida na parte (a) poderia ser calculada com as informações dadas?
- 51.** O tempo que um candidato a uma hipoteca, selecionado aleatoriamente, gasta para preencher um formulário tem distribuição normal com valor médio de 10 min e desvio padrão de 2 min. Se cinco indivíduos preenchem um formulário em um dia, e seis, em outro, qual é a probabilidade de a quantidade média de tempo da amostra em cada dia ser de no máximo 11 min?
- 52.** A vida útil de certo tipo de bateria é distribuída normalmente com valor médio de 10 horas e desvio padrão de 1 hora. Há quatro baterias em um pacote. Que valor

de vida útil faz com que a vida total de todas as baterias em um pacote exceda tal valor em apenas 5% de todos os pacotes?

53. É sabido que a dureza de Rockwell de certo tipo de pinos tem valor médio de 50 e desvio padrão de 1,2.
- Se a distribuição for normal, qual é a probabilidade de a dureza média de uma amostra aleatória de 9 (nove) pinos ser no mínimo 51?
  - Qual é a probabilidade (aproximada) de a dureza média de uma amostra aleatória de 40 pinos ser de no mínimo 51?
54. Suponha que a densidade sedimentar (g/cm) de um espécime selecionado aleatoriamente de uma determinada região tenha distribuição normal com média de 2,65 e desvio padrão de 0,85 (sugerido em “Modeling Sediment and Water Column Interactions for Hydrophobic Pollutants” (Modelo de Sedimentos e Interações da Coluna de Água para Poluentes Hidrofóbicos) *Water Research*, 1984, p. 1169-1174).
- Se uma amostra aleatória de 25 espécimes é selecionada, qual é a probabilidade de a densidade sedimentar média amostral ser de no máximo 3,00? Entre 2,65 e 3,00?
  - Qual deve ser o tamanho de uma amostra para garantir que a primeira probabilidade da parte (a) seja no mínimo 0,99?
55. O primeiro trabalho de uma classe de cálculo estatístico envolve a execução de um programa curto. Se a experiência anterior indica que 40% de todos os alunos não cometerão erros de programação, calcule a probabilidade (aproximada) de que em uma classe de 50 alunos:
- pelo menos 25 não cometerão erros. (*Sugestão*: aproximação da normal para a binomial.)
  - entre 15 e 25 (inclusive) não cometerão erros.
56. O número de multas por estacionamento irregular emitidas em uma determinada cidade em um dia de semana qualquer possui distribuição de Poisson com parâmetro  $\lambda = 50$ . Qual é a probabilidade aproximada de
- entre 35 e 70 multas serem dadas em um dia específico? (*Sugestão*: quando  $\lambda$  é grande, uma *va* de Poisson possui aproximadamente distribuição normal.)
  - o número total de multas dadas durante os cinco dias da semana estar entre 225 e 275?
57. Suponha que a distribuição do tempo  $X$  (em horas) gasto pelos alunos de certa universidade em um projeto específico seja gama com parâmetros  $\alpha = 50$  e  $\beta = 2$ . Em virtude de  $\alpha$  ser grande, pode-se demonstrar que  $X$  possui distribuição aproximadamente normal. Use esse fato para calcular a probabilidade de um aluno selecionado aleatoriamente gastar no máximo 125 horas no projeto.

## 5.5 | Distribuição de uma Combinação Linear

A  $\bar{X}$  média amostral e o  $T_o$  total da amostra são casos especiais de um tipo de variável aleatória que surge com bastante frequência em aplicações estatísticas.

### DEFINIÇÃO

Dado um conjunto de  $n$  variáveis aleatórias  $X_1, \dots, X_n$  e  $n$  constantes numéricas  $a_1, \dots, a_n$ , a *va*

$$Y = a_1X_1 + \dots + a_nX_n = \sum_{i=1}^n a_iX_i \quad (5.7)$$

é denominada **combinação linear** dos  $X_i$ 's.

Assumindo  $a_1 = a_2 = \dots = a_n = 1$  fornece  $Y = X_1 + \dots + X_n = T_o$ , e  $a_1 = a_2 = \dots = a_n = \frac{1}{n}$  produz  $Y = \frac{1}{n}X_1 + \dots + \frac{1}{n}X_n = \frac{1}{n}(X_1 + \dots + X_n) = \frac{1}{n}T_o = \bar{X}$ . Observe que não é requerido que os  $X_i$ s sejam independentes ou identicamente distribuídos. Todos os  $X_i$ s podem ter distribuições diferentes e, dessa forma, valores médios e variâncias distintas. Consideramos primeiro o valor esperado e a variância de uma combinação linear.

PROPOSIÇÃO

Suponha que  $X_1, X_2, \dots, X_n$  tenham valores médios  $\mu_1, \dots, \mu_n$ , respectivamente, e variâncias  $\sigma_1^2, \dots, \sigma_n^2$ , respectivamente.

1. Se os  $X_i$ s são ou não independentes,

$$\begin{aligned} E(a_1X_1 + a_2X_2 + \dots + a_nX_n) &= a_1E(X_1) + a_2E(X_2) + \dots + a_nE(X_n) \\ &= a_1\mu_1 + \dots + a_n\mu_n \end{aligned} \tag{5.8}$$

2. Se  $X_1, \dots, X_n$  são independentes,

$$\begin{aligned} V(a_1X_1 + a_2X_2 + \dots + a_nX_n) &= a_1^2V(X_1) + a_2^2V(X_2) + \dots + a_n^2V(X_n) \\ &= a_1^2\sigma_1^2 + \dots + a_n^2\sigma_n^2 \end{aligned} \tag{5.9}$$

e

$$\sigma_{a_1X_1 + \dots + a_nX_n} = \sqrt{a_1^2\sigma_1^2 + \dots + a_n^2\sigma_n^2} \tag{5.10}$$

3. Para qualquer  $X_1, \dots, X_n$ ,

$$V(a_1X_1 + \dots + a_nX_n) = \sum_{i=1}^n \sum_{j=1}^n a_i a_j \text{Cov}(X_i, X_j) \tag{5.11}$$

As demonstrações estão esquematicamente no final da seção. Uma paráfrase de (5.8) é que o valor desejado de uma combinação linear é a mesma combinação linear dos valores esperados – por exemplo:  $E(2X_1 + 5X_2) = 2\mu_1 + 5\mu_2$ . O resultado (5.9) da Proposição 2 é um caso especial de (5.11) da Proposição 3; quando os  $X_i$ s são independentes,  $\text{Cov}(X_i, X_j) = 0$  para  $i \neq j$  e  $= V(X_i)$  para  $i = j$  (essa simplificação realmente ocorre quando os  $X_i$ s são não-correlacionados, uma condição mais fraca que a independência). Generalizando para o caso de uma amostra aleatória ( $X_i$ s iid) com  $a_i = 1/n$  para cada  $i$  fornece  $E(\bar{X}) = \mu$  e  $V(\bar{X}) = \sigma^2/n$ , conforme discutido na Seção 5.4. Aplica-se um comentário semelhante às regras de  $T_o$ .

Exemplo 5.28

Um posto de gasolina vende três tipos do produto: gasolina comum, gasolina aditivada e gasolina super. Elas custam US\$ 1,20, US\$ 1,35 e US\$ 1,50 por galão, respectivamente. Sejam  $X_1, X_2$ , e  $X_3$  os valores desses galões comprados em um dia específico. Suponha que os  $X_i$ s sejam independentes com  $\mu_1 = 1000, \mu_2 = 500, \mu_3 = 300, \sigma_1 = 100, \sigma_2 = 80$ , e  $\sigma_3 = 50$ . A receita das vendas é  $Y = 1,2X_1 + 1,35X_2 + 1,5X_3$ , e

$$\begin{aligned} E(Y) &= 1,2\mu_1 + 1,35\mu_2 + 1,5\mu_3 = \$ 2325 \\ V(Y) &= (1,2)^2\sigma_1^2 + (1,35)^2\sigma_2^2 + (1,5)^2\sigma_3^2 = 31.689 \\ \sigma_Y &= \sqrt{31.689} = \$178,01 \end{aligned} \quad \blacksquare$$

Diferença Entre Duas Variáveis Aleatórias

Um caso especial importante de combinação linear resulta de  $n = 2, a_1 = 1$ , e  $a_2 = -1$ :

$$Y = a_1X_1 + a_2X_2 = X_1 - X_2$$

Temos, então, a seguinte conclusão da proposição.

COROLÁRIO

$E(X_1 - X_2) = E(X_1) - E(X_2)$  e, se  $X_1$  e  $X_2$  são independentes,  $V(X_1 - X_2) = V(X_1) + V(X_2)$ .

O valor esperado de uma diferença é a diferença entre os dois valores esperados, mas a variância da diferença entre duas variáveis independentes é a soma, não a diferença, das duas variâncias. Há tanta variabilidade em  $X_1 - X_2$  como em  $X_1 + X_2$  [escrever  $X_1 - X_2 = X_1 + (-1)X_2, (-1)X_2$  varia tanto quanto o próprio  $X_2$ ].

**Exemplo 5.29**

Um fabricante de automóveis equipa um determinado modelo com um motor de seis cilindros ou de quatro cilindros. Sejam por  $X_1$  e  $X_2$  o consumo de combustível de carros com seis e quatro cilindros selecionados de modo independente e aleatório, respectivamente. Com  $\mu_1 = 22$ ,  $\mu_2 = 26$ ,  $\sigma_1 = 1,2$ , e  $\sigma_2 = 1,5$ ,

$$\begin{aligned} E(X_1 - X_2) &= \mu_1 - \mu_2 = 22 - 26 = -4 \\ V(X_1 - X_2) &= \sigma_1^2 + \sigma_2^2 = (1,2)^2 + (1,5)^2 = 3,69 \\ \sigma_{X_1 - X_2} &= \sqrt{3,69} = 1,92 \end{aligned}$$

Se fizermos nova indicação de que  $X_1$  se refere ao carro de quatro cilindros, então  $E(X_1 - X_2) = 4$ , mas a variância da diferença ainda será 3,69. ■

**Caso de Variáveis Aleatórias Normais**

Quando os  $X_i$ s formam uma amostra aleatória de uma distribuição normal,  $\bar{X}$  e  $T_o$  são ambos distribuídos normalmente. Eis um resultado mais geral sobre combinações lineares.

**PROPOSIÇÃO**

Se  $X_1, X_2, \dots, X_n$  são variáveis independentes, distribuídas normalmente (com médias e/ou variâncias possivelmente diferentes), então qualquer combinação linear dos  $X_i$ s também tem distribuição normal. Em particular, a própria diferença  $X_1 - X_2$  entre duas variáveis independentes e normalmente distribuídas é normalmente distribuída.

**Exemplo 5.30** (Continuação do Exemplo 5.28)

A receita total da venda dos três tipos de gasolina em um dia específico foi  $Y = 1,2X_1 + 1,35X_2 + 1,5X_3$ , e calculamos  $\mu_Y = 2325$  e (assumindo a independência)  $\sigma_Y = 178,01$ . Se os  $X_i$ s são distribuídos normalmente, a probabilidade de a receita exceder 2500 é

$$\begin{aligned} P(Y > 2500) &= P\left(Z > \frac{2500 - 2325}{178,01}\right) \\ &= P(Z > 0,98) = 1 - \Phi(0,98) = 0,1635 \end{aligned} \quad \blacksquare$$

O TLC também pode ser generalizado de modo que se aplique a certas combinações lineares. *Grosso modo*, se  $n$  for grande e nenhum termo individual igualmente provável de contribuir muito para o valor total, então  $Y$  possui aproximadamente uma distribuição normal.

**Provas para o caso de  $n = 2$** 

Para o resultado dos valores esperados, suponha que  $X_1$  e  $X_2$  sejam contínuas com fdp conjunta  $f(x_1, x_2)$ . Então

$$\begin{aligned} E(a_1X_1 + a_2X_2) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (a_1x_1 + a_2x_2)f(x_1, x_2) dx_1 dx_2 \\ &= a_1 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 f(x_1, x_2) dx_2 dx_1 \\ &\quad + a_2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_2 f(x_1, x_2) dx_1 dx_2 \\ &= a_1 \int_{-\infty}^{\infty} x_1 f_{X_1}(x_1) dx_1 + a_2 \int_{-\infty}^{\infty} x_2 f_{X_2}(x_2) dx_2 \\ &= a_1 E(X_1) + a_2 E(X_2) \end{aligned}$$

A soma substitui a integração no caso discreto. O argumento para o resultado da variância não requer que seja especificado se a variável é discreta ou contínua. Lembrando que  $V(Y) = E[(Y - \mu_Y)^2]$ ,

$$\begin{aligned} V(a_1X_1 + a_2X_2) &= E\{[a_1X_1 + a_2X_2 - (a_1\mu_1 + a_2\mu_2)]^2\} \\ &= E\{a_1^2(X_1 - \mu_1)^2 + a_2^2(X_2 - \mu_2)^2 + 2a_1a_2(X_1 - \mu_1)(X_2 - \mu_2)\} \end{aligned}$$

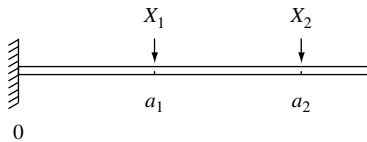
A expressão dentro das chaves é uma combinação linear das variáveis  $Y_1 = (X_1 - \mu_1)^2$ ,  $Y_2 = (X_2 - \mu_2)^2$ , e  $Y_3 = (X_1 - \mu_1)(X_2 - \mu_2)$ , assim, fazer a operação de  $E$  para os três termos resulta em  $a_1^2V(X_1) + a_2^2V(X_2) + 2a_1a_2\text{Cov}(X_1, X_2)$ , como requerido. ■

## Exercícios | Seção 5.5 (58–74)

- 58.** Uma empresa de remessa trabalha com contêineres de três tamanhos diferentes: (1) 27 pé<sup>3</sup> ( $3 \times 3 \times 3$ ), (2) 125 pé<sup>3</sup> e (3) 512 pé<sup>3</sup>. Seja  $X_i$  ( $i = 1, 2, 3$ ) o número de contêineres de tipo  $i$  enviados durante uma determinada semana. Com  $\mu_i = E(X_i)$  e  $\sigma_i^2 = V(X_i)$ , suponha que os valores médios e os desvios padrão sejam como segue:
- $$\begin{array}{lll} \mu_1 = 200 & \mu_2 = 250 & \mu_3 = 100 \\ \sigma_1 = 10 & \sigma_2 = 12 & \sigma_3 = 8 \end{array}$$
- a.** Assumindo que  $X_1, X_2, X_3$  sejam independentes, calcule o valor esperado e a variância do volume total enviado. [Sugestão: volume =  $27X_1 + 125X_2 + 512X_3$ .]
- b.** Seus cálculos estariam necessariamente corretos se os  $X_i$ s não fossem independentes? Explique.
- 59.** Sejam  $X_1, X_2$ , e  $X_3$  os tempos necessários para realizar três reparos sucessivos em determinada oficina. Suponha que sejam *vas normais independentes* com valores esperados  $\mu_1, \mu_2$ , e  $\mu_3$  e variâncias  $\sigma_1^2, \sigma_2^2$ , e  $\sigma_3^2$ , respectivamente.
- a.** Se  $\mu = \mu_2 = \mu_3 = 60$  e  $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 15$ , calcule  $P(X_1 + X_2 + X_3 \leq 200)$ . Qual é  $P(150 \leq X_1 + X_2 + X_3 \leq 200)$ ?
- b.** Usando os  $\mu_i$ s e  $\sigma_i$ s dados na parte (a), calcule  $P(55 \leq \bar{X})$  e  $P(58 \leq \bar{X} \leq 62)$ .
- c.** Usando os  $\mu_i$ s e  $\sigma_i$ s dados na parte (a), calcule  $P(-10 \leq X_1 - 0,5X_2 - 0,5X_3 \leq 5)$ .
- d.** Se  $\mu_1 = 40, \mu_2 = 50, \mu_3 = 60, \sigma_1^2 = 10, \sigma_2^2 = 12$  e  $\sigma_3^2 = 14$ , calcule  $P(X_1 + X_2 + X_3 \leq 160)$  e  $P(X_1 + X_2 \geq 2X_3)$ .
- 60.** Cinco automóveis do mesmo tipo farão uma viagem de 300 milhas. Os dois primeiros usarão uma marca de gasolina econômica, e os outros três, uma marca renomada. Sejam  $X_1, X_2, X_3, X_4$ , e  $X_5$  os consumos de combustível observados (milhas/galão) dos cinco carros. Suponha que essas variáveis sejam independentes e normalmente distribuídas com  $\mu_1 = \mu_2 = 20, \mu_3 = \mu_4 = \mu_5 = 21$  e  $\sigma^2 = 4$  para a marca econômica e 3,5 para a marca renomada. Defina a *va*  $Y$  como
- $$Y = \frac{X_1 + X_2}{2} - \frac{X_3 + X_4 + X_5}{3}$$
- de modo que  $Y$  seja a medida da diferença do consumo entre a gasolina econômica e a renomada. Calcule  $P(0 \leq Y)$  e  $P(-1 \leq Y \leq 1)$ . [Sugestão:  $Y = a_1X_1 + \dots + a_5X_5$ , com  $a_1 = \frac{1}{2}, \dots, a_5 = -\frac{1}{3}$ .]
- 61.** O Exercício 26 introduziu as variáveis aleatórias  $X$  e  $Y$ , número de carros e de ônibus, respectivamente, transportados por uma balsa em uma única viagem. A *fmp* conjunta de  $X$  e  $Y$  é dada na tabela do Exercício 7. Demonstra-se com facilidade que  $X$  e  $Y$  são independentes.
- a.** Calcule o valor esperado, a variância e o desvio padrão do número total de veículos em uma única viagem.
- b.** Se cada carro pagar um pedágio de US\$ 3 e cada ônibus de US\$ 10, calcule o valor esperado, a variância e o desvio padrão da receita resultante de uma única viagem.
- 62.** A fabricação de um componente requer três operações diferentes de usinagem. O tempo de usinagem de cada operação possui distribuição normal e os três tempos são independentes um do outro. Os valores médios são 15, 30 e 20 min, respectivamente, e os desvios padrão são 1, 2 e 1,5 min, respectivamente. Qual é a probabilidade de a usinagem levar no máximo 1 hora para produzir um componente selecionado aleatoriamente?
- 63.** Consulte o Exercício 3.
- a.** Calcule a covariância entre  $X_1 =$  o número de clientes na caixa expressa e  $X_2 =$  o número de clientes na caixa superexpressa.
- b.** Calcule  $V(X_1 + X_2)$ . Como este valor se compara a  $V(X_1) + V(X_2)$ ?
- 64.** Suponha que seu tempo de espera por um ônibus, de manhã, seja uniformemente distribuído em  $[0, 8]$ , enquanto à noite, é uniformemente distribuído em  $[0, 10]$  independente do tempo de espera pela manhã.
- a.** Se você tomar o ônibus toda manhã e toda noite por uma semana, qual é seu tempo de espera total? (Sugestão: Defina as *vas*  $X_1, \dots, X_{10}$  e use uma regra de valor esperado.)
- b.** Qual é a variância do seu tempo de espera total?
- c.** Quais são o valor esperado e a variância da diferença entre os tempos de espera pela manhã e à noite em um dia?

d. Quais são o valor esperado e a variância da diferença entre o tempo de espera total pela manhã e o tempo de espera total à noite, durante uma semana específica?

65. Suponha que, quando o pH de certo composto químico for 5,00, o pH medido por um aluno iniciante de química selecionado ao acaso é uma variável aleatória com média 5,00 e desvio padrão 0,2. Um grande lote do composto é subdividido e uma amostra foi dada a cada aluno do laboratório da manhã e do laboratório da tarde. Seja  $\bar{X}$  = pH médio determinado pelos alunos da manhã e  $\bar{Y}$  = o pH médio determinado pelos alunos da tarde.
- a. Se o pH for uma variável normal e houver 25 alunos em cada laboratório, calcule  $P(-0,1 \leq \bar{X} - \bar{Y} \leq 0,1)$ . (Sugestão:  $\bar{X} - \bar{Y}$  é uma combinação linear de variáveis normais, portanto, é normalmente distribuída. Calcule  $\mu_{\bar{X}-\bar{Y}}$  e  $\sigma_{\bar{X}-\bar{Y}}$ .)
- b. Se há 36 alunos em cada laboratório, mas as determinações do pH não são consideradas normais, calcule (aproximadamente)  $P(-0,1 \leq \bar{X} - \bar{Y} \leq 0,1)$ .
66. Se duas cargas são aplicadas a uma viga em balanço, conforme mostra o desenho a seguir, o momento fletor em 0 devido às cargas é  $a_1X_1 + a_2X_2$ .



- a. Suponha que  $X_1$  e  $X_2$  sejam variáveis independentes com médias de 2 e 4 kips / 2000 e 4000 lb, respectivamente, e desvios padrão de 0,5 e 1,0 kip / 500 e 100 lb, respectivamente. Se  $a_1 = 5$  ft / 1,52 m e  $a_2 = 10$  ft / 3,04 m, qual é o momento fletor esperado e qual é o seu desvio padrão?
- b. Se  $X_1$  e  $X_2$  são normalmente distribuídos, qual é a probabilidade de o momento fletor exceder 75 kip-ft?
- c. Suponha que as posições das duas cargas sejam variáveis aleatórias. Representando-as por  $A_1$  e  $A_2$ , assumamos que essas variáveis tenham médias de 5 ft e 10 ft, respectivamente, que cada uma tenha desvio padrão de 0,5 e que todos os  $A_i$ s e  $X_i$ s sejam independentes uns dos outros. Qual é o momento esperado, agora?
- d. Para a situação da parte (c), qual é a variância do momento fletor?
- e. Se a situação for a descrita na parte (a), exceto pelo fato de que  $\text{Corr}(X_1, X_2) = 0,5$  (de modo que as duas cargas não sejam independentes), qual é a variância do momento fletor?
67. Um pedaço de cano de PVC será colocado dentro de outro pedaço. O comprimento do primeiro pedaço é distribuído normalmente com valor médio de 20 pol e desvio padrão de 0,5 pol. O comprimento do segundo pedaço é uma variável normal com média e desvio padrão de 15 pol e 0,4 pol, respectivamente. O valor da superposição é distribuído normalmente com valor médio de 1 pol e desvio padrão de 0,1 pol. Assumindo que os comprimentos e o valor da superposição sejam inde-

pendentes um do outro, qual é a probabilidade de o comprimento total após a inserção estar entre 34,5 pol e 35 pol?

68. Dois aviões estão voando na mesma direção em corredores paralelos adjacentes. No tempo  $t = 0$ , o primeiro avião está 10 km à frente do segundo. Suponha que a velocidade do primeiro (km/h) seja distribuída normalmente com média de 520 e desvio padrão de 10, e que a velocidade do segundo avião também seja normalmente distribuída com média e desvio padrão de 500 e 10, respectivamente.
- a. Qual é a probabilidade de, após 2 h de voo, o segundo avião não ter alcançado o primeiro?
- b. Determine a probabilidade de os aviões estarem separados por no máximo 10 km após 2 h.
69. Três estradas diferentes se encontram na entrada de uma determinada rodovia. Suponha que durante um período de tempo estabelecido, o número de carros que vêm de cada estrada para a rodovia seja uma variável aleatória, com valor esperado e desvio padrão dados na tabela.

	Estrada 1	Estrada 2	Estrada 3
Valor esperado	800	1000	600
Desvio padrão	16	25	18

- a. Qual é o número total esperado de carros que entram na rodovia nesse ponto durante o período? (Sugestão: represente por  $X_i$  = o número da estrada  $i$ .)
- b. Qual é a variância do número total de carros que entram? Você fez alguma hipótese sobre a relação entre os números de carros nas diferentes estradas?
- c. Com  $X_i$  representando o número de carros que entram da estrada  $i$  durante o período, suponha que  $\text{Cov}(X_1, X_2) = 80$ ,  $\text{Cov}(X_1, X_3) = 90$ , e  $\text{Cov}(X_2, X_3) = 100$  (de modo que os três fluxos de trânsito não sejam independentes). Calcule o número total esperado de carros que entram e o desvio padrão do total.
70. Suponha que tomemos uma amostra aleatória de tamanho  $n$  de uma distribuição contínua que possui mediana 0, de modo que a probabilidade de qualquer observação ser positiva seja de 0,5. Agora, desprezamos os sinais das observações, classificamos-nas da menor para a maior em valor absoluto e, então, representamos por  $W$  = soma das classificações das observações que têm sinal positivo. Por exemplo: se as observações forem  $-0,3$ ,  $+0,7$ ,  $+2,1$ , e  $-2,5$ , então as classificações das observações positivas serão 2 e 3, assim,  $W = 5$ . No Capítulo 15,  $W$  será denominada estatística de postos de Wilcoxon.  $W$  pode ser representada como segue:

$$W = 1 \cdot Y_1 + 2 \cdot Y_2 + 3 \cdot Y_3 + \dots + n \cdot Y_n$$

$$= \sum_{i=1}^n i \cdot Y_i$$

onde os  $Y_i$ s são variáveis de Bernoulli independentes, cada uma com  $p = 0,5$  ( $Y_i = 1$  corresponde à observação com classificação  $i$  positiva). Calcule os dados a seguir:

a.  $E(Y_i)$  e, então,  $E(W)$ , usando a equação para  $W$  [Sugestão: a soma dos  $n$  primeiros inteiros positivos é  $n(n + 1)/2$ .]



- b.  $V(Y_i)$  e, então  $V(W)$  [Sugestão: a soma dos quadrados dos  $n$  primeiros inteiros positivos é  $n(n+1)(2n+1)/6$ .]
71. No Exercício 66, o peso próprio da viga contribui para o momento fletor. Assuma que a viga seja de espessura e densidade uniformes, de modo que a carga resultante seja distribuída uniformemente. Se o peso da viga for aleatório, a carga resultante do peso também será; represente essa carga por  $W$  (kip-ft).
- a. Se a viga tem 12 pés de comprimento,  $W$  tem média de 1,5 e desvio padrão de 0,25 e as cargas fixas são as descritas na parte (a) do Exercício 66, quais são o valor esperado e a variância do momento fletor? (Sugestão: se a carga, devido à viga, fosse  $w$  kip-pé, a contribuição para o momento fletor seria  $w \int_0^{12} x dx$ .)
- b. Se as três variáveis ( $X_1$ ,  $X_2$ , e  $W$ ) são normalmente distribuídas, qual é a probabilidade de o momento fletor ser de no máximo 200 kip-ft?
72. Sou responsável por três tarefas no prédio administrativo. Seja  $X_i$  = o tempo que leva a  $i$ -ésima tarefa ( $i = 1, 2, 3$ ), e  $X_4$  = o tempo total, em minutos, que eu percorro, andando pelo prédio e entre cada tarefa. Suponha que os  $X_i$ s sejam independentes, normalmente distribuídos, com as seguintes médias e desvios padrão:  $\mu_1 = 15$ ,  $\sigma_1 = 4$ ,  $\mu_2 = 5$ ,  $\sigma_2 = 1$ ,  $\mu_3 = 8$ ,  $\sigma_3 = 2$ ,  $\mu_4 = 12$ ,  $\sigma_4 = 3$ . Pretendo sair do escritório exatamente às 10 h e deixar um bilhete na porta, dizendo: “Retornarei à  $t$  h”. A que horas  $t$  devo escrever, se quiser que a probabilidade da minha chegada depois de  $t$  seja 0,01?
73. Suponha que a resistência de tração esperada do aço tipo A seja 105 ksi e que o desvio padrão da resistência de tração seja 8 ksi. Para o aço tipo B, suponha que a resistência de tração esperada e o desvio padrão da resistência de tração sejam 100 ksi e 6 ksi, respectivamente. Seja  $\bar{X}$  = a resistência de tração média de uma amostra aleatória de 40 espécimes tipo A, e por  $\bar{Y}$  = a resistência de tração média de uma amostra aleatória de 35 espécimes tipo B.
- a. Qual é a distribuição aproximada de  $\bar{X}$ ? De  $\bar{Y}$ ?
- b. Qual é a distribuição aproximada de  $\bar{X} - \bar{Y}$ ? Justifique sua resposta.
- c. Calcule (aproximadamente)  $P(-1 \leq \bar{X} - \bar{Y} \leq 1)$ .
- d. Calcule  $P(\bar{X} - \bar{Y} \geq 10)$ . Se realmente observou  $\bar{X} - \bar{Y} \geq 10$ , você duvidaria que  $\mu_1 - \mu_2 = 5$ ?
74. Em uma área que possui solo arenoso, foram plantadas 50 árvores pequenas de um determinado tipo e outras 50 foram plantadas em uma área de solo argiloso. Sejam  $X$  = número de árvores plantadas no solo arenoso que sobrevivem 1 ano, e  $Y$  = número de árvores plantadas no solo argiloso que sobrevivem 1 ano. Se a probabilidade de a árvore plantada no solo arenoso sobreviver 1 ano for de 0,7 e de sobreviver 1 ano no solo argiloso for de 0,6, calcule uma aproximação a  $P(-5 \leq X - Y \leq 5)$  (não se preocupe com a correção de continuidade).

---

### Exercícios Suplementares (75–94)

75. Um restaurante serve três pratos ao preço fixo de US\$ 12, US\$ 15 e US\$ 20. Para um casal selecionado aleatoriamente que janta nesse restaurante, sejam  $X$  = custo do jantar do homem e  $Y$  = custo do jantar da mulher. A fmp conjunta de  $X$  e  $Y$  é dada na tabela a seguir:

$p(x, y)$		$y$		
		12	15	20
$x$	12	0,05	0,05	0,10
	15	0,05	0,10	0,35
	20	0	0,20	0,10

- a. Calcule as fmps marginais de  $X$  e  $Y$ .
- b. Qual é a probabilidade de o jantar do homem e da mulher custar no máximo US\$ 15 cada?
- c.  $X$  e  $Y$  são independentes? Justifique sua resposta.
- d. Qual é o custo total esperado do jantar para as duas pessoas?
- e. Suponha que, quando um casal abre biscoitos da sorte no final da refeição, encontra a mensagem “Você receberão de volta a diferença entre o custo da refeição mais cara e o da mais barata que escolheram”. Quanto o restaurante espera devolver?
76. Na estimativa de custo, o preço total de um projeto é a soma dos gastos das tarefas componentes. Cada custo é uma variável aleatória com uma distribuição de probabilidade. É costume obter informações sobre a

distribuição do custo total, somando-se características das distribuições de custo do componente individual; isso é chamado procedimento de *roll-up*. Por exemplo:  $E(X_1 + \dots + X_n) = E(X_1) + \dots + E(X_n)$ , assim, o procedimento de *roll-up* é válido para o custo médio. Suponha que existam duas tarefas componentes e que  $X_1$  e  $X_2$  sejam variáveis aleatórias independentes e normalmente distribuídas. O procedimento de *roll-up* é válido para o 75º percentil? Isto é, o 75º percentil da distribuição de  $X_1 + X_2$  é o mesmo que a soma dos 75º percentis das duas distribuições individuais? Se não, qual é a relação entre o percentil da soma e a soma dos percentis? Para quais percentis o procedimento de *roll-up* é válido, nesse caso?

77. Uma loja de comida saudável possui duas marcas diferentes de um certo tipo de grão. Sejam  $X$  = quantidade (lb) da marca A em estoque e  $Y$  = quantidade da marca B em estoque. Suponha que a fdp conjunta de  $X$  e  $Y$  seja

$$f(x, y) = \begin{cases} kxy & x \geq 0, y \geq 0, 20 \leq x + y \leq 30 \\ 0 & \text{caso contrário} \end{cases}$$

- a. Desenhe a região de densidade positiva e determine o valor de  $k$ .
- b.  $X$  e  $Y$  são independentes? Responda, encontrando primeiro a fdp marginal de cada variável.

- c. Calcule  $P(X + Y \leq 25)$ .
- d. Qual é a quantidade total esperada desse grão em estoque?
- e. Calcule  $\text{Cov}(X, Y)$  e  $\text{Corr}(X, Y)$ .
- f. Qual é a variância da quantidade total de grão em estoque?
78. Sejam  $X_1, X_2, \dots, X_n$  variáveis aleatórias que representam  $n$  propostas independentes de um item que está à venda. Suponha que cada  $X_i$  seja uniformemente distribuído no intervalo  $[100, 200]$ . Se o vendedor vende pela oferta maior, quanto pode esperar ganhar com a venda? [Sugestão: seja  $Y = \max(X_1, X_2, \dots, X_n)$ . Encontre primeiro  $F_Y(y)$ , observando que  $Y \leq y$  se e somente se cada  $X_i$  for  $\leq y$ . Então, obtenha a fdp e  $E(Y)$ .]
79. Suponha que, para certo indivíduo, a ingestão de calorias no café da manhã seja uma variável aleatória com valor esperado de 500 e desvio padrão de 50, a ingestão de calorias no almoço seja aleatória com valor esperado de 900 e desvio padrão de 100, e a ingestão de calorias no jantar seja uma variável aleatória com valor esperado de 2000 e desvio padrão de 180. Assumindo que os consumos nas diferentes refeições sejam independentes, qual é a probabilidade de a ingestão média diária de calorias durante o ano seguinte (365 dias) ser de no máximo 3500? [Sugestão: sejam  $X_i, Y_i$ , e  $Z_i$  as três ingestões de calorias no dia  $i$ . Então, a ingestão total é dada por  $\sum(X_i + Y_i + Z_i)$ .]
80. O peso médio da bagagem de um passageiro da classe turística selecionado aleatoriamente, que voa entre duas cidades em uma determinada linha aérea, é 40 lb, e o desvio padrão é 10 lb. A média e o desvio padrão para um passageiro da classe executiva são 30 lb e 6 lb, respectivamente.
- a. Se há 12 passageiros na classe executiva e 50 na classe turística em certo voo, quais são o valor esperado e o desvio padrão do peso total da bagagem?
- b. Se os pesos individuais da bagagem forem independentes e normalmente distribuídas, qual é a probabilidade de o peso total da bagagem ser de no máximo 2500 lb?
81. Vimos que, se  $E(X_1) = E(X_2) = \dots = E(X_n) = \mu$ , então,  $E(X_1 + \dots + X_n) = n\mu$ . Em algumas aplicações, o número de  $X_i$ s considerados não é um número fixo  $n$ , mas uma  $N$ . Por exemplo: sejam  $N =$  o número de componentes levados a uma oficina em certo dia, e  $X_i$  o tempo de reparo do  $i$ -ésimo componente. Então, o tempo de reparo total é  $X_1 + X_2 + \dots + X_N$ , a soma de um número aleatório de variáveis aleatórias. Quando  $N$  for independente dos  $X_i$ s, pode-se mostrar que
- $$E(X_1 + \dots + X_N) = E(N) \cdot \mu$$
- a. Se o número esperado de componentes levados em um dia for 10 e o tempo de conserto esperado de um componente enviado aleatoriamente for 40 min, qual é o tempo total esperado de conserto dos componentes enviados em um dia qualquer?
- b. Suponha que certo tipo de componente chegue para conserto, de acordo com um processo de Poisson, com um índice de 5 por hora. O número esperado de defeitos por componente é 3,5. Qual é o valor esperado do número total de defeitos em componentes enviados para conserto durante um período de 4 horas? Certifique-se de indicar como sua resposta se desenvolve a partir do resultado geral dado.
82. Suponha que a proporção de eleitores rurais em um determinado estado que favorece certo candidato a governador seja de 0,45 e que a proporção de eleitores urbanos e suburbanos que o favorecem seja de 0,60. Se uma amostra de 200 eleitores rurais e 300 urbanos e suburbanos for obtida, qual será a probabilidade aproximada de pelo menos 250 desses eleitores favorecerem o candidato?
83. Seja  $\mu$  o pH real de um composto químico. Será feita uma seqüência de  $n$  determinações independentes do pH da amostra. Suponha que o pH de cada amostra seja uma variável aleatória com valor esperado  $\mu$  e desvio padrão 0,1. Quantas determinações serão necessárias se quisermos que a probabilidade da média amostral esteja entre 0,02 do pH real seja de no mínimo 0,95? Que teorema justifica seu cálculo de probabilidade?
84. Se a quantidade de refrigerante que consumo em certo dia é independente do consumo em qualquer outro dia e é normalmente distribuída com  $\mu = 13$  oz e  $\sigma = 2$ , e se eu atualmente tenho duas caixas com seis garrafas de 16 oz, qual é a probabilidade de eu ainda ter refrigerante ao fim de duas semanas (14 dias)?
85. Consulte o Exercício 58 e suponha que os  $X_i$ s sejam independentes com cada um tendo distribuição normal. Qual é a probabilidade de o volume total embarcado ser de no máximo 100.000 pé<sup>3</sup>?
86. Um aluno tem uma aula que supostamente termina às 9 h e outra que começa às 9h10. Suponha que a hora de término real da aula das 9 h seja uma variável normalmente distribuída  $X_1$  com média de 9h02 e desvio padrão de 1,5 min e que a hora de início da aula seguinte também seja uma variável normalmente distribuída  $X_2$  com média de 9h10 e desvio padrão de 1 min. Suponha também que o tempo necessário para sair de uma sala e chegar à outra seja uma variável normalmente distribuída  $X_3$  com média de 6 min e desvio padrão de 1 min. Qual é a probabilidade de o aluno chegar à segunda aula antes de ela começar? (Assuma independência de  $X_1, X_2$ , e  $X_3$ , que é razoável se o aluno não prestar atenção à hora do término da primeira aula.)
87. a. Use a fórmula geral da variância de uma combinação linear para escrever uma expressão para  $V(aX + Y)$ . Então, seja  $a = \sigma_Y/\sigma_X$ , e mostre que  $\rho \geq -1$ . [Sugestão: a variância sempre é  $\geq 0$ , e  $\text{Cov}(X, Y) = \sigma_X \cdot \sigma_Y \cdot \rho$ .]
- b. Considerando que  $V(aX - Y)$ , conclua que  $\rho \leq 1$ .
- c. Use o fato de  $V(W) = 0$  somente se  $W$  for uma constante para mostrar que  $\rho = 1$  somente se  $Y = aX + b$ .

88. Suponha que a pontuação verbal  $X$  e a pontuação quantitativa  $Y$  de um indivíduo escolhido aleatoriamente em um exame de aptidão administrado nacionalmente tenham fdp conjunta.

$$f(x, y) = \begin{cases} \frac{2}{5}(2x + 3y) & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{caso contrário} \end{cases}$$

Você é solicitado a fornecer uma previsão  $t$  da pontuação total do indivíduo  $X + Y$ . O erro de previsão é o erro médio quadrado  $E[(X + Y - t)^2]$ . Que valor de  $t$  minimiza o erro de previsão?

89. a.  $X_1$  tem distribuição qui-quadrado com parâmetro  $\nu_1$  (veja a Seção 4.4),  $X_2$  é independente de  $X_1$  e tem uma distribuição qui-quadrado com parâmetro  $\nu_2$ . Use a técnica do Exemplo 5.21 para mostrar que  $X_1 + X_2$  tem uma distribuição qui-quadrado com parâmetro  $\nu_1 + \nu_2$ .
- b. No Exercício 65 do Capítulo 4, você foi solicitado a mostrar que, se  $Z$  é uma va normal-padrão, então  $Z^2$  possui uma distribuição qui-quadrado com  $\nu = 1$ . Sejam  $Z_1, Z_2, \dots, Z_n$ , as vas normais-padrão independentes de  $n$ . Qual é a distribuição de  $Z_1^2 + \dots + Z_n^2$ ? Justifique sua resposta.
- c. Seja  $X_1, \dots, X_n$  uma amostra aleatória de uma distribuição normal com média  $\mu$  e variância  $\sigma^2$ . Qual é a distribuição da soma  $Y = \sum_{i=1}^n [(X_i - \mu)/\sigma]^2$ ? Justifique sua resposta.
90. a. Mostre que  $\text{Cov}(X, Y + Z) = \text{Cov}(X, Y) + \text{Cov}(X, Z)$ .
- b. Sejam  $X_1$  e  $X_2$  as pontuações verbal e quantitativa em um exame de aptidão, e  $Y_1$  e  $Y_2$  as pontuações correspondentes em outro exame. Se  $\text{Cov}(X_1, Y_1) = 5$ ,  $\text{Cov}(X_1, Y_2) = 1$ ,  $\text{Cov}(X_2, Y_1) = 2$ , e  $\text{Cov}(X_2, Y_2) = 8$ , qual é a covariância entre as duas pontuações totais  $X_1 + X_2$  e  $Y_1 + Y_2$ ?

91. Um espécime de rocha de uma área específica é selecionado aleatoriamente e pesado duas vezes. Represente por  $W$  o peso real e por  $X_1$  e  $X_2$  os dois pesos medidos. Então,  $X_1 = W + E_1$  e  $X_2 = W + E_2$ , onde  $E_1$  e  $E_2$  são dois erros de medição. Suponha que os  $E_i$ s sejam independentes entre si e de  $W$  e que  $V(E_1) = V(E_2) = \sigma_E^2$ .

- a. Expresse por  $\rho$  o coeficiente de correlação entre os dois pesos medidos  $X_1$  e  $X_2$ , em termos de  $\sigma_W^2$ , a variância do peso real, e por  $\sigma_X^2$ , a variância do peso medido.
- b. Calcule  $\rho$  quando  $\sigma_W = 1$  kg e  $\sigma_E = 0,01$  kg.

92. Sejam  $A$  a porcentagem de um elemento de um espécime de rocha selecionado aleatoriamente, e  $B$  a por-

centagem de um segundo elemento desse mesmo espécime. Suponha que  $D$  e  $E$  sejam erros de medição ao determinar os valores de  $A$  e  $B$ , de modo que os valores medidos sejam  $X = A + D$  e  $Y = B + E$ , respectivamente. Assuma que os erros de medição sejam independentes um do outro e dos valores reais.

- a. Mostre que

$$\text{Corr}(X, Y) = \text{Corr}(A, B) \cdot \sqrt{\text{Corr}(X_1, X_2)} \cdot \sqrt{\text{Corr}(Y_1, Y_2)}$$

onde  $X_1$  e  $X_2$  são medições duplas do valor de  $A$  e  $Y_1$  e  $Y_2$  são definidos analogamente com relação a  $B$ . Que efeito a presença de erro de medição tem sobre a correlação?

- b. Qual é o valor máximo de  $\text{Corr}(X, Y)$  quando  $\text{Corr}(X_1, X_2) = 0,8100$  e  $\text{Corr}(Y_1, Y_2) = 0,9025$ ? Isso é preocupante?

93. Sejam  $X_1, \dots, X_n$  vas independentes com valores médios  $\mu_1, \dots, \mu_n$  e variâncias  $\sigma_1^2, \dots, \sigma_n^2$ . Considere uma função  $h(x_1, \dots, x_n)$  e use-a para definir uma nova va  $Y = h(X_1, \dots, X_n)$ . Sob condições bastante gerais sobre a função  $h$ , se os  $\sigma_i$ s forem pequenos em relação aos correspondentes  $\mu_i$ s, pode-se mostrar que  $E(Y) \approx h(\mu_1, \dots, \mu_n)$  e

$$V(Y) \approx \left(\frac{\partial h}{\partial x_1}\right)^2 \cdot \sigma_1^2 + \dots + \left(\frac{\partial h}{\partial x_n}\right)^2 \cdot \sigma_n^2$$

onde cada derivada parcial é avaliada em  $(x_1, \dots, x_n) = (\mu_1, \dots, \mu_n)$ . Suponha que três resistores com resistências  $X_1, X_2, X_3$  sejam conectados em paralelo a uma bateria com voltagem  $X_4$ . Então, pela lei de Ohm, a corrente é

$$Y = X_4 \left[ \frac{1}{X_1} + \frac{1}{X_2} + \frac{1}{X_3} \right]$$

Sejam  $\mu_1 = 10$  ohms,  $\sigma_1 = 1,0$  ohm,  $\mu_2 = 15$  ohms,  $\sigma_2 = 1,0$  ohm,  $\mu_3 = 20$  ohms,  $\sigma_3 = 1,5$  ohms,  $\mu_4 = 120$  V,  $\sigma_4 = 4,0$  V. Calcule o valor esperado aproximado e o desvio padrão da corrente (sugerido por "Random Samplings" (Amostras Aleatórias) CHEMTECH, 1984, p. 696-697).

94. Uma aproximação mais precisa de  $E[h(X_1, \dots, X_n)]$  do Exercício 93 é

$$h(\mu_1, \dots, \mu_n) + \frac{1}{2} \sigma_1^2 \left(\frac{\partial^2 h}{\partial x_1^2}\right) + \dots + \frac{1}{2} \sigma_n^2 \left(\frac{\partial^2 h}{\partial x_n^2}\right)$$

Calcule-a para  $Y = h(X_1, X_2, X_3, X_4)$ , dado no Exercício 93, e compare-a ao termo principal  $h(\mu_1, \dots, \mu_n)$ .

## Bibliografia

LARSEN, Richard e MARX Morris, *An Introduction to Mathematical Statistics and its Applications* (3. ed.). Prentice Hall, Englewood Cliffs, NJ, 2000. Cobertura mais limitada que a do livro de Olkin et al., porém mais bem escrita e agradável.

OLKIN, Ingram, DERMAN, Cyrus e GLENER, Leon *Probability Models and Application* (2. ed.), Macmillan, Nova York, 1994. Contém uma exposição cuidadosa e abrangente das distribuições conjuntas, regras de valor esperado e teoremas de limite.

## Estimativa Pontual

### Introdução

Dado um parâmetro de interesse, como a média da população  $\mu$  ou proporção da população  $p$ , o objetivo da estimativa pontual é usar uma amostra para calcular um número que represente, de certo modo, uma boa previsão do valor real do parâmetro. O número resultante é chamado estimativa pontual. Na Seção 6.1, apresentamos alguns conceitos gerais de estimativa pontual. Na Seção 6.2, descrevemos e ilustramos dois métodos importantes para se obter estimativas pontuais: o método dos momentos e o método da máxima verossimilhança.

### 6.1 Alguns Conceitos Gerais sobre Estimativa Pontual

A inferência estatística é quase sempre direcionada à obtenção de algum tipo de conclusão sobre um ou mais parâmetros (características da população). O processo requer que o pesquisador obtenha dados de amostras de cada população em estudo. As conclusões baseiam-se, então, nos valores calculados das várias quantidades da amostra. Por exemplo: seja  $\mu$  (um parâmetro) a tensão média de quebra real das conexões do fio usado na união dos *wafers* semicondutores. Uma amostra aleatória de  $n = 10$  conexões pode ser feita e a tensão de quebra de cada uma, determinada, resultando em tensões observadas  $x_1, x_2, \dots, x_{10}$ . A tensão média de quebra da amostra  $\bar{x}$  é, então, usada para tirar conclusões sobre o valor de  $\mu$ . De forma semelhante, se  $\sigma^2$  for a variância da distribuição da tensão de quebra (variância da população, outro parâmetro), o valor da variância da amostra  $s^2$  é usado para inferir alguma coisa sobre  $\sigma^2$ .

Ao discutir conceitos gerais e métodos de inferência, é conveniente definir um símbolo genérico para o parâmetro de interesse. Usaremos a letra grega  $\theta$  para esse fim. O objetivo da estimativa pontual é selecionar um único valor, com base nos dados da amostra, que represente um valor sensato para  $\theta$ . Suponha, por exemplo, que

o parâmetro de interesse seja  $\mu$ , a vida útil média real de certo tipo de pilhas. Uma amostra aleatória de  $n = 3$  pilhas resulta nos valores de vida útil observados (horas)  $x_1 = 5,0$ ,  $x_2 = 6,4$ ,  $x_3 = 5,9$ . O valor calculado da vida útil média amostral é  $\bar{x} = 5,77$ , e é razoável considerar  $5,77$  um valor bastante plausível de  $\mu$  — nossa “melhor previsão” para o valor de  $\mu$  com base nas informações disponíveis da amostra.

Suponha que queiramos estimar um parâmetro de uma única população (por exemplo,  $\mu$  ou  $\sigma$ ) com base em uma amostra aleatória de tamanho  $n$ . Lembre-se do capítulo anterior que, antes de os dados estarem disponíveis, as observações da amostra devem ser consideradas como variáveis aleatórias (vas)  $X_1, X_2, \dots, X_n$ . Acontece que qualquer função dos  $X_i$ s — isto é, qualquer estatística — como a média amostral  $\bar{X}$  ou o desvio padrão amostral  $S$  também é uma variável aleatória. O mesmo acontece se os dados disponíveis consistirem de mais de uma amostra. Por exemplo: podemos representar as resistências à tração de  $m$  espécimes tipo 1 e de  $n$  espécimes tipo 2 por  $X_1, \dots, X_m$  e  $Y_1, \dots, Y_n$ , respectivamente. A diferença entre as duas resistências médias amostrais  $\bar{X} - \bar{Y}$  é a estatística natural para fazer inferências sobre  $\mu_1 - \mu_2$ , a diferença entre as resistências médias da população.

## DEFINIÇÃO

Uma **estimativa pontual** de um parâmetro  $\theta$  é um único número que pode ser considerado um valor sensato para  $\theta$ . Obtém-se uma estimativa pontual selecionando uma estatística adequada e calculando seu valor pelos dados da amostra. A estatística selecionada é chamada **estimador pontual** de  $\theta$ .

No exemplo da pilha, o estimador usado para obter a estimativa pontual de  $\mu$  foi  $\bar{X}$ , e a estimativa pontual de  $\mu$  foi  $5,77$ . Se os três valores de vida útil observados forem  $x_1 = 5,6$ ,  $x_2 = 4,5$ , e  $x_3 = 6,1$ , o uso do estimador  $\bar{X}$  teria resultado na estimativa  $\bar{x} = (5,6 + 4,5 + 6,1)/3 = 5,40$ . O símbolo  $\hat{\theta}$  é habitualmente usado para representar como o estimador de  $\theta$  como a estimativa pontual resultante de uma amostra dada.<sup>1</sup> Dessa forma,  $\hat{\mu} = \bar{X}$  é lido como “estimador pontual de  $\mu$  é a média amostral  $\bar{X}$ ”. A afirmativa “a estimativa pontual de  $\mu$  é  $5,77$ ” é escrita resumidamente como  $\hat{\mu} = 5,77$ . Observe que, ao escrever  $\hat{\theta} = 72,5$ , não há indicação de como a estimativa pontual foi obtida (que estatística foi usada). Assim, recomenda-se que o estimador e a estimativa resultante sejam descritos.

## Exemplo 6.1

Um fabricante de automóveis desenvolveu um novo tipo de pára-choque que supostamente absorve impactos com menos danos que os pára-choques anteriores. O fabricante usou pára-choques novos em uma seqüência de 25 colisões controladas contra uma parede, cada uma a 10 mph, usando um de seus modelos de carro compacto. Seja  $X$  = número de colisões que resultam em danos imperceptíveis ao automóvel. O parâmetro a ser estimado é  $p$  = proporção de todas as colisões que não provocam danos [de maneira alternativa,  $p = P(\text{nenhum dano em uma única colisão})$ ]. Se  $X$  observado é  $x = 15$ , o estimador e a estimativa mais razoáveis são

$$\text{estimador } \hat{p} = \frac{X}{n} \quad \text{estimativa} = \frac{x}{n} = \frac{15}{25} = 0,60 \quad \blacksquare$$

Se, para cada parâmetro de interesse houvesse somente um estimador pontual razoável, não haveria muito o que discutir sobre a estimativa pontual. Na maioria dos problemas, entretanto, há mais de um estimador razoável.

## Exemplo 6.2

Reconsidere as 20 observações que seguem na voltagem dielétrica de quebra de peças de resina de epóxi introduzida primeiro no Exemplo 4.29 (Seção 4.6).

24,46	25,61	26,25	26,42	26,66	27,15	27,31	27,54	27,74	27,94
27,98	28,04	28,28	28,49	28,50	28,87	29,11	29,13	29,50	30,88

<sup>1</sup> Seguindo as notações anteriores, poderíamos usar  $\hat{\theta}$  ( $\theta$  maiúsculo) para o estimador, mas é trabalhoso escrevê-lo.

O padrão no gráfico de probabilidade normal mostrado é bastante reto; assim, assumimos que a distribuição da voltagem de quebra é normal com valor médio  $\mu$ . Em virtude de as distribuições normais serem simétricas,  $\mu$  também é a mediana da vida útil da distribuição. As observações dadas são, então, assumidas como o resultado de uma amostra aleatória  $X_1, X_2, \dots, X_{20}$  dessa distribuição normal. Considere os seguintes estimadores e estimativas resultantes para  $\mu$ :

- Estimador =  $\bar{X}$ , estimativa =  $\bar{x} = \sum x_i/n = 555,86/20 = 27,793$
- Estimador =  $\tilde{X}$ , estimativa =  $\tilde{x} = (27,94 + 27,98)/2 = 27,960$
- Estimador =  $[\min(X_i) + \max(X_i)]/2$  = a média das duas vidas úteis extremas, estimativa =  $[\min(x_i) + \max(x_i)]/2 = (24,46 + 30,88)/2 = 27,670$
- Estimador =  $\bar{X}_{tr(10)}$ , a média aparada de 10% (despreze os 10% menores e maiores da amostra e calcule a média),

$$\begin{aligned} \text{estimativa} &= \bar{x}_{tr(10)} \\ &= \frac{555,86 - 24,46 - 25,61 - 29,50 - 30,88}{16} \\ &= 27,838 \end{aligned}$$

Cada um dos estimadores (a)–(d) utiliza uma medida diferente do centro da amostra para estimar  $\mu$ . Qual das estimativas está mais próxima do valor real? Não podemos responder sem saber o valor real. Uma pergunta que pode ser respondida é: “Que estimador, quando usado em outras amostras dos  $X_i$ s, tenderá a produzir estimativas mais próximas do valor real?” Consideraremos em breve esse tipo de pergunta. ■

### Exemplo 6.3

No futuro próximo haverá um interesse crescente no desenvolvimento de ligas à base de Mg de baixo custo para diversos processos de fundição. Por isso, é importante ter maneiras práticas de determinar as várias propriedades mecânicas de tais ligas. O artigo “On the Development of a New Approach for the Determination of Yield Strength in Mg-based Alloys” (*Light Metal Age*, out. 1998, p. 50-53) propôs um método ultra-sônico para esse fim. Considere a seguinte amostra de observações no módulo de elasticidade (GPa) de espécimes da liga AZ91D de um processo de fundição:

44,2    43,9    44,7    44,2    44,0    43,8    44,6    43,1

Assuma que essas observações sejam o resultado de uma amostra aleatória  $X_1, \dots, X_8$  da distribuição da população de módulo elástico sob tais circunstâncias. Queremos estimar a variância da população  $\sigma^2$ . Um estimador natural é a variância amostral:

$$\hat{\sigma}^2 = S^2 = \frac{\sum (X_i - \bar{X})^2}{n - 1} = \frac{\sum X_i^2 - (\sum X_i)^2/n}{n - 1}$$

A estimativa correspondente é

$$\begin{aligned} \hat{\sigma}^2 = s^2 &= \frac{\sum x_i^2 - (\sum x_i)^2/8}{7} = \frac{15.533,79 - (352,5)^2/8}{7} \\ &= 0,25125 \approx 0,251 \end{aligned}$$

A estimativa de  $\sigma$  seria, então,  $\hat{\sigma} = s = \sqrt{0,25125} = 0,501$ .

Um estimador alternativo seria resultado do uso do divisor  $n$  em vez de  $n - 1$  (isto é, o desvio elevado ao quadrado médio):

$$\hat{\sigma}^2 = \frac{\sum (X_i - \bar{X})^2}{n} \quad \text{estimativa} = \frac{1,75875}{8} = 0,220$$

Indicaremos em breve o porquê de muitos estatísticos preferirem  $S^2$  ao estimador com divisor  $n$ . ■

No melhor dos mundos, podemos encontrar um estimador  $\hat{\theta}$  para o qual  $\hat{\theta} = \theta$ , sempre. Entretanto,  $\hat{\theta}$  é uma função da amostra dos  $X_i$ s, portanto, é uma variável aleatória. Para algumas amostras,  $\hat{\theta}$  produzirá um valor maior que  $\theta$ , enquanto, para outras,  $\hat{\theta}$  subestimar $\hat{\theta}$   $\theta$ . Se escrevermos

$$\hat{\theta} = \theta + \text{erro de estimativa}$$

então, um estimador preciso seria aquele que resultasse em pequenos erros de estimativa, de modo que os valores estimados estarão próximos do valor real. Um estimador que possui as propriedades de não-tendenciosidade e variância mínima geralmente será preciso nesse sentido.

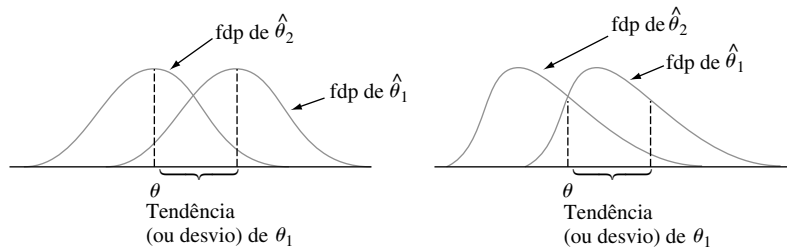
### Estimadores Não-tendenciosos

Suponha que tenhamos dois instrumentos de medidas: um foi calibrado com precisão, mas o outro fornece sistematicamente leituras menores do que o valor real que está sendo medido. Quando cada instrumento for utilizado repetidas vezes no mesmo objeto devido ao erro de medição, as medições observadas não serão idênticas. Entretanto, as medições produzidas pelo primeiro instrumento serão distribuídas com relação ao valor real de tal maneira que, na média, esse instrumento mede o que se propõe a medir; assim, é chamado de instrumento não-tendencioso. O segundo instrumento produz observações que têm um componente de erro sistemático ou desvio.

**DEFINIÇÃO**

Um estimador pontual  $\hat{\theta}$  é considerado um **estimador não-tendencioso** de  $\theta$  se  $E(\hat{\theta}) = \theta$  para cada valor possível de  $\theta$ . Se  $\hat{\theta}$  não for não-tendencioso, a diferença  $E(\hat{\theta}) - \theta$  é chamada **desvio** ou **tendenciosidade** de  $\hat{\theta}$ .

Isto é,  $\hat{\theta}$  é não-tendencioso se sua distribuição de probabilidade (isto é, amostragem) estiver sempre “centrada” no valor real do parâmetro. Suponha que  $\hat{\theta}$  seja um estimador não-tendencioso; então, se  $\theta = 100$ , a distribuição amostral de  $\hat{\theta}$  está centralizada em 100; se  $\theta = 27,5$ , então, a distribuição amostral de  $\hat{\theta}$  está centrada em 27,5 e assim por diante. A Figura 6.1 ilustra as distribuições de diversos estimadores tendenciosos e não-tendenciosos. Observe que “centrado”, aqui, significa que o valor esperado, não a mediana, da distribuição de  $\hat{\theta}$  é igual a  $\theta$ .



**Figura 6.1** As fdps de um estimador tendencioso  $\hat{\theta}_1$  e um estimador não-tendencioso  $\hat{\theta}_2$ , para o parâmetro  $\theta$

Talvez seja necessário saber o valor de  $\theta$  (caso em que a estimativa é desnecessária) para ver se  $\hat{\theta}$  é não-tendencioso. Esse geralmente não é o caso, pois muitas vezes o argumento de um valor esperado geral pode ser usado para verificar a não-tendenciosidade.

No Exemplo 6.1, a proporção amostral  $X/n$  foi usada como um estimador de  $p$ , onde  $X$ , o número de sucessos da amostra, tinha uma distribuição binomial com parâmetros  $n$  e  $p$ . Assim,

$$E(\hat{p}) = E\left(\frac{X}{n}\right) = \frac{1}{n} E(X) = \frac{1}{n} (np) = p$$

**PROPOSIÇÃO**

Quando  $X$  é uma va binomial com parâmetros  $n$  e  $p$ , a proporção amostral  $\hat{p} = X/n$  é um estimador não-tendencioso de  $p$ .

Qualquer que seja o valor real de  $p$ , a distribuição do estimador  $\hat{p}$  estará centrada no valor real.

#### Exemplo 6.4

Suponha que  $X$ , tempo de reação a certo estímulo, possua distribuição uniforme no intervalo de 0 a um limite superior desconhecido  $\theta$  (assim, a função de densidade de  $X$  tem formato retangular com altura  $1/\theta$  para  $0 \leq x \leq \theta$ ). Deseja-se estimar  $\theta$  com base em uma amostra aleatória  $X_1, X_2, \dots, X_n$  de tempos de reação. Uma vez que  $\theta$  é o maior tempo possível em toda a população de tempos de reação, considere como um primeiro estimador o maior tempo de reação da amostra:  $\hat{\theta}_1 = \max(X_1, \dots, X_n)$ . Se  $n = 5$  e  $x_1 = 4,2, x_2 = 1,7, x_3 = 2,4, x_4 = 3,9, x_5 = 1,3$ , a estimativa pontual de  $\theta$  é  $\hat{\theta}_1 = \max(4,2, 1,7, 2,4, 3,9, 1,3) = 4,2$ .

A não-tendenciosidade implica que algumas amostras produzirão estimativas que excedem  $\theta$  e que outras produzirão estimativas menores que  $\theta$ . Caso contrário,  $\theta$  possivelmente não poderia ser o centro (ponto de equilíbrio) da distribuição dos  $\hat{\theta}_1$ . Entretanto, nosso estimador proposto nunca superestimar  $\theta$  (o maior valor da amostra não pode exceder o maior valor da população) e subestimar  $\theta$ , a menos que o maior valor amostral seja igual a  $\theta$ . Este argumento intuitivo mostra que  $\hat{\theta}_1$  é um estimador tendencioso. Mais precisamente, pode-se demonstrar (veja o Exercício 32) que

$$E(\hat{\theta}_1) = \frac{n}{n+1} \cdot \theta < \theta \quad \left( \text{uma vez que } \frac{n}{n+1} < 1 \right)$$

O desvio de  $\hat{\theta}_1$  é dado por  $n\theta/(n+1) - \theta = -\theta/(n+1)$ , que se aproxima de 0 à medida que  $n$  aumenta. É fácil modificar  $\hat{\theta}_1$  para obter um estimador não-tendencioso de  $\theta$ . Considere o estimador

$$\hat{\theta}_2 = \frac{n+1}{n} \cdot \max(X_1, \dots, X_n)$$

Aplicar este estimador nos dados fornece a estimativa  $(6/5)(4,2) = 5,04$ . O fato de  $(n+1)/n > 1$  implica que  $\hat{\theta}_2$  superestimar  $\theta$  para algumas amostras e o subestimar para outras. O valor médio desse estimador é

$$\begin{aligned} E(\hat{\theta}_2) &= E\left[\frac{n+1}{n} \max(X_1, \dots, X_n)\right] = \frac{n+1}{n} \cdot E[\max(X_1, \dots, X_n)] \\ &= \frac{n+1}{n} \cdot \frac{n}{n+1} \theta = \theta \end{aligned}$$

Se  $\hat{\theta}_2$  for usado repetidamente em diferentes amostras para estimar  $\theta$ , algumas estimativas serão muito grandes e outras muito pequenas, mas, em geral, não haverá tendência sistemática para subestimar ou superestimar  $\theta$ . ■

#### Princípio da Estimação Não-tendenciosa

Ao escolher dentre diversos estimadores diferentes de  $\theta$ , selecione um que seja não-tendencioso.

De acordo com esse princípio, deve ser dada preferência ao estimador não-tendencioso  $\hat{\theta}_2$  no Exemplo 6.4 com relação ao estimador tendencioso  $\hat{\theta}_1$ . Considere agora o problema de se estimar  $\sigma^2$ .

#### PROPOSIÇÃO

Sejam  $X_1, X_2, \dots, X_n$  uma amostra aleatória de uma distribuição com média  $\mu$  e variância  $\sigma^2$ . Então, o estimador

$$\hat{\sigma}^2 = S^2 = \frac{\sum (X_i - \bar{X})^2}{n-1}$$

é um estimador não-tendencioso de  $\sigma^2$ .



**Demonstração** Para qualquer va  $Y$ ,  $V(Y) = E(Y^2) - [E(Y)]^2$ , assim,  $E(Y^2) = V(Y) + [E(Y)]^2$ . Aplicando essa expressão em

$$S^2 = \frac{1}{n-1} \left[ \sum X_i^2 - \frac{(\sum X_i)^2}{n} \right]$$

fornece

$$\begin{aligned} E(S^2) &= \frac{1}{n-1} \left\{ \sum E(X_i^2) - \frac{1}{n} E[(\sum X_i)^2] \right\} \\ &= \frac{1}{n-1} \left\{ \sum (\sigma^2 + \mu^2) - \frac{1}{n} \{ V(\sum X_i) + [E(\sum X_i)]^2 \} \right\} \\ &= \frac{1}{n-1} \left\{ n\sigma^2 + n\mu^2 - \frac{1}{n} n\sigma^2 - \frac{1}{n} (n\mu)^2 \right\} \\ &= \frac{1}{n-1} \{ n\sigma^2 - \sigma^2 \} = \sigma^2 \quad (\text{conforme desejado}) \quad \blacksquare \end{aligned}$$

O estimador que utiliza divisor  $n$  pode ser expresso como  $(n-1)S^2/n$ , assim,

$$E\left[\frac{(n-1)S^2}{n}\right] = \frac{n-1}{n} E(S^2) = \frac{n-1}{n} \sigma^2$$

Esse estimador, portanto, é tendencioso. O desvio é  $(n-1)\sigma^2/n - \sigma^2 = -\sigma^2/n$ . Em virtude de o desvio ser negativo, o estimador com divisor  $n$  tende a subestimar  $\sigma^2$ , e é por esse motivo que o divisor  $n-1$  é preferido por muitos estatísticos (mesmo quando  $n$  é grande, o desvio é pequeno e há pouca diferença entre os dois).

Embora  $S^2$  seja não-tendencioso para  $\sigma^2$ ,  $S$  é um estimador tendencioso de sigma (seu desvio é pequeno, a menos que  $n$  seja bem pequeno). Entretanto, existem outros bons motivos para usar  $S$  como um estimador, especialmente quando a distribuição da população é normal. Isso ficará mais visível quando discutirmos os intervalos de confiança e os testes de hipóteses nos capítulos seguintes.

No Exemplo 6.2, propusemos vários estimadores diferentes para a média  $\mu$  de uma distribuição normal. Se houvesse um único estimador não-tendencioso para  $\mu$ , o problema da estimação seria resolvido usando-se tal estimador. Infelizmente, esse não é o caso.

#### PROPOSIÇÃO

Se  $X_1, X_2, \dots, X_n$  forem uma amostra aleatória de uma distribuição com média  $\mu$ , então  $\bar{X}$  é um estimador não-tendencioso de  $\mu$ . Se, além disso, a distribuição for contínua e simétrica, então  $\tilde{X}$  e qualquer média aparada também serão estimadores não-tendenciosos de  $\mu$ .

O fato de  $\bar{X}$  ser não-tendencioso é simplesmente uma reafirmação de uma de nossas regras do valor esperado:  $E(\bar{X}) = \mu$  para cada valor possível de  $\mu$  (para distribuições discretas e contínuas). A não-tendenciosidade dos outros estimadores é mais difícil de ser verificada.

De acordo com essa proposição, apenas o princípio da não-tendenciosidade nem sempre nos permite selecionar um único estimador. Quando a população em estudo é normal, mesmo o terceiro estimador do Exemplo 6.2 é não-tendencioso, e há muitos outros estimadores não-tendenciosos. Precisamos agora de um meio de selecionar os estimadores não-tendenciosos.

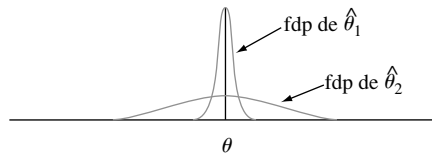
### Estimadores com Mínima Variância

Suponha que  $\hat{\theta}_1$  e  $\hat{\theta}_2$  sejam dois estimadores de  $\theta$ , ambos não-tendenciosos. Então, embora a distribuição de cada estimador esteja centrada no valor real de  $\theta$ , as dispersões das distribuições em torno do valor real podem ser diferentes.

**Princípio da Estimação Não-tendenciosa de Mínima Variância**

Dentre todos os estimadores não-tendenciosos de  $\theta$ , escolha aquele que tiver variância mínima. O  $\hat{\theta}$  resultante é chamado **estimador não-tendencioso de mínima variância (ENTMV)** de  $\theta$ .

A Figura 6.2 ilustra as fdp de dois estimadores não-tendenciosos, com  $\hat{\theta}_1$  tendo variância menor que  $\hat{\theta}_2$ . Então, é mais provável que  $\hat{\theta}_1$  produza uma estimativa mais próxima do  $\theta$  real do que  $\hat{\theta}_2$ . O ENTMV é, de certo modo, dentre todos os estimadores não-tendenciosos, o que produzirá, mais provavelmente, uma estimativa próxima de  $\theta$  real.



**Figura 6.2** Gráficos das fdp de dois estimadores não-tendenciosos diferentes

**Exemplo 6.5**

Afirmamos no Exemplo 6.4 que, quando  $X_1, \dots, X_n$  é uma amostra aleatória de uma distribuição uniforme em  $[0, \theta]$ , o estimador

$$\hat{\theta}_1 = \frac{n + 1}{n} \cdot \max(X_1, \dots, X_n)$$

é não-tendencioso para  $\theta$  (representamos anteriormente esse estimador por  $\hat{\theta}_2$ ). Esse não é o único estimador não-tendencioso de  $\theta$ . O valor esperado de uma va uniformemente distribuída é exatamente o ponto central do intervalo de densidade positiva, assim  $E(X_i) = \theta/2$ . Isso implica que  $E(\bar{X}) = \theta/2$ , de onde  $E(2\bar{X}) = \theta$ . Isto é, o estimador  $\hat{\theta}_2 = 2\bar{X}$  é não-tendencioso para  $\theta$ .

Se  $X$  for uniformemente distribuído no intervalo  $[A, B]$ , então,  $V(X) = \sigma^2 = (B - A)^2/12$ . Dessa forma, para o nosso caso,  $V(X_i) = \theta^2/12$ ,  $V(\bar{X}) = \sigma^2/n = \theta^2/(12n)$ , e  $V(\hat{\theta}_2) = V(2\bar{X}) = 4V(\bar{X}) = \theta^2/(3n)$ . Os resultados do Exercício 32 podem ser usados para mostrar que  $V(\hat{\theta}_1) = \theta^2/[n(n + 2)]$ . O estimador  $\hat{\theta}_1$  tem variância menor que  $\hat{\theta}_2$  se  $3n < n(n + 2)$ —isso é, se  $0 < n^2 - n = n(n - 1)$ . Como  $n > 1$ ,  $V(\hat{\theta}_1) < V(\hat{\theta}_2)$ , então,  $\hat{\theta}_1$  é um estimador melhor que  $\hat{\theta}_2$ . Métodos mais avançados podem ser usados para mostrar que  $\hat{\theta}_1$  é o ENTMV de  $\theta$  — qualquer outro estimador não-tendencioso de  $\theta$  tem variância que excede  $\theta^2/[n(n + 2)]$ . ■

Um dos triunfos da estatística matemática foi o desenvolvimento da metodologia para identificar o ENTMV em uma ampla variedade de situações. O resultado mais importante desse tipo para nossos propósitos diz respeito a estimar a média  $\mu$  de uma distribuição normal.

**TEOREMA**

Sejam  $X_1, \dots, X_n$  uma amostra aleatória de uma distribuição normal com parâmetros  $\mu$  e  $\sigma$ . Então, o estimador  $\hat{\mu} = \bar{X}$  é o ENTMV de  $\mu$ .

Quando estamos convencidos de que a população que está sendo amostrada é normal, o resultado diz que  $\bar{X}$  deve ser usado para estimar  $\mu$ . No Exemplo 6.2, então, nossa estimativa seria  $\bar{x} = 27,793$ .

Em algumas situações, é possível obter um estimador com pequena tendenciosidade preferível ao melhor estimador não-tendencioso. Isso é ilustrado na Figura 6.3. Entretanto, os ENTMVs geralmente são mais fáceis de ser obtidos do que o tipo de estimador tendencioso cuja distribuição é ilustrada.

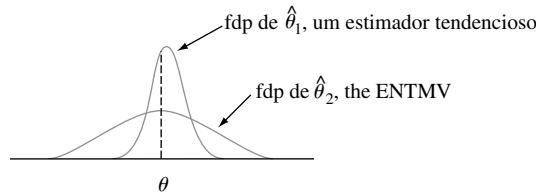


Figura 6.3 Um estimador tendencioso que é preferível ao ENTMV

### Algumas Complicações

O último teorema não afirma que, ao estimar a média da população  $\mu$ , o estimador  $\bar{X}$  deve ser usado independentemente da distribuição que está sendo amostrada.

#### Exemplo 6.6

Suponha que queiramos estimar a condutividade térmica  $\mu$  de certo material. Usando técnicas de medição-padrão, obteremos uma amostra aleatória  $X_1, \dots, X_n$  de  $n$  medições de condutividade térmica. Vamos assumir que a distribuição da população pertença a um membro de uma das três famílias a seguir:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/(2\sigma^2)} \quad -\infty < x < \infty \tag{6.1}$$

$$f(x) = \frac{1}{\pi[1 + (x - \mu)^2]} \quad -\infty < x < \infty \tag{6.2}$$

$$f(x) = \begin{cases} \frac{1}{2c} & -c \leq x - \mu \leq c \\ 0 & \text{caso contrário} \end{cases} \tag{6.3}$$

A fdp (6.1) é a distribuição normal, (6.2) é denominada distribuição de Cauchy e (6.3) é uma distribuição uniforme. As três distribuições são simétricas com relação a  $\mu$ , e, de fato, a distribuição de Cauchy tem o formato de sino, mas com caudas muito mais pesadas (mais probabilidades distantes) que a curva normal. A distribuição uniforme não possui caudas. Os quatro estimadores de  $\mu$  considerados anteriormente são  $\bar{X}$ ,  $\tilde{X}$ ,  $\bar{X}_e$  (a média das duas observações extremas) e  $\bar{X}_{tr(10)}$ , uma média aparada.

A conclusão mais importante aqui é que o melhor estimador de  $\mu$  depende crucialmente da distribuição que está sendo amostrada. Em particular,

1. Se a amostra aleatória foi tirada de uma distribuição normal, então  $\bar{X}$  é o melhor dos quatro estimadores, uma vez que possui mínima variância dentre todos os estimadores não-tendenciosos.
2. Se a amostra aleatória foi tirada de uma distribuição de Cauchy, então  $\bar{X}$  e  $\bar{X}_e$  são péssimos estimadores de  $\mu$ , enquanto  $\tilde{X}$  é relativamente bom (o ENTMV não é conhecido);  $\bar{X}$  é ruim porque é muito sensível às observações aberrantes, e as caudas pesadas da distribuição de Cauchy tornam provável a aparição de tais observações em qualquer amostra.
3. Se a distribuição em estudo for uniforme, o melhor estimador é  $\bar{X}_e$ , pois é fortemente influenciado pelas observações aberrantes; mas a ausência de caudas torna tais observações impossíveis para essa distribuição.
4. A média aparada não é melhor em nenhuma dessas três situações, mas funciona razoavelmente bem em todas elas. Isto é,  $\bar{X}_{tr(10)}$  não sofre muito em comparação ao melhor procedimento em qualquer uma das três situações. ■

Uma pesquisa estatística recente estabeleceu que, ao estimar um ponto de simetria  $\mu$  de uma distribuição de probabilidade contínua, a média aparada com proporção de aparagem de 10% ou 20% (de cada

extremidade da amostra) produz estimativas razoavelmente comportadas sobre uma variedade bastante grande de modelos possíveis. Por essa razão, a média aparada com pequena porcentagem de aparagem é considerada um **estimador robusto**.

Em algumas situações, a escolha não está entre dois estimadores diferentes construídos a partir da mesma amostra, mas entre os estimadores baseados em dois experimentos diferentes.

### Exemplo 6.7

Suponha que certo tipo de componente possua uma distribuição de vida útil exponencial com parâmetro  $\lambda$ , de modo que a vida útil esperada seja  $\mu = 1/\lambda$ . Uma amostra de  $n$  componentes é selecionada e cada um é colocado em funcionamento. Se o experimento for continuado até que todas as  $n$  vidas úteis,  $X_1, \dots, X_n$ , tenham sido observadas, então  $\bar{X}$  será um estimador não-tendencioso de  $\mu$ .

Em alguns experimentos, entretanto, os componentes são deixados em funcionamento somente até o momento da  $r$ -ésima falha, onde  $r < n$ . Esse procedimento é denominado **censo**. Seja  $Y_1$  a vida útil da primeira falha (a vida útil mínima entre os  $n$  componentes),  $Y_2$  o tempo em que ocorre a segunda falha (a segunda menor vida útil) e assim por diante. Uma vez que o experimento termina na hora  $Y_r$ , a vida útil total acumulada no término é

$$T_r = \sum_{i=1}^r Y_i + (n - r)Y_r$$

Agora, demonstramos que  $\hat{\mu} = T_r/r$  é um estimador não-tendencioso de  $\mu$ . Para isso, precisamos de duas propriedades das variáveis exponenciais:

1. A propriedade de ausência de memória (veja a Seção 4.4), a qual diz que, em qualquer instante, a vida útil remanescente possui a mesma distribuição exponencial que a vida útil original.
2. Se  $X_1, \dots, X_k$  são independentes, cada exponencial com parâmetro  $\lambda$ , então,  $\min(X_1, \dots, X_k)$  é exponencial com parâmetro  $k\lambda$  e possui valor esperado  $1/(k\lambda)$ .

Uma vez que todos os  $n$  componentes duram até  $Y_1$ ,  $n - 1$  dura uma quantidade de tempo adicional,  $Y_2 - Y_1$ ,  $n - 2$  uma quantidade de tempo adicional  $Y_3 - Y_2$  e assim por diante, outra expressão para  $T_r$  é

$$T_r = nY_1 + (n - 1)(Y_2 - Y_1) + (n - 2)(Y_3 - Y_2) + \dots + (n - r + 1)(Y_r - Y_{r-1})$$

Mas  $Y_1$  é o mínimo de  $n$  variáveis exponenciais, assim,  $E(Y_1) = 1/(n\lambda)$ . De forma semelhante,  $Y_2 - Y_1$  é o menor das  $n - 1$  vidas úteis remanescentes, cada exponencial com parâmetro  $\lambda$  (pela propriedade de ausência de memória), assim,  $E(Y_2 - Y_1) = 1/[(n - 1)\lambda]$ . Continuando,  $E(Y_{i+1} - Y_i) = 1/[(n - i)\lambda]$ , então,

$$\begin{aligned} E(T_r) &= nE(Y_1) + (n - 1)E(Y_2 - Y_1) + \dots + (n - r + 1)E(Y_r - Y_{r-1}) \\ &= n \cdot \frac{1}{n\lambda} + (n - 1) \cdot \frac{1}{(n - 1)\lambda} + \dots + (n - r + 1) \cdot \frac{1}{(n - r + 1)\lambda} \\ &= \frac{r}{\lambda} \end{aligned}$$

Dessa forma,  $E(T_r/r) = (1/r)E(T_r) = (1/r) \cdot (r/\lambda) = 1/\lambda = \mu$ , conforme afirmado.

Como exemplo, suponha que 20 componentes sejam testados e  $r = 10$ . Então, se os primeiros 10 tempos de falha forem 11, 15, 29, 33, 35, 40, 47, 55, 58 e 72, a estimativa de  $\mu$  é

$$\hat{\mu} = \frac{11 + 15 + \dots + 72 + (10)(72)}{10} = 111,5$$

A vantagem do experimento com censo é que termina mais rápido que o experimento sem censo. Entretanto, pode-se mostrar que  $V(T_r/r) = 1/(\lambda^2 r)$  é maior que  $1/(\lambda^2 n)$ , a variância de  $\bar{X}$  no experimento sem censo. ■

## Relatando uma Estimativa Pontual: O Erro-Padrão

Além de relatar o valor de uma estimativa pontual, deve-se indicar a sua precisão. A medida comum de precisão é o erro-padrão do estimador usado.

### DEFINIÇÃO

O **erro-padrão** de um estimador  $\theta$  é seu desvio padrão  $\sigma_{\hat{\theta}} = \sqrt{V(\hat{\theta})}$ . Se aquele envolver parâmetros desconhecidos cujos valores possam ser estimados, a substituição dessas estimativas em  $\sigma_{\hat{\theta}}$  produz o **erro-padrão estimado** (desvio padrão estimado) do estimador que pode ser representado por  $\hat{\sigma}_{\hat{\theta}}$  (o  $\hat{\ }$  sobre  $\sigma$  enfatiza que  $\sigma_{\hat{\theta}}$  está sendo estimado) ou por  $s_{\hat{\theta}}$ .

### Exemplo 6.8 (continuação do Exemplo 6.2)

Assumindo que a voltagem de quebra é normalmente distribuída,  $\hat{\mu} = \bar{X}$  é o melhor estimador de  $\mu$ . Se o valor de  $\sigma$  for conhecido como 1,5, o erro-padrão de  $\bar{X}$  é  $\sigma_{\bar{X}} = \sigma/\sqrt{n} = 1,5/\sqrt{20} = 0,335$ . Se, como geralmente é o caso, o valor de  $\sigma$  for desconhecido, a estimativa  $\hat{\sigma} = s = 1,462$  é substituída em  $\sigma_{\bar{X}}$  para se obter o erro-padrão estimado  $\hat{\sigma}_{\bar{X}} = s_{\bar{X}} = s/\sqrt{n} = 1,462/\sqrt{20} = 0,327$ . ■

### Exemplo 6.9 (continuação do Exemplo 6.1)

O erro-padrão de  $\hat{p} = X/n$  é

$$\sigma_{\hat{p}} = \sqrt{V(X/n)} = \sqrt{\frac{V(X)}{n^2}} = \sqrt{\frac{npq}{n^2}} = \sqrt{\frac{pq}{n}}$$

Uma vez que  $p$  e  $q = 1 - p$  são desconhecidos, substituímos  $\hat{p} = x/n$  e  $\hat{q} = 1 - x/n$  em  $\sigma_{\hat{p}}$ , produzindo o erro-padrão estimado  $\hat{\sigma}_{\hat{p}} = \sqrt{\hat{p}\hat{q}/n} = \sqrt{(0,6)(0,4)/25} = 0,098$ . De maneira alternativa, uma vez que o maior valor de  $pq$  é obtido quando  $p = q = 0,5$ , um limite superior no erro-padrão é  $\sqrt{1/(4n)} = 0,10$ . ■

Quando o estimador pontual  $\hat{\theta}$  possui uma distribuição aproximadamente normal, que geralmente será o caso quando  $n$  for grande, então podemos ficar razoavelmente seguros de que o valor real de  $\theta$  está dentro de aproximadamente 2 erros-padrão (desvios padrão) de  $\hat{\theta}$ . Dessa forma, se uma amostra de  $n = 36$  vidas úteis de componentes fornecer  $\hat{\mu} = \bar{x} = 28,50$  e  $s = 3,60$ , então  $s/\sqrt{n} = 0,60$ , assim, dentro de 2 erros-padrão estimados de  $\hat{\mu}$ , transforma-se no intervalo  $28,50 \pm (2)(0,60) = (27,30, 29,70)$ .

Se  $\hat{\theta}$  for aproximadamente normal (não necessariamente) mas não-tendencioso, então pode-se mostrar que a estimativa desviará de  $\theta$  por 4 erros-padrão no máximo 6% das vezes. Esperaríamos, então, que o valor real estivesse dentro de 4 erros-padrão de  $\hat{\theta}$  (e essa é uma afirmação bastante conservadora, uma vez que se aplica a qualquer  $\hat{\theta}$  não-tendencioso). Resumindo, o erro-padrão nos diz grosseiramente dentro de qual distância de  $\hat{\theta}$  podemos esperar que o valor real de  $\theta$  esteja.

A forma do estimador  $\hat{\theta}$  pode ser suficientemente complicada, de modo que a teoria estatística-padrão não pode ser aplicada a fim de se obter uma expressão para  $\sigma_{\hat{\theta}}$ . Isso é verdade, por exemplo, no caso  $\theta = \sigma$ ,  $\hat{\theta} = S$ ; o desvio padrão da estatística  $S$ ,  $\sigma_S$ , geralmente não pode ser determinado. Recentemente, um novo método computacional chamado **bootstrap** foi desenvolvido para tratar desse problema. Suponha que a fdp da população seja  $f(x; \theta)$ , um membro de uma família paramétrica, e que os dados  $x_1, x_2, \dots, x_n$  fornecem  $\hat{\theta} = 21,7$ . Agora, usamos o computador para obter “amostras do *bootstrap*” da fdp  $f(x; 21,7)$  e, para cada amostra, calculamos uma “estimativa de *bootstrap*”  $\hat{\theta}^*$ :

Primeira amostra do *bootstrap*:  $x_1^*, x_2^*, \dots, x_n^*$ ; estimativa =  $\hat{\theta}_1^*$

Segunda amostra do *bootstrap*:  $x_1^*, x_2^*, \dots, x_n^*$ ; estimativa =  $\hat{\theta}_2^*$

⋮

$B$ -ésima amostra do *bootstrap*:  $x_1^*, x_2^*, \dots, x_n^*$ ; estimativa =  $\hat{\theta}_B^*$

Geralmente é usado  $B = 100$  ou  $200$ . Agora, seja  $\bar{\theta}^* = \sum \hat{\theta}_i^*/B$  a média amostral das estimativas do *bootstrap*. A **estimativa do bootstrap** do erro-padrão de  $\hat{\theta}$ 's é agora apenas o desvio padrão amostral dos  $\hat{\theta}_i^*$ 's:

$$S_{\hat{\theta}} = \sqrt{\frac{1}{B-1} \sum (\hat{\theta}_i^* - \bar{\theta}^*)^2}$$

(Na literatura sobre *bootstrap*,  $B$  geralmente é usado no lugar de  $B - 1$ ; para valores típicos de  $B$ , há, em geral, pouca diferença entre as estimativas resultantes.)

**Exemplo 6.10**

Um modelo teórico sugere que  $X$ , o tempo de quebra de um fluido isolante entre os elétrodos em uma voltagem específica, possui  $f(x; \lambda) = \lambda e^{-\lambda x}$ , uma distribuição exponencial. Uma amostra aleatória de  $n = 10$  tempos de quebra (min) fornece os seguintes dados:

41,53 18,73 2,99 30,34 12,33 117,52 73,02 223,63 4,00 26,78

Uma vez que  $E(X) = 1/\lambda$ ,  $E(\bar{X}) = 1/\lambda$ , então uma estimativa razoável de  $\lambda$  é  $\hat{\lambda} = 1/\bar{x} = 1/55,087 = 0,018153$ . Usamos um pacote estatístico de computador para obter  $B = 100$  amostras do *bootstrap*, cada uma de tamanho 10, a partir de  $f(x; 0,018153)$ . A primeira de tais amostras foi 41,00, 109,70, 16,78, 6,31, 6,76, 5,62, 60,96, 78,81, 192,25, 27,61, da qual  $\sum x_i^* = 545,8$  e  $\hat{\lambda}_1^* = 1/54,58 = 0,01832$ . A média das 100 estimativas do *bootstrap* é  $\bar{\lambda}^* = 0,02153$ , e o desvio padrão amostral dessas 100 estimativas é  $s_{\hat{\lambda}} = 0,0091$ , a estimativa do *bootstrap* do erro-padrão dos  $\hat{\lambda}$ 's. Um histograma dos 100  $\hat{\lambda}_i^*$ 's estava positivamente inclinado, sugerindo que a distribuição amostral de  $\hat{\lambda}$  também possuía essa propriedade. ■

Às vezes, o pesquisador deseja estimar a característica de uma população sem assumir que sua distribuição pertence a uma família paramétrica particular. Um caso desses ocorreu no Exemplo 6.6, em que uma média aparada 10% foi proposta para estimar o centro da distribuição simétrica da população  $\theta$ . Os dados do Exemplo 6.2 forneceram  $\hat{\theta} = \bar{x}_{tr(10)} = 27,838$ , mas agora não há  $f(x; \theta)$  assumido. Então, como podemos obter uma amostra do *bootstrap*? A resposta é considerar a própria amostra como constituinte da população (as  $n = 10$  observações no Exemplo 6.2) e tirar  $B$  amostras diferentes, cada uma de tamanho  $n$ , com substituição da referida população. O livro de Bradley Efron e Robert Tibshirani ou o de John Rice, listados na bibliografia, fornecem mais informações.

**Exercícios** | Seção 6.1 (1–19)

1. Os dados que seguem sobre a resistência à flexão (MPa) das vigas de um tipo de concreto foram introduzidos no Exemplo 1.2.
 

5,9	7,2	7,3	6,3	8,1	6,8	7,0
7,6	6,8	6,5	7,0	6,3	7,9	9,0
8,2	8,7	7,8	9,7	7,4	7,7	9,7
7,8	7,7	11,6	11,3	11,8	10,7	

  - a. Calcule uma estimativa pontual do valor médio da resistência para a população conceitual de todas as vigas fabricadas dessa forma e diga qual estimador você utilizou. (*Sugestão*:  $\sum x_i = 219,8$ .)
  - b. Calcule uma estimativa pontual do valor da resistência que separa as 50% mais fracas de todas as vigas das 50% mais fortes e diga qual estimador você utilizou.
  - c. Calcule e interprete uma estimativa pontual do desvio padrão da população  $\sigma$ . Que estimador você utilizou? (*Sugestão*:  $\sum x_i^2 = 1860,94$ .)
  - d. Calcule uma estimativa pontual da proporção de todas as vigas cuja resistência à flexão exceda 10 MPa. (*Sugestão*: considere como observação um “sucesso” se exceder 10.)
  - e. Calcule uma estimativa pontual do coeficiente de variação  $\sigma/\mu$  da população e diga qual estimador você utilizou.
2. Uma amostra de 20 alunos que recentemente tiveram estatística básica descobriu as seguintes informações sobre a marca das suas calculadoras (T = Texas Instruments, H = Hewlett Packard, C = Cassio, S = Sharp)
 

T	T	H	T	C	T	T	S	C	H
S	S	T	H	C	T	T	T	H	T

- a. Calcule a proporção real de todos os alunos que possuem uma calculadora Texas Instruments.
  - b. Dos 10 alunos que tinham uma calculadora TI, 4 possuíam calculadoras com gráfico. Calcule a proporção de alunos que não possuem uma calculadora com gráfico TI.
3. Considere a seguinte amostra de observações sobre a espessura da camada da pintura de baixa viscosidade (“Achieving a Target Value for a Manufacturing Process: A Case Study,” *J. of Quality Technology*, 1992, p. 22-26):
- 0,83 0,88 0,88 1,04 1,09 1,12 1,29 1,31  
1,48 1,49 1,59 1,62 1,65 1,71 1,76 1,83

Assuma que a distribuição da espessura da camada seja normal (um gráfico de probabilidade normal suporta fortemente essa hipótese).

- a. Calcule uma estimativa pontual do valor médio da espessura da camada e diga qual estimador você utilizou.
  - b. Calcule uma estimativa pontual da mediana da distribuição da espessura da camada e diga qual estimador você utilizou.
  - c. Calcule uma estimativa pontual do valor que separa os 10% maiores de todos os valores na distribuição da espessura dos 90% remanescentes e diga qual estimador você utilizou. (*Sugestão*: mostre o que você está tentando estimar em termos de  $\mu$  e  $\sigma$ .)
  - d. Calcule  $P(X < 1,5)$ , isto é, a proporção de todos os valores da espessura menores que 1,5. (*Sugestão*: Se você conhecesse os valores de  $\mu$  e  $\sigma$ , poderia calcular essa probabilidade. Esses valores não estão disponíveis, mas podem ser estimados.)
  - e. Qual é o erro-padrão estimado do estimador que você utilizou na parte (b)?
4. O artigo do qual os dados do Exercício 1 foram tirados também forneceu as observações de resistência a seguir para os cilindros:

6,1 5,8 7,8 7,1 7,2 9,2 6,6 8,3 7,0 8,3  
7,8 8,1 7,4 8,5 8,9 9,8 9,7 14,1 12,6 11,2

Antes de obter os dados, represente as resistências da viga por  $X_1, \dots, X_m$  e as resistências do cilindro por  $Y_1, \dots, Y_n$ . Suponha que os  $X_s$  constituem uma amostra aleatória de uma distribuição com média  $\mu_1$  e desvio padrão  $\sigma_1$  e que os  $Y_s$  formem uma amostra aleatória (independente dos  $X_s$ ) de outra distribuição com média  $\mu_2$  e desvio padrão  $\sigma_2$ .

- a. Use as regras do valor esperado para mostrar que  $\bar{X} - \bar{Y}$  é um estimador não-tendencioso de  $\mu_1 - \mu_2$ . Calcule a estimativa para os dados fornecidos.
- b. Use as regras de variância do Capítulo 5 para obter uma expressão para a variância e o desvio padrão (erro-padrão) do estimador da parte (a) e calcule o erro-padrão estimado.
- c. Calcule uma estimativa pontual da relação  $\sigma_1/\sigma_2$  dos dois desvios padrão.

- d. Suponha que uma única viga e um único cilindro sejam selecionados aleatoriamente. Calcule uma estimativa pontual da variância da diferença  $X - Y$  entre a resistência da viga e a resistência do cilindro.
5. Como exemplo de uma situação em que várias estatísticas diferentes podiam razoavelmente ser usadas para calcular uma estimativa pontual, considere uma população de  $N$  faturas. Associado a cada fatura está seu “valor contábil”, o valor registrado dessa fatura. Seja por  $T$  o valor contábil total, um valor conhecido. Alguns desses valores são errôneos. Uma auditoria será realizada, selecionando aleatoriamente  $n$  faturas e determinando o valor (correto) examinado de cada uma. Suponha que a amostra forneça os seguintes resultados (em dólares).

	Fatura				
	1	2	3	4	5
Valor contábil	300	720	526	200	127
Valor examinado	300	520	526	200	157
Erro	0	200	0	0	-30

Sejam

$\bar{Y}$  = valor contábil médio amostral

$\bar{X}$  = valor médio amostral examinado

$\bar{D}$  = erro médio amostral

Diversas estatísticas diferentes para estimar o valor (correto) examinado total foram propostas (veja “Statistical Models and Analysis in Auditing”, *Statistical Science*, 1989, p. 2-33). Elas incluem

Média por estatística da unidade =  $\bar{N}\bar{X}$

Estatística da diferença =  $T - N\bar{D}$

Estatística da razão =  $T \cdot (\bar{X}/\bar{Y})$

Se  $N = 5000$  e  $T = 1.761,300$ , calcule as três estimativas pontuais correspondentes. (O artigo mencionado discute as propriedades desses estimadores.)

- 6. Considere as observações que seguem sobre o fluxo direto (1000 acre-pés) registrado em um posto em Colorado de 1º de abril a 31 de agosto, por um período de 31 anos (Dados extraídos de um artigo do volume de 1974 de *Water Resources Research*).

127,96	210,07	203,24	108,91	178,21
285,37	100,85	89,59	185,36	126,94
200,19	66,24	247,11	299,87	109,64
125,86	114,79	109,11	330,33	85,54
117,64	302,74	280,55	145,11	95,36
204,91	311,13	150,58	262,09	477,08
94,33				

Um gráfico de probabilidade apropriado confirma o uso da distribuição lognormal (veja Seção 4.5) como um modelo razoável do fluxo dos rios.

- a. Calcule os parâmetros da distribuição. [*Sugestão*: lembre-se que  $X$  possui distribuição lognormal com parâmetros  $\mu$  e  $\sigma^2$  se  $\ln(X)$  for normalmente distribuído com média  $\mu$  e variância  $\sigma^2$ .]

- b. Use as estimativas da parte (a) para calcular uma estimativa do valor esperado de fluxo dos rios. [Sugestão: qual é o  $E(X)$ ?]
7. a. Uma amostra aleatória de 10 casas em uma área específica, todas aquecidas com gás natural, é selecionada; a quantidade de gás (termos) usada durante o mês de janeiro é determinada para cada casa. As observações resultantes são: 103, 156, 118, 89, 125, 147, 122, 109, 138, 99. Seja  $\mu$  o consumo médio de gás durante janeiro por todas as casas dessa área. Calcule a estimativa pontual de  $\mu$ .
- b. Suponha que existam 10.000 casas nessa área que utilizem gás natural para aquecimento. Seja  $\tau$  a quantidade total de gás usado por todas essas casas no mês de janeiro. Calcule  $\tau$ , usando os dados da parte (a). Que estimador você utilizou ao calcular sua estimativa?
- c. Use os dados na parte (a) para calcular  $p$ , a proporção de todas as casas que usaram no mínimo 100 termos.
- d. Forneça uma estimativa pontual do uso mediano da população (o valor médio na população de todas as casas) com base na amostra da parte (a). Que estimador você utilizou?
8. Em uma amostra aleatória de certo tipo de 80 componentes, 12 são considerados defeituosos.
- a. Forneça uma estimativa pontual da proporção de todos os componentes que *não* são defeituosos.
- b. Um sistema será construído, selecionando-se aleatoriamente dois desses componentes e conectando-os em série, conforme ilustramos abaixo:



A conexão em série implica que o sistema funcionará somente se nenhum componente estiver com defeito (isto é, se ambos funcionarem adequadamente). Calcule a proporção de todos os sistemas que funcionam adequadamente. [Sugestão: se  $p$  representa a probabilidade de um componente funcionar corretamente, como pode  $P$  (funcionamentos do sistema) ser expresso em termos de  $p$ ?]

9. Todos os 150 itens fabricados recentemente são examinados e o número de arranhões por item é registrado (os itens são supostamente livres de arranhões), produzindo os seguintes dados:

Número de arranhões por item	0	1	2	3	4	5	6	7
Frequência observada	18	37	42	30	13	7	2	1

Seja  $X =$  o número de arranhões em um item escolhido aleatoriamente, assuma que  $X$  possui uma distribuição de Poisson com parâmetro  $\lambda$ .

- a. Determine um estimador não-tendencioso de  $\lambda$  e calcule a estimativa dos dados. [Sugestão:  $E(X) = \lambda$  para Poisson de  $X$ , então  $E(\bar{X}) = ?$ ]
- b. Qual é o desvio padrão (erro-padrão) do seu estimador? Calcule o erro-padrão estimado. (Sugestão:  $\sigma_X^2 = \lambda$  para Poisson de  $X$ .)
10. Usando um bastão longo de comprimento  $\mu$ , você traçará um gráfico quadrado em que o comprimento de cada lado é  $\mu$ . Assim, a área do gráfico será  $\mu^2$ . Entretanto, você não conhece o valor de  $\mu$ , então decide fazer  $n$  medições independentes  $X_1, X_2, \dots, X_n$  do comprimento. Assuma que cada  $X_i$  tenha média  $\mu$  (medições não-tendenciosas) e variância  $\sigma^2$ .
- a. Mostre que  $\bar{X}^2$  não é um estimador não-tendencioso para  $\mu^2$ . [Sugestão: para qualquer  $Y$ ,  $E(Y^2) = V(Y) + [E(Y)]^2$ . Aplique a expressão com  $Y = \bar{X}$ .]
- b. Para que valor de  $k$  o estimador  $\bar{X}^2 - kS^2$  é não-tendencioso para  $\mu^2$ ? [Sugestão: calcule  $E(\bar{X}^2 - kS^2)$ .]
11. De  $n_1$  fumantes do sexo masculino escolhidos aleatoriamente,  $X_1$  fumavam cigarro com filtro, enquanto de  $n_2$  fumantes do sexo feminino selecionadas aleatoriamente,  $X_2$  fumavam cigarro com filtro. Represente por  $p_1$  e  $p_2$  as probabilidades de um homem e uma mulher selecionados de modo aleatório, respectivamente, fumarem cigarro com filtro.
- a. Mostre que  $(X_1/n_1) - (X_2/n_2)$  é um estimador não-tendencioso para  $p_1 - p_2$ . [Sugestão:  $E(X_i) = n_i p_i$  para  $i = 1, 2$ .]
- b. Qual é o erro-padrão do estimador da parte (a)?
- c. Como você usaria os valores observados de  $x_1$  e  $x_2$  para calcular o erro-padrão do seu estimador?
- d. Se  $n_1 = n_2 = 200$ ,  $x_1 = 127$  e  $x_2 = 176$ , use o estimador da parte (a) para obter uma estimativa de  $p_1 - p_2$ .
- e. Use o resultado da parte (c) e os dados da parte (d) para calcular o erro-padrão do estimador.
12. Suponha que certo tipo de fertilizante possua um rendimento esperado por acre de  $\mu_1$ , com variância  $\sigma^2$ , enquanto o rendimento esperado de um segundo tipo de fertilizante seja  $\mu_2$  com a mesma variância  $\sigma^2$ . Represente por  $S_1^2$  e  $S_2^2$  as variâncias amostrais dos rendimentos com base nos tamanhos de amostra  $n_1$  e  $n_2$ , respectivamente, dos dois fertilizantes. Mostre que o estimador combinado

$$\hat{\sigma}^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

é um estimador não-tendencioso de  $\sigma^2$ .

13. Considere uma amostra aleatória  $X_1, \dots, X_n$  da fdp

$$f(x; \theta) = 0,5(1 + \theta x) \quad -1 \leq x \leq 1$$



em que  $-1 \leq \theta \leq 1$  (essa distribuição surge na física de partículas). Mostre que  $\hat{\theta} = 3\bar{X}$  é um estimador não-tendencioso de  $\theta$ . [Sugestão: determine primeiro  $\mu = E(X) = E(\bar{X})$ .]

14. Uma amostra de  $n$  aeronaves *Pandemonium jet fighters* capturadas resulta em números de série  $x_1, x_2, x_3, \dots, x_n$ . A CIA sabe que cada aeronave foi numerada consecutivamente na fábrica, começando com  $\alpha$  e terminando com  $\beta$ , de modo que o número total de aviões fabricados é  $\beta - \alpha + 1$  (por exemplo: se  $\alpha = 17$  e  $\beta = 29$ , então,  $29 - 17 + 1 = 13$  aviões que têm números de série 17, 18, 19, ..., 28, 29, foram fabricados). Entretanto, a CIA não conhece os valores de  $\alpha$  ou  $\beta$ . Um estatístico dessa instituição sugere o uso do estimador  $\max(X_i) - \min(X_i) + 1$  para calcular o número total de aviões fabricados.
- a. Se  $n = 5, x_1 = 237, x_2 = 375, x_3 = 202, x_4 = 525$  e  $x_5 = 418$ , qual é a estimativa correspondente?
- b. Sob quais condições na amostra o valor da estimativa será exatamente igual ao número total real de aviões? A estimativa sempre será maior que o total real? Você acha que o estimador é não-tendencioso para calcular  $\beta - \alpha + 1$ ? Explique em uma ou duas frases.
15. Seja  $X_1, X_2, \dots, X_n$  uma amostra aleatória de uma distribuição de Rayleigh com fdp

$$f(x; \theta) = \frac{x}{\theta} e^{-x^2/(2\theta)} \quad x > 0$$

- a. Pode-se mostrar que  $E(X^2) = 2\theta$ . Use esse fato para construir um estimador não-tendencioso de  $\theta$  com base em  $\sum X_i^2$  (use regras de valor esperado para mostrar que é não-tendencioso).
- b. Calcule  $\theta$  a partir das  $n = 10$  observações a seguir sobre a força vibratória de uma palheta de turbina sob condições especificadas:
- |       |       |      |      |       |
|-------|-------|------|------|-------|
| 16,88 | 10,23 | 4,59 | 6,66 | 13,68 |
| 14,23 | 19,87 | 9,40 | 6,51 | 10,95 |
16. Suponha que o crescimento médio real  $\mu$  de um tipo de planta durante o período de um ano seja idêntico ao de um segundo tipo, mas a variância de crescimento do primeiro tipo seja  $\sigma^2$ , enquanto a do segundo é  $4\sigma^2$ . Sejam  $X_1, \dots, X_m, m$  observações independentes de crescimento para o primeiro tipo [assim,  $E(X_i) = \mu, V(X_i) = \sigma^2$ ], e  $Y_1, \dots, Y_n, n$  observações independentes de crescimento para o segundo tipo [ $E(Y_i) = \mu, V(Y_i) = 4\sigma^2$ ].
- a. Mostre que, para qualquer  $\delta$  entre 0 e 1, o estimador  $\hat{\mu} = \delta\bar{X} + (1 - \delta)\bar{Y}$  é não-tendencioso para  $\mu$ .
- b. Para  $m$  e  $n$  fixos, calcule  $V(\hat{\mu})$  e depois determine o valor de  $\delta$  que minimiza  $V(\hat{\mu})$ . [Sugestão: diferencie  $V(\hat{\mu})$  em relação a  $\delta$ .]
17. No Capítulo 3, definimos uma va binomial negativa como o número de falhas que ocorrem antes do  $r$ -ésimo sucesso em uma seqüência de tentativas de sucesso / falha idênticas e independentes. A função de massa de probabilidade (fmp) de  $X$  é

$$nb(x; r, p) = \begin{cases} \binom{x+r-1}{r-1} p^r (1-p)^x & x = 0, 1, 2, \dots \\ 0 & \text{caso contrário} \end{cases}$$

- a. Suponha que  $r \geq 2$ . Mostre que
- $$\hat{p} = (r-1)/(X+r-1)$$
- é um estimador não-tendencioso para  $p$ . [Sugestão: Desenvolva  $E(\hat{p})$  e cancele  $x+r-1$  dentro da soma.]
- b. Um repórter que deseja entrevistar cinco pessoas que apóiam um candidato começa perguntando a elas se apóiam ( $S$ ) ou não ( $F$ ) tal candidato. Se a seqüência das respostas é *SFFSFFSSS*, estime  $p =$  proporção real que apóia o candidato.
18. Seja  $X_1, X_2, \dots, X_n$  uma amostra aleatória de uma fdp  $f(x)$  simétrica em relação a  $\mu$ , de modo que  $\bar{X}$  seja um estimador não-tendencioso de  $\mu$ . Se  $n$  for grande, pode-se mostrar que  $V(\bar{X}) \approx 1/(4n[f(\mu)]^2)$ .
- a. Compare  $V(\bar{X})$  a  $V(\bar{X})$  quando a distribuição subjacente for normal.
- b. Quando a fdp subjacente for Cauchy (veja o Exemplo 6.6),  $V(\bar{X}) = \infty$ , então  $\bar{X}$  será um terrível estimador. Qual é  $V(\bar{X})$ , neste caso, quando  $n$  for grande?
19. Uma pesquisadora deseja estimar a proporção de alunos que violaram o código de honra em certa universidade. Ao obter uma amostra aleatória de  $n$  alunos, ela percebe que perguntar a cada um: "Você violou o código de honra?" provavelmente resultará em algumas respostas falsas. Considere o seguinte esquema, chamado de técnica de **resposta aleatória**. A pesquisadora monta um baralho de 100 cartas, das quais 50 são do tipo I e 50 do tipo II.
- Tipo I: Você violou o código de honra (sim ou não)?
- Tipo II: O último dígito de seu número de telefone é 0, 1 ou 2 (sim ou não)?
- Cada aluno da amostra aleatória é solicitado a misturar o baralho, tirar uma carta e responder à pergunta resultante com sinceridade. Devido à pergunta irrelevante nas cartas do tipo II, a resposta sim não estigmatiza mais o entrevistado, de modo que assumimos que as respostas sejam verdadeiras. Seja  $p$  a proporção dos violadores do código de honra (isto é, a probabilidade de um aluno selecionado aleatoriamente ser um violador) e por  $\lambda = P$  (resposta sim). Então,  $\lambda$  e  $p$  são relacionados por  $\lambda = 0,5p + (0,5)(0,3)$ .
- a. Seja  $Y$  o número de respostas afirmativas, de modo que  $Y \sim \text{Bin}(n, \lambda)$ . Dessa forma,  $Y/n$  é um estimador não-tendencioso de  $\lambda$ . Deduza um estimador para  $p$  com base em  $Y$ . Se  $n = 80$  e  $y = 20$ , qual é sua estimativa? (Sugestão: resolva  $\lambda = 0,5p + 0,15$  para  $p$  e então substitua  $Y/n$  por  $\lambda$ .)
- b. Use o fato de  $E(Y/n) = \lambda$  para mostrar que seu estimador  $\hat{p}$  é não-tendencioso.
- c. Se havia 70 cartas tipo I e 30 tipo II, qual seria seu estimador para  $p$ ?

## 6.2 | Métodos de Estimativa Pontual

A definição de não-tendenciosidade geralmente não indica como os estimadores não-tendenciosos podem ser deduzidos. Agora, discutimos dois métodos “construtivos” para obter estimadores pontuais: o método dos momentos e o método da máxima verossimilhança. Por “construtivo” queremos dizer que a definição geral de cada tipo de estimador sugere explicitamente como obter o estimador em qualquer problema específico. Embora os estimadores da máxima verossimilhança sejam geralmente preferíveis aos estimadores de momentos devido a certas propriedades de eficiência, eles usualmente precisam significativamente de mais cálculos que os estimadores de momentos. Às vezes, é o caso de esses métodos produzirem estimadores não-tendenciosos.

### O Método dos Momentos

A idéia básica deste método é igualar certas características da amostra, como a média, aos valores esperados correspondentes da população. Então, resolvendo essas equações para os valores dos parâmetros desconhecidos produz os estimadores.

#### DEFINIÇÃO

Seja  $X_1, \dots, X_n$  uma amostra aleatória de uma fmp ou fdp  $f(x)$ . Para  $k = 1, 2, 3, \dots$ , o  **$k$ -ésimo momento da população** ou  **$k$ -ésimo momento da distribuição  $f(x)$**  é  $E(X^k)$ . O  **$k$ -ésimo momento amostral** é  $(1/n)\sum_{i=1}^n X_i^k$ .

Dessa forma, o primeiro momento da população é  $E(X) = \mu$  e o primeiro momento amostral é  $\sum X_i/n = \bar{X}$ . Os segundos momentos da população e da amostra são  $E(X^2)$  e  $\sum X_i^2/n$ , respectivamente. Os momentos da população serão funções de quaisquer parâmetros desconhecidos  $\theta_1, \theta_2, \dots$

#### DEFINIÇÃO

Seja  $X_1, X_2, \dots, X_n$  uma amostra aleatória de uma distribuição com fmp ou fdp  $f(x; \theta_1, \dots, \theta_m)$ , onde  $\theta_1, \dots, \theta_m$  são parâmetros cujos valores são desconhecidos. Então, os **estimadores pelo método dos momentos**  $\hat{\theta}_1, \dots, \hat{\theta}_m$  são obtidos igualando-se os primeiros  $m$  momentos da amostra aos primeiros  $m$  momentos da população correspondente e resolvendo para  $\theta_1, \dots, \theta_m$ .

Se, por exemplo,  $m = 2$ ,  $E(X)$  e  $E(X^2)$  serão funções de  $\theta_1$  e  $\theta_2$ . Definindo  $E(X) = (1/n) \sum X_i (= \bar{X})$  e  $E(X^2) = (1/n) \sum X_i^2$  fornece duas equações em  $\theta_1$  e  $\theta_2$ . A solução, então, define os estimadores. Para estimar a média da população  $\mu$ , o método fornece  $\mu = \bar{X}$ , assim, o estimador é a média amostral.

#### Exemplo 6.11

Seja  $X_1, X_2, \dots, X_n$  uma amostra aleatória de tempos de serviço de  $n$  clientes em uma determinada instalação, onde a distribuição subjacente é considerada exponencial com parâmetro  $\lambda$ . Uma vez que há somente um parâmetro a ser estimado, o estimador é obtido igualando  $E(X)$  a  $\bar{X}$ . Desde que  $E(X) = 1/\lambda$  para uma distribuição exponencial, temos  $1/\lambda = \bar{X}$  ou  $\lambda = 1/\bar{X}$ . O estimador pelo método dos momentos de  $\lambda$  é, então,  $\lambda = 1/\bar{X}$ . ■

#### Exemplo 6.12

Seja  $X_1, \dots, X_n$  uma amostra aleatória de uma distribuição gama com parâmetros  $\alpha$  e  $\beta$ . Pela Seção 4.4,  $E(X) = \alpha\beta$  e  $E(X^2) = \beta^2\Gamma(\alpha + 2)/\Gamma(\alpha) = \beta^2(\alpha + 1)\alpha$ . Os estimadores pelo método dos momentos de  $\alpha$  e  $\beta$  são obtidos, resolvendo-se

$$\bar{X} = \alpha\beta \quad \frac{1}{n} \sum X_i^2 = \alpha(\alpha + 1)\beta^2$$

Uma vez que  $\alpha(\alpha + 1)\beta^2 = \alpha^2\beta^2 + \alpha\beta^2$  e a primeira equação implica  $\alpha^2\beta^2 = \bar{X}^2$ , a segunda equação se torna

$$\frac{1}{n} \sum X_i^2 = \bar{X}^2 + \alpha\beta^2$$

Dividindo-se cada membro dessa segunda equação pelo lado correspondente da primeira e substituindo-se de volta, são fornecidos os estimadores

$$\hat{\alpha} = \frac{\bar{X}^2}{(1/n)\sum X_i^2 - \bar{X}^2} \quad \hat{\beta} = \frac{(1/n)\sum X_i^2 - \bar{X}^2}{\bar{X}}$$

Para ilustrar, os dados do tempo de sobrevivida mencionados no Exemplo 4.21 são

152	115	109	94	88	137	152	77	160	165
125	40	128	123	136	101	62	153	83	69

com  $\bar{x} = 113,5$  e  $(1/20)\sum x_i^2 = 14,087.8$ . As estimativas são

$$\hat{\alpha} = \frac{(113,5)^2}{14,087,8 - (113,5)^2} = 10,7 \quad \hat{\beta} = \frac{14,087,8 - (113,5)^2}{113,5} = 10,6$$

Essas estimativas de  $\alpha$  e  $\beta$  diferem dos valores sugeridos por Gross e Clark porque eles usaram uma técnica de estimativa diferente. ■

### Exemplo 6.13

Seja  $X_1, \dots, X_n$  uma amostra aleatória de uma distribuição binomial negativa geral com parâmetros  $r$  e  $p$  (Seção 3.5). Uma vez que  $E(X) = r(1-p)/p$  e  $V(X) = r(1-p)/p^2$ ,  $E(X^2) = V(X) + [E(X)]^2 = r(1-p)(r-rp+1)/p^2$ . Igualando  $E(X)$  a  $\bar{X}$  e  $E(X^2)$  a  $(1/n)\sum X_i^2$  finalmente fornece

$$\hat{p} = \frac{\bar{X}}{(1/n)\sum X_i^2 - \bar{X}^2} \quad \hat{r} = \frac{\bar{X}^2}{(1/n)\sum X_i^2 - \bar{X}^2 - \bar{X}}$$

Como ilustração, Reep, Pollard e Benjamin ("Skill and Chance in Ball Games," *J. Royal Stat. Soc.*, 1971, p. 623-629) consideram a distribuição binomial negativa como um modelo para o número de gols por jogo marcados pelos times da Liga Nacional de Hóquei. Os dados de 1966-1967 aparecem a seguir (420 jogos):

<i>Gols</i>	0	1	2	3	4	5	6	7	8	9	10
<i>Frequência</i>	29	71	82	89	65	45	24	7	4	1	3

Então,

$$\bar{x} = \sum x_i/420 = [(0)(29) + (1)(71) + \dots + (10)(3)]/420 = 2,98$$

e

$$\sum x_i^2/420 = [(0)^2(29) + (1)^2(71) + \dots + (10)^2(3)]/420 = 12,40$$

Dessa forma,

$$\hat{p} = \frac{2,98}{12,40 - (2,98)^2} = 0,85 \quad \hat{r} = \frac{(2,98)^2}{12,40 - (2,98)^2 - 2,98} = 16,5$$

Embora  $r$  por definição deva ser positivo, o denominador de  $\hat{r}$  pode ser negativo, indicando que a distribuição binomial negativa não é apropriada (ou que o estimador pelo método dos momentos é inadequado). ■

### Estimadores de Máxima Verossimilhança

O método da máxima verossimilhança foi introduzido primeiro por R. A. Fisher, um geneticista e estatístico, na década de 1920. A maioria dos estatísticos recomenda esse método, pelo menos quando o tamanho da amostra for grande, uma vez que os estimadores resultantes têm certas propriedades de eficiência desejáveis (veja a segunda proposição da página 240).

#### Exemplo 6.14

Uma amostra de 10 capacetes de ciclista fabricados por certa empresa é obtida. Em teste, descobre-se que o primeiro, o terceiro e o décimo capacetes estão com defeito, enquanto os demais não. Seja  $p = P(\text{capacete com defeito})$  e defina  $X_1, \dots, X_{10}$  por  $X_i = 1$  se o  $i$ -ésimo capacete estiver com defeito e por zero, caso contrário. Então, os  $x_i$ s observados são 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, de modo que a fmp conjunta da amostra é

$$f(x_1, x_2, \dots, x_{10}; p) = p(1 - p)p \cdots p = p^3(1 - p)^7 \tag{6.4}$$

Perguntamos agora: “Para que valor de  $p$  a amostra observada é mais provável de ter ocorrido?” Isto é, queremos determinar o valor de  $p$  que maximiza a fmp (6.4) ou, de forma equivalente, que maximiza o log natural de (6.4).<sup>2</sup> Uma vez que

$$\ln[f(x_1, \dots, x_{10}; p)] = 3 \ln(p) + 7 \ln(1 - p) \tag{6.5}$$

a qual é uma função diferenciável de  $p$ , igualando a derivada de (6.5) a zero, resultando o valor máximo<sup>3</sup>

$$\frac{d}{dp} \ln[f(x_1, \dots, x_{10}; p)] = \frac{3}{p} - \frac{7}{1 - p} = 0 \quad p = \frac{3}{10} = \frac{x}{n}$$

onde  $x$  é o número observado de sucessos (capacetes com defeito). A estimativa de  $p$  é agora  $\hat{p} = \frac{3}{10}$ . Chama-se estimativa da máxima verossimilhança, pois, para valores fixos de  $x_1, \dots, x_{10}$ , esse é o valor do parâmetro que maximiza a verossimilhança (fmp conjunta) da amostra observada.

Observe que, se nos dissessem somente que entre os 10 capacetes havia três que estavam com defeito, a Equação (6.4) seria substituída pela fmp binomial  $\binom{10}{3} p^3(1 - p)^7$ , que também é maximizada para  $\hat{p} = \frac{3}{10}$ . ■

**DEFINIÇÃO**

Sejam  $X_1, X_2, \dots, X_n$  a fmp ou fdp conjunta

$$f(x_1, x_2, \dots, x_n; \theta_1, \dots, \theta_m) \tag{6.6}$$

onde os parâmetros  $\theta_1, \dots, \theta_m$  têm valores desconhecidos. Quando  $x_1, \dots, x_n$  são os valores observados da amostra e (6.6) é considerada uma função de  $\theta_1, \dots, \theta_m$ , é denominada **função de verossimilhança**. As estimativas de máxima verossimilhança (emvs)  $\hat{\theta}_1, \dots, \hat{\theta}_m$  são os valores dos  $\theta_i$ s que maximizam a função de verossimilhança, de modo que

$$f(x_1, \dots, x_n; \hat{\theta}_1, \dots, \hat{\theta}_m) \geq f(x_1, \dots, x_n; \theta_1, \dots, \theta_m) \quad \text{para todos os } \theta_1, \dots, \theta_m$$

Quando os  $X_i$ s são substituídos no lugar de  $x_i$ s, resultam os **estimadores de máxima verossimilhança**.

<sup>2</sup> Uma vez que  $\ln[g(x)]$  é uma função monotônica de  $g(x)$ , determinar  $x$  que maximiza  $\ln[g(x)]$  é equivalente a maximizar a própria  $g(x)$ . Em estatística, calcular o logaritmo geralmente muda um produto para uma soma, com a qual é mais fácil trabalhar.

<sup>3</sup> Esta conclusão requer a verificação da segunda derivada, mas os detalhes são omitidos.

A função verossimilhança nos diz quão provável a amostra observada é uma função dos valores dos parâmetros possíveis. Maximizar a verossimilhança fornece os valores dos parâmetros para os quais a amostra observada é mais provável de ter sido gerada – isto é, os valores de parâmetro que “mais aproximadamente concordam” com os dados observados.

### Exemplo 6.15

Suponha que  $X_1, X_2, \dots, X_n$  seja a amostra aleatória de uma distribuição exponencial com parâmetro  $\lambda$ . Devido à independência, a função verossimilhança é o produto das fdps individuais:

$$f(x_1, \dots, x_n; \lambda) = (\lambda e^{-\lambda x_1}) \cdot \dots \cdot (\lambda e^{-\lambda x_n}) = \lambda^n e^{-\lambda \sum x_i}$$

O ln(verossimilhança) é

$$\ln[f(x_1, \dots, x_n; \lambda)] = n \ln(\lambda) - \lambda \sum x_i$$

Igualar  $(d/d\lambda)[\ln(\text{verossimilhança})]$  a zero resulta  $n/\lambda - \sum x_i = 0$ , ou  $\lambda = n/\sum x_i = 1/\bar{x}$ . Dessa forma, a emv é  $\lambda = 1/\bar{X}$ ; é idêntica ao método de estimadores pelos momentos [mas não é um estimador não-tendencioso, uma vez que  $E(1/\bar{X}) \neq 1/E(\bar{X})$ ]. ■

### Exemplo 6.16

Seja  $X_1, \dots, X_n$  a amostra aleatória de uma distribuição normal. A função verossimilhança é

$$\begin{aligned} f(x_1, \dots, x_n; \mu, \sigma^2) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x_1-\mu)^2/(2\sigma^2)} \cdot \dots \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x_n-\mu)^2/(2\sigma^2)} \\ &= \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} e^{-\sum(x_i-\mu)^2/(2\sigma^2)} \end{aligned}$$

então

$$\ln[f(x_1, \dots, x_n; \mu, \sigma^2)] = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum (x_i - \mu)^2$$

Para determinar os valores máximos de  $\mu$  e  $\sigma^2$ , devemos calcular as derivadas parciais de  $\ln(f)$  em relação a  $\mu$  e  $\sigma^2$ , igualá-las a zero e resolver as duas equações resultantes. Omitindo os detalhes, as emvs resultantes são

$$\hat{\mu} = \bar{X} \quad \hat{\sigma}^2 = \frac{\sum (X_i - \bar{X})^2}{n}$$

A emv de  $\sigma^2$  não é o estimador não-tendencioso; então, dois princípios diferentes de estimativa (não-tendenciosidade e máxima verossimilhança) produzem dois estimadores diferentes. ■

### Exemplo 6.17

No Capítulo 3, discutimos o uso da distribuição de Poisson para modelar o número de “eventos” que ocorrem em uma região bidimensional. Assuma que, quando a região  $R$ , que está servindo como amostra, possui área  $a(R)$ , o número  $X$  de eventos que ocorrem em  $R$  tem distribuição de Poisson com parâmetro  $\lambda a(R)$  (em que  $\lambda$  é o número esperado de eventos por unidade de área) que as regiões não-sobrepostas produzem  $X$ s independentes.

Suponha que um ecologista selecione  $n$  regiões não-sobrepostas  $R_1, \dots, R_n$  e conte o número de certa espécie de plantas encontrada em cada região. A fmp conjunta (verossimilhança) é então

$$\begin{aligned} p(x_1, \dots, x_n; \lambda) &= \frac{[\lambda \cdot a(R_1)]^{x_1} e^{-\lambda \cdot a(R_1)}}{x_1!} \cdot \dots \cdot \frac{[\lambda \cdot a(R_n)]^{x_n} e^{-\lambda \cdot a(R_n)}}{x_n!} \\ &= \frac{[a(R_1)]^{x_1} \cdot \dots \cdot [a(R_n)]^{x_n} \cdot \lambda^{\sum x_i} \cdot e^{-\lambda \sum a(R_i)}}{x_1! \cdot \dots \cdot x_n!} \end{aligned}$$

O  $\ln(\text{verossimilhança})$  é

$$\ln[p(x_1, \dots, x_n; \lambda)] = \sum x_i \cdot \ln[a(R_i)] + \ln(\lambda) \cdot \sum x_i - \lambda \sum a(R_i) - \sum \ln(x_i!)$$

Calcular  $d/d\lambda \ln(p)$  e igualá-la a zero produz

$$\frac{\sum x_i}{\lambda} - \sum a(R_i) = 0$$

então

$$\lambda = \frac{\sum x_i}{\sum a(R_i)}$$

A emv é então  $\hat{\lambda} = \sum X_i / \sum a(R_i)$ . Isto é intuitivamente razoável, pois  $\lambda$  é a densidade real (plantas por área de unidade), enquanto  $\hat{\lambda}$  é a densidade da amostra, uma vez que  $\sum a(R_i)$  é apenas a área total que serviu como amostra. Em virtude de  $E(X_i) = \lambda \cdot a(R_i)$ , o estimador é não-tendencioso.

Às vezes, é usado um procedimento alternativo de amostragem. Em vez de fixar regiões a servirem como amostra, o ecologista selecionará  $n$  pontos em toda a área de interesse e representará por  $y_i$  a distância do  $i$ -ésimo ponto para a planta mais próxima. A função distribuição acumulada (FDA) de  $Y =$  distância para a planta mais próxima é

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = 1 - P(Y > y) = 1 - P\left(\begin{array}{l} \text{nenhuma planta em um} \\ \text{círculo de raio } y \end{array}\right) \\ &= 1 - \frac{e^{-\lambda \pi y^2} (\lambda \pi y^2)^0}{0!} = 1 - e^{-\lambda \cdot \pi y^2} \end{aligned}$$

Calculando a derivada de  $F_Y(y)$  em relação a  $y$  produz

$$f_Y(y; \lambda) = \begin{cases} 2\pi\lambda y e^{-\lambda \pi y^2} & y \geq 0 \\ 0 & \text{caso contrário} \end{cases}$$

Se agora montarmos a verossimilhança  $f_Y(y_1; \lambda) \cdot \dots \cdot f_Y(y_n; \lambda)$ , diferenciando  $\ln(\text{verossimilhança})$ , e assim por diante, a emv resultante é

$$\hat{\lambda} = \frac{n}{\pi \sum Y_i^2} = \frac{\text{número de plantas observadas}}{\text{área total amostrada}}$$

que também é a densidade amostral. Pode-se demonstrar que, em um ambiente esparsos ( $\lambda$  pequeno), o método da distância é, de certo modo, melhor, enquanto em um ambiente denso, o primeiro método de amostragem é melhor. ■

### Exemplo 6.18

Seja  $X_1, \dots, X_n$  a amostra aleatória de uma fdp de Weibull

$$f(x; \alpha, \beta) = \begin{cases} \frac{\alpha}{\beta^\alpha} \cdot x^{\alpha-1} \cdot e^{-(x/\beta)^\alpha} & x \geq 0 \\ 0 & \text{caso contrário} \end{cases}$$

Escrever a verossimilhança e  $\ln(\text{verossimilhança})$ , depois, definir  $(\partial/\partial\alpha)[\ln(f)] = 0$  e  $(\partial/\partial\beta)[\ln(f)] = 0$  produz as equações

$$\alpha = \left[ \frac{\sum x_i^\alpha \cdot \ln(x_i)}{\sum x_i^\alpha} - \frac{\sum \ln(x_i)}{n} \right]^{-1} \quad \beta = \left( \frac{\sum x_i^\alpha}{n} \right)^{1/\alpha}$$

Essas duas equações não podem ser resolvidas explicitamente para fornecer fórmulas gerais para as funções  $\hat{\alpha}$  e  $\hat{\beta}$ . Ao contrário, para cada amostra  $x_1, \dots, x_n$ , as equações devem ser resolvidas por um procedimento numérico iterativo. Mesmo os estimadores pelo método dos momentos para  $\alpha$  e  $\beta$  são um tanto complicados (veja o Exercício 21). ■

## Estimando as Funções de Parâmetros

No Exemplo 6.16, obtivemos a emv de  $\sigma^2$  quando a distribuição subjacente era normal. A emv de  $\sigma = \sqrt{\sigma^2}$ , bem como tantas outras emvs, pode ser facilmente deduzida, usando-se a seguinte proposição.

### PROPOSIÇÃO

#### Princípio de Invariância

Sejam  $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_m$  as emvs dos parâmetros  $\theta_1, \theta_2, \dots, \theta_m$ . Então, a emv de qualquer função  $h(\theta_1, \theta_2, \dots, \theta_m)$  desses parâmetros é a função  $h(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_m)$  das emvs.

**Exemplo 6.19** (continuação do Exemplo 6.16)

No caso normal, as emvs de  $\mu$  e  $\sigma^2$  são  $\hat{\mu} = \bar{X}$  e  $\hat{\sigma}^2 = \sum(X_i - \bar{X})^2/n$ . Para obter a emv da função  $h(\mu, \sigma^2) = \sqrt{\sigma^2} = \sigma$ , substitua as emvs na função:

$$\hat{\sigma} = \sqrt{\hat{\sigma}^2} = \left[ \frac{1}{n} \sum(X_i - \bar{X})^2 \right]^{1/2}$$

A emv de  $\sigma$  não é o desvio padrão amostral  $S$ , embora sejam próximos, a menos que  $n$  seja relativamente pequeno. ■

**Exemplo 6.20** (continuação do Exemplo 6.18)

O valor médio de uma va  $X$  que tenha distribuição de Weibull é

$$\mu = \beta \cdot \Gamma(1 + 1/\alpha)$$

A emv de  $\mu$  é, portanto,  $\hat{\mu} = \hat{\beta}\Gamma(1 + 1/\hat{\alpha})$ , onde  $\hat{\alpha}$  e  $\hat{\beta}$  são emvs de  $\alpha$  e  $\beta$ . Em particular,  $\bar{X}$  não é a emv de  $\mu$ , embora seja um estimador não-tendencioso. Pelo menos para  $n$  grande,  $\hat{\mu}$  é um estimador melhor que  $\bar{X}$ . ■

## Comportamento de Amostra Grande da EMV

Embora o princípio da estimativa da máxima verossimilhança tenha apelação intuitiva considerável, a seguinte proposição fornece base adicional para o uso das emvs.

### PROPOSIÇÃO

Sob condições muito genéricas, a distribuição conjunta da amostra, quando o tamanho amostral  $n$  é grande, o estimador da máxima verossimilhança de qualquer parâmetro  $\theta$  é aproximadamente não-tendencioso [ $E(\hat{\theta}) \approx \theta$ ] e tem variância tão pequena quanto a que pode ser atingida por qualquer estimador. Dito de outra maneira, a emv  $\hat{\theta}$  é aproximadamente o ENTMV de  $\theta$ .

Em virtude desse resultado e do fato de as técnicas com base no cálculo geralmente poderem ser usados para deduzir as emvs (embora com freqüência sejam necessários métodos numéricos, como o método de

Newton), a estimativa de máxima verossimilhança é a técnica de estimativa mais amplamente usada entre os estatísticos. Muitos dos estimadores usados no restante do livro são emvs. Obter uma emv, entretanto, requer que a distribuição subjacente seja especificada.

### Algumas Complicações

Às vezes, o cálculo não pode ser usado para obter as emvs.

#### Exemplo 6.21

Suponha que meu tempo de espera por um ônibus seja uniformemente distribuído em  $[0, \theta]$  e que os resultados  $x_1, \dots, x_n$  de uma amostra aleatória dessa distribuição tenham sido observados. Uma vez que  $f(x; \theta) = 1/\theta$  para  $0 \leq x \leq \theta$  e 0, caso contrário,

$$f(x_1, \dots, x_n; \theta) = \begin{cases} \frac{1}{\theta^n} & 0 \leq x_1 \leq \theta, \dots, 0 \leq x_n \leq \theta \\ 0 & \text{caso contrário} \end{cases}$$

Contanto que  $\max(x_i) \leq \theta$ , a verossimilhança é  $1/\theta^n$ , que é positiva, mas logo que  $\theta < \max(x_i)$ , a verossimilhança cai para 0. Este fato é ilustrado na Figura 6.4. O cálculo não funcionará porque o máximo da verossimilhança ocorre em um ponto de descontinuidade, mas a figura mostra que  $\theta = \max(X_i)$ . Assim, se meus tempos de espera são 2,3; 3,7; 1,5; 0,4 e 3,2, então a emv é  $\hat{\theta} = 3,7$ .

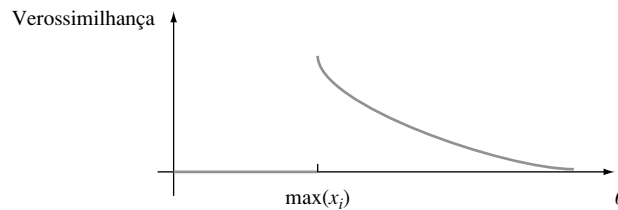


Figura 6.4 A função verossimilhança do Exemplo 6.21

#### Exemplo 6.22

Um método geralmente usado para calcular o tamanho de uma população de animais selvagens envolve a realização de um experimento de captura/recaptura no qual uma amostra inicial de  $M$  animais é capturada, todos são identificados e devolvidos à população. Após permitir tempo suficiente para os animais marcados se misturarem na população, outra amostra de tamanho  $n$  é capturada. Com  $X$  = número de animais marcados na segunda amostra, o objetivo é usar o  $x$  observado para estimar o tamanho da população  $N$ .

O parâmetro de interesse é  $\theta = N$ , que pode assumir somente valores inteiros; assim, mesmo após determinar a função verossimilhança (fmp de  $X$  aqui), usar o cálculo para obter  $N$  apresentaria dificuldades. Se considerarmos um sucesso um animal anteriormente marcado que está sendo recapturado, então a amostragem é sem reposição de uma população que contém  $M$  sucessos e  $N - M$  falhas, de modo que  $X$  é uma va hipergeométrica e a função verossimilhança é

$$p(x; N) = h(x; n, M, N) = \frac{\binom{M}{x} \cdot \binom{N - M}{n - x}}{\binom{N}{n}}$$

Com a natureza de  $N$  estimada em valor inteiro, entretanto, seria difícil calcular a derivada de  $p(x; N)$ . No entanto, se considerarmos a relação entre  $p(x; N)$  a  $p(x; N - 1)$ , teremos

$$\frac{p(x; N)}{p(x; N - 1)} = \frac{(N - M) \cdot (N - n)}{N(N - M - n + x)}$$



Essa relação é maior que 1 se e somente se  $N < Mn/x$ . O valor de  $N$  para o qual  $p(x; N)$  é maximizado é, portanto, o maior inteiro inferior a  $Mn/x$ . Se usarmos notação matemática-padrão  $[r]$  para o maior inteiro menor ou igual a  $r$ , a emv de  $N$  é  $\hat{N} = [Mn/x]$ . Como ilustração, se  $M = 200$  peixes que são tirados de um lago e marcados, subseqüentemente,  $n = 100$  peixes que são recapturados, e entre os 100 há  $x = 11$  peixes marcados, então  $\hat{N} = [(200)(100)/11] = [1818,18] = 1818$ . A estimativa é realmente bastante intuitiva;  $x/n$  é a proporção da amostra recapturada que é marcada, enquanto  $M/N$  é a proporção da população inteira marcada. A estimativa é obtida igualando-se essas duas proporções (estimando a proporção de uma população pela proporção de uma amostra). ■

Suponha que  $X_1, X_2, \dots, X_n$  seja uma amostra aleatória de uma fdp  $f(x; \theta)$  simétrica em relação a  $\theta$ , mas que o investigador esteja incerto quanto à forma da função  $f$ . Então, é desejável usar um estimador  $\hat{\theta}$  que seja robusto – isto é, que tenha bom desempenho para uma variedade ampla de fdps subjacentes. Tal estimador é uma média aparada. Nos últimos anos, os estatísticos propuseram outro tipo de estimador, chamado *Estimador M*, com base em uma generalização da estimativa de máxima verossimilhança. Em vez de maximizar o logaritmo da verossimilhança  $\sum \ln[f(x; \theta)]$  para um  $f$  especificado, maximiza-se  $\sum \rho(x_i; \theta)$ . A “função objetivo”  $\rho$  é selecionada para produzir um estimador com boas propriedades de robustez. O livro de David Hoaglin et al. (veja a bibliografia) contém uma boa exposição sobre esse assunto.

## Exercícios | Seção 6.2 (20–30)

- 20.** Uma amostra aleatória de  $n$  capacetes de ciclista fabricados por uma empresa é selecionada. Seja  $X =$  número entre os  $n$  que esteja com defeito, e  $p = P(\text{com defeito})$ . Assuma que somente  $X$  seja observado, em vez da seqüência de  $S$ s e  $F$ s.
- Encontre o estimador de máxima verossimilhança de  $p$ . Se  $n = 20$  e  $x = 3$ , qual é a estimativa?
  - O estimador da parte (a) é não-tendencioso?
  - Se  $n = 20$  e  $x = 3$ , qual é a emv da probabilidade  $(1 - p)^5$  de nenhum dos cinco capacetes seguintes examinados estar com defeito?
- 21.** Seja  $X$  uma distribuição de Weibull com parâmetros  $\alpha$  e  $\beta$ , então

$$E(X) = \beta \cdot \Gamma(1 + 1/\alpha)$$

$$V(X) = \beta^2 \{ \Gamma(1 + 2/\alpha) - [\Gamma(1 + 1/\alpha)]^2 \}$$

- Com base em uma amostra aleatória  $X_1, \dots, X_n$ , escreva equações para o método dos estimadores pelos momentos de  $\beta$  e  $\alpha$ . Mostre que, uma vez obtida a estimativa de  $\alpha$ , a estimativa de  $\beta$  é determinada numa tabela da função gama, e que a estimativa de  $\alpha$  é a solução para uma equação complicada que envolve a função gama.
  - Se  $n = 20$ ,  $\bar{x} = 28,0$ , e  $\sum x_i^2 = 16.500$ , calcule as estimativas. [Sugestão:  $[\Gamma(1,2)]^2/\Gamma(1,4) = 0,95$ .]
- 22.** Seja  $X$  a proporção de tempo distribuído que um aluno selecionado aleatoriamente gasta trabalhando em um teste de aptidão. Suponha que a fdp de  $X$  seja

$$f(x; \theta) = \begin{cases} (\theta + 1)x^\theta & 0 \leq x \leq 1 \\ 0 & \text{caso contrário} \end{cases}$$

onde  $-1 < \theta$ . Uma amostra aleatória de 10 alunos produz dados  $x_1 = 0,92$ ,  $x_2 = 0,79$ ,  $x_3 = 0,90$ ,  $x_4 = 0,65$ ,  $x_5 = 0,86$ ,  $x_6 = 0,47$ ,  $x_7 = 0,73$ ,  $x_8 = 0,97$ ,  $x_9 = 0,94$ ,  $x_{10} = 0,77$ .

- Use o método dos momentos para obter um estimador de  $\theta$  e calcule a estimativa para esses dados.
  - Obtenha o estimador de máxima verossimilhança de  $\theta$  e calcule a estimativa dos dados fornecidos.
- 23.** Dois sistemas diferentes de computador são monitorados durante um total de  $n$  semanas. Seja  $X_i$  o número de falhas do primeiro sistema durante a  $i$ -ésima semana e suponha que os  $X_i$ 's sejam independentes e tirados de uma distribuição de Poisson com parâmetro  $\lambda_1$ . De forma semelhante, seja  $Y_i$  o número de falhas do segundo sistema durante a  $i$ -ésima semana, e assumamos independência em que cada Poisson de  $Y_i$  possua parâmetro  $\lambda_2$ . Derive as emvs de  $\lambda_1$ ,  $\lambda_2$ , e  $\lambda_1 - \lambda_2$ . [Sugestão: usando a independência, escreva a fmp conjunta (verossimilhança) dos  $X_i$ 's e  $Y_i$ 's juntos.]
- 24.** Consulte o Exercício 20. Em vez de selecionar  $n = 20$  capacetes para examinar, suponha que eu examine os capacetes em sucessão até que encontre  $r = 3$  defeituosos. Se o 20º capacete for o terceiro com defeito (de modo que o número de capacetes examinados que não tinham defeito fosse  $x = 17$ ), qual é a emv de  $p$ ? Isso é o mesmo que a estimativa do Exercício 20? Por que ou por que não? É o mesmo que a estimativa calculada pelo estimador não-tendencioso do Exercício 17?
- 25.** A resistência de corte das 10 soldas de ponto de teste é determinada, produzindo os seguintes dados (psi):
- 392 376 401 367 389 362 409 415 358 375

- a. Assumindo que a resistência de corte seja normalmente distribuída, estime a resistência de corte média real e o desvio padrão da resistência de corte, usando o método da máxima verossimilhança.
- b. Novamente assumindo uma distribuição normal, estime o valor da resistência abaixo do qual 95% das soldas terão a resistência especificada. (*Sugestão*: qual é o 95º percentil em termos de  $\mu$  e  $\sigma$ ? Agora, use o princípio da invariância.)
26. Consulte o Exercício 25. Suponha que decidamos examinar outra solda de ponto de teste. Seja  $X$  = a resistência de corte da solda. Use os dados fornecidos para obter a emv de  $P(X \leq 400)$ . [*Sugestão*:  $P(X \leq 400) = \Phi((400 - \mu)/\sigma)$ .]
27. Seja  $X_1, \dots, X_n$  a amostra aleatória de uma distribuição gama com parâmetros  $\alpha$  e  $\beta$ .
- a. Encontre as equações cuja solução produz os estimadores de máxima verossimilhança de  $\alpha$  e  $\beta$ . Você acha que eles podem ser resolvidos explicitamente?
- b. Mostre que a emv de  $\mu = \alpha\beta$  é  $\hat{\mu} = \bar{X}$ .
28. Seja  $X_1, X_2, \dots, X_n$  a amostra aleatória da distribuição de Rayleigh com função de densidade dada no Exercício 15. Determine
- a. o estimador de máxima verossimilhança de  $\theta$  e depois calcule a estimativa dos dados da força vibratória fornecida nesse exercício. Esse estimador é o mesmo que o estimador não-tendencioso sugerido no Exercício 15?
- b. a emv da mediana da distribuição da força vibratória. (*Sugestão*: expresse primeiro a mediana em termos de  $\theta$ .)
29. Considere uma amostra aleatória  $X_1, X_2, \dots, X_n$  da fdp exponencial deslocada
- $$f(x; \lambda, \theta) = \begin{cases} \lambda e^{-\lambda(x-\theta)} & x \geq \theta \\ 0 & \text{caso contrário} \end{cases}$$
- Considerando  $\theta = 0$  fornece a fdp da distribuição exponencial considerada anteriormente (com densidade positiva à direita de zero). Um exemplo da distribuição exponencial deslocada apareceu no Exemplo 4.4, em que a variável de interesse era o tempo de avanço no fluxo do tráfego, e  $\theta = 0,5$  era o tempo de avanço mínimo possível.
- a. Obtenha os estimadores de máxima verossimilhança de  $\theta$  e  $\lambda$ .
- b. Se são feitas  $n = 10$  observações do tempo de avanço, resultando nos valores 3,11, 0,64, 2,55, 2,20, 5,44, 3,42, 10,39, 8,93, 17,82 e 1,30, calcule as estimativas de  $\theta$  e  $\lambda$ .
30. No tempo  $t = 0$ ; 20 componentes idênticos são examinados. A distribuição da vida útil de cada um é exponencial com parâmetro  $\lambda$ . O investigador deixa a instalação de teste sem monitoramento. Retornando 24 horas depois, o pesquisador termina imediatamente o teste, após perceber que  $y = 15$  dos 20 componentes ainda estão funcionando (então, 5 falharam). Encontre a emv de  $k$ . (*Sugestão*: seja  $Y$  = número que sobrevive 24 horas. Então,  $Y \sim \text{Bin}(n, p)$ . Qual é a emv de  $p$ ? Agora, observe que  $p = P(X_i \geq 24)$ , em que  $X_i$  é exponencialmente distribuído. O que relaciona  $\lambda$  a  $p$ , de modo que o primeiro pode ser estimado, uma vez que o último o foi.]

---

Exercícios Suplementares (31–38)

31. Um estimador  $\theta$  é dito ser **consistente** se para qualquer  $\hat{\epsilon} > 0$ ,  $P(|\hat{\theta} - \theta| \geq \epsilon) \leftrightarrow 0$  quando  $n \leftrightarrow \infty$ . Isto é,  $\hat{\theta}$  é consistente se, à medida que o tamanho da amostra aumenta, menor é a probabilidade de  $\hat{\theta}$  ser maior que  $\epsilon$  a partir do valor real de  $\theta$ . Mostre que  $\bar{X}$  é um estimador consistente de  $\mu$  quando  $\sigma^2 < \infty$  usando a desigualdade de Chebyshev do Exercício 43, do Capítulo 3. (*Sugestão*: a desigualdade pode ser reescrita na forma

$$P(|Y - \mu_Y| \geq \epsilon) \leq \sigma_Y^2/\epsilon$$

Agora, identifique  $Y$  com  $\bar{X}$ .)

32. a. Seja  $X_1, \dots, X_n$  a amostra aleatória de uma distribuição uniforme em  $[0, \theta]$  Então, a emv de  $\theta$  é  $\hat{\theta} = Y = \max(X_i)$ . Use o fato de  $Y \leq y$  se e somente se cada  $X_i \leq y$  para encontrar a fdc de  $Y$ . Então, mostre que a fdp de  $Y = \max(X_i)$  é

$$f_Y(y) = \begin{cases} \frac{ny^{n-1}}{\theta^n} & 0 \leq y \leq \theta \\ 0 & \text{caso contrário} \end{cases}$$

- b. Use o resultado da parte (a) para mostrar que a emv é tendenciosa, mas que  $(n + 1)\max(X_i)/n$  é não-tendenciosa.
33. No tempo  $t = 0$ , há um indivíduo com vida em certa população. Um **processo de nascimento puro**, então, desenvolve-se como segue. O tempo até o primeiro nascimento é exponencialmente distribuído com parâmetro  $\lambda$ . Após o primeiro nascimento, há dois indivíduos com vida. O tempo até que o primeiro dê à luz novamente é exponencial com parâmetro  $\lambda$  e acontece de maneira semelhante com o segundo indivíduo. Dessa forma, o tempo até o próximo nascimento é o mínimo de duas variáveis exponenciais ( $\lambda$ ), que é exponencial com parâmetro  $2\lambda$ . De forma semelhante, uma vez que

tenha ocorrido o segundo nascimento, há três indivíduos com vida, de modo que o tempo até o nascimento seguinte é uma va exponencial com parâmetro  $3\lambda$  e assim por diante (a propriedade de ausência de memória da distribuição exponencial está sendo usada aqui). Suponha que o processo seja observado até que o sexto nascimento tenha ocorrido e que os tempos de nascimento sucessivos sejam 25,2, 41,7, 51,2, 55,5, 59,5, 61,8 (a partir dos quais você deve calcular os tempos entre os nascimentos sucessivos). Encontre a emv de  $\lambda$ . (*Sugestão*: a verossimilhança é um produto de termos exponenciais.)

34. A **média dos erros quadráticos** de um estimador  $\hat{\theta}$  é  $MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2$ . Se  $\hat{\theta}$  é não-tendencioso, então  $MSE(\hat{\theta}) = V(\hat{\theta})$ , mas em geral  $MSE(\hat{\theta}) = V(\hat{\theta}) + (\text{desvio})^2$ . Considere o estimador  $\hat{\sigma}^2 = KS^2$ , em que  $S^2 = \text{variância da amostra}$ . Que valor de  $K$  minimiza a média dos erros quadráticos desse estimador quando a distribuição da população é normal? [*Sugestão*: pode-se mostrar que

$$E[(S^2)^2] = (n + 1)\sigma^4/(n - 1)$$

Em geral, é difícil achar que  $\hat{\theta}$  minimiza  $MSE(\hat{\theta})$ , motivo pelo qual olhamos somente para estimadores não-tendenciosos e minimizamos  $V(\hat{\theta})$ .]

35. Seja  $X_1, \dots, X_n$  a amostra aleatória de uma fdp que seja simétrica com relação a  $\mu$ . Um estimador de  $\mu$  considerado por realizar bem uma variedade de distribuições subjacentes é o *estimador de Hodges-Lehmann*. Para defini-lo, calcule primeiro para cada  $i \leq j$  e cada  $j = 1, 2, \dots, n$  a média de pareados  $\bar{X}_{i,j} = (X_i + X_j)/2$ . Então, o estimador é  $\hat{\mu} =$  a mediana dos  $\bar{X}_{i,j}$ s. Calcule o valor dessa estimativa, usando os dados do Exercício 44, do

Capítulo 1. (*Sugestão*: construa uma tabela quadrada com os  $x_{i,j}$  listados na margem esquerda e na parte superior. Depois, calcule as médias na diagonal e acima dela.)

36. Quando a distribuição da população é normal, a mediana estatística  $\{|X_1 - \bar{X}|, \dots, |X_n - \bar{X}|\}/0,6745$  é usada para estimar  $\sigma$ . O estimador é mais resistente aos efeitos de *outliers* (observações distantes da parte principal dos dados) que o desvio padrão da amostra. Calcule a estimativa pontual correspondente e  $s$  para os dados do Exemplo 6.2.
37. Quando o desvio padrão da amostra  $S$  se baseia em uma amostra aleatória de uma distribuição normal da população, pode ser demonstrado que

$$E(S) = \sqrt{2/(n-1)}\Gamma(n/2)\sigma/\Gamma((n-1)/2)$$

Use essa equação para obter um estimador não-tendencioso para  $\sigma$  da forma  $cS$ . Qual é  $c$  quando  $n = 20$ ?

38. Todos os  $n$  espécimes serão pesados duas vezes na mesma escala. Represente por  $X_i$  e  $Y_i$  os dois pesos observados do  $i$ -ésimo espécime. Suponha que  $X_i$  e  $Y_i$  sejam independentes entre si, cada um normalmente distribuído com valor médio  $\mu_i$  (o peso real do espécime  $i$ ) e variância  $\sigma^2$ .
- a. Mostre que o estimador de máxima verossimilhança de  $\sigma^2$  é  $\hat{\sigma}^2 = \sum(X_i - Y_i)^2/(4n)$ . [*Sugestão*: se  $\bar{z} = (z_1 + z_2)/2$ , então,  $\sum(z_i - \bar{z})^2 = (z_1 - z_2)^2/2$ .]
- b. A emv  $\hat{\sigma}^2$  é um estimador não-tendencioso de  $\sigma^2$ . Determine um estimador não-tendencioso de  $\sigma^2$ . [*Sugestão*: para qualquer va  $Z$ ,  $E(Z^2) = V(Z) + [E(Z)]^2$ . Aplique a equação em  $Z = X_i - Y_i$ .]

---

## Bibliografia

- DEGROOT, Morris e SCHERVISH, Mark. *Probability and Statistics* (3. ed.). Addison-Wesley: Boston, MA, 2002. Inclui uma excelente discussão das propriedades e métodos gerais da estimativa pontual; para interesse particular, existem exemplos que mostram como os princípios e métodos gerais podem produzir estimadores insatisfatórios em situações específicas.
- EFRON, Bradley e TIBSHIRANI, Robert. *An Introduction to the Bootstrap*, Chapman e Hall, Nova York, 1993. A bíblia do *bootstrap*.
- HOAGLIN, David, MOSTELLER, Frederick e TUKEY, John. *Understanding Robust and Exploratory Data Analysis*. Wiley, Nova York, 1983. Contém diversos

capítulos bons sobre estimativa pontual robusta, incluindo um sobre Estimativa  $M$ .

- HOGG, Robert e CRAIG, Allen. *Introduction to Mathematical Statistics*. 5. ed., Prentice Hall, Englewood Cliffs, NJ, 1995. Uma boa discussão sobre não-tendenciosidade.
- LARSEN, Richard e MARX, Morris. *Introduction to Mathematical Statistics*. 2. ed., Prentice Hall, Englewood Cliffs, NJ, 1985. Uma discussão muito boa sobre estimativa pontual a partir de uma perspectiva levemente mais matemática que o texto atual.
- RICE, John. *Mathematical Statistics and Data Analysis*. 2. ed., Duxbury Press, Belmont, CA, 1994. Uma boa mistura de teoria estatística e dados.

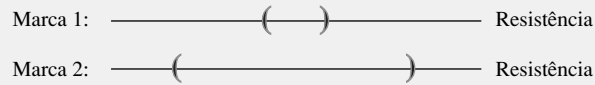
# Intervalos Estatísticos Baseados em uma Única Amostra

## Introdução

Uma estimativa pontual, por ser um único número, não fornece por si mesma qualquer informação sobre a precisão e a confiabilidade da estimativa. Considere, por exemplo, o uso da estatística  $\bar{X}$  para calcular a estimativa pontual da tensão de quebra média real ( $\mu$ ) de toalhas de papel de certa marca e suponha que  $\bar{x} = 9322,7$ . Devido à variabilidade da amostragem, virtualmente nunca é o caso de  $\bar{x} = \mu$ . A estimativa pontual não diz nada sobre o quanto pode estar próxima de  $\mu$ . Uma alternativa para apresentar um único valor sensato para o parâmetro que está sendo estimado é calcular e relatar um intervalo completo de valores plausíveis — uma *estimativa de intervalo* ou *intervalo de confiança* (IC). Um intervalo de confiança sempre é calculado selecionando-se primeiro o *nível de confiança*, que é uma medida do grau de confiabilidade do intervalo. Um intervalo de confiança com nível de confiança de 95% da tensão de quebra média real pode ter limite inferior de 9162,5 e limite superior de 9482,9. Então, com nível de confiança de 95%, qualquer valor de  $\mu$  entre 9162,5 e 9482,9 é plausível. O nível de confiança de 95% implica que 95% de todas as amostras forneceriam um intervalo que inclui  $\mu$ , ou qualquer outro parâmetro que esteja sendo estimado, e apenas 5% de todas as amostras dariam um intervalo errôneo. Os níveis de confiança usados com mais frequência são 95%, 99% e 90%. Quanto maior o nível de confiança, mais fortemente acreditamos que o valor do parâmetro que está sendo estimado está dentro do intervalo (a interpretação de qualquer nível de confiança específico será dada em breve).

As informações sobre a precisão de uma estimativa de intervalo são transmitidas pela sua extensão. Se o nível de confiança for alto e o intervalo resultante, bastante restrito, nosso conhecimento do valor do parâmetro será razoavelmente preciso. Um intervalo de confiança muito amplo, entretanto, passa a idéia de que há muita incerteza com relação ao valor do que estamos estimando. A Figura 7.1 mostra intervalos de confiança de 95% das tensões de quebra médias reais de duas marcas diferentes de toalhas de papel. Um

desses intervalos sugere conhecimento exato de  $\mu$ , enquanto o outro sugere uma variedade bastante grande de valores plausíveis.



**Figura 7.1** Intervalos de confiança que indicam informações precisas (marca 1) e imprecisas (marca 2) sobre  $\mu$

## 7.1 | Propriedades Básicas de Intervalos de Confiança

Os conceitos básicos e as propriedades dos intervalos de confiança (ICs) são introduzidos mais facilmente enfocando primeiro uma situação de problema simples, embora um pouco irreal. Suponha que o parâmetro de interesse seja a média de uma população  $\mu$  e que

1. A distribuição da população é normal
2. O valor do desvio padrão da população  $\sigma$  é conhecido

A normalidade da distribuição populacional é geralmente uma suposição razoável. Entretanto, se o valor de  $\mu$  é desconhecido, é improvável que o valor de  $\sigma$  esteja disponível (o conhecimento do centro de uma população precede tipicamente informações relacionadas à dispersão). Nas seções posteriores, desenvolveremos métodos com base em suposições menos restritivas.

### Exemplo 7.1

Os engenheiros industriais que se especializam em ergonomia estão preocupados em projetar espaços e dispositivos operados por trabalhadores, de modo a obter maior produtividade e conforto. O artigo “Studies on Ergonomically Designed Alphanumeric Keyboards” (*Human Factors*, 1985, p. 175-187) relata o estudo de altura preferida de um teclado experimental com grande apoio para o pulso e o antebraço. Uma amostra de  $n = 31$  digitadores treinados foi selecionada, e a altura preferida do teclado foi determinada para cada digitador. A altura preferida média resultante da amostra foi  $\bar{x} = 80$  cm. Assumindo que a altura preferida seja normalmente distribuída com  $\sigma = 2$  cm (valor sugerido pelos dados no artigo), obtenha um IC para  $\mu$ , a altura média real preferida pela população de todos os digitadores experientes. ■

As observações reais da amostra  $x_1, x_2, \dots, x_n$  são consideradas o resultado de uma amostra aleatória  $X_1, \dots, X_n$  de uma distribuição normal com valor médio  $\mu$  e desvio padrão  $\sigma$ . Os resultados do Capítulo 5, então, implicam que, independentemente do tamanho da amostra  $n$ , a média amostral  $\bar{X}$  é normalmente distribuída com valor esperado  $\mu$  e desvio padrão  $\sigma/\sqrt{n}$ . Padronizar  $\bar{X}$  subtraindo primeiro seu valor esperado e dividindo-o por seu desvio padrão produz a variável

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \quad (7.1)$$

que tem distribuição normal padronizada. Em virtude de a área sob a curva normal padrão entre  $-1,96$  e  $1,96$  ser  $0,95$ ,

$$P\left(-1,96 < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < 1,96\right) = 0,95 \quad (7.2)$$

O passo seguinte no desenvolvimento é manipular as desigualdades dentro dos parênteses em (7.2), para que apareçam sob a forma equivalente  $l < \mu < u$ , em que os pontos finais  $l$  e  $u$  envolvam  $\bar{X}$  e  $\sigma/\sqrt{n}$ . Isso é alcançado por meio da seguinte seqüência de operações, cada uma produzindo desigualdades equivalentes às que iniciamos com:

1. Multiplique por  $\sigma/\sqrt{n}$  para obter

$$-1,96 \cdot \frac{\sigma}{\sqrt{n}} < \bar{X} - \mu < 1,96 \cdot \frac{\sigma}{\sqrt{n}}$$

2. Subtraia  $\bar{X}$  de cada termo para obter

$$-\bar{X} - 1,96 \cdot \frac{\sigma}{\sqrt{n}} < -\mu < -\bar{X} + 1,96 \cdot \frac{\sigma}{\sqrt{n}}$$

3. Multiplique por  $-1$  para eliminar o sinal negativo de  $\mu$  (que inverte a direção de cada desigualdade) para obter

$$\bar{X} + 1,96 \cdot \frac{\sigma}{\sqrt{n}} > \mu > \bar{X} - 1,96 \cdot \frac{\sigma}{\sqrt{n}}$$

isto é,

$$\bar{X} - 1,96 \cdot \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1,96 \cdot \frac{\sigma}{\sqrt{n}}$$

Em virtude de cada conjunto de desigualdades da seqüência ser equivalente ao original, a probabilidade associada a cada um é 0,95. Em particular

$$P\left(\bar{X} - 1,96 \cdot \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1,96 \cdot \frac{\sigma}{\sqrt{n}}\right) = 0,95 \quad (7.3)$$

O evento dentro dos parênteses em (7.3) possui uma aparência um tanto desconhecida; antes a quantidade aleatória sempre apareceu no meio com constantes em ambas as extremidades, como em  $a \leq Y \leq b$ . Em (7.3), a quantidade aleatória aparece nas duas extremidades, enquanto a constante desconhecida  $\mu$  aparece no meio. Para interpretar (7.3), imagine um **intervalo aleatório** que tem ponto final esquerdo  $\bar{X} - 1,96 \cdot \sigma/\sqrt{n}$  e ponto final direito  $\bar{X} + 1,96 \cdot \sigma/\sqrt{n}$  que, na notação de intervalo, é

$$\left(\bar{X} - 1,96 \cdot \frac{\sigma}{\sqrt{n}}, \bar{X} + 1,96 \cdot \frac{\sigma}{\sqrt{n}}\right) \quad (7.4)$$

O intervalo (7.4) é aleatório porque seus dois pontos finais envolvem uma variável aleatória (va). Observe que o intervalo está centrado na média amostral  $\bar{X}$  e se estende  $1,96\sigma/\sqrt{n}$  de cada lado de  $\bar{X}$ . Dessa forma, a amplitude do intervalo é  $2 \cdot (1,96) \cdot \sigma/\sqrt{n}$ , que não é aleatório; somente a posição do intervalo (seu ponto central  $\bar{X}$ ) é aleatória (Figura 7.2). Agora, (7.3) pode ser parafraseado como “a probabilidade de que o intervalo aleatório (7.4) inclua o valor real de  $\mu$  é 0,95”. Antes de qualquer experimento ser realizado e de quaisquer dados serem coletados, é bastante provável (probabilidade de 0,95) que  $\mu$  esteja dentro do intervalo na Expressão (7.4).

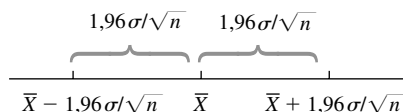


Figura 7.2 O intervalo aleatório (7.4) centralizado em  $\bar{X}$

## DEFINIÇÃO

Após observar  $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$ , calculamos a média amostral observada  $\bar{x}$  e substituímos  $\bar{x}$  em (7.4) no lugar de  $\bar{X}$ ; o intervalo fixo resultante é chamado **intervalo de confiança de 95% de  $\mu$** . Esse IC pode ser expresso tanto como

$$\left( \bar{x} - 1,96 \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + 1,96 \cdot \frac{\sigma}{\sqrt{n}} \right) \text{ é um IC de 95\% de } \mu$$

ou como

$$\bar{x} - 1,96 \cdot \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + 1,96 \cdot \frac{\sigma}{\sqrt{n}} \text{ com 95\% de confiança}$$

Uma expressão concisa para o intervalo é  $\bar{x} \pm 1,96 \cdot \sigma/\sqrt{n}$ , onde  $-$  é o ponto final esquerdo (limite inferior) e  $+$  é o ponto final direito (limite superior).

**Exemplo 7.2** (continuação do Exemplo 7.1)

As quantidades requeridas para o cálculo do IC de 95% da altura média real preferida são  $\sigma = 2,0$ ,  $n = 31$  e  $\bar{x} = 80,0$ . O intervalo resultante é

$$\bar{x} \pm 1,96 \cdot \frac{\sigma}{\sqrt{n}} = 80,0 \pm (1,96) \frac{2,0}{\sqrt{31}} = 80,0 \pm 0,7 = (79,3, 80,7)$$

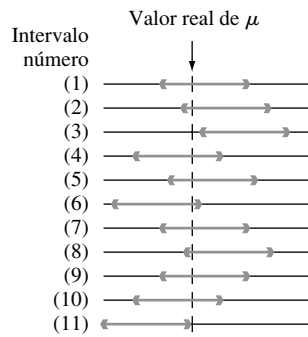
Isto é, podemos estar altamente confiantes, no nível de confiança de 95%, de que  $79,3 < \mu < 80,7$ . Esse intervalo é relativamente restrito, indicando que  $\mu$  foi estimado com precisão. ■

**Interpretando um Intervalo de Confiança**

O nível de confiança de 95% do intervalo definido foi herdado da probabilidade de 0,95 do intervalo aleatório (7.4). Os intervalos que têm outros níveis de confiança serão introduzidos em breve. Agora, no entanto, considere como a confiança de 95% pode ser interpretada.

Em virtude de termos começado com um evento cuja probabilidade era 0,95 – que o intervalo aleatório (7.4) incluiria o valor real de  $\mu$  – e depois usado os dados do Exemplo 7.1 para calcular o intervalo estabelecido (79,3, 80,7), é tentador concluir que  $\mu$  está dentro desse intervalo estabelecido com probabilidade de 0,95. Mas, substituindo  $\bar{x} = 80,0$  por  $\bar{X}$ , toda a aleatoriedade desaparece; o intervalo (79,3, 80,7) não é aleatório e  $\mu$  é uma constante (infelizmente desconhecida para nós), de modo que é *errado* escrever a expressão  $P(\mu \text{ está em } (79,3, 80,7)) = 0,95$ .

A interpretação correta da “confiança de 95%” baseia-se na interpretação de frequência relativa de longo prazo da probabilidade: dizer que um evento  $A$  tem probabilidade de 0,95 significa dizer que, se o experimento no qual  $A$  é definido for repetido várias vezes, em longo prazo,  $A$  ocorrerá 95% das vezes. Suponha que obtemos outra amostra das alturas preferidas dos digitadores e calculemos outro intervalo de 95%. Então, consideramos repetir o experimento para uma terceira amostra, para uma quarta e assim por diante. Seja  $A$  o evento em que  $\bar{X} - 1,96 \cdot \sigma/\sqrt{n} < \mu < \bar{X} + 1,96 \cdot \sigma/\sqrt{n}$ . Uma vez que  $P(A) = 0,95$ , em longo prazo, 95% de nossos ICs calculados conterão  $\mu$ . Esta afirmativa é mostrada na Figura 7.3, onde a linha vertical corta o eixo de medição no valor real (embora desconhecido) de  $\mu$ . Observe que, dos 11 intervalos ilustrados, somente os intervalos 3 e 11 não contêm  $\mu$ . Em longo prazo, somente 5% dos intervalos construídos desse modo não contêm  $\mu$ .



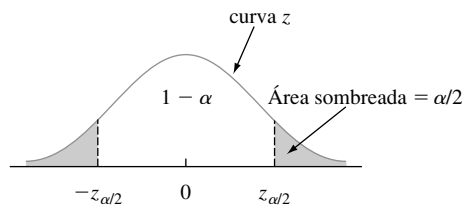
**Figura 7.3** Construção repetida dos ICs de 95%

Conforme essa interpretação, o nível de confiança de 95% não é tanto uma declaração sobre qualquer intervalo específico, como (79,3, 80,7), mas se refere ao que aconteceria se um número muito grande de intervalos semelhantes fosse construído. Embora isso possa parecer insatisfatório, a origem da dificuldade permanece com nossa interpretação de probabilidade; aplica-se a uma seqüência longa de replicações de um experimento em vez de apenas uma única replicação. Existe outra abordagem para a construção e interpretação de ICs que utiliza a noção de probabilidade subjetiva e o teorema de Bayes, mas os detalhes técnicos estão além do escopo deste texto; o livro de Winkler (veja a bibliografia do Capítulo 2) é uma boa fonte. O intervalo apresentado aqui (bem como cada intervalo apresentado subsequentemente) é chamado IC “clássico”, pois sua interpretação depende da noção clássica de probabilidade (embora as idéias principais tenham sido desenvolvidas em um passado relativamente recente, na década de 1930).

### Outros Níveis de Confiança

O nível de confiança de 95% foi herdado da probabilidade de 0,95 das desigualdades iniciais em (7.2). Se for desejado um nível de confiança de 99%, a probabilidade inicial de 0,95 deve ser substituída por 0,99, o que exige a mudança do valor crítico de  $z$  de 1,96 para 2,58. Um IC de 99% resulta, então, do uso de 2,58 no lugar de 1,96 na fórmula do IC de 95%.

Essa afirmativa sugere que qualquer nível de confiança desejado pode ser alcançado substituindo-se 1,96 ou 2,58 pelo valor crítico normal padronizado apropriado. Conforme mostra a Figura 7.4, a probabilidade de  $1 - \alpha$  é alcançada usando-se  $z_{\alpha/2}$  no lugar de 1,96.



**Figura 7.4**  $P(-z_{\alpha/2} \leq Z < z_{\alpha/2}) = 1 - \alpha$

**DEFINIÇÃO**

O **intervalo de confiança  $100(1 - \alpha)\%$**  da média  $\mu$  de uma população normal, quando o valor de  $\sigma$  é conhecido, é dado por

$$\left( \bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \right) \tag{7.5}$$

ou, de forma equivalente, por  $\bar{x} \pm z_{\alpha/2} \cdot \sigma/\sqrt{n}$ .



**Exemplo 7.3**

O processo de produção das unidades de caixa de controle de um tipo específico de motor foi modificado recentemente. Antes dessa modificação, os dados históricos sugeriam que a distribuição dos diâmetros do orifício dos mancais nas caixas eram normais, com um desvio padrão de 0,100 mm. Acredita-se que a modificação não tenha afetado o formato da distribuição ou o desvio padrão, mas que o valor do diâmetro médio possa ter mudado. Uma amostra de 40 unidades da caixa é selecionada e o diâmetro do orifício é determinado para cada uma, resultando em um diâmetro médio da amostra de 5,426 mm. Vamos calcular um intervalo de confiança para o diâmetro médio real do orifício usando um nível de confiança de 90%. O procedimento requer que  $100(1 - \alpha) = 90$ , do qual  $\alpha = 0,10$  e  $z_{\alpha/2} = z_{0,05} = 1,645$  (correspondendo a uma área acumulada da curva  $z$  de 0,9500). O intervalo desejado é, então,

$$5,426 \pm (1,645) \frac{0,100}{\sqrt{40}} = 5,426 \pm 0,026 = (5,400, 5,452)$$

Com um grau razoavelmente alto de confiança, podemos dizer que  $5,400 < \mu < 5,452$ . Esse intervalo é bastante restrito, devido ao pequeno valor da variabilidade do diâmetro do orifício ( $\sigma = 0,100$ ). ■

**Nível de Confiança, Precisão e Escolha do Tamanho da Amostra**

Por que seguir para um nível de confiança de 95% quando se pode chegar a um nível de 99%? Porque o preço pago pelo nível de confiança maior é um intervalo mais largo. Em virtude de o intervalo de 95% estender-se  $1,96 \cdot \sigma/\sqrt{n}$  a cada lado de  $\bar{x}$ , a amplitude do intervalo é  $2(1,96) \cdot \sigma/\sqrt{n} = 3,92 \cdot \sigma/\sqrt{n}$ . De maneira semelhante, a amplitude do intervalo de 99% é  $2(2,58) \cdot \sigma/\sqrt{n} = 5,16 \cdot \sigma/\sqrt{n}$ . Isto é, temos mais confiança no intervalo de 99% exatamente porque é mais largo. Assim, quanto maior o grau de confiança desejado, mais largo é o intervalo resultante. Na verdade, o único IC de 100% para  $\mu$  é  $(-\infty, \infty)$ , que não acrescenta grandes informações, pois, mesmo antes da amostragem, sabíamos que esse intervalo cobria  $\mu$ .

Se imaginarmos a amplitude do intervalo como a especificação de sua precisão ou acurácia, então seu nível de confiança (ou confiabilidade) estará inversamente relacionado à sua precisão. Uma estimativa altamente confiável do intervalo pode ser imprecisa, quando os pontos finais do intervalo estiverem muito distantes, enquanto um intervalo preciso pode exigir confiabilidade relativamente baixa. Dessa forma, não se pode dizer inequivocamente que um intervalo de 99% será preferível a um intervalo de 95%; o ganho na confiabilidade exige uma perda na precisão.

Uma estratégia interessante é especificar o nível de confiança desejado e a amplitude do intervalo e então determinar o tamanho necessário da amostra.

**Exemplo 7.4**

O monitoramento extensivo de um sistema de computador de compartilhamento de tempo sugeriu que o tempo de resposta a um comando de edição específico é normalmente distribuído com desvio padrão de 25 milissegundos. Um novo sistema operacional foi instalado e desejamos estimar o tempo de resposta médio real  $\mu$  do novo ambiente. Assumindo que os tempos de resposta ainda sejam normalmente distribuídos com  $\sigma = 25$ , que tamanho de amostra é necessário para garantir que o IC de 95% resultante tenha uma amplitude de (no máximo) 10? O tamanho da amostra  $n$  deve satisfazer

$$10 = 2 \cdot (1,96)(25/\sqrt{n})$$

O reajuste dessa equação fornece

$$\sqrt{n} = 2 \cdot (1,96)(25)/10 = 9,80$$

então

$$n = (9,80)^2 = 96,04$$

Uma vez que  $n$  deve ser um número inteiro, é necessário um tamanho de amostra de 97. ■

A fórmula geral do tamanho da amostra  $n$  necessária para garantir uma amplitude de intervalo  $w$  é obtida de  $w = 2 \cdot z_{\alpha/2} \cdot \sigma/\sqrt{n}$ , já que

$$n = \left( 2z_{\alpha/2} \cdot \frac{\sigma}{w} \right)^2$$

Quanto menor a amplitude desejada  $w$ , maior deve ser  $n$ . Além disso,  $n$  é uma função crescente de  $\sigma$  (maior variabilidade da população precisa de um tamanho de amostra maior) e do nível de confiança  $100(1 - \alpha)$  (à medida que  $\alpha$  diminui,  $z_{\alpha/2}$  aumenta).

A metade da amplitude  $1,96\sigma/\sqrt{n}$  do IC de 95% é chamada às vezes de **limite do erro de estimação** associado a um nível de confiança de 95%; isto é, com confiança de 95%, a estimativa pontual  $\bar{x}$  não estará mais distante de  $\mu$  do que este valor. Antes de obter os dados, o investigador pode desejar determinar o tamanho de uma amostra para o qual é obtido um valor específico do limite. Por exemplo: com  $\mu$  representando o consumo médio de combustível (milhas/galão) de todos os carros de um determinado tipo, o objetivo de uma investigação pode ser estimar  $\mu$  dentro de um intervalo de confiança de 95% de 1 milha/galão. Geralmente, se desejamos estimar  $\mu$  para um valor  $B$  (o limite especificado do erro de estimação) com confiança de  $100(1 - \alpha)\%$ , o tamanho da amostra necessário resulta da substituição de  $2/w$  por  $1/B$  na fórmula da caixa precedente.

## Derivando um Intervalo de Confiança

Sejam  $X_1, X_2, \dots, X_n$ , a amostra na qual o IC de um parâmetro  $\theta$  se baseará. Suponha que uma variável aleatória que satisfaça às duas propriedades a seguir possa ser encontrada:

1. A variável depende funcionalmente de  $X_1, \dots, X_n$  e  $\theta$ .
2. A distribuição de probabilidade da variável não depende de  $\theta$  ou de quaisquer outros parâmetros desconhecidos.

Seja  $h(X_1, X_2, \dots, X_n; \theta)$  essa variável aleatória. Por exemplo: se a distribuição da população for normal com  $\sigma$  conhecido e  $\theta = \mu$ , a variável  $h(X_1, \dots, X_n; \mu) = (\bar{X} - \mu)/(\sigma/\sqrt{n})$  satisfaz ambas as propriedades; sua forma funcional claramente depende de  $\mu$ , ainda que tenha distribuição de probabilidade normal padronizada, que não depende de  $\mu$ . Em geral, a forma da função  $h$  é sugerida examinando-se a distribuição de um estimador apropriado  $\theta$ .

Para qualquer  $\alpha$  entre 0 e 1, podem ser determinadas constantes  $a$  e  $b$  que satisfaçam a

$$P(a < h(X_1, \dots, X_n; \theta) < b) = 1 - \alpha \quad (7.6)$$

Devido à segunda propriedade,  $a$  e  $b$  não dependem de  $\theta$ . No exemplo normal,  $a = -z_{\alpha/2}$  e  $b = z_{\alpha/2}$ . Agora, suponha que as desigualdades em (7.6) possam ser manipuladas para isolar  $\theta$ , fornecendo a probabilidade equivalente

$$P(l(X_1, X_2, \dots, X_n) < \theta < u(X_1, X_2, \dots, X_n)) = 1 - \alpha$$

Então,  $l(x_1, x_2, \dots, x_n)$  e  $u(x_1, \dots, x_n)$  são os limites de confiança inferior e superior, respectivamente, de um IC de  $100(1 - \alpha)\%$ . No exemplo normal, vimos que  $l(X_1, \dots, X_n) = \bar{X} - z_{\alpha/2} \cdot \sigma/\sqrt{n}$  e  $u(X_1, \dots, X_n) = \bar{X} + z_{\alpha/2} \cdot \sigma/\sqrt{n}$ .

### Exemplo 7.5

Um modelo teórico sugere que o tempo de quebra de um líquido isolante entre eletrodos em uma voltagem específica tem distribuição exponencial com parâmetro  $\lambda$  (veja a Seção 4.4). Uma amostra aleatória de  $n = 10$  tempos de quebra produz os seguintes dados amostrais (em min):  $x_1 = 41,53$ ,  $x_2 = 18,73$ ,  $x_3 = 2,99$ ,  $x_4 = 30,34$ ,  $x_5 = 12,33$ ,  $x_6 = 117,52$ ,  $x_7 = 73,02$ ,  $x_8 = 223,63$ ,  $x_9 = 4,00$ ,  $x_{10} = 26,78$ . Um IC de 95% de  $\lambda$  e do tempo de quebra médio real são desejados.

Seja  $h(X_1, X_2, \dots, X_n; \lambda) = 2\lambda \sum X_i$ . Demonstra-se que essa variável aleatória possui uma distribuição de probabilidade chamada distribuição qui-quadrada com  $2n$  graus de liberdade (gl) ( $\nu = 2n$ , onde  $\nu$  é o parâmetro da distribuição qui-quadrada, conforme discutido na Seção 4.4). A Tabela A.7 do Apêndice ilustra uma curva de densidade qui-quadrada típica e tabula valores críticos que incluem as áreas de cauda especificadas. O número relevante de graus de liberdade neste caso é  $2(10) = 20$ . A linha  $\nu = 20$  da tabela mostra que 34,170 inclui a área sob a cauda superior 0,025 e 9,591 inclui a área sob a cauda inferior 0,025 (área sob a cauda superior 0,975). Assim, para  $n = 10$ ,

$$P(9,591 < 2\lambda \sum X_i < 34,170) = 0,95$$

Divisão por  $2\sum X_i$  isola  $\lambda$ , produzindo

$$P(9,591/(2\sum X_i) < \lambda < (34,170/(2\sum X_i)) = 0,95$$

O limite inferior do IC de 95% de  $\lambda$  é  $9,591/(2\sum x_i)$ , e o limite superior é  $34,170/(2\sum x_i)$ . Para os dados fornecidos,  $\sum x_i = 550,87$ , dando o intervalo (0,00871, 0,03101).

O valor esperado de uma va exponencial é  $\mu = 1/\lambda$ . Uma vez que

$$P(2\sum X_i/34,170 < 1/\lambda < 2\sum X_i/9,591) = 0,95$$

o IC de 95% do tempo de quebra médio real é  $(2\sum x_i/34,170, 2\sum x_i/9,591) = (32,24, 114,87)$ . Esse intervalo é evidentemente bastante largo, refletindo variabilidade substancial nos tempos de quebra e o tamanho de uma amostra pequena. ■

Em geral, os limites de confiança superior e inferior resultam da substituição de cada  $<$  em (7.6) por  $=$  e da resolução de  $\theta$ . No exemplo do fluido isolante considerado,  $2\lambda \sum x_i = 34,170$  fornece  $\lambda = 34,170/(2\sum x_i)$  como o limite de confiança superior e o limite inferior, obtidos da outra equação. Observe que os dois limites do intervalo não são equidistantes da estimativa pontual, uma vez que o intervalo não é da forma  $\hat{\theta} \pm c$ .

## Intervalos de Confiança *Bootstrap*

A técnica *bootstrap* foi introduzida no Capítulo 6 como uma forma de estimar  $\sigma_{\hat{\theta}}$ . Também pode ser aplicada para se obter um IC de  $\theta$ . Considere novamente a estimação da média  $\mu$  de uma distribuição normal quando  $\sigma$  é conhecido. Vamos substituir  $\mu$  por  $\theta$  e usar  $\hat{\theta} = \bar{X}$  como o estimador pontual. Observe que  $1,96\sigma/\sqrt{n}$  é o 97,5º percentil da distribuição de  $\hat{\theta} - \theta$  [isto é,  $P(\bar{X} - \mu < 1,96\sigma/\sqrt{n}) = P(Z < 1,96) = 0,9750$ ]. De forma semelhante,  $-1,96\sigma/\sqrt{n}$  é o 2,5º percentil, assim

$$\begin{aligned} 0,95 &= P(2,5^\circ \text{ percentil} < \hat{\theta} - \theta < 97,5^\circ \text{ percentil}) \\ &= P(\hat{\theta} - 2,5^\circ \text{ percentil} > \theta > \hat{\theta} - 97,5^\circ \text{ percentil}) \end{aligned}$$

Isto é, com

$$\begin{aligned} l &= \hat{\theta} - 97,5^\circ \text{ percentil de } \hat{\theta} - \theta \\ u &= \hat{\theta} - 2,5^\circ \text{ percentil de } \hat{\theta} - \theta \end{aligned} \tag{7.7}$$

o IC de  $\theta$  é  $(l, u)$ . Em muitos casos, os percentis em (7.7) não podem ser calculados, mas *podem* ser estimados a partir das amostras *bootstrap*. Suponha que obtenhamos  $B = 1000$  amostras *bootstrap* e calcule  $\hat{\theta}_1^*, \dots, \hat{\theta}_{1000}^*$ , e  $\bar{\theta}^*$  e, depois, as 1000 diferenças  $\hat{\theta}_1^* - \bar{\theta}^*, \dots, \hat{\theta}_{1000}^* - \bar{\theta}^*$ . A 25ª maior e a 25ª menor dessas diferenças são estimativas dos percentis desconhecidos em (7.7). Consulte os livros de Rice ou Efron, citados no Capítulo 6, para mais informações.

## Exercícios | Seção 7.1 (1–11)

1. Considere a distribuição de uma população normal com o valor de  $\sigma$  conhecido.
  - a. Qual é o nível de confiança do intervalo  $\bar{x} \pm 2,81\sigma/\sqrt{n}$ ?
  - b. Qual é o nível de confiança do intervalo  $\bar{x} \pm 1,44\sigma/\sqrt{n}$ ?
  - c. Que valor de  $z_{\alpha/2}$  na fórmula do IC (7.5) resulta em um nível de confiança de 99,7%?
  - d. Responda à pergunta proposta no item (c) para um nível de confiança de 75%.
2. Cada um dos intervalos a seguir é um intervalo de confiança de  $\mu =$  média real (isso é, média da população) da frequência de ressonância (Hz) de todas as raquetes de tênis de um determinado tipo:
 

(114,4, 115,6)      (114,1, 115,9)

  - a. Qual é o valor da frequência de ressonância da média amostral?
  - b. Ambos os intervalos foram calculados a partir dos mesmos dados amostrais. O nível de confiança de um desses intervalos é 90% e do outro é 99%. Qual dos intervalos possui o nível de confiança de 90% e por quê?
3. Suponha que uma amostra aleatória de 50 garrafas de uma marca específica de xarope para tosse seja selecionada e o teor alcoólico de cada garrafa seja determinado. Seja  $\mu$  o teor médio de álcool da população de todas as garrafas da marca em estudo. Suponha que o intervalo de confiança de 95% resultante seja (7,8, 9,4).
  - a. Um intervalo de confiança de 90% calculado dessa mesma amostra teria sido mais estreito ou mais largo que o intervalo mencionado acima? Explique seu raciocínio.
  - b. Considere a afirmação a seguir: existe 95% de chance de  $\mu$  estar entre 7,8 e 9,4. Essa afirmação está correta? Por quê?
  - c. Considere a afirmação a seguir: podemos estar certos de que 95% de todas as garrafas desse tipo de xarope têm um conteúdo alcoólico que está entre 7,8 e 9,4. A afirmação está correta? Por quê?
  - d. Considere a afirmação a seguir: se o processo de seleção de uma amostra de tamanho 50 e de cálculo do intervalo de 95% correspondente for repetido 100 vezes, 95 dos intervalos resultantes incluirão  $\mu$ . Essa afirmação está correta? Por quê?
4. Deseja-se um IC para a média real da perda de carga por dispersão  $\mu$  (watts) de um tipo de motor a indução, quando a corrente da linha é mantida em 10 amps para uma velocidade de 1500 rpm. Assuma que a perda de carga por dispersão seja normalmente distribuída com  $\sigma = 3,0$ .
  - a. Calcule um IC de 95% de  $\mu$  quando  $n = 25$  e  $\bar{x} = 58,3$ .
  - b. Calcule um IC de 95% de  $\mu$  quando  $n = 100$  e  $\bar{x} = 58,3$ .
  - c. Calcule um IC de 99% de  $\mu$  quando  $n = 100$  e  $\bar{x} = 58,3$ .
  - d. Calcule um IC de 82% de  $\mu$  quando  $n = 100$  e  $\bar{x} = 58,3$ .
  - e. Quão grande deve ser  $n$  se a largura do intervalo de 99% de  $\mu$  for 1,0?
5. Assuma que a porosidade do hélio (em porcentagem) das amostras de carvão tiradas de qualquer junta específica seja normalmente distribuída com desvio padrão real de 0,75.
  - a. Calcule um IC de 95% da porosidade média real de uma junta, caso a porosidade média de 20 de seus espécimes seja 4,85.
  - b. Calcule um IC de 98% da porosidade média real de outra junta com base nos 16 espécimes com média amostral de porosidade de 4,56.
  - c. Quão grande o tamanho de uma amostra deve ser se a amplitude do intervalo de 95% for 0,40?
  - d. Que tamanho de amostra é necessário para estimar a porosidade média real dentro de 0,2 com confiança de 99%?
6. Com base em testes extensivos, o ponto de escoamento de um tipo específico de barra de aço reforçado é conhecido por ser normalmente distribuído com  $\sigma = 100$ . A composição da barra foi levemente modificada, mas acredita-se que essa modificação não tenha afetado nem a normalidade e nem o valor de  $\sigma$ .
  - a. Assumindo que esse seja o caso, se uma amostra de 25 barras modificadas tiver resultado em um ponto de escoamento médio amostral de 8439 lb, calcule um IC de 90% para o ponto de escoamento médio real da barra modificada.
  - b. Como você modificaria o intervalo no item (a) para obter um nível de confiança de 92%?
7. Em quanto o tamanho da amostra  $n$  deve ser aumentado se a amplitude do IC (7.5) for reduzida à metade? Se o tamanho da amostra for aumentado por um fator de 25, que efeito terá sobre a amplitude do intervalo? Justifique suas afirmações.
8. Seja  $\alpha_1 > 0$ ,  $\alpha_2 > 0$ , com  $\alpha_1 + \alpha_2 = \alpha$ . Então,
 
$$P\left(-z_{\alpha_1} < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < z_{\alpha_2}\right) = 1 - \alpha$$
  - a. Use essa equação para deduzir uma expressão mais geral para um IC de  $100(1 - \alpha)\%$  de  $\mu$  do qual o intervalo (7.5) é um caso especial.
  - b. Seja  $\alpha = 0,05$  e  $\alpha_1 = \alpha/4$ ,  $\alpha_2 = 3\alpha/4$ . Essa condição resulta em um intervalo mais estreito ou mais largo que o intervalo (7.5)?

9. a. Sob as mesmas condições que as que levam ao intervalo (7.5),  $P[(\bar{X} - \mu)/(\sigma/\sqrt{n}) < 1,645] = 0,95$ . Use isso para deduzir um intervalo unilateral de  $\mu$  que possui amplitude infinita e fornece um limite de confiança inferior em  $\mu$ . Que intervalo é esse para os dados do Exercício 5(a)?
- b. Generalize o resultado do item (a) para obter um limite inferior com nível de confiança  $100(1 - \alpha)\%$ .
- c. Qual é o intervalo análogo ao do item (b) que fornece um limite superior em  $\mu$ ? Calcule esse intervalo de 99% para os dados do Exercício 4(a).
10. Uma amostra aleatória de  $n = 15$  determinados tipos de bombas de aquecimento produziu as seguintes observações sobre a vida útil (em anos):
- |      |     |     |     |      |     |     |     |
|------|-----|-----|-----|------|-----|-----|-----|
| 2,0  | 1,3 | 6,0 | 1,9 | 5,1  | 0,4 | 1,0 | 5,3 |
| 15,7 | 0,7 | 4,8 | 0,9 | 12,2 | 5,3 | 0,6 |     |
- a. Assuma que a distribuição da vida útil seja exponencial e use um argumento paralelo ao do Exemplo 7.5 para obter um IC de 95% da vida útil esperada (médio real).
- b. Como o intervalo do item (a) deve ser alterado para chegar a um nível de confiança de 99%?
- c. O que é o IC de 95% para o desvio padrão da distribuição da vida útil? (*Sugestão*: qual é o desvio padrão de uma variável aleatória exponencial?)
11. Considere os próximos 1000 ICs de 95% de  $\mu$  que um consultor estatístico obterá para vários clientes. Suponha que os conjuntos de dados em que os intervalos se baseiam sejam selecionados independentemente um do outro. Quantos desses 1000 intervalos capturam os valores correspondentes de  $\mu$ ? Qual é a probabilidade, de que 940 e 960 desses intervalos contenham o valor correspondente de  $\mu$ ? (*Sugestão*: seja  $Y =$  número de intervalos entre os 1000 que contém  $\mu$ . Que tipo de variável aleatória é  $Y$ ?)

## 7.2 Intervalos de Confiança para Amostras Grandes para uma Média e Proporção da População

O IC de  $\mu$  dado na seção anterior assumiu que a distribuição da população é normal e que o valor de  $\sigma$  é conhecido. Apresentamos agora o IC de uma amostra grande, cuja validade não exige essas hipóteses. Depois de mostrar como o argumento que leva a esse intervalo é generalizado para produzir outros de amostra grande, abordamos o intervalo da proporção de uma população  $p$ .

### Um Intervalo de Amostra Grande para $\mu$

Sejam  $X_1, X_2, \dots, X_n$  uma amostra aleatória de uma população que tem média  $\mu$  e desvio padrão  $\sigma$ . Desde que  $n$  seja grande, o Teorema do Limite Central (TLC) implica que  $\bar{X}$  tem aproximadamente distribuição normal, qualquer que seja a natureza da distribuição da população. Segue então que  $Z = (\bar{X} - \mu)/(\sigma/\sqrt{n})$  possui aproximadamente distribuição normal padrão, de modo que

$$P\left(-z_{\alpha/2} < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < z_{\alpha/2}\right) \approx 1 - \alpha$$

Um argumento paralelo ao fornecido na Seção 7.1 produz  $\bar{x} \pm z_{\alpha/2} \cdot \sigma/\sqrt{n}$  como IC de amostra grande de  $\mu$  com nível de confiança de *aproximadamente*  $100(1 - \alpha)\%$ . Isto é, quando  $n$  é grande, o IC de  $\mu$  fornecido anteriormente permanece válido, qualquer que seja a distribuição da população, contanto que o adjetivo “aproximadamente” se insira na frente do nível de confiança.

Uma dificuldade prática com esse desenvolvimento é que o cálculo do intervalo exige o valor de  $\sigma$ , que quase nunca será conhecido. Considere a variável padronizada

$$Z = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

em que o desvio padrão amostral  $S$  substitui  $\sigma$ . Havia anteriormente aleatoriedade somente no numerador de  $Z$  (em virtude de  $\bar{X}$ ). Agora, existe aleatoriedade no numerador e no denominador – os valores de  $\bar{X}$  e  $S$  variam de amostra para amostra. Entretanto, quando  $n$  é grande, o uso de  $S$  em vez de  $\sigma$  acrescenta pouca variabilidade extra a  $Z$ . Mais especificamente, nesse caso, o novo  $Z$  também possui aproximadamente uma distribuição normal padronizada. A manipulação das desigualdades para probabilidades que utilizam esse novo  $Z$  produz um intervalo geral de amostras grandes para  $\mu$ .

## PROPOSIÇÃO

Se  $n$  é suficientemente grande, a variável padronizada

$$Z = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

possui aproximadamente uma distribuição normal padronizada. Isso implica que

$$\bar{x} \pm z_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \quad (7.8)$$

é um **intervalo de confiança de amostra grande para  $\mu$**  com nível de confiança de aproximadamente  $100(1 - \alpha)\%$ . Essa fórmula é válida independentemente do formato da distribuição da população.

Geralmente,  $n > 40$  será suficiente para justificar o uso desse intervalo. Essa afirmação é mais conservadora que a regra prática para o TLC, devido à variabilidade adicional introduzida, usando-se  $S$  em vez de  $\sigma$ .

## Exemplo 7.6

A voltagem de quebra da Corrente Alternada (CA) de um líquido isolante indica sua resistência dielétrica. O artigo “Testing Practices for the AC Breakdown Voltage Testing of Insulation Liquids” (*IEEE Electrical Insulation Magazine*, 1995, p. 21-26) forneceu as seguintes observações amostrais da voltagem de quebra (kV) de um circuito específico sob certas condições.

62 50 53 57 41 53 55 61 59 64 50 53 64 62 50 68  
 54 55 57 50 55 50 56 55 46 55 53 54 52 47 47 55  
 57 48 63 57 57 55 53 59 53 52 50 55 60 50 56 58

Um *boxplot* dos dados (Figura 7.5) mostra alta concentração central dos dados (largura estreita da caixa). Há um único *outlier* na extremidade superior, mas esse valor está realmente um pouco mais próximo da mediana (55) que a observação da amostra menor.

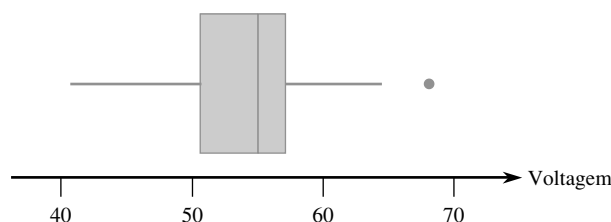


Figura 7.5 Um *boxplot* para os dados da voltagem de quebra do Exemplo 7.6

As quantidades resumidas incluem  $n = 48$ ,  $\sum x_i = 2626$  e  $\sum x_i^2 = 144,950$ , pelas quais  $\bar{x} = 54,7$  e  $s = 5,23$ . O intervalo de confiança de 95% é então

$$54,7 \pm 1,96 \frac{5,23}{\sqrt{48}} = 54,7 \pm 1,5 = (53,2, 56,2)$$

Isto é,

$$53,2 < \mu < 56,2$$

com um nível de confiança de aproximadamente 95%. O intervalo é razoavelmente estreito, indicando que estimamos  $\mu$  com precisão. ■

Infelizmente, a escolha do tamanho da amostra para produzir a largura desejada de um intervalo não é tão direta aqui como foi para o caso de sigma conhecido. Isso se deve ao fato de a largura de (7.8) ser  $2z_{\alpha/2}s/\sqrt{n}$ . Uma vez que o valor de  $s$  não está disponível antes da coleta dos dados, a amplitude do intervalo não pode ser determinada somente pela escolha de  $n$ . A única opção para um investigador que quer especificar uma amplitude desejada é fazer uma estimativa sensata sobre qual pode ser o valor de  $s$ . Sendo conservador e estimando um valor maior de  $s$ , será escolhido um  $n$  maior do que o necessário. O investigador pode ser capaz de especificar um valor razoavelmente preciso para a amplitude da população (a diferença entre os valores menor e maior). Então, se a distribuição da população não estiver muito distorcida, dividir a amplitude por 4 proporciona um valor aproximado de  $s$ .

### Exemplo 7.7

Consulte o Exemplo 7.6 sobre a voltagem de quebra. Suponha que o investigador acredite que, virtualmente, todos os valores da população estejam entre 40 e 70. Então,  $(70 - 40)/4 = 7,5$  é um valor razoável de  $s$ . O tamanho apropriado da amostra para estimar a média real da voltagem de quebra para dentro de 1 kV com nível de confiança de 95% — isto é, para o IC de 95% ter amplitude de 2 kV — é

$$n = [(1,96)(7,5)/1]^2 \approx 217 \quad \blacksquare$$

## Um Intervalo de Confiança Geral para Amostra Grande

Os intervalos de amostra grande  $\bar{x} \pm z_{\alpha/2} \cdot \sigma/\sqrt{n}$  e  $\bar{x} \pm z_{\alpha/2} \cdot s/\sqrt{n}$  são casos especiais de IC geral para amostra grande para um parâmetro  $\theta$ . Suponha que  $\hat{\theta}$  seja um estimador que satisfaça às seguintes propriedades: (1) Possui distribuição aproximadamente normal; (2) é (pelo menos aproximadamente) não-tendencioso; e (3) está disponível uma expressão de  $\sigma_{\hat{\theta}}$ , o desvio padrão de  $\hat{\theta}$ . Por exemplo: no caso  $\theta = \mu$ ,  $\hat{\mu} = \bar{X}$  é um estimador não-tendencioso cuja distribuição é aproximadamente normal quando  $n$  é grande e  $\sigma_{\hat{\mu}} = \sigma_{\bar{X}} = \sigma/\sqrt{n}$ . Padronizar  $\hat{\theta}$  produz a va  $Z = (\hat{\theta} - \theta)/\sigma_{\hat{\theta}}$  que possui distribuição aproximadamente normal padronizada. Isso justifica a declaração de probabilidade.

$$P\left(-z_{\alpha/2} < \frac{\hat{\theta} - \theta}{\sigma_{\hat{\theta}}} < z_{\alpha/2}\right) \approx 1 - \alpha \quad (7.9)$$

Suponha, primeiro, que  $\sigma_{\hat{\theta}}$  não utilize quaisquer parâmetros desconhecidos (por exemplo:  $\sigma$  conhecido no caso  $\theta = \mu$ ). Então, substituir cada  $<$  em (7.9) por  $=$  resulta em  $\theta = \hat{\theta} \pm z_{\alpha/2} \cdot \sigma_{\hat{\theta}}$ , de modo que os limites de confiança inferior e superior são  $\hat{\theta} - z_{\alpha/2} \cdot \sigma_{\hat{\theta}}$  e  $\hat{\theta} + z_{\alpha/2} \cdot \sigma_{\hat{\theta}}$ , respectivamente. Agora, suponha que  $\sigma_{\hat{\theta}}$  não utilize  $\theta$ , mas que use pelo menos outro parâmetro desconhecido. Considere que  $s_{\hat{\theta}}$  seja a estimativa de  $\sigma_{\hat{\theta}}$  obtida usando-se estimativas em vez de parâmetros desconhecidos (por exemplo:  $s/\sqrt{n}$  estima  $\sigma/\sqrt{n}$ ). Sob condições gerais (essencialmente que  $s_{\hat{\theta}}$  esteja próximo de  $\sigma_{\hat{\theta}}$  para a maioria das amostras), um IC válido é  $\hat{\theta} \pm z_{\alpha/2} \cdot s_{\hat{\theta}}$ . O intervalo  $\bar{x} \pm z_{\alpha/2} \cdot s/\sqrt{n}$  é um exemplo.

Finalmente, suponha que  $\sigma_{\hat{\theta}}$  utilize o  $\theta$  desconhecido. Esse é o caso, por exemplo, quando  $\theta = p$ , uma proporção da população. Então,  $(\hat{\theta} - \theta)/\sigma_{\hat{\theta}} = z_{\alpha/2}$  pode ser difícil de resolver. Uma solução aproximada geralmente pode ser obtida substituindo-se  $\theta$  em  $\sigma_{\hat{\theta}}$  por sua estimativa  $\hat{\theta}$ . Resulta, pois, um desvio padrão estimado  $s_{\hat{\theta}}$  e o intervalo correspondente é novamente  $\hat{\theta} \pm z_{\alpha/2} \cdot s_{\hat{\theta}}$ .

## Um Intervalo de Confiança para uma Proporção da População

Seja  $p$  a proporção de “sucessos” de uma população em que *sucesso* identifica um indivíduo ou objeto que tenha uma propriedade especificada. Uma amostra aleatória de  $n$  indivíduos será selecionada e  $X$  é o número de sucessos na amostra. Contanto que  $n$  seja pequeno, em comparação ao tamanho da população,  $X$  pode ser considerado uma va binomial com  $E(X) = np$  e  $\sigma_X = \sqrt{np(1-p)}$ . Além disso, se  $n$  for grande ( $np \geq 10$  e  $nq \geq 10$ ),  $X$  possui distribuição aproximadamente normal.

O estimador natural de  $p$  é  $\hat{p} = X/n$ , a fração de sucessos da amostra. Uma vez que  $\hat{p}$  é apenas  $X$  multiplicado pela constante  $1/n$ ,  $\hat{p}$  também tem distribuição aproximadamente normal. Como mostrado na Seção 6.1,

$E(\hat{p}) = p$  (não-tendencioso) e  $\sigma_{\hat{p}} = \sqrt{p(1-p)/n}$ . O desvio padrão  $\sigma_{\hat{p}}$  utiliza o parâmetro desconhecido  $p$ . Padronizar  $\hat{p}$ , pela subtração de  $p$  e divisão por  $\sigma_{\hat{p}}$ , implica, então, que

$$P\left(-z_{\alpha/2} < \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} < z_{\alpha/2}\right) \approx 1 - \alpha$$

Prosseguindo conforme foi sugerido na subseção “Derivando um Intervalo de Confiança” (Seção 7.1), os limites de confiança resultam da substituição de cada  $<$  por  $=$  e resolvendo a equação quadrática resultante de  $p$ . Obtêm-se as duas raízes

$$p = \frac{\hat{p} + \frac{z_{\alpha/2}^2}{2n} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n} + \frac{z_{\alpha/2}^2}{4n^2}}}{1 + (z_{\alpha/2}^2)/n}$$

#### PROPOSIÇÃO

O intervalo de confiança para uma proporção da população  $p$  com nível de confiança de aproximadamente  $100(1 - \alpha)\%$  tem

$$\text{limite de confiança inferior} = \frac{\hat{p} + \frac{z_{\alpha/2}^2}{2n} - z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n} + \frac{z_{\alpha/2}^2}{4n^2}}}{1 + (z_{\alpha/2}^2)/n}$$

e

$$\text{limite de confiança superior} = \frac{\hat{p} + \frac{z_{\alpha/2}^2}{2n} + z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n} + \frac{z_{\alpha/2}^2}{4n^2}}}{1 + (z_{\alpha/2}^2)/n}$$

(7.10)

Se o tamanho da amostra for muito grande,  $z^2/(2n)$  é desprezível, comparado a  $\hat{p}$ ,  $z^2/(4n^2)$  sob a raiz quadrada é desprezível se comparado a  $\hat{p}\hat{q}/n$ , e  $z^2/n$  é desprezível se comparado a 1. Desconsiderando esses termos desprezíveis obtemos os limites de confiança aproximados

$$\hat{p} \pm z_{\alpha/2} \sqrt{\hat{p}\hat{q}/n} \quad (7.11)$$

Essa expressão tem a forma geral  $\hat{\theta} \pm z_{\alpha/2} \hat{\sigma}_{\hat{\theta}}$  de um intervalo de amostra grande sugerido na última subseção. Durante décadas, o intervalo acima foi recomendado, desde que a aproximação normal de  $\hat{p}$  fosse justificada. Entretanto uma pesquisa recente mostrou que o intervalo um tanto mais complicado dado na proposição tem um nível de confiança real que tende a estar mais próximo do nível nominal do que o intervalo tradicional (Agresti, Alan e Coull, “Approximate Is Better Than ‘Exact’ for Interval Estimation of a Binomial Proportion,” *The American Statistician*, 1998, p. 119-126). Isto é, se  $z_{\alpha/2} = 1,96$  for usado, o nível de confiança do “novo” intervalo tende a estar mais próximo de 95% para quase todos os valores de  $p$  que é o caso do intervalo tradicional; isto também acontece com outros níveis de confiança. Além disso, Agresti e Coull afirmam que o intervalo “pode ser recomendado para uso em aproximadamente todos os tamanhos de amostra e valores de parâmetro”, de modo que as condições  $n\hat{p} \geq 10$  e  $n\hat{q} \geq 10$  não precisem ser verificadas.

#### Exemplo 7.8

O artigo “Repeatability and Reproducibility for Pass/Fail Data” (*J. of Testing and Eval.*, 1997, p. 151-153) relatou que em  $n = 48$  tentativas em um laboratório específico, 16 resultaram em ignição de um tipo específico de substrato por um cigarro aceso. Seja  $p$  a proporção no longo prazo de todas as tentativas que resultariam em ignição. A estimativa pontual de  $p$  é  $\hat{p} = 16/48 = 0,333$ . O intervalo de confiança de  $p$  com nível de confiança de aproximadamente 95% é



$$\frac{0,333 + (1,96)^2/96 \pm 1,96\sqrt{(0,333)(0,667)/48 + (1,96)^2/9216}}{1 + (1,96)^2/48} = \frac{0,373 \pm 0,139}{1,08} = (0,217, 0,474)$$

O intervalo tradicional é

$$0,333 \pm 1,96\sqrt{(0,333)(0,667)/48} = 0,333 \pm 1,33 = (0,200, 0,466)$$

Esses dois intervalos estariam muito mais próximos entre si se o tamanho da amostra fosse substancialmente maior. ■

Igualar a amplitude do IC de  $p$  a uma amplitude pré-especificada  $w$  fornece uma equação quadrática para o tamanho da amostra  $n$  necessário para fornecer um intervalo com grau de precisão desejado. Eliminando-se o índice de  $z_{\alpha/2}$ , a solução é

$$n = \frac{2z^2\hat{p}\hat{q} - z^2w^2 \pm \sqrt{4z^4\hat{p}\hat{q}(\hat{p}\hat{q} - w^2) + w^2z^4}}{w^2} \quad (7.12)$$

Desprezando os termos do numerador que envolvem  $w^2$ , temos

$$n \approx \frac{4z^2\hat{p}\hat{q}}{w^2}$$

Essa última expressão é o resultado da igualdade da amplitude do intervalo tradicional com  $w$ .

Essas fórmulas infelizmente envolvem  $\hat{p}$  desconhecido. A abordagem mais conservadora é tirar proveito do fato de  $\hat{p}\hat{q}$  [=  $\hat{p}(1 - \hat{p})$ ] ser um máximo quando  $\hat{p} = 0,5$ . Assim, se  $\hat{p} = \hat{q} = 0,5$  for usado em (7.12), a amplitude será no máximo  $w$ , independentemente do valor de  $\hat{p}$  que resulte da amostra. De maneira alternativa, se o investigador acredita fortemente, com base em informações anteriores, que  $p \leq p_0 \leq 0,5$ , então  $p_0$  pode ser usado no lugar de  $\hat{p}$ . Um comentário semelhante se aplica quando  $p \geq p_0 \geq 0,5$ .

### Exemplo 7.9

A amplitude do IC de 95% do Exemplo 7.8 é 0,257. O valor de  $n$  necessário para garantir uma amplitude de 0,10 sem considerar o valor de  $\hat{p}$  é

$$n = \frac{2(1,96)^2(0,25) - (1,96)^2(0,01) \pm \sqrt{4(1,96)^4(0,25)(0,25 - 0,01) + (0,01)(1,96)^4}}{0,01} = 380,3$$

Dessa forma, deve ser usado um tamanho de amostra de 381. A expressão de  $n$  com base no IC tradicional fornece um valor levemente maior de 385. ■

## Intervalos de Confiança Monocaudais (Limites de Confiança)

Os intervalos de confiança discutidos até agora fornecem tanto um limite de confiança inferior como um limite de confiança superior para o parâmetro que está sendo estimado. Em algumas circunstâncias, o investigador vai querer somente um desses dois tipos de limites. Por exemplo: um psicólogo pode querer calcular o limite de confiança superior de 95% para o tempo de reação médio real a um estímulo específico, ou um engenheiro de confiabilidade pode querer somente o limite de confiança inferior da vida útil média real de um determinado tipo de componente. Em virtude de a área acumulada sob a curva normal padrão à esquerda de 1,645 ser 0,95,

$$P\left(\frac{\bar{X} - \mu}{S/\sqrt{n}} < 1,645\right) \approx 0,95$$

Manipular a desigualdade dentro dos parênteses para isolar  $\mu$  em um lado e substituir as  $s$  pelos valores calculados fornece a desigualdade  $\mu > \bar{x} - 1,645s/\sqrt{n}$ ; a expressão à direita é o limite de confiança inferior desejado. Começar com  $P(-1,645 < Z) \approx 0,95$  e manipular a desigualdade resulta no limite de confiança superior. Um argumento semelhante fornece o limite monocaudal associado a qualquer outro nível de confiança.

## PROPOSIÇÃO

O limite de confiança superior de amostra grande para  $\mu$  é

$$\mu < \bar{x} + z_{\alpha} \cdot \frac{s}{\sqrt{n}}$$

e o limite de confiança inferior de amostra grande para  $\mu$  é

$$\mu > \bar{x} - z_{\alpha} \cdot \frac{s}{\sqrt{n}}$$

O limite de confiança monocaudal para  $p$  resulta da substituição de  $z_{\alpha/2}$  por  $z_{\alpha}$  e  $\pm$  por  $+$  ou  $-$  na fórmula do IC (7.10) para  $p$ . Em todos os casos, o nível de confiança é de aproximadamente  $100(1 - \alpha)\%$ .

## Exemplo 7.10

O teste de corte inclinado é o procedimento mais amplamente aceito para avaliar a qualidade de uma ligação entre um material de conserto e seu substrato de concreto. O artigo “Testing the Bond Between Repair Materials and Concrete Substrate” (*ACI Materials J.*, 1996, p. 553-558) relatou que, em uma investigação específica, uma amostra de 48 observações de resistência de corte forneceu uma resistência média amostral de  $17,17 \text{ N/mm}^2$  e um desvio padrão da amostra de  $3,28 \text{ N/mm}^2$ . O limite de confiança inferior para a resistência de corte média real  $\mu$  com nível de confiança de 95% é

$$17,17 - (1,645) \frac{(3,28)}{\sqrt{48}} = 17,17 - 0,78 = 16,39$$

Isto é, com um nível de confiança de 95%, o valor de  $\mu$  está no intervalo  $(16,39, \infty)$ . ■

## Exercícios | Seção 7.2 (12–27)

12. Uma amostra aleatória de 110 relâmpagos, em certa região, resultou em uma duração média do eco do radar da amostra de 0,81 s. e um desvio padrão da amostra de 0,34 s. (“Lightning Strikes to an Airplane in a Thunderstorm,” *J. of Aircraft*, 1984, p. 607-611). Calcule o intervalo de confiança (bicaudal) de 99% para a duração média real do eco  $\mu$ , e interprete o intervalo resultante.
13. O artigo “Extravisual Damage Detection? Defining the Standard Normal Tree” (*Photogrammetric Engr. and Remote Sensing*, 1981, p. 515-522) discute o uso de fotografia infravermelha colorida na identificação de árvores normais nos grupos de pinheiros Douglas. Dentre os dados relatados, estava a estatística resumida das medições densitométricas ópticas analíticas de filtro verde nas amostras de árvores doentes e saudáveis. Para uma amostra de 69 árvores saudáveis, a densidade da camada colorida média amostral era 1,028, e o desvio padrão da amostra, 0,163.
  - a. Calcule um IC (bicaudal) de 95% para a densidade da camada colorida média real de todas as árvores.
  - b. Suponha que os investigadores tenham feito uma adivinhação grosseira de 0,16 para o valor de  $s$  antes de coletar os dados. Que tamanho de amostra seria necessário para obter uma amplitude do intervalo de 0,05, para um nível de confiança de 95%?
14. O artigo “Evaluating Tunnel Kiln Performance” (*Amer. Ceramic Soc. Bull.*, ago. 1997, p. 59-63) forneceu as seguintes informações resumidas das resistências a fraturas (Mpa) de  $n = 169$  barras cerâmicas cozidas em um determinado forno:  $\bar{x} = 89,10$ ,  $s = 3,73$ .

- a. Calcule o intervalo de confiança (bicaudal) para a resistência à fratura média real, usando um nível de confiança de 95%. Parece que a resistência à fratura média real foi estimada com precisão?
- b. Suponha que os investigadores acreditaram *a priori* que o desvio padrão da população era cerca de 4 MPa. Com base nessa suposição, quão grande deve ser uma amostra para calcular  $\mu$  dentro de 0,5 MPa com confiança de 95%?
15. Determine o nível de confiança para cada um dos limites de confiança monocaudais de amostras grandes a seguir:
- a. Limite superior:  $\bar{x} + 0,84s/\sqrt{n}$
- b. Limite inferior:  $\bar{x} - 2,05s/\sqrt{n}$
- c. Limite superior:  $\bar{x} + 0,67s/\sqrt{n}$
16. O tempo entre a carga e o final do processo (min) de um aço carbono em um tipo de fornalha aberta foi determinado para cada aquecimento em uma amostra de tamanho 46, resultando um tempo médio amostral de 382,1 e um desvio padrão da amostra de 31,5. Calcule o limite de confiança superior de 95% do tempo médio real entre a carga e o final do processo.
17. O Exercício 1.13 forneceu uma amostra de observações da resistência extrema à tensão (ksi). Use a saída da estatística descritiva do MINITAB a seguir para calcular o limite de confiança inferior de 99% para a resistência extrema à tensão média real e interpretar o resultado.
- | N      | Média  | Mediana | Média Ap | Desv Padrão | Média SE |
|--------|--------|---------|----------|-------------|----------|
| 153    | 135,39 | 135,40  | 135,41   | 4,59        | 0,37     |
| Mínimo | Máximo | Q1      | Q3       |             |          |
| 122,20 | 147,70 | 132,95  | 138,25   |             |          |
18. O artigo “Ultimate Load Capacities of Expansion Anchor Bolts” (*J. of Energy Engr.*, 1993, p. 139-158) forneceu os seguintes dados resumidos sobre a resistência de corte (kip) de uma amostra de pinos de ferro de 3/8-pol.:  $n = 78$ ,  $\bar{x} = 4,25$ ,  $s = 1,30$ . Calcule o limite de confiança inferior, usando um nível de confiança de 90% para a resistência de corte média real.
19. O artigo “Limited Yield Estimation for Visual Defect Sources” (*IEEE Trans. on Semiconductor Manuf.*, 1997, p. 17-23) relatou que, em um estudo do processo de inspeção de um *wafers* específico, 356 dados foram examinados por uma inspeção e 201 deles foram aprovados. Assumindo um processo estável, calcule o intervalo de confiança (bicaudal) de 95% para a proporção de todos os dados que foram aprovados pela inspeção.
20. A Associated Press (9 de outubro de 2002) relatou que em uma inspeção de 4722 jovens americanos, com idade entre 6 e 19 anos, 15% estavam seriamente acima do peso (índice de massa corporal de pelo menos 30; esse índice é uma medida do peso em relação à altura). Calcule e interprete um intervalo de confiança, usando um nível de confiança de 99% para a proporção de todos os jovens americanos que estão seriamente acima do peso.
21. Uma amostra aleatória de 539 lares de certa cidade do Meio-Oeste foi selecionada e determinou-se que em 133 deles havia pelo menos uma arma de fogo (“The Social Determinants of Gun Ownership: Self-Protection in an Urban Environment”, *Criminology*, 1997, p. 629-640). Usando um nível de confiança de 95%, calcule o limite de confiança inferior para a proporção de todos os lares dessa cidade que possuam no mínimo uma arma de fogo.
22. Uma amostra aleatória de 487 mulheres não-fumantes de peso normal (índice de massa corporal entre 19,8 e 26,0) que deram à luz em um grande centro médico metropolitano foi selecionada (“The Effects of Cigarette Smoking and Gestational Weight Change on Birth Outcomes in Obese and Normal-Weight Women”, *Amer. J. of Public Health*, 1997, p. 591-596). Determinou-se que 7,2% desses partos resultaram em crianças com baixo peso (menos de 2500 g). Calcule o limite de confiança superior, usando um nível de confiança de 99% para a proporção de todos os nascimentos que resultaram em crianças com baixo peso.
23. O artigo “An Evaluation of Football Helmets Under Impact Conditions” (*Amer. J. Sports Medicine*, 1984, p. 233-237) relata que, quando cada capacete de futebol americano, em uma amostra aleatória de 37 capacetes do tipo suspensão, foi submetida a certo teste de impacto, 24 mostraram estar com defeito. Seja  $p$  a proporção de todos os capacetes desse tipo que mostrariam algum dano quando testados da maneira prescrita.
- a. Calcule o IC de 99% para  $p$ .
- b. Que tamanho de amostra seria necessário para a amplitude de um IC de 99% ser no máximo de 0,10 sem considerar  $\hat{p}$ ?
24. Uma de 56 amostras de algodão de pesquisa resultou em uma porcentagem de alongamento médio da amostra de 8,17 e desvio padrão amostral de 1,42 (“An Apparent Relation Between the Spiral Angle  $\phi$ , the Percent Elongation  $E_1$ , and the Dimensions of the Cotton Fiber”, *Textile Research J.*, 1978, p. 407-410). Calcule o IC de amostra grande de 95% para a porcentagem de alongamento médio real  $\mu$ . Que suposições você está fazendo com relação à distribuição da porcentagem de alongamento?
25. Um deputado estadual deseja pesquisar os residentes de seu distrito para ver que proporção do eleitorado está ciente de sua posição sobre o uso de fundos do Estado para pagar abortos.
- a. Que tamanho de amostra é necessário, se o IC de 95% de  $p$  tiver amplitude de no máximo 0,10 sem considerar  $p$ ?
- b. Se o deputado tem forte razão para acreditar que pelo menos  $\frac{2}{3}$  do eleitorado sabe de sua posição, quão grande você recomendaria que o tamanho da amostra fosse?
26. O superintendente do distrito de uma grande escola, que fez um curso de probabilidade e estatística, acredita

que o número de professores ausentes em um dia qualquer possua distribuição de Poisson com parâmetro  $\lambda$ . Use os dados a seguir sobre as ausências durante 50 dias para deduzir o IC de amostra grande para  $\lambda$ . [Sugestão: a média e a variância de uma variável de Poisson são ambos iguais a  $\lambda$ , assim

$$Z = \frac{\bar{X} - \lambda}{\sqrt{\lambda/n}}$$

possui aproximadamente distribuição normal padronizada. Prossiga como na derivação do intervalo de  $p$ , definindo a situação (com probabilidade  $1 - \alpha$ ) e resolvendo as desigualdades resultantes de  $\lambda$  (veja a explicação abaixo de (7.10)).]

Número de ausentes	0	1	2	3	4	5	6	7	8	9	10
Frequência	1	4	8	10	8	7	5	3	2	1	1

27. Reconsidere o IC (7.10) para  $p$  e enfoque um nível de confiança de 95%. Mostre que os limites de confiança estão de acordo com os do intervalo tradicional (7.11), uma vez que dois sucessos e duas falhas foram adicionados à amostra [isto é, (7.11) com base em  $x + 2$  S's em  $n + 4$  tentativas]. (Sugestão:  $1,96 \approx 2$ . Nota: Agresti e Coull mostraram que esse ajuste do intervalo tradicional também possui nível de confiança real próximo ao nível nominal.)

### 7.3 Intervalos Baseados em uma Distribuição Normal da População

O IC de  $\mu$  apresentado na Seção 7.2 é válido, desde que  $n$  seja grande. O intervalo resultante é usado, qualquer que seja a natureza da distribuição da população. O TLC não poderá ser usado, entretanto, quando  $n$  for pequeno. Nesse caso, uma maneira de continuar é fazer uma suposição específica sobre a forma da distribuição da população e deduzir um IC feito sob medida a essa suposição. Por exemplo: poderíamos desenvolver um IC de  $\mu$  quando a população for descrita por uma distribuição gama, outro intervalo para o caso de uma população de Weibull, e assim por diante. Os estatísticos realmente fizeram esse programa para um número de famílias de distribuição diferentes. Em virtude de a distribuição normal ser apropriada mais frequentemente como um modelo da população que qualquer outro tipo de distribuição, enfocaremos aqui um IC para essa situação.

#### SUPosição

A população de interesse é normal, de modo que  $X_1, \dots, X_n$  constitui uma amostra aleatória de uma distribuição normal com  $\mu$  e  $\sigma$  desconhecidos.

O resultado-chave subordinado ao intervalo na Seção 7.2 foi que, para  $n$  grande, a va  $Z = (\bar{X} - \mu)/(S/\sqrt{n})$  possui aproximadamente uma distribuição normal padronizada. Quando  $n$  é pequeno,  $S$  provavelmente não está mais próximo de  $\sigma$ , de modo que a variabilidade na distribuição de  $Z$  surge da aleatoriedade no numerador e no denominador. Isso implica que a distribuição de probabilidade de  $(\bar{X} - \mu)/(S/\sqrt{n})$  será mais dispersa que a distribuição normal padronizada. O resultado no qual as inferências se baseiam introduz uma nova família de distribuições de probabilidade chamada família de *distribuições t*.

#### TEOREMA

Quando  $\bar{X}$  é a média amostral aleatória de tamanho  $n$  de uma distribuição normal com média  $\mu$ , a va

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \quad (7.13)$$

possui uma distribuição de probabilidade chamada distribuição  $t$  com  $n - 1$  graus de liberdade (gl).

## Propriedades das Distribuições $t$

Antes de aplicar este teorema, uma discussão das propriedades das distribuições  $t$  é solicitada. Embora a variável de interesse ainda seja  $(\bar{X} - \mu)/(S/\sqrt{n})$ , nós a representamos agora por  $T$ , para enfatizar que ela não possui uma distribuição normal padronizada quando  $n$  é pequeno. Lembre-se de que uma distribuição normal é determinada por dois parâmetros, a média  $\mu$  e o desvio padrão  $\sigma$ . A distribuição  $t$  é determinada apenas por um parâmetro, chamado **número de graus de liberdade** da distribuição, abreviado por gl. Nós o representamos pela letra grega  $\nu$ . Valores possíveis de  $\nu$  são os inteiros positivos 1, 2, 3, ... Cada valor diferente de  $\nu$  corresponde a uma distribuição  $t$  diferente.

Para qualquer valor fixo do parâmetro  $\nu$ , a função densidade que especifica a curva  $t$  associada possui uma aparência ainda mais complicada que a função densidade normal. Felizmente, precisamos nos preocupar somente com algumas das características mais importantes dessas curvas.

### Propriedades das distribuições $t$

Seja  $t_\nu$  a curva da função densidade dos gl de  $\nu$ .

1. Cada curva  $t_\nu$  possui formato de sino e está centrada em 0.
2. Toda curva  $t_\nu$  é mais dispersa que a curva normal padronizada ( $z$ ).
3. À medida que  $\nu$  aumenta, a dispersão da curva  $t_\nu$  correspondente diminui.
4. À medida que  $\nu \rightarrow \infty$  a seqüência das curvas  $t_\nu$  se aproxima da curva normal padronizada (de modo que a curva  $z$  é chamada geralmente de curva  $t$  com gl =  $\infty$ ).

A Figura 7.6 ilustra algumas dessas propriedades dos valores selecionados de  $\nu$ .

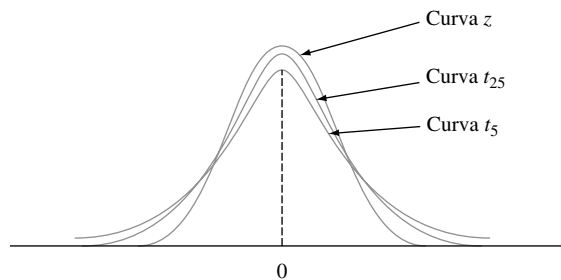


Figura 7.6 Curvas  $t_\nu$  e  $z$

O número de gl de  $T$  em (7.13) é  $n - 1$ , pois, embora  $S$  se baseie nos  $n$  desvios  $X_1 - \bar{X}, \dots, X_n - \bar{X}$ ,  $\sum(X_i - \bar{X}) = 0$  implica que somente  $n - 1$  deles são “livremente determinados”. A quantidade de gl de uma variável  $t$  é a quantidade de desvios determinados livremente onde o desvio padrão estimado no denominador de  $T$  se baseia.

Uma vez que queremos usar  $T$  para obter um IC da mesma maneira que  $Z$  foi usado anteriormente, é necessário estabelecer notação análoga a  $z_\alpha$  para a distribuição  $t$ .

### Notações

Seja  $t_{\alpha,\nu}$  = o número no eixo de medição para o qual a área sob a curva  $t$  com gl  $\nu$  à direita de  $t_{\alpha,\nu}$  é  $\alpha$ ;  $t_{\alpha,\nu}$  é chamado **valor crítico  $t$** .

Essa notação é ilustrada na Figura 7.7. A Tabela A.5 do Apêndice fornece  $t_{\alpha,\nu}$  para valores selecionados de  $\alpha$  e  $\nu$ . Essa tabela também aparece na contracapa. As colunas da tabela correspondem a valores diferentes de  $\alpha$ . Para obter  $t_{0,05,15}$ , vá para a coluna  $\alpha = 0,05$ , procure a linha  $\nu = 15$  e leia  $t_{0,05,15} = 1,753$ . De maneira semelhante,  $t_{0,05,22} = 1,717$  (coluna 0,05, linha  $\nu = 22$ ) e  $t_{0,01,22} = 2,508$ .

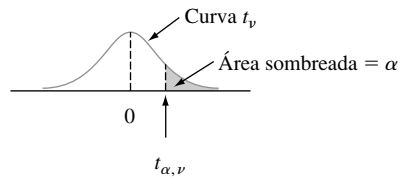


Figura 7.7 Definição ilustrada de  $t_{\alpha, \nu}$

Os valores de  $t_{\alpha, \nu}$  exibem comportamento regular, à medida que mudamos de linha ou coluna. Para  $\nu$  fixo,  $t_{\alpha, \nu}$  aumenta à medida que  $\alpha$  diminui, uma vez que temos de mover à direita de zero para encontrar a área  $\alpha$  na cauda. Para  $\alpha$  fixo à medida que  $\nu$  é aumentado (isto é, quando procuramos em qualquer coluna específica da tabela  $t$ ), o valor de  $t_{\alpha, \nu}$  diminui, porque um valor maior de  $\nu$  implica uma distribuição  $t$  com dispersão menor, de modo que não é necessário se afastar muito de zero para encontrar a área  $\alpha$  da cauda. Além disso,  $t_{\alpha, \nu}$  diminui mais lentamente à medida que  $\nu$  aumenta. Conseqüentemente, os valores da tabela são mostrados em incrementos de 2 entre 30 e 40 gl e depois pulam para  $\nu = 50, 60, 120$  e finalmente  $\infty$ . Em virtude de  $t_{\infty}$  ser a curva normal padronizada, os valores  $z_{\alpha}$  familiares aparecem na última linha da tabela. A regra prática sugerida anteriormente para uso do IC de amostra grande (se  $n > 40$ ) surge da igualdade aproximada das distribuições  $t$  e normal padronizada de  $\nu \geq 40$ .

## O Intervalo de Confiança $t$ de uma Amostra

A variável padronizada  $T$  possui distribuição  $t$  com  $n - 1$  gl e a área sob a curva de densidade  $t$  correspondente entre  $-t_{\alpha/2, n-1}$  e  $t_{\alpha/2, n-1}$  é  $1 - \alpha$  (a área  $\alpha/2$  está em cada cauda), assim

$$P(-t_{\alpha/2, n-1} < T < t_{\alpha/2, n-1}) = 1 - \alpha \quad (7.14)$$

A expressão (7.14) difere-se das expressões nas seções anteriores porque  $T$  e  $t_{\alpha/2, n-1}$  são usados no lugar de  $Z$  e  $z_{\alpha/2}$  mas pode ser manipulada da mesma forma para obter um intervalo de confiança de  $\mu$ .

### PROPOSIÇÃO

Sejam  $\bar{x}$  e  $s$  a média e o desvio padrão amostrais calculados a partir dos resultados de uma amostra aleatória de uma população normal com média  $\mu$ . Então, o **intervalo de confiança de  $100(1 - \alpha)\%$  para  $\mu$**  é

$$\left( \bar{x} - t_{\alpha/2, n-1} \cdot \frac{s}{\sqrt{n}}, \bar{x} + t_{\alpha/2, n-1} \cdot \frac{s}{\sqrt{n}} \right) \quad (7.15)$$

ou, mais compactamente,  $\bar{x} \pm t_{\alpha/2, n-1} \cdot s/\sqrt{n}$ .

O **limite de confiança superior de  $\mu$**  é

$$\bar{x} + t_{\alpha, n-1} \cdot \frac{s}{\sqrt{n}}$$

e substituir  $+$  por  $-$  nessa última expressão fornece o **limite de confiança inferior de  $\mu$** , com nível de confiança de  $100(1 - \alpha)\%$ .

### Exemplo 7.11

Como parte de um projeto maior para estudar o comportamento de painéis de revestimento tencionado, componente estrutural que está sendo usado extensivamente nos Estados Unidos, o artigo “Time-Dependent Bending Properties of Lumber” (*J. of Testing and Eval.*, 1996, p. 187-193) relatou diversas propriedades mecânicas de espécimes de madeira serrada de pinho da Escócia. Considere as seguintes observações sobre o módulo de elasticidade (MPa) obtido 1 minuto depois da aplicação de carga em uma determinada configuração:

10.490 16.620 17.300 15.480 12.970 17.260 13.400 13.900  
 13.630 13.260 14.370 11.700 15.470 17.840 14.070 14.760

A Figura 7.8 mostra um gráfico de probabilidade normal obtido do MINITAB. O padrão reto no gráfico fornece forte apoio para assumir que a distribuição da população do módulo de elasticidade é pelo menos aproximadamente normal.

O cálculo manual da média amostral e do desvio padrão é simplificado subtraindo 10.000 de cada observação:  $y_i = x_i - 10.000$ . Verifica-se com facilidade que  $\sum y_i = 72.520$  e  $\sum y_i^2 = 392.083.800$ , pelo que  $\bar{y} = 4532,5$  e  $s_y = 2055,67$ . Dessa forma,  $\bar{x} = 14.532,5$  e  $s_x = 2055,67$  (adicionar ou subtrair a mesma constante de cada observação não afeta a variabilidade). O tamanho da amostra é 16, de modo que o intervalo de confiança do módulo de elasticidade médio da população se baseia em 15 gl. Um nível de confiança de 95% de um intervalo bicaudal exige o valor crítico  $t$  de 2,131. O intervalo resultante é

$$\begin{aligned} \bar{x} \pm t_{0,025,15} \cdot \frac{s}{\sqrt{n}} &= 14.532,5 \pm (2,131) \frac{2055,67}{\sqrt{16}} \\ &= 14.532,5 \pm 1095,2 = (13.437,3, 15.627,7) \end{aligned}$$

O intervalo é bastante amplo devido tanto ao tamanho pequeno da amostra, como ao valor grande da variabilidade. O limite de confiança inferior de 95% é obtido usando-se  $-$  e 1,753 em vez de  $\pm$  e 2,131, respectivamente.

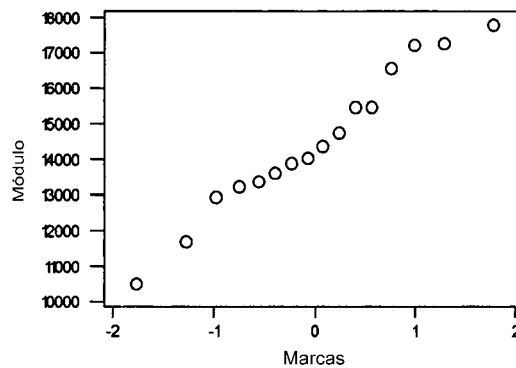


Figura 7.8 Gráfico de probabilidade normal dos dados do módulo de elasticidade

Infelizmente, não é fácil selecionar  $n$  para controlar a amplitude do intervalo  $t$ , porque a amplitude envolve o  $s$  desconhecido (antes de os dados serem coletados) e porque  $n$  entra não somente por meio de  $1/\sqrt{n}$  como também por meio de  $t_{\alpha/2, n-1}$ . Como resultado,  $n$  apropriado pode ser obtido somente por tentativa e erro.

No Capítulo 15, discutiremos um IC de amostra pequena para  $\mu$  que seja válido somente se a distribuição da população for simétrica, uma suposição mais fraca que a normalidade. Entretanto, quando a distribuição da população é normal, o intervalo  $t$  tende a ser menor do que *qualquer* outro intervalo com o mesmo nível de confiança.

### Um Intervalo de Previsão para um Único Valor Futuro

Em muitas aplicações, o investigador deseja *prever* um único valor de uma variável a ser observado futuramente, em vez de *estimar* o valor médio dessa variável.

#### Exemplo 7.12

Considere a seguinte amostra do teor de gordura (em porcentagem) de  $n = 10$  cachorros-quentes selecionados aleatoriamente (“Sensory and Mechanical Assessment of the Quality of Frankfurters”, *J. Texture Studies*, 1990, p. 395-409):

25,2 21,3 22,8 17,0 29,8 21,0 25,5 16,0 20,9 19,5

Assumindo que foram selecionados, de uma distribuição normal da população, um IC de 95% (estimativa do intervalo de) do teor de gordura médio da população é

$$\begin{aligned}\bar{x} \pm t_{0,025,9} \cdot \frac{s}{\sqrt{n}} &= 21,90 \pm 2,262 \cdot \frac{4,134}{\sqrt{10}} = 21,90 \pm 2,96 \\ &= (18,94, 24,86)\end{aligned}$$

Suponha, entretanto, que você comerá um único cachorro-quente desse tipo e queira uma *previsão* do teor de gordura resultante. A previsão *pontual*, análoga à estimativa *pontual*, é apenas  $\bar{x} = 21,90$ . Essa previsão infelizmente não fornece informações sobre confiabilidade ou precisão. ■

A configuração geral é a seguinte: teremos disponível a amostra aleatória  $X_1, X_2, \dots, X_n$  de uma distribuição normal da população, e queremos prever o valor de  $X_{n+1}$  de uma única observação futura. Um previsor pontual é  $\bar{X}$ , e o erro de previsão resultante é  $\bar{X} - X_{n+1}$ . O valor esperado do erro de previsão é

$$E(\bar{X} - X_{n+1}) = E(\bar{X}) - E(X_{n+1}) = \mu - \mu = 0$$

Uma vez que  $X_{n+1}$  é independente de  $X_1, \dots, X_n$ , *ele* é independente de  $\bar{X}$ , de modo que a variância do erro de previsão é

$$V(\bar{X} - X_{n+1}) = V(\bar{X}) + V(X_{n+1}) = \frac{\sigma^2}{n} + \sigma^2 = \sigma^2 \left(1 + \frac{1}{n}\right)$$

O erro de previsão é uma combinação linear de *vas* independentes normalmente distribuídas, de modo que ele próprio seja normalmente distribuído. Dessa forma,

$$Z = \frac{(\bar{X} - X_{n+1}) - 0}{\sqrt{\sigma^2 \left(1 + \frac{1}{n}\right)}} = \frac{\bar{X} - X_{n+1}}{\sqrt{\sigma^2 \left(1 + \frac{1}{n}\right)}}$$

possui uma distribuição normal padronizada. É possível mostrar que substituindo  $\sigma$  pelo desvio-padrão da amostra  $S$  (de  $X_1, \dots, X_n$ ), resulta em

$$T = \frac{\bar{X} - X_{n+1}}{S \sqrt{1 + \frac{1}{n}}} \sim \text{distribuição } t \text{ com } n - 1 \text{ gl}$$

Manipular a variável  $T$ , como  $T = (\bar{X} - \mu)/(S/\sqrt{n})$  foi manipulado no desenvolvimento de um IC, fornece o seguinte resultado.

#### PROPOSIÇÃO

O **intervalo de previsão** (IP) de uma única observação a ser selecionada de uma distribuição normal da população é

$$\bar{x} \pm t_{\alpha/2, n-1} \cdot s \sqrt{1 + \frac{1}{n}} \quad (7.16)$$

O nível de previsão é  $100(1 - \alpha)\%$ .

A interpretação de um nível de previsão de 95% é semelhante à de um nível de confiança de 95%; se o intervalo (7.16) for calculado para amostra após amostra, os 95% de longo prazo desses intervalos incluirão os valores futuros correspondentes de  $X$ .



**Exemplo 7.13** (continuação do Exemplo 7.12)

Com  $n = 10$ ,  $\bar{x} = 21,90$ ,  $s = 4,134$  e  $t_{0,025,9} = 2,262$ , um IP de 95% do teor de gordura de um único cachorro-quente é

$$\begin{aligned} 21,90 \pm (2,262)(4,134) \sqrt{1 + \frac{1}{10}} &= 21,90 \pm 9,81 \\ &= (12,09, 31,71) \end{aligned}$$

Esse intervalo é bastante amplo, indicando incerteza substancial sobre o teor de gordura. Observe que a amplitude do IP é mais que três vezes a do IC. ■

O erro de previsão é  $\bar{X} - X_{n+1}$ , diferença entre duas variáveis aleatórias, enquanto o erro de estimativa é  $\bar{X} - \mu$ , diferença entre uma variável aleatória e um valor fixo (mas desconhecido). O IP é mais amplo que o IC, pois há mais variabilidade no erro de previsão (devido a  $X_{n+1}$ ) que no erro de estimativa. De fato, à medida que  $n$  fica arbitrariamente maior, o IC diminui para o único valor  $\mu$  e o IP se aproxima de  $\mu \pm z_{\alpha/2} \cdot \sigma$ . Há incerteza sobre um único valor de  $X$ , mesmo quando não há necessidade de fazer estimativa.

**Intervalos de Tolerância**

Considere uma população de determinado tipo de automóveis e suponha que, sob condições especificadas, o consumo de combustível (milhas/galão) possui distribuição normal com  $\mu = 30$  e  $\sigma = 2$ . Então, uma vez que o intervalo de  $-1,645$  a  $1,645$  contendo 90% da área sob a curva  $z$ , 90% de todos esses automóveis terão valores de consumo de combustível entre  $\mu - 1,645\sigma = 26,71$  e  $\mu + 1,645\sigma = 33,29$ . Mas o que acontece se os valores de  $\mu$  e  $\sigma$  não são conhecidos? Podemos tomar uma amostra de tamanho  $n$ , determinar os consumos de combustível,  $\bar{x}$  e  $s$ , e formar o intervalo cujo limite inferior seja  $\bar{x} - 1,645s$  e o limite superior seja  $\bar{x} + 1,645s$ . Entretanto, devido à variabilidade da amostragem nas estimativas de  $\mu$  e  $\sigma$ , existe uma boa chance de o intervalo resultante incluir menos de 90% dos valores da população. Intuitivamente, para ter uma chance *a priori* de 95%, de o intervalo resultante incluir pelo menos 90% dos valores da população, quando  $\bar{x}$  e  $s$  são usados no lugar de  $\mu$  e  $\sigma$ , devemos também substituir 1,645 por algum número maior. Por exemplo: quando  $n = 20$ , o valor 2,310 é tal que podemos ter 95% de confiança de que o intervalo  $\bar{x} \pm 2,310s$  incluirá pelo menos 90% dos valores de consumo de combustível na população.

Seja  $k$  um número entre 0 e 100. Um **intervalo de tolerância** para incluir pelo menos  $k\%$  dos valores de uma distribuição normal da população com nível de confiança de 95% possui a forma

$$\bar{x} \pm (\text{valor crítico de tolerância}) \cdot s$$

Os valores críticos de tolerância para  $k = 90, 95$  e  $99$  em combinação com vários tamanhos de amostra são dados na Tabela A.6 do Apêndice. Essa tabela também inclui valores críticos para um nível de confiança de 99% (esses valores são maiores que os valores de 95% correspondentes). Substituir  $\pm$  por  $+$  fornece um limite de tolerância superior, e usar  $-$  no lugar de  $\pm$  resulta em um limite de tolerância inferior. Os valores críticos para obter esses limites monocaudais também aparecem na Tabela A.6 do Apêndice.

**Exemplo 7.14**

Vamos voltar aos dados do módulo de elasticidade discutidos no Exemplo 7.11, onde  $n = 16$ ,  $\bar{x} = 14.532,5$ ,  $s = 2055,67$  e um gráfico de probabilidade normal dos dados indicou que a normalidade da população era bastante plausível. Para um nível de confiança de 95%, o intervalo de tolerância bicaudal para incluir pelo menos 95%

dos valores do módulo de elasticidade de espécimes de madeira serrada na população que serviu como amostra utiliza o valor crítico de tolerância de 2,903. O intervalo resultante é

$$14.532,5 \pm (2,903)(2055,67) = 14.532,5 \pm 5967,6 = (8.564,9, 20.500,1)$$

Podemos estar altamente confiantes de que pelo menos 95% de todos os espécimes de madeira serrada têm valores de módulo de elasticidade entre 8.564,9 e 20.500,1.

O IC de 95% de  $\mu$  era (13.437,3, 15.627,7) e o intervalo de previsão de 95% do módulo de elasticidade de um único espécime de madeira serrada é (10.017,0, 19.048,0). O intervalo de previsão e o intervalo de tolerância são substancialmente mais amplos que o intervalo de confiança. ■

## Intervalos Baseados em Distribuições Não-normais de População

O IC de  $t$  de uma amostra para  $\mu$  é robusto, para desvios pequenos ou mesmo moderados da hipótese de normalidade, a menos que  $n$  seja muito pequeno. Por isso, queremos dizer que, se um valor crítico de confiança de 95%, por exemplo, for usado para calcular o intervalo, o nível de confiança real estará razoavelmente próximo do nível de 95% nominal. Entretanto, se  $n$  for pequeno e a distribuição da população for altamente não-normal, então o nível de confiança real pode ser consideravelmente diferente do que você pensa que está utilizando quando obtém um valor crítico específico da tabela de  $t$ . Certamente, seria frustrante acreditar que seu nível de confiança é cerca de 95% quando, na verdade, era realmente mais provável ser de 88%! A técnica de *bootstrap*, introduzida na Seção 7.1, tem muito sucesso na estimativa de parâmetros em uma grande variedade de situações não-normais.

Em comparação com o intervalo de confiança, a validade dos intervalos de tolerância e previsão descritos nesta seção está ligada mais proximamente à suposição de normalidade. Esses últimos intervalos não devem ser utilizados na ausência de evidência constrangedora à normalidade. A excelente referência *Statistical Intervals*, citada na bibliografia no final deste capítulo, discute procedimentos alternativos desse tipo para várias outras situações.

---

### Exercícios | Seção 7.3 (28–41)

28. Determine os valores das seguintes quantidades:  
 a.  $t_{0,1,15}$    b.  $t_{0,05,15}$    c.  $t_{0,05,25}$    d.  $t_{0,05,40}$    e.  $t_{0,005,40}$
29. Determine o valor crítico  $t$  que incluirá a área da curva  $t$  desejada em cada um dos casos a seguir:  
 a. Área central = 0,95, gl = 10  
 b. Área central = 0,95, gl = 20  
 c. Área central = 0,99, gl = 20  
 d. Área central = 0,99, gl = 50  
 e. Área da cauda superior = 0,01, gl = 25  
 f. Área da cauda inferior = 0,025, gl = 5
30. Determine o valor crítico  $t$  de um intervalo de confiança bicaudal em cada uma das situações a seguir:  
 a. Nível de confiança = 95%, gl = 10  
 b. Nível de confiança = 95%, gl = 15  
 c. Nível de confiança = 99%, gl = 15  
 d. Nível de confiança = 99%,  $n = 5$   
 e. Nível de confiança = 98%, gl = 24  
 f. Nível de confiança = 99%,  $n = 38$
31. Determine o valor crítico  $t$  de um limite de confiança inferior ou superior para cada uma das situações descritas no Exercício 30.
32. Uma amostra aleatória de  $n = 8$  espécimes de teste de certo tipo de fibra de vidro E produziu uma força do rendimento de corte interfacial média amostral de 30,2 e desvio padrão da amostra de 3,1 (“On Interfacial Failure in Notched Unidirectional Glass/Epoxy Composites”, *J. of Composite Materials*, 1985, p. 276-286). Assuma que a força do rendimento de corte interfacial é normalmente distribuída, calcule o IC de 95% para a força média real (como fizeram os autores do artigo citado).
33. O artigo “Measuring and Understanding the Aging of Kraft Insulating Paper in Power Transformers” (*IEEE Electrical Insul. Mag.*, 1996, p. 28-34) continha as seguintes observações sobre o grau de polimerização dos espécimes de papel para os quais a viscosidade vezes a concentração caíram em certo intervalo intermediário:
- |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|
| 418 | 421 | 421 | 422 | 425 | 427 | 431 |
| 434 | 437 | 439 | 446 | 447 | 448 | 453 |
| 454 | 463 | 465 |     |     |     |     |
- a. Construa um *boxplot* dos dados e comente quaisquer características interessantes.  
 b. É plausível que as observações da amostra fornecidas tenham sido selecionadas de uma distribuição normal?

- c. Calcule o intervalo de confiança de 95% bicaudal para o grau médio real de polimerização (como fizeram os autores do artigo). O intervalo sugere que 440 seja um valor plausível para o grau médio real de polimerização? E quanto a 450?
34. Uma amostra de 14 espécimes de juntas de um tipo específico forneceu o limite de carga proporcional médio da amostra de 8,48 MPa e desvio padrão da amostra de 0,79 MPa (“Characterization of Bearing Strength Factors in Pegged Timber Connections”, *J. of Structural Engr.*, 1997, p. 326-332).
- a. Calcule e interprete o limite de confiança inferior de 95% para o limite de carga proporcional médio real de todas as juntas. Que hipótese, se houver, você faria sobre a distribuição do limite de carga proporcional?
- b. Calcule e interprete o limite de previsão inferior de 95% para o limite de carga proporcional de uma única junta desse tipo.
35. O Exercício 46 no Capítulo 1 introduziu as seguintes observações da amostra sobre viscosidade estabilizada de espécimes de asfalto: 2781, 2900, 3013, 2856, 2888. Um gráfico de probabilidade normal suporta a hipótese de a viscosidade ser distribuída, pelo menos aproximadamente, de forma normal.
- a. Calcule a viscosidade média real, de forma que transmita informações sobre precisão e confiabilidade.
- b. Preveja a viscosidade de um único espécime de asfalto, de modo que transmita informações sobre precisão e confiabilidade. Como a previsão se compara com a estimativa calculada na parte (a)?
36. As  $n = 26$  observações sobre o tempo de fuga dado no Exercício 36 do Capítulo 1 fornecem a média e o desvio padrão da amostra de 370,69 e 24,36, respectivamente.
- a. Calcule o limite de confiança superior para o tempo médio de fuga da população, usando um nível de confiança de 95%.
- b. Calcule o limite de previsão superior para o tempo de fuga de um único trabalhador adicional, usando um nível de previsão de 95%. Como esse limite se compara ao limite de confiança da parte (a)?
- c. Suponha que dois trabalhadores adicionais sejam escolhidos para participar do exercício simulado de fuga. Represente seus tempos de fuga por  $X_{27}$  e  $X_{28}$ , e por  $\bar{X}_{\text{nov0}}$  a média desses dois valores. Modifique a fórmula de um IP para um único valor de  $x$ , a fim de obter o IP de  $\bar{X}_{\text{nov0}}$  e calcule o intervalo bicaudal de 95%, com base nos dados de fuga fornecidos.
37. Um estudo da habilidade de as pessoas andarem em linha reta (“Can We Really Walk Straight?” *Amer. J. of Physical Anthro.*, 1992, p. 19-27) relatou os dados que seguem sobre a cadência (passos por segundo) de uma amostra de  $n = 20$  homens saudáveis, aleatoriamente selecionados.

0,95 0,85 0,92 0,95 0,93 0,86 1,00 0,92 0,85 0,81  
0,78 0,93 0,93 1,05 0,93 1,06 1,06 0,96 0,81 0,96

Um gráfico de probabilidade normal fornece suporte substancial à hipótese de que a distribuição da população da cadência seja aproximadamente normal. Segue um resumo descritivo dos dados do MINITAB:

Variável	N	Média	Mediana	Média Ap	Desv Padrão	Média SE
cadência	20	0,9255	0,9300	0,9261	0,0809	0,0181
Variável	Min	Max	Q1	Q3		
cadência	0,7800	1,0600	0,8525	0,9600		

- a. Calcule e interprete o intervalo de confiança de 95% para a cadência média da população.
- b. Calcule e interprete o intervalo de previsão de 95% para a cadência de uma única pessoa selecionada aleatoriamente dessa população.
- c. Calcule o intervalo que inclua pelo menos 99% das cadências da distribuição da população, usando um nível de confiança de 95%.
38. Uma amostra de 25 peças de laminado usado na fabricação de placas de circuito foi selecionada, e a quantidade de deformação (in.) sob condições específicas foi determinada para cada peça, resultando em uma deformação média amostral de 0,0635 e um desvio padrão de 0,0065.
- a. Calcule uma previsão para a quantidade de deformação de uma única peça de laminado, de modo que forneça informações sobre precisão e confiabilidade.
- b. Calcule o intervalo no qual você possa ter alto grau de confiança de que pelo menos 95% de todas as peças de laminado resultem em quantidades de deformação que estejam entre os dois limites do intervalo.
39. O Exercício 72 do Capítulo 1 forneceu as seguintes observações sobre a medida de absorção do receptor (volume de distribuição ajustado) de uma amostra de 13 pessoas saudáveis: 23, 39, 40, 41, 43, 47, 51, 58, 63, 66, 67, 69, 72.
- a. É plausível que a distribuição da população da qual essa amostra foi selecionada seja normal?
- b. Calcule o intervalo para o qual possamos ter 95% de confiança de que pelo menos 95% de todas as pessoas saudáveis da população tenham ajustado os volumes da distribuição ajustados incluídos entre os limites do intervalo.
- c. Preveja o volume da distribuição ajustado de uma única pessoa saudável, calculando o intervalo de previsão de 95%. Como a amplitude desse intervalo se compara com o intervalo calculado na parte (b)?
40. O Exercício 13 do Capítulo 1 apresentou uma amostra de  $n = 153$  observações sobre a resistência à tensão máxima, e o Exercício 17 da seção anterior forneceu quantidades resumidas e pediu o intervalo de confiança de amostra grande. Em virtude de o tamanho da amostra ser

grande, nenhuma suposição sobre a distribuição da população é necessária para a validade do IC.

- a. Qualquer suposição sobre a distribuição da resistência à tensão é necessária, antes de calcular um limite de previsão inferior para a resistência à tensão do próximo espécime selecionado usando o método descrito nesta seção? Explique.
- b. Use um programa estatístico para investigar a plausibilidade de uma distribuição normal da população.
- c. Calcule o limite de previsão inferior com um nível de previsão de 95% para a resistência extrema à tensão do próximo espécime selecionado.

41. Uma tabulação mais extensiva dos valores críticos  $t$  que a mostrada neste livro indica que, para a distribuição  $t$  com gl 20, as áreas à direita dos valores 0,687, 0,860 e 1,064 são 0,25, 0,20 e 0,15, respectivamente. Qual é o nível de confiança de cada um dos três intervalos de confiança a seguir para a média  $\mu$  de uma distribuição normal da população? Qual dos três intervalos você recomendaria para ser usado e por quê?

- a.  $(\bar{x} - 0,687s/\sqrt{21}, \bar{x} + 1,725s/\sqrt{21})$
- b.  $(\bar{x} - 0,860s/\sqrt{21}, \bar{x} + 1,325s/\sqrt{21})$
- c.  $(\bar{x} - 1,064s/\sqrt{21}, \bar{x} + 1,064s/\sqrt{21})$

## 7.4 Intervalos de Confiança para Variância e Desvio Padrão de uma População Normal

Embora as inferências relacionadas à variância de uma população  $\sigma^2$  ou ao desvio padrão  $\sigma$  sejam geralmente de menor interesse que as referentes à média ou proporção, há ocasiões em que tais procedimentos são necessários. No caso de uma distribuição normal da população, as inferências são feitas com base no seguinte resultado relacionado à variância da amostra  $S^2$ .

### TEOREMA

Seja  $X_1, X_2, \dots, X_n$  a amostra aleatória de uma distribuição normal com parâmetros  $\mu$  e  $\sigma^2$ . Então, a va

$$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum(X_i - \bar{X})^2}{\sigma^2}$$

possui distribuição de probabilidade qui-quadrada ( $\chi^2$ ) com  $n - 1$  gl.

Conforme discutido nas seções 4.4 e 7.1, a distribuição qui-quadrada é uma distribuição de probabilidade contínua com um único parâmetro  $\nu$ , chamado número de graus de liberdade, com valores possíveis 1, 2, 3, ... Os gráficos de diversas funções-distribuição de probabilidade de  $\chi^2$  (fdps) são ilustrados na Figura 7.9. Cada fdp  $f(x; \nu)$  é positiva somente para  $x > 0$ , e cada uma possui inclinação positiva (cauda superior longa), embora a distribuição se mova à direita e se torne mais simétrica à medida que  $\nu$  aumenta. Para especificar os procedimentos inferenciais que utilizam a distribuição qui-quadrada, precisamos de notação análoga à de um valor crítico  $t$  de  $t_{\alpha, \nu}$

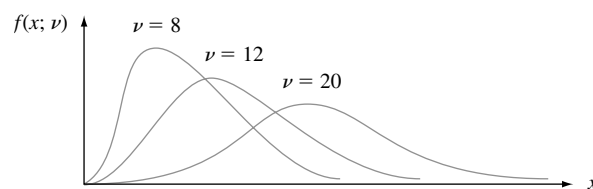


Figura 7.9 Gráficos das funções-densidade qui-quadrada

### Notação

Seja  $\chi_{\alpha,\nu}^2$ , chamado **valor crítico qui-quadrado**, o número no eixo da medição, de modo que  $\alpha$  da área sob a curva qui-quadrada com gl  $\nu$  esteja à direita de  $\chi_{\alpha,\nu}^2$ .

Em virtude de a distribuição  $t$  ser simétrica, foi necessário tabular somente os valores críticos da cauda superior ( $t_{\alpha,\nu}$  para valores pequenos de  $\alpha$ ). A distribuição qui-quadrada não é simétrica; de modo que, a Tabela A.7 do Apêndice contém valores de  $\chi_{\alpha,\nu}^2$ , tanto para  $\alpha$  próximo de 0 e de 1, conforme ilustrado na Figura 7.10(b). Por exemplo:  $\chi_{0,025,14}^2 = 26,119$   $\chi_{0,95,20}^2$  (o 5º percentil) = 10,851.

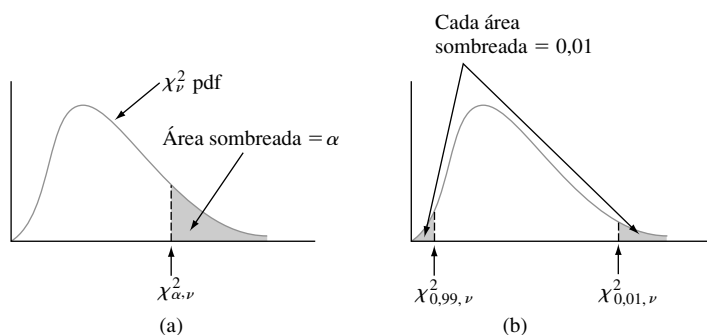


Figura 7.10 Notação ilustrada de  $\chi_{\alpha,\nu}^2$

A  $\nu(n-1)S^2/\sigma^2$  satisfaz as duas propriedades nas quais o método geral para obter IC se baseia: é uma função do parâmetro de interesse  $\sigma^2$ , ainda que sua distribuição de probabilidade (qui-quadrada) não dependa desse parâmetro. A área sob uma curva qui-quadrada com gl  $\nu$  à direita de  $\chi_{\alpha/2,\nu}^2$  é  $\alpha/2$ , como a área à esquerda de  $\chi_{1-\alpha/2,\nu}^2$ . Dessa forma, a área incluída entre esses dois valores críticos é  $1 - \alpha$ . Como consequência disso e do teorema que acabamos de enunciar,

$$P\left(\chi_{1-\alpha/2,n-1}^2 < \frac{(n-1)S^2}{\sigma^2} < \chi_{\alpha/2,n-1}^2\right) = 1 - \alpha \quad (7.17)$$

As desigualdades de (7.17) são equivalentes a

$$\frac{(n-1)S^2}{\chi_{\alpha/2,n-1}^2} < \sigma^2 < \frac{(n-1)S^2}{\chi_{1-\alpha/2,n-1}^2}$$

Substituindo o valor calculado  $s^2$  nos limites fornece o IC para  $\sigma^2$  e extraíndo as raízes quadradas obtém-se o intervalo de  $\sigma$ .

**O intervalo de confiança de  $100(1 - \alpha)\%$  para a variância  $\sigma^2$  de uma população normal** possui limite inferior

$$(n-1)s^2/\chi_{\alpha/2,n-1}^2$$

e limite superior

$$(n-1)s^2/\chi_{1-\alpha/2,n-1}^2$$

**O intervalo de confiança para  $\sigma$**  possui limites superior e inferior que são as raízes quadradas dos limites correspondentes no intervalo para  $\sigma^2$ .

**Exemplo 7.15**

Os dados que seguem sobre a voltagem de quebra de circuitos carregados eletricamente foram obtidos de um gráfico de probabilidade normal que apareceu no artigo “Damage of Flexible Printed Wiring Boards Associated with Lightning-Induced Voltage Surges” (*IEEE Transactions on Components, Hybrids, and Manuf. Tech.*, 1985, p. 214-220). A disposição em reta do gráfico forneceu forte apoio à suposição de que a voltagem de quebra é aproximadamente distribuída de forma normal.

1470 1510 1690 1740 1900 2000 2030 2100 2190  
2200 2290 2380 2390 2480 2500 2580 2700

Seja  $\sigma^2$  a variância da distribuição da voltagem de quebra. O valor calculado da variância da amostra é  $s^2 = 137.324,3$ , a estimativa pontual de  $\sigma^2$ . Com  $gl = n - 1 = 16$ , o IC de 95% requer que  $\chi_{0,975,16}^2 = 6,908$  e  $\chi_{0,025,16}^2 = 28,845$ . O intervalo é

$$\left( \frac{16(137.324,3)}{28,845}, \frac{16(137.324,3)}{6,908} \right) = (76.172,3, 318.064,4)$$

Extraindo a raiz quadrada de cada ponto final produz (276,0, 564,0), IC de 95% para  $\sigma$ . Esses intervalos são bastante amplos, refletindo a variabilidade substancial na voltagem de quebra em combinação com o tamanho pequeno de uma amostra. ■

Os ICs de  $\sigma^2$  e  $\sigma$ , quando a distribuição da população não é normal, podem ser difíceis de se obter, mesmo quando o tamanho da amostra é grande. Para esses casos, consulte um estatístico de renome.

---

**Exercícios** | Seção 7.4 (42–46)

42. Determine os valores das seguintes quantidades:

- a.  $\chi_{0,1,15}^2$     b.  $\chi_{0,1,25}^2$   
c.  $\chi_{0,01,25}^2$     d.  $\chi_{0,005,25}^2$   
e.  $\chi_{0,99,25}^2$     f.  $\chi_{0,995,25}^2$

43. Determine o seguinte:

- a. O 95º percentil da distribuição qui-quadrada com  $\nu = 10$   
b. O 5º percentil da distribuição qui-quadrada com  $\nu = 10$   
c.  $P(10,98 \leq \chi^2 \leq 36,78)$ , onde  $\chi^2$  é uma va qui-quadrada com  $\nu = 22$   
d.  $P(\chi^2 < 14,611 \text{ ou } \chi^2 > 37,652)$ , onde  $\chi^2$  é uma va qui-quadrada com  $\nu = 25$

44. A quantidade de expansão lateral (mils) foi determinada para uma amostra de  $n = 9$  soldas de arco de gás metálico acionado por pulsos usados em tanques de carga de navio LNG. O desvio padrão da amostra resultante foi  $s = 2,81$  mils. Assumindo a normalidade, encontre o IC de 95% para  $\sigma^2$  e para  $\sigma$ .

45. As seguintes observações foram feitas sobre a dureza de fratura de uma placa de base de 18% de níquel em aço

inoxidável [“Fracture Testing of Weldments”, *ASTM Special Publ.* nº 381, 1965, p. 328-356 (in ksi  $\sqrt{\text{in.}}$ , dado em ordem crescente)]:

69,5 71,9 72,6 73,1 73,3 73,5 75,5 75,7  
75,8 76,1 76,2 76,2 77,0 77,9 78,1 79,6  
79,7 79,9 80,1 82,2 83,7 93,7

Calcule o IC de 99% para o desvio padrão da distribuição de dureza de fratura. Esse intervalo é válido qualquer que seja a natureza da distribuição? Explique.

46. Os resultados de um teste de turvação de Wagner realizado em 15 amostras de areia de teste-padrão de Ottawa foram (em microampères):

26,7 25,8 24,0 24,9 26,4 25,9 24,4 21,7  
24,1 25,9 27,3 26,9 27,3 24,8 23,6

- a. É plausível que essa amostra tenha sido selecionada de uma distribuição normal da população?  
b. Calcule o limite de confiança superior com nível de confiança de 95% para o desvio padrão da população de turvação.

## Exercícios Suplementares (47–62)

47. O Exemplo 1.11 introduziu as observações que seguem sobre resistência de contato.

11,5 12,1 9,9 9,3 7,8 6,2 6,6 7,0  
 13,4 17,1 9,3 5,6 5,7 5,4 5,2 5,1  
 4,9 10,7 15,2 8,5 4,2 4,0 3,9 3,8  
 3,6 3,4 20,6 25,5 13,8 12,6 13,1 8,9  
 8,2 10,7 14,2 7,6 5,2 5,5 5,1 5,0  
 5,2 4,8 4,1 3,8 3,7 3,6 3,6 3,6

- a. Estime a resistência de contato média real, de modo que transmita informações sobre precisão e confiabilidade. (*Sugestão:*  $\sum x_i = 387,8$  e  $\sum x_i^2 = 4247,08$ .)  
 b. Calcule o IC de 95% para a proporção de todas as junções cujos valores de resistência excederiam 10.

48. Um triatlo que consiste em nadar, andar de bicicleta e correr é um dos eventos de esporte amador mais áduos. O artigo “Cardiovascular and Thermal Response of Triathlon Performance” (*Medicine and Science in Sports and Exercise*, 1988, p. 385-389) relata um estudo de pesquisa que envolve nove triatletas do sexo masculino. O índice cardíaco máximo (batimentos/min) foi registrado durante a realização de cada um dos três eventos. Para a natação, a média e o desvio padrão amostrais foram 188,0 e 7,2, respectivamente. Assumindo que a distribuição do índice cardíaco seja (aproximadamente) normal, construa um IC de 98% para o índice cardíaco médio real dos triatletas durante a natação.

49. Para cada um dos 18 centros preservados de reservatórios de carbonato de óleo, a quantidade de saturação de gás residual após a injeção de um solvente foi medida na saída do fluxo de água. As observações, em porcentagem de volume de poros, foram

23,5 31,5 34,0 46,7 45,6 32,5  
 41,4 37,2 42,5 46,9 51,5 36,4  
 44,5 35,7 33,5 39,3 22,0 51,2

(Consulte “Relative Permeability Studies of Gas-Water Flow Following Solvent Injection in Carbonate Rocks,” *Soc. Petroleum Engineers J.*, 1976, p. 23-30.)

- a. Construa um *boxplot* desses dados e faça comentários sobre quaisquer características interessantes.  
 b. É plausível que a amostra fosse selecionada de uma distribuição normal da população?  
 c. Calcule o IC de 98% para a quantidade média real de saturação de gás residual.

50. Um artigo de um periódico relata que uma amostra de tamanho 5 foi usada como base para calcular o IC de 95% para a frequência natural média real (Hz) de vigas laminadas de determinado tipo. O intervalo resultante foi (229,764, 233,504). Você decide que um nível

de confiança de 99% é mais apropriado que o nível de 95% utilizado. Quais são os limites do intervalo de 99%? (*Sugestão:* use o centro do intervalo e sua amplitude para determinar  $\bar{x}$  e  $s$ .)

51. O gerente financeiro de uma grande rede de lojas de departamentos selecionou uma amostra aleatória de 200 de seus clientes com cartão de crédito e descobriu que 136 pagaram juros no ano anterior devido a saldo não-pago.
- a. Calcule o IC de 90% para a proporção real dos clientes com cartão de crédito que pagaram juros no ano anterior.  
 b. Se a amplitude desejada do intervalo de 90% for 0,05, que tamanho de amostra é necessário para garantir isso?  
 c. O limite superior do intervalo na parte (a) especifica um limite de confiança superior de 90% para a proporção que está sendo estimada? Explique.
52. O tempo de reação a um estímulo compreende o intervalo de tempo entre o início e o fim desse estímulo, sinalizado pelo primeiro movimento visível. O artigo “Relationship of Reaction Time and Movement Time in a Gross Motor Skill” (*Perceptual and Motor Skills*, 1973, p. 453-454) relata que o TR médio da amostra de 16 nadadores experientes a um tiro inicial foi de 0,214 s. e o desvio padrão da amostra foi de 0,036 s.
- a. Fazendo quaisquer suposições necessárias, encontre o IC de 90% do TR médio real de todos os nadadores experientes.  
 b. Calcule o limite de confiança superior de 90% para o desvio padrão da distribuição do tempo de reação.  
 c. Preveja o TR para outro indivíduo, de forma que transmita informações sobre precisão e confiabilidade.
53. A infestação de pulgão em árvores frutíferas pode ser controlada pela pulverização de pesticida ou inundação de joaninhas. Em uma área específica, quatro pomares de diferentes árvores frutíferas são selecionados para experiência. Os três primeiros pomares são pulverizados com pesticidas 1, 2 e 3, respectivamente, e o quarto é tratado com joaninhas, produzindo os seguintes resultados:

Tratamento	$n_i =$	$\bar{x}_i$	$s_i$
	número de árvores	(alqueires/árvore)	
1	100	10,5	1,5
2	90	10,0	1,3
3	100	10,1	1,8
4	120	10,7	1,6

Seja  $\mu_i$  = o rendimento médio real (alqueires/árvore) depois de receber o  $i$ -ésimo tratamento. Então

$$\theta = \frac{1}{3}(\mu_1 + \mu_2 + \mu_3) - \mu_4$$

meça a diferença nos rendimentos médios reais entre o tratamento com pesticidas e o tratamento com joaninhas. Quando  $n_1, n_2, n_3$  e  $n_4$  são grandes, o estimador  $\hat{\theta}$  obtido pela substituição de cada  $\mu_i$  por  $\bar{X}_i$  é aproximadamente normal. Use esses dados para determinar o IC de  $100(1 - \alpha)\%$  de amostra grande para  $\theta$  e calcule o intervalo de 95% para os dados fornecidos.

54. É importante que as máscaras usadas pelos bombeiros sejam capazes de resistir a altas temperaturas, pois esses profissionais trabalham com frequência em temperaturas de 200-500°F. Em um teste de um tipo de máscara, 11 dos 55 equipamentos tiveram as lentes estouradas a 250°. Construa o IC de 90% para a proporção real de máscaras desse tipo, cujas lentes estourariam a 250°.

55. Um fabricante de livros didáticos está interessado em estimar a resistência das capas produzidas por uma máquina de encadernação específica. A resistência pode ser medida registrando-se a força necessária para arrancar as páginas da encadernação. Se essa força for medida em libras, quantos livros devem ser testados para estimar a força média necessária para quebrar a encadernação dentro de 0,1 lb com confiança de 95%? Assuma que  $\sigma$  seja 0,8.

56. A exposição crônica à fibra de amianto é um perigo à saúde bastante conhecido. O artigo “The Acute Effects of Chrysotile Asbestos Exposure on Lung Function” (*Environ. Research*, 1978, p. 360-372) relata resultados de um estudo com base em uma amostra de operários de construção expostos ao amianto durante um longo período. Dentre os dados fornecidos no artigo, 8 meses após o período de exposição, os seguintes valores (ordenados) de complacência pulmonar ( $\text{cm}^3/\text{cm H}_2\text{O}$ ) foram apresentados em cada uma das 16 pessoas (a complacência pulmonar é uma medida da elasticidade do pulmão, ou como os pulmões, de maneira eficaz, são capazes de inalar e exalar):

167,9 180,8 184,8 189,8 194,8 200,2  
201,9 206,9 207,2 208,4 226,3 227,7  
228,5 232,4 239,8 258,6

- É plausível que a distribuição da população seja normal?
- Calcule o IC de 95% para a complacência pulmonar média real após a exposição.
- Calcule o intervalo que, com um nível de confiança de 95%, inclua pelo menos 95% dos valores de complacência pulmonar na distribuição da população.

57. No Exemplo 6.7, introduzimos o conceito de um experimento censurado em que  $n$  componentes são testados e que o experimento termina logo que  $r$  dos componentes tiver falhado. Suponha que a vida útil dos componentes seja independente, cada um com uma distribuição exponencial com parâmetro  $\lambda$ . Represente por  $Y_1$  a hora em que ocorre a primeira falha, por  $Y_2$  a hora

em que ocorre a segunda falha, e assim por diante, de modo que  $T_r = Y_1 + \dots + Y_r + (n - r)Y_r$  seja a vida útil total acumulada no fim. Então, é possível mostrar que  $2\lambda T_r$  possui uma distribuição qui-quadrada com gl de  $2r$ . Use esse fato para desenvolver uma fórmula de IC de  $100(1 - \alpha)\%$  da vida útil média real de  $1/\lambda$ . Calcule o IC de 95% a partir dos dados no Exemplo 6.7.

58. Seja por  $X_1, X_2, \dots, X_n$  a amostra aleatória de uma distribuição de probabilidade contínua que tem mediana  $\tilde{\mu}$  (de modo que  $P(X_i \leq \tilde{\mu}) = P(X_i \geq \tilde{\mu}) = 0,5$ ).

a. Mostre que

$$P(\min(X_i) < \tilde{\mu} < \max(X_i)) = 1 - \left(\frac{1}{2}\right)^{n-1}$$

para que  $(\min(x_i), \max(x_i))$  seja um intervalo de confiança de  $100(1 - \alpha)\%$  de  $\tilde{\mu}$  com  $\alpha = \left(\frac{1}{2}\right)^{n-1}$ . [Sugestão: o complemento do evento  $\{\min(X_i) < \tilde{\mu} < \max(X_i)\}$  é  $\{\max(X_i) \leq \tilde{\mu}\} \cup \{\min(X_i) \geq \tilde{\mu}\}$ . Mas  $\max(X_i) \leq \tilde{\mu}$  se e somente se  $X_i \leq \tilde{\mu}$  para todos os  $i$ .]

b. Para cada um dos seis meninos normais, a quantidade do aminoácido alanina (mg/100 mL) foi determinada enquanto as crianças estavam em uma dieta sem isoleucina, resultando nos seguintes dados:

2,84 3,54 2,80 1,44 2,94 2,70

Calcule o IC de 97% da quantidade mediana real de alanina das crianças nessa dieta (“The Essential Amino Acid Requirements of Infants,” *Amer. J. Nutrition*, 1964, p. 322-330).

c. Seja  $x_{(2)}$  o segundo menor  $x_i$  e  $x_{(n-1)}$  o segundo maior  $x_i$ . Qual é o coeficiente de confiança do intervalo  $(x_{(2)}, x_{(n-1)})$  para  $\tilde{\mu}$ ?

59. Seja  $X_1, X_2, \dots, X_n$  a amostra aleatória de uma distribuição uniforme no intervalo  $[0, \theta]$ , de modo que

$$f(x) = \begin{cases} \frac{1}{\theta} & 0 \leq x \leq \theta \\ 0 & \text{caso contrário} \end{cases}$$

Então, se  $Y = \max(X_i)$ , é possível mostrar que a  $U = Y/\theta$  possui função densidade

$$f_U(u) = \begin{cases} nu^{n-1} & 0 \leq u \leq 1 \\ 0 & \text{caso contrário} \end{cases}$$

a. Use  $f_U(u)$  para verificar que

$$P\left((\alpha/2)^{1/n} < \frac{Y}{\theta} \leq (1 - \alpha/2)^{1/n}\right) = 1 - \alpha$$

e use a expressão para encontrar o IC de  $100(1 - \alpha)\%$  para  $\theta$ .

b. Verifique se  $P(\alpha^{1/n} \leq Y/\theta \leq 1) = 1 - \alpha$  e encontre o IC de  $100(1 - \alpha)\%$  para  $\theta$  com base nessa declaração de probabilidade.

c. Qual dos dois intervalos deduzidos anteriormente é menor? Se meu tempo de espera por um ônibus, pela manhã, for uniformemente distribuído e os tempos de espera observados forem  $x_1 = 4,2$ ,  $x_2 = 3,5$ ,  $x_3 = 1,7$ ,  $x_4 = 1,2$ , e  $x_5 = 2,4$ , encontre o IC de 95% para  $\theta$  usando o intervalo menor.



60. Seja  $0 \leq \gamma \leq \alpha$ . Então, o IC de  $100(1 - \alpha) \%$  para  $\mu$  quando  $n$  for grande é

$$\left( \bar{x} - z_\gamma \cdot \frac{s}{\sqrt{n}}, \bar{x} + z_{\alpha-\gamma} \cdot \frac{s}{\sqrt{n}} \right)$$

A escolha  $\gamma = \alpha/2$  produz o intervalo usual encontrado na Seção 7.2; se  $\gamma \neq \alpha/2$ , esse intervalo não é simétrico em torno de  $\bar{x}$ . A amplitude desse intervalo é  $w = s(z_\gamma + z_{\alpha-\gamma})/\sqrt{n}$ . Mostre que  $w$  é minimizado para  $\gamma = \alpha/2$ , de modo que o intervalo simétrico é o menor. [Sugestões: (a) pela definição de  $z_\alpha$ ,  $\Phi(z_\alpha) = 1 - \alpha$ , de modo que  $z_\alpha = \Phi^{-1}(1 - \alpha)$ ; (b) a relação entre a derivada de uma função  $y = f(x)$  e a função inversa  $x = f^{-1}(y)$  é  $(d/dy)f^{-1}(y) = 1/f'(x)$ .]

61. Suponha que  $x_1, x_2, \dots, x_n$  sejam valores observados que resultam de uma amostra aleatória de distribuição simétrica, porém possivelmente de cauda longa. Sejam  $\bar{x}$  e  $f_s$  a mediana da amostra e a quarta dispersão, respectivamente. O Capítulo 11 de *Understanding Robust and Exploratory Data Analysis* (consulte a bibliografia

do Capítulo 6) sugere o seguinte IC robusto de 95% para a média da população (ponto de simetria):

$$\bar{x} \pm \left( \frac{\text{valor crítico de } t \text{ conservativo}}{1,075} \right) \cdot \frac{f_s}{\sqrt{n}}$$

O valor da quantidade entre parênteses é 2,10 para  $n = 10$ , 1,94 para  $n = 20$  e 1,91 para  $n = 30$ . Calcule esse IC para os dados do Exercício 45, e compare ao IC de  $t$  apropriado para a distribuição de uma população normal.

62. a. Use os resultados do Exemplo 7.5 para obter o limite de confiança inferior de 95% do parâmetro  $\lambda$  de uma distribuição exponencial, e calcule o limite com base nos dados fornecidos no exemplo.
- b. Se a vida útil  $X$  tem uma distribuição exponencial, a probabilidade de a vida útil exceder  $t$  é  $P(X > t) = e^{-\lambda t}$ . Use o resultado da parte (a) para obter o limite de confiança inferior de 95% para a probabilidade de o tempo de quebra exceder 100 min.

---

## Bibliografia

- DEGROOT, Morris e SCHERVISH, Mark, *Probability and Statistics*, Reading, MA: Addison-Wesley, 2002. (3. ed.) Uma exposição muito boa dos princípios gerais de inferência estatística.
- HAHN, Gerald e MEEKER, William, *Statistical Intervals*, Wiley: Nova York, 1991. Tudo que você sempre quis

saber sobre os intervalos estatísticos (confiança, previsão, tolerância e outros).

- LARSEN, Richard e MARX, Morris, *Introduction to Mathematical Statistics* (2. ed.), Prentice Hall, Englewood Cliffs, NJ: 1986. Semelhante à apresentação de DeGroot, mas levemente menos matemático.

# Testes de Hipóteses com Base em uma Única Amostra

## Introdução

Um parâmetro pode ser estimado a partir dos dados da amostra tanto por um único número (uma estimativa pontual) como por um intervalo inteiro de valores plausíveis (um intervalo de confiança). Frequentemente, entretanto, o objetivo de uma investigação não é estimar um parâmetro, mas decidir qual das duas alegações contraditórias sobre o parâmetro está correta. Os métodos de decisão compreendem a parte da inferência estatística chamada *teste de hipóteses*. Neste capítulo, discutimos primeiro alguns dos conceitos básicos e a terminologia usada no teste de hipóteses e depois desenvolvemos os procedimentos de decisão dos problemas de teste de hipóteses encontrados com mais frequência com base em uma amostra de uma única população.

---

## 8.1 | Hipóteses e procedimentos de teste

Uma **hipótese estatística**, ou simplesmente hipótese, é uma alegação ou afirmação sobre o valor de um único parâmetro (característica da população ou característica de uma distribuição de probabilidade), sobre os valores de vários parâmetros ou sobre a forma de uma distribuição de probabilidade inteira. Um exemplo de uma hipótese é a expressão  $\mu = 0,75$ , onde  $\mu$  é o diâmetro interno médio real de certo tipo de cano de PVC. Outro exemplo é a expressão  $p < 0,10$ , onde  $p$  é a proporção de placas de circuito com defeito dentre todas as placas de circuito produzidas por um determinado fabricante. Se  $\mu_1$  e  $\mu_2$  representam as tensões de quebra médias reais de dois tipos diferentes de barbante, uma hipótese é a expressão  $\mu_1 - \mu_2 = 0$ , e outra é a de que  $\mu_1 - \mu_2 > 5$ . Ainda outro exemplo de hipótese é a afirmação de que a distância de freada sob condições específicas tem distribuição normal. Hipóteses desse tipo serão consideradas no Capítulo 14. Neste e em alguns próximos nos concentraremos nas hipóteses sobre parâmetros.

Em qualquer problema de teste de hipóteses, existem duas suposições contraditórias em consideração. Uma hipótese pode ser a definição  $\mu = 0,75$ , e a outra,  $\mu \neq 0,75$ , ou as duas expressões contraditórias podem

ser  $p \geq 0,10$  e  $p < 0,10$ . O objetivo é decidir, com base nas informações da amostra, qual das duas hipóteses está correta. Há uma analogia familiar a isso em um processo criminal. Uma alegação é a afirmação de que o acusado é inocente. No sistema judiciário dos Estados Unidos, essa justificativa é a considerada inicialmente verdadeira. Somente com a presença de forte evidência do contrário é que o júri deve desprezar tal alegação em favor da afirmação alternativa de que o acusado é culpado. Nesse sentido, a alegação de inocência é a hipótese favorecida ou protegida, e o ônus da prova recai sobre aquele que acredita na explicação alternativa.

De maneira semelhante, no teste de hipóteses estatísticas, o problema será formulado de modo que uma das alegações seja inicialmente favorecida. Tal alegação não será rejeitada em favor da alegação alternativa, a menos que a evidência da amostra a contradiga e forneça forte apoio à afirmação alternativa.

#### DEFINIÇÃO

A **hipótese nula**, representada por  $H_0$ , é a alegação inicialmente assumida como verdadeira (a suposição de afirmação de prioridade). A **hipótese alternativa** representada por  $H_a$  é a afirmação contraditória a  $H_0$ .

A hipótese nula será rejeitada em favor da hipótese alternativa somente se a evidência da amostra sugerir que  $H_0$  seja falsa. Se a amostra não contradisser fortemente  $H_0$ , continuaremos a acreditar na verdade da hipótese nula. As duas conclusões possíveis de uma análise do teste de hipóteses são, então, *rejeitar  $H_0$*  ou *não rejeitar  $H_0$* .

Um **teste de hipóteses** é um método para usar os dados da amostra para decidir se a hipótese nula deve ser desprezada. Dessa forma, podemos testar  $H_0: \mu = 0,75$  contra a alternativa  $H_a: \mu \neq 0,75$ . Somente se os dados da amostra sugerirem fortemente que  $\mu$  é qualquer coisa exceto 0,75 é que a hipótese nula deve ser rejeitada. Na ausência de tal evidência,  $H_0$  não deve ser rejeitada, uma vez que ainda é bastante plausível.

Às vezes, um investigador não quer aceitar uma afirmação específica a menos e até que os dados possam fornecer forte apoio à afirmação. Como exemplo, suponha que uma empresa planeja colocar um novo tipo de revestimento nos mancais que produz. A durabilidade média real com o revestimento atual é de 1000 horas. Com  $\mu$  representando a duração média real do novo revestimento, a empresa não gostaria de fazer uma mudança, a menos que uma evidência sugerisse fortemente que  $\mu$  excede 1000. A formulação apropriada de um problema envolveria o teste  $H_0: \mu = 1000$  contra  $H_a: \mu > 1000$ . A conclusão de que uma mudança é justificada é identificada por  $H_a$ , e necessitaria evidência conclusiva para justificar a recusa de  $H_0$  e a mudança para o novo revestimento.

A pesquisa científica geralmente compreende a tentativa de decidir se uma teoria atual deve ser substituída por uma explicação mais satisfatória e plausível do fenômeno sob investigação. Uma abordagem conservadora é identificar a teoria atual com  $H_0$  e a explicação alternativa do investigador com  $H_a$ . A rejeição da teoria atual ocorrerá, então, somente quando a evidência for muito mais consistente com a nova teoria. Em muitas situações,  $H_a$  é considerada a “hipótese do investigador”, uma vez que é a alegação que o investigador realmente gostaria de confirmar. A palavra *nula* significa “sem valor, efeito ou conseqüência”, que sugere que  $H_0$  deve ser identificada com a hipótese de não haver mudança (da opinião atual), diferença, melhora e assim por diante. Suponha, por exemplo, que 10% de todas as placas de circuito produzidas por um certo fabricante durante um período recente estivessem com defeito. Um engenheiro sugeriu uma mudança no processo de produção, acreditando que resultará em um índice reduzido de defeito. Seja  $p$  a proporção real de placas com defeito resultante do processo alterado. Então, a hipótese da pesquisa, na qual o ônus da prova é colocado, é a afirmação de que  $p < 0,10$ . Dessa forma, a hipótese alternativa é  $H_a: p < 0,10$ .

Em nosso tratamento do teste de hipóteses,  $H_0$  sempre será definida como uma expressão de igualdade. Se  $\theta$  representar o parâmetro de interesse, a hipótese nula terá a forma  $H_0: \theta = \theta_0$ , onde  $\theta_0$  é um valor especificado chamado *valor nulo* do parâmetro (valor definido para  $\theta$  pela hipótese nula). Como exemplo, considere a situação da placa de circuito recém-discutida. A hipótese alternativa sugerida foi  $H_a: p < 0,10$ , a alegação de que o índice de defeitos foi reduzido pela modificação do processo. A escolha natural de  $H_0$  nesta situação é a alegação de que  $p \geq 0,10$ , de acordo com a qual o novo processo ou não é melhor ou é pior do que o utilizado atualmente. Em vez disso, consideraremos  $H_0: p = 0,10$  versus  $H_a: p < 0,10$ . O raciocínio para usar essa hipótese nula simplificada é que qualquer procedimento de decisão razoável para escolher entre  $H_0: p = 0,10$  e

$H_a: p < 0,10$  também será razoável para escolher entre  $p \geq 0,10$  e  $H_a$ . O uso de  $H_0$  simplificada é preferido porque tem certos benefícios técnicos que serão mostrados mais à frente.

A alternativa para a hipótese nula  $H_0: \theta = \theta_0$  será semelhante a uma das três afirmações a seguir: (1)  $H_a: \theta > \theta_0$  (caso em que a hipótese nula implícita é  $\theta \leq \theta_0$ ), (2)  $H_a: \theta < \theta_0$  (assim, a hipótese nula implícita afirma que  $\theta \geq \theta_0$ ), ou (3)  $H_a: \theta \neq \theta_0$ . Por exemplo, seja  $\sigma$  o desvio padrão da distribuição dos diâmetros internos (polegadas) de um certo tipo de manga metálica. Se a decisão tomada foi usar a manga, a menos que a evidência da amostra demonstrasse conclusivamente que  $\sigma > 0,001$ , a hipótese apropriada seria  $H_0: \sigma = 0,001$  versus  $H_a: \sigma > 0,001$ . O número  $\theta_0$  que aparece em  $H_0$  e  $H_a$  (separa a alternativa da nula) é chamado de **valor nulo**.

## Procedimentos do teste

O procedimento de um teste é uma regra, com base nos dados da amostra, para decidir se  $H_0$  será rejeitada. Um teste do tipo  $H_0: p = 0,10$  versus  $H_a: p < 0,10$  no problema da placa de circuito pode ser feito com base no exame de uma amostra aleatória de  $n = 200$  placas. Seja  $X$  o número de placas com defeito na amostra, uma variável aleatória binomial,  $x$  representa o valor observado de  $X$ . Se  $H_0$  for verdadeira,  $E(X) = np = 200(0,10) = 20$ , enquanto podemos esperar menos de 20 placas com defeito se  $H_a$  for real. Um valor  $x$  apenas um pouco abaixo de 20 não contradiz fortemente  $H_0$ , de modo que é razoável rejeitar  $H_0$  somente se  $x$  for substancialmente menor que 20. Um procedimento de tal teste é rejeitar  $H_0$  se  $x \leq 15$  e, por outro lado, não rejeitar  $H_0$ . Esse procedimento tem dois elementos: (1) uma estatística de teste ou função dos dados da amostra usada para tomar uma decisão e (2) uma região de rejeição que consiste nesses valores de  $x$  para os quais  $H_0$  será rejeitada em favor de  $H_a$ . Pela regra que acabamos de sugerir, a região de rejeição consiste em  $x = 0, 1, 2, \dots$  e 15.  $H_0$  não será rejeitada se  $x = 16, 17, \dots, 199$  ou 200.

Um procedimento do teste é especificado pelo seguinte:

1. Uma **estatística de teste**, uma função dos dados da amostra na qual a decisão (rejeitar  $H_0$  ou não rejeitar  $H_0$ ) se baseia
2. Uma **região de rejeição**, o conjunto de todos os valores estatísticos do teste para os quais  $H_0$  será rejeitada

A hipótese nula será então rejeitada se, e somente se, o valor estatístico do teste calculado ou observado cair na região de rejeição.

Como outro exemplo, suponha que um fabricante de cigarros alegue que o teor médio de nicotina  $\mu$  dos cigarros da marca B é de (no máximo) 1,5 mg. Seria insensato rejeitar a alegação do fabricante sem ter forte evidência contraditória, de modo que a formulação apropriada do problema é testar  $H_0: \mu = 1,5$  versus  $H_a: \mu > 1,5$ . Considere a regra de decisão com base na análise de uma amostra aleatória de 32 cigarros. Seja  $\bar{X}$  o teor médio de nicotina da amostra. Se  $H_0$  for verdadeira,  $E(\bar{X}) = \mu = 1,5$ , enquanto se  $H_0$  for falsa, esperamos que  $\bar{X}$  exceda 1,5. Uma forte evidência contra  $H_0$  é fornecida por um valor  $\bar{x}$  que excede consideravelmente 1,5. Dessa forma, podemos usar  $\bar{X}$  como uma estatística de teste com a região de rejeição  $\bar{x} \geq 1,6$ .

Nos exemplos da placa de circuito e da nicotina, a escolha da estatística do teste e a forma da região de rejeição fazem sentido intuitivamente. Entretanto, a opção pelo valor de corte usado para especificar a região de rejeição é um tanto arbitrária. Em vez de rejeitar  $H_0: p = 0,10$  em favor de  $H_a: p < 0,10$  quando  $x \leq 15$ , podemos usar a região de rejeição  $x \leq 14$ . Para essa região,  $H_0$  não seria rejeitada se 15 placas com defeito fossem observadas; essa ocorrência levaria à rejeição de  $H_0$  se a região inicialmente sugerida fosse empregada. De maneira semelhante, a região de rejeição  $\bar{x} \geq 1,55$  pode ser usada no problema de nicotina no lugar da região  $\bar{x} \geq 1,60$ .

## Erros em Testes de Hipóteses

A base para escolher uma região de rejeição específica está na compreensão dos erros que alguém pode encontrar ao tirar uma conclusão. Considere a região de rejeição  $x \leq 15$  no problema da placa de circuito. Mesmo

quando  $H_0: p = 0,10$  for verdadeira, pode acontecer de uma amostra incomum resultar em  $x = 13$ , de modo que  $H_0$  seja erroneamente rejeitada. Por outro lado, mesmo quando  $H_a: p < 0,10$  é verdadeira, uma amostra incomum pode produzir  $x = 20$ , e nesse caso  $H_0$  não seria recusada, novamente uma conclusão incorreta. Dessa forma, é possível que  $H_0$  possa ser rejeitada quando for verdadeira ou que  $H_0$  possa não ser rejeitada quando for falsa. Esses erros possíveis não são consequência de uma região de rejeição insensatamente escolhida. Qualquer um desses dois erros pode acontecer quando a região  $x \leq 14$  é empregada ou na verdade quando qualquer outra região for utilizada.

## DEFINIÇÃO

Um **erro tipo I** consiste em rejeitar a hipótese nula  $H_0$  quando ela é verdadeira.

Um **erro tipo II** envolve a não-rejeição de  $H_0$  quando  $H_0$  é falsa.

No problema da nicotina, um erro tipo I consiste em rejeitar a alegação do fabricante de que  $\mu = 1,5$  quando ele for realmente verdadeiro. Se a região de rejeição  $\bar{x} \geq 1,6$  for empregada, pode acontecer de  $\bar{x} = 1,63$  mesmo quando  $\mu = 1,5$ , resultando em um erro tipo I. De forma alternativa, pode ser que  $H_0$  seja falsa e que ainda  $\bar{x} = 1,52$  seja observado, levando à não-rejeição de  $H_0$  (um erro tipo II).

No melhor de todos os mundos possíveis, os procedimentos de teste, para os quais nenhum tipo de erro é possível, poderiam ser desenvolvidos. Entretanto esse ideal pode ser alcançado somente tomando uma decisão com base em um exame da população inteira, que é quase sempre impraticável. A dificuldade com o uso de um procedimento baseado nos dados da amostra é que, devido à variabilidade da amostragem, pode resultar uma amostra atípica. Apesar de  $E(\bar{X}) = \mu$ , o valor observado  $\bar{x}$  pode diferir substancialmente de  $\mu$  (pelo menos se  $n$  for pequeno). Dessa forma, quando  $\mu = 1,5$  na situação da nicotina,  $\bar{x}$  pode ser muito maior que 1,5, resultando em rejeição errônea de  $H_0$ . De maneira alternativa, pode ser que  $\mu = 1,6$ , ainda que seja observado um  $\bar{x}$  muito menor que esse, levando a um erro tipo II.

Em vez de exigir procedimentos sem erro, devemos procurar procedimentos em que os dois tipos de erro sejam improváveis de ocorrer. Isto é, um bom procedimento é aquele para o qual a probabilidade de cometer qualquer tipo de erro é pequena. A escolha do valor de corte de uma região de rejeição específica determina as probabilidades de erros dos tipos I e II. Essas probabilidades de erro são tradicionalmente representadas por  $\alpha$  e  $\beta$ , respectivamente. Em virtude de  $H_0$  especificar um único valor do parâmetro, existe um único valor de  $\alpha$ . Entretanto, existe um valor diferente de  $\beta$  para cada valor do parâmetro consistente com  $H_a$ .

## Exemplo 8.1

Sabe-se que certo tipo de automóvel não sofre nenhum dano visível 25% das vezes nos testes de colisão a 10 mph. Foi proposto um modelo modificado de pára-choque a fim de aumentar essa porcentagem. Seja  $p$  a proporção de todas as colisões a 10 mph com esse novo pára-choque que resultam em nenhum dano visível. As hipóteses a serem testadas são  $H_0: p = 0,25$  (sem melhoria) versus  $H_a: p > 0,25$ . O teste será feito com base em um experimento que envolve  $n = 20$  colisões independentes com protótipos do novo modelo. Intuitivamente,  $H_0$  deve ser rejeitada, se grande número de colisões não mostrar danos. Considere o seguinte procedimento de teste:

Estatística de teste:  $X =$  o número de colisões sem dano visível

Região de rejeição:  $R_8 = \{8, 9, 10, \dots, 19, 20\}$ ; isto é, rejeitar  $H_0$  se  $x \geq 8$ , onde  $x$  é o valor observado da estatística de teste.

Essa região de rejeição é chamada de *cauda superior*, pois consiste somente em valores elevados da estatística de teste.

Quando  $H_0$  é verdadeira,  $X$  possui distribuição de probabilidade binomial com  $n = 20$  e  $p = 0,25$ . Então,

$$\begin{aligned}\alpha &= P(\text{erro tipo I}) = P(H_0 \text{ é rejeitada quando for verdadeira}) \\ &= P(X \geq 8 \text{ quando } X \sim \text{Bin}(20, 0,25)) = 1 - B(7; 20, 0,25) \\ &= 1 - 0,898 = 0,102\end{aligned}$$

Isto é, quando  $H_0$  é realmente verdadeira, aproximadamente 10% de todos os experimentos que consistem em 20 colisões resultariam na rejeição incorreta de  $H_0$  (um erro tipo I).

Em contraste a  $\alpha$ , não existe um único  $\beta$ . Ao contrário, existe um  $\beta$  diferente para cada  $p$  diferente que excede 0,25. Dessa forma, existe um valor de  $\beta$  para  $p = 0,3$  (caso em que  $X \sim \text{Bin}(20, 0,3)$ ), outro valor de  $\beta$  para  $p = 0,5$ , e assim por diante. Por exemplo,

$$\begin{aligned} \beta(0,3) &= P(\text{erro tipo II quando } p = 0,3) \\ &= P(H_0 \text{ não é rejeitada quando for falsa porque } p = 0,3) \\ &= P(X \leq 7 \text{ quando } X \sim \text{Bin}(20,0,3)) = B(7; 20,0,3) = 0,772 \end{aligned}$$

Quando  $p$  for realmente 0,3 em vez de 0,25 (um “pequeno” desvio de  $H_0$ ), aproximadamente 77% de todos os experimentos desse tipo resultarão na não-rejeição, incorreta, de  $H_0$ !

A tabela a seguir exhibe  $\beta$  para valores selecionados de  $p$  (cada um calculado para a região de rejeição  $R_\alpha$ ). Evidentemente,  $\beta$  diminui à medida que o valor de  $p$  muda à direita do valor nulo 0,25. Intuitivamente, quanto maior o desvio de  $H_0$ , menor a probabilidade de tal desvio não ser detectado.

$p$	0,3	0,4	0,5	0,6	0,7	0,8
$\beta(p)$	0,772	0,416	0,132	0,021	0,001	0,000

O procedimento de teste proposto ainda é razoável para testar a hipótese nula mais realística de  $p \leq 0,25$ . Neste caso, não existe mais um único  $\alpha$ , mas sim um  $\alpha$  para cada  $p$  de no máximo 0,25:  $\alpha(0,25)$ ,  $\alpha(0,23)$ ,  $\alpha(0,20)$ ,  $\alpha(0,15)$ , e assim por diante. Verifica-se com facilidade, no entanto, que  $\alpha(p) < \alpha(0,25) = 0,102$  se  $p < 0,25$ . Isto é, o maior valor de  $\alpha$  ocorre para o valor limite de 0,25 entre  $H_0$  e  $H_a$ . Desta forma, se  $\alpha$  for pequeno para a hipótese nula simplificada, também será tão pequeno (ou menor) para a  $H_0$  mais realística. ■

### Exemplo 8.2

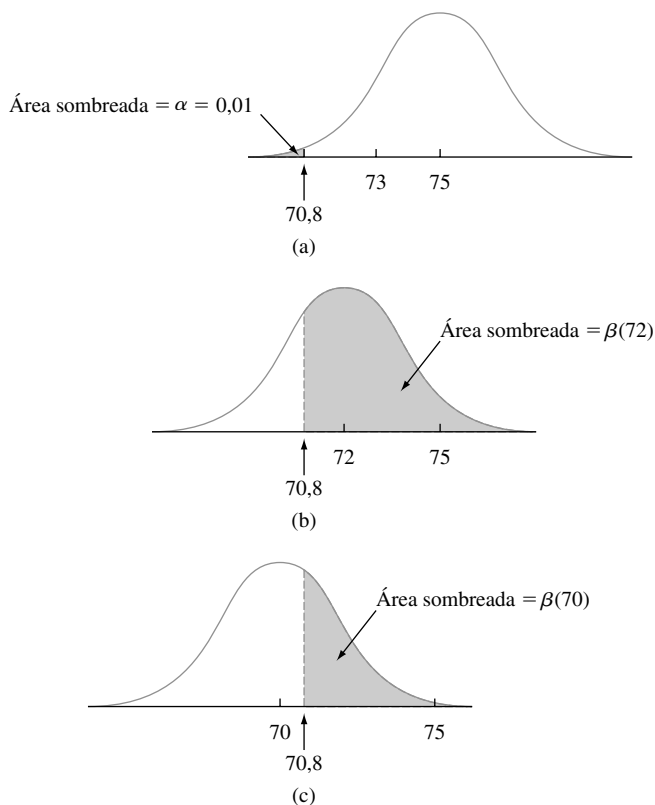
O tempo de secagem de certo tipo de pintura sob condições de teste especificadas é conhecido por ser normalmente distribuído com valor médio de 75 min. e desvio padrão de 9 min. Químicos propuseram um novo aditivo projetado para diminuir o tempo médio de secagem. Acredita-se que os tempos de secagem com esse aditivo permanecerão normalmente distribuídos com  $\sigma = 9$ . Devido à despesa associada ao aditivo, a evidência deve sugerir fortemente uma melhora no tempo médio de secagem antes de tal conclusão ser adotada. Seja  $\mu$  o tempo de secagem médio real quando o aditivo é usado. As hipóteses apropriadas são  $H_0: \mu = 75$  versus  $H_a: \mu < 75$ . Somente se  $H_0$  puder ser recusada é que o aditivo será considerado bem-sucedido e utilizado.

Os dados experimentais consistem em tempos de secagem de  $n = 25$  espécimes de teste. Sejam  $X_1, \dots, X_{25}$  os 25 tempos de secagem – uma amostra aleatória de tamanho 25 de uma distribuição normal com valor médio  $\mu$  e desvio padrão  $\sigma = 9$ . O tempo de secagem médio da amostra  $\bar{X}$ , então, possui distribuição normal com valor esperado  $\mu_{\bar{X}} = \mu$  e desvio padrão  $\sigma_{\bar{X}} = \sigma/\sqrt{n} = 9/\sqrt{25} = 1,80$ . Quando  $H_0$  for verdadeira,  $\mu_{\bar{X}} = 75$ , de modo que um valor de  $\bar{x}$  um tanto menor que 75 não contradiria fortemente  $H_0$ . Uma região de rejeição razoável possui a forma  $\bar{X} \leq c$ , em que o valor de corte  $c$  é escolhido adequadamente. Considere a escolha  $c = 70,8$ , de modo que o procedimento de teste consista em estatística de teste  $\bar{X}$  e região de rejeição  $\bar{x} \leq 70,8$ . Em virtude de a região de rejeição consistir somente em valores pequenos da estatística de teste, o teste é chamado de *cauda inferior*. Os cálculos de  $\alpha$  e  $\beta$  agora envolvem a padronização habitual de  $\bar{X}$  seguida pela referência às probabilidades normais padrão da Tabela A.3 do Apêndice:

$$\begin{aligned} \alpha &= P(\text{erro tipo I}) = P(H_0 \text{ é rejeitada quando for verdadeira}) \\ &= P(\bar{X} \leq 70,8 \text{ quando } \bar{X} \sim \text{normal com } \mu_{\bar{X}} = 75, \sigma_{\bar{X}} = 1,8) \\ &= \Phi\left(\frac{70,8 - 75}{1,8}\right) = \Phi(-2,33) = 0,01 \end{aligned}$$

$$\begin{aligned}
 \beta(72) &= P(\text{erro tipo II quando } \mu = 72) \\
 &= P(H_0 \text{ não é rejeitada quando é falsa porque } \mu = 72) \\
 &= P(\bar{X} > 70,8 \text{ quando } \bar{X} \sim \text{normal com } \mu_{\bar{X}} = 72 \text{ e } \sigma_{\bar{X}} = 1,8) \\
 &= 1 - \Phi\left(\frac{70,8 - 72}{1,8}\right) = 1 - \Phi(-0,67) = 1 - 0,2514 = 0,7486 \\
 \\
 \beta(70) &= 1 - \Phi\left(\frac{70,8 - 70}{1,8}\right) = 0,3300 \quad \beta(67) = 0,0174
 \end{aligned}$$

Para o procedimento de teste especificado, somente 1% de todos os experimentos realizados conforme descritos resultará na rejeição de  $H_0$  quando ela for realmente verdadeira. Entretanto, a chance de um erro tipo II é bastante grande quando  $\mu = 72$  (somente um desvio pequeno de  $H_0$ ), um tanto menor quando  $\mu = 70$ , e bastante pequeno quando  $\mu = 67$  (um desvio muito substancial de  $H_0$ ). Essas probabilidades de erro são ilustradas na Figura 8.1. Observe que  $\alpha$  é calculado usando a distribuição de probabilidade da estatística de teste quando  $H_0$  for verdadeira, enquanto a determinação de  $\beta$  demanda o conhecimento da distribuição da estatística de teste quando  $H_0$  for falsa.



**Figura 8.1**  $\alpha$  e  $\beta$  ilustrados para o Exemplo 8.2: (a) a distribuição de  $\bar{X}$  quando  $\mu = 75$  ( $H_0$  verdadeira); (b) a distribuição de  $\bar{X}$  quando  $\mu = 72$  ( $H_0$  falsa); (c) a distribuição de  $\bar{X}$  quando  $\mu = 70$  ( $H_0$  falsa)

Como no Exemplo 8.1, se a hipótese nula mais realística  $\mu \geq 75$  for considerada, há um  $\alpha$  para cada valor de parâmetro para o qual  $H_0$  é verdadeira:  $\alpha(75)$ ,  $\alpha(75,8)$ ,  $\alpha(76,5)$ , e assim por diante. Verifica-se com facilidade, no entanto, que  $\alpha(75)$  é a maior de todas essas probabilidades de erro tipo I. Enfocar o valor limite equivale a trabalhar explicitamente com o “pior caso”. ■

A especificação de um valor de corte para a região de rejeição nos exemplos apenas considerados foi um tanto arbitrária. O uso da região de rejeição  $R_8 = \{8, 9, \dots, 20\}$  no Exemplo 8.1 resultou em  $\alpha = 0,102$ ,

$\beta(0,3) = 0,772$ , e  $\beta(0,5) = 0,132$ . Muitos considerariam essas probabilidades de erro intoleravelmente grandes. Talvez possam ser reduzidas pela mudança do valor de corte.

**Exemplo 8.3** (continuação do Exemplo 8.1)

Vamos usar o mesmo experimento e testar a estatística  $X$ , conforme descrito anteriormente no problema do pára-choque do automóvel. Agora, entretanto, considere a região de rejeição  $R_9 = \{9, 10, \dots, 20\}$ . Uma vez que  $X$  ainda possua uma distribuição binomial com parâmetros  $n = 20$  e  $p$ ,

$$\begin{aligned}\alpha &= P(H_0 \text{ é rejeitada quando } p = 0,25) \\ &= P(X \geq 9 \text{ quando } X \sim \text{Bin}(20, 0,25)) = 1 - B(8; 20, 0,25) = 0,041\end{aligned}$$

A probabilidade de erro tipo I foi reduzida, usando a nova região de rejeição. Entretanto, pagou-se um preço por essa diminuição:

$$\begin{aligned}\beta(0,3) &= P(H_0 \text{ não é rejeitada quando } p = 0,3) \\ &= P(X \leq 8 \text{ quando } X \sim \text{Bin}(20, 0,3)) = B(8; 20, 0,3) = 0,887 \\ \beta(0,5) &= B(8; 20, 0,5) = 0,252\end{aligned}$$

Ambos  $\beta$ s são maiores que as probabilidades de erro correspondentes 0,772 e 0,132 para a região  $R_8$ . Retrospectivamente, isso não é surpreendente;  $\alpha$  é calculado pela soma das probabilidades dos valores da estatística de teste na *região de rejeição*, enquanto  $\beta$  é a probabilidade de  $X$  estar *no complemento* da região de rejeição. Tornar a região de rejeição menor deve, portanto, diminuir  $\alpha$  enquanto aumenta  $\beta$  para qualquer valor alternativo fixo do parâmetro. ■

**Exemplo 8.4** (continuação do Exemplo 8.2)

O uso do valor de corte  $c = 70,8$  no exemplo da secagem da pintura resultou em um valor muito pequeno de  $\alpha(0,01)$ , mas muito grande de  $\beta$ . Considere o mesmo experimento e teste a estatística  $\bar{X}$  com a nova região de rejeição  $\bar{x} \leq 72$ . Em virtude de  $\bar{X}$  ainda ser normalmente distribuído com valor médio  $\mu_{\bar{X}} = \mu$  e  $\sigma_{\bar{X}} = 1,8$ ,

$$\begin{aligned}\alpha &= P(H_0 \text{ é rejeitada quando é verdadeira}) \\ &= P(\bar{X} \leq 72 \text{ quando } \bar{X} \sim N(75, 1,8^2)) \\ &= \Phi\left(\frac{72 - 75}{1,8}\right) = \Phi(-1,67) = 0,0475 \approx 0,05\end{aligned}$$

$$\begin{aligned}\beta(72) &= P(H_0 \text{ não é rejeitada quando } \mu = 72) \\ &= P(\bar{X} > 72 \text{ quando } \bar{X} \text{ é uma va normal com média } 72 \text{ e desvio padrão } 1,8) \\ &= 1 - \Phi\left(\frac{72 - 72}{1,8}\right) = 1 - \Phi(0) = 0,5\end{aligned}$$

$$\beta(70) = 1 - \Phi\left(\frac{72 - 70}{1,8}\right) = 0,1335 \quad \beta(67) = 0,0027$$

A mudança no valor de corte tornou a região de rejeição maior (inclui mais valores de  $\bar{x}$ ), resultando em uma diminuição de  $\beta$  para cada  $\mu$  fixo menor que 75. Entretanto,  $\alpha$  para essa nova região aumentou do valor anterior 0,01 para aproximadamente 0,05. Se uma probabilidade de erro tipo I dessa grandeza puder ser tolerada, apesar da segunda região ( $c = 72$ ), é preferível à primeira ( $c = 70,8$ ), devido aos  $\beta$ s menores. ■

Os resultados desses exemplos podem ser generalizados da seguinte maneira:



## PROPOSIÇÃO

Suponha que um experimento e o tamanho de uma amostra sejam fixos, e que seja escolhida uma estatística de teste. Então, reduzir o tamanho da região de rejeição para obter um valor menor de  $\alpha$  resulta em um valor maior de  $\beta$  para qualquer valor de parâmetro específico consistente com  $H_a$ .

Essa proposição diz que, uma vez que a estatística de teste e  $n$  forem fixados, não há região de rejeição que tornará, simultaneamente,  $\alpha$  e todos os  $\beta$ s pequenos. Uma região deve ser escolhida para cumprir um compromisso entre  $\alpha$  e  $\beta$ .

Devido às diretrizes sugeridas para especificar  $H_0$  e  $H_a$  um erro tipo I geralmente é mais sério que um erro tipo II (isso sempre pode ser alcançado pela escolha apropriada das hipóteses). A abordagem adotada pela maioria dos estatísticos é especificar o valor maior de  $\alpha$  que pode ser tolerado e encontrar uma região de rejeição que tenha esse valor de  $\alpha$  em vez de qualquer outro menor. Isso torna  $\beta$  o menor possível, sujeito ao limite em  $\alpha$ . O valor resultante de  $\alpha$  geralmente é denominado **nível de significância** do teste. Os níveis tradicionais de significância são 0,10, 0,05 e 0,01, embora o nível em qualquer problema específico dependa da seriedade de um erro tipo I – quanto mais sério for esse erro, menor deve ser o nível de significância. O procedimento de teste correspondente é chamado de **teste nível  $\alpha$**  (por exemplo, um teste nível 0,05 ou um teste nível 0,01). Um teste com nível de significância  $\alpha$  é aquele para o qual a probabilidade de erro tipo I é controlada no nível especificado.

## Exemplo 8.5

Considere a situação mencionada anteriormente em que  $\mu$  era o teor de nicotina médio real de cigarros da marca B. O objetivo é testar  $H_0: \mu = 1,5$  versus  $H_a: \mu > 1,5$  com base em uma amostra aleatória  $X_1, X_2, \dots, X_{32}$  de teor de nicotina. Suponha que a distribuição do teor de nicotina seja conhecida como normal com  $\sigma = 0,20$ . Então,  $\bar{X}$  é normalmente distribuído com valor médio  $\mu_{\bar{X}} = \mu$  e desvio padrão  $\sigma_{\bar{X}} = 0,20/\sqrt{32} = 0,0354$ .

Em vez de usar o próprio  $\bar{X}$  como a estatística de teste, vamos padronizar  $\bar{X}$ , assumindo que  $H_0$  seja verdadeira.

$$\text{Estatística de teste: } Z = \frac{\bar{X} - 1,5}{\sigma/\sqrt{n}} = \frac{\bar{X} - 1,5}{0,0354}$$

$Z$  expressa a distância entre  $\bar{X}$  e seu valor esperado quando  $H_0$  é verdadeira como certo número de desvios padrão. Por exemplo,  $z = 3$  resulta de um  $\bar{x}$  que é 3 desvios padrão maior que o valor que esperávamos, caso  $H_0$  fosse verdadeira.

Rejeitar  $H_0$  quando  $\bar{x}$  excede “consideravelmente” 1,5 equivale a rejeitar  $H_0$  quando  $z$  excede “consideravelmente” 0. Isto é, a forma da região de rejeição é  $z \geq c$ . Vamos agora determinar  $c$ , de modo que  $\alpha = 0,05$ . Quando  $H_0$  é verdadeira,  $Z$  possui uma distribuição normal padronizada. Assim,

$$\begin{aligned} \alpha &= P(\text{erro tipo I}) = P(\text{rejeitar } H_0 \text{ quando } H_0 \text{ é verdadeira}) \\ &= P(Z \geq c \text{ quando } Z \sim N(0, 1)) \end{aligned}$$

O valor  $c$  deve incluir a área da cauda superior 0,05 sob a curva  $z$ . Tanto da Seção 4.3 como diretamente da Tabela A.3 do Apêndice,  $c = z_{0,05} = 1,645$ .

Observe que  $z \geq 1,645$  é equivalente a  $\bar{x} - 1,5 \geq (0,0354)(1,645)$ , isto é,  $\bar{x} \geq 1,56$ . Então,  $\beta$  é a probabilidade de  $\bar{X} < 1,56$  e pode ser calculado para qualquer  $\mu$  maior do que 1,5. ■

**Exercícios** | Seção 8.1 (1–14)

1. Para cada afirmação a seguir, determine se é uma hipótese estatística legítima e por quê:
  - a.  $H: \sigma > 100$       b.  $H: \tilde{x} = 45$
  - c.  $H: s \leq 0,20$       d.  $H: \sigma_1/\sigma_2 < 1$
  - e.  $H: \bar{X} - \bar{Y} = 5$
  - f.  $H: \lambda \leq 0,01$ , onde  $\lambda$  é o parâmetro de uma distribuição exponencial usada para modelar a vida útil do componente.
2. Para os pares de afirmações a seguir, indique qual não satisfaz às nossas regras de definição de hipóteses e por quê. (Os subscritos 1 e 2 diferenciam-se entre as quantidades das duas populações ou amostras diferentes.)
  - a.  $H_0: \mu = 100, H_a: \mu > 100$
  - b.  $H_0: \sigma = 20, H_a: \sigma \leq 20$
  - c.  $H_0: p \neq 0,25, H_a: p = 0,25$
  - d.  $H_0: \mu_1 - \mu_2 = 25, H_a: \mu_1 - \mu_2 > 100$
  - e.  $H_0: S_1^2 = S_2^2, H_a: S_1^2 \neq S_2^2$
  - f.  $H_0: \mu = 120, H_a: \mu = 150$
  - g.  $H_0: \sigma_1/\sigma_2 = 1, H_a: \sigma_1/\sigma_2 \neq 1$
  - h.  $H_0: p_1 - p_2 = -0,1, H_a: p_1 - p_2 < -0,1$
3. Para determinar se as soldas dos canos em uma usina de energia nuclear atendem às especificações, uma amostra aleatória de soldas é selecionada e os testes são conduzidos em cada solda na amostra. A resistência da solda é medida como a força necessária para quebrá-la. Suponha que as especificações determinem que a resistência média das soldas deve exceder 100 lb/in<sup>2</sup>; a equipe de inspeção decide testar  $H_0: \mu = 100$  versus  $H_a: \mu > 100$ . Explique o motivo pelo qual se pode preferir usar esta  $H_a$  em vez de  $\mu < 100$ .
4. Seja  $\mu$  o nível de radioatividade médio real (picocuries por litro). O valor 5 pCi/L é considerado a linha divisória entre a água segura e a perigosa. Você recomendaria testar  $H_0: \mu = 5$  versus  $H_a: \mu > 5$  ou  $H_0: \mu = 5$  versus  $H_a: \mu < 5$ ? Explique o seu raciocínio. (*Sugestão*: Pense nas conseqüências de um erro tipo I e tipo II para cada possibilidade.)
5. Antes de concordar em fazer um grande pedido de revestimentos de polietileno para um tipo específico de cabo de força submarino preenchido com óleo de alta pressão, uma empresa deseja ter evidências conclusivas de que o desvio padrão real da espessura do revestimento é menor que 0,05 mm. Quais hipóteses devem ser testadas e por quê? Neste contexto, quais são os erros tipo I e tipo II?
6. Muitas casas antigas possuem sistemas elétricos que utilizam fusíveis em vez de disjuntores. Um fabricante de fusíveis de 40amp quer ter certeza de que a corrente média na qual seus fusíveis queimam é de fato 40. Se a corrente média for menor que esse número, os clientes se queixarão, pois os fusíveis precisam ser substituídos muito cedo. Se a corrente média for maior que 40, o fabricante pode ser responsável pelo dano a um sistema elétrico devido ao mau funcionamento do fusível. Para verificar a corrente dos fusíveis, uma amostra sua será selecionada e inspecionada. Se um teste de hipóteses foi realizado sobre os dados resultantes, quais hipóteses nula e alternativa interessariam ao fabricante? Descreva os erros tipo I e tipo II no contexto da situação deste problema.
7. Amostras de água são coletadas da água usada para resfriamento quando está sendo despejada de uma usina de energia em um rio. Determinou-se que, desde que a temperatura média da água despejada seja de no máximo 150°F, não haverá efeitos negativos sobre o ecossistema do rio. Para investigar se a usina está em conformidade com as regulamentações que proíbem uma temperatura média de água de descarga acima de 150°, 50 amostras de água serão tiradas em tempos selecionados aleatoriamente e a temperatura de cada amostra será registrada. Os dados resultantes serão usados para testar as hipóteses  $H_0: \mu = 150^\circ$  versus  $H_a: \mu > 150^\circ$ . No contexto dessa situação, descreva os erros tipo I e tipo II. Que tipo de erro você consideraria mais sério? Explique.
8. Um tipo regular de laminado está sendo usado atualmente por um fabricante de placas de circuito. Um laminado especial foi desenvolvido para reduzir a deformação. O laminado regular será utilizado em uma amostra de espécimes e o laminado especial em outra amostra; a quantidade de deformação será, então, determinada para cada espécime. O fabricante mudará para o laminado especial somente se puder ser demonstrado que a quantidade média real de deformação desse laminado for menor que a do laminado regular. Determine as hipóteses relevantes e descreva os erros tipo I e tipo II no contexto dessa situação.
9. Duas empresas diferentes concentraram-se em oferecer serviços de televisão a cabo em uma determinada região. Seja  $p$  a proporção de todos os assinantes em potencial que favorecem a primeira empresa com relação à segunda. Considere o teste de  $H_0: p = 0,5$  versus  $H_a: p \neq 0,5$  com base em uma amostra aleatória de 25 indivíduos. Seja  $X$  o número na amostra que favorece a primeira empresa e  $x$  o valor observado de  $X$ .
  - a. Qual das regiões de rejeição a seguir é mais apropriada e por quê?
 
$$R_1 = \{x: x \leq 7 \text{ ou } x \geq 18\}, R_2 = \{x: x \leq 8\},$$

$$R_3 = \{x: x \geq 17\}$$
  - b. No contexto da situação deste problema, descreva quais são os erros tipo I e tipo II.
  - c. Qual é a distribuição de probabilidades da estatística do teste  $X$  quando  $H_0$  for verdadeira? Use-a para calcular a probabilidade de um erro tipo I.

- d. Calcule a probabilidade de um erro tipo II para a região selecionada quando  $p = 0,3$ , novamente quando  $p = 0,4$ , e também para  $p = 0,6$  e  $p = 0,7$ .
- e. Usando a região selecionada, a que conclusão você chegaria se 6 dos 25 questionados favorecessem a empresa 1?
10. Uma mistura de resíduo de combustível pulverizado e cimento Portland a ser usada para reboco deve ter uma resistência à compressão de mais de 1300 KN/m<sup>2</sup>. A mistura não será utilizada a menos que a evidência experimental indique conclusivamente que a especificação da resistência seja satisfeita. Suponha que a resistência à compressão dos espécimes dessa mistura seja normalmente distribuída com  $\sigma = 60$ . Seja  $\mu$  a resistência à compressão média real.
- a. Quais são as hipóteses nula e alternativa apropriadas?
- b. Seja  $\bar{X}$  a resistência à compressão média amostral de  $n = 20$  espécimes selecionados aleatoriamente. Considere o procedimento de teste com estatística de teste  $\bar{X}$  e região de rejeição  $\bar{x} \geq 1331,26$ . Qual é a distribuição de probabilidades da estatística de teste quando  $H_0$  for verdadeira? Qual é a probabilidade de um erro tipo I para o procedimento de teste?
- c. Qual é a distribuição de probabilidades da estatística de teste quando  $\mu = 1350$ ? Usando o procedimento de teste do item (b), qual é a probabilidade de a mistura ser considerada insatisfatória quando de fato  $\mu = 1350$  (um erro tipo II)?
- d. Como você mudaria o procedimento de teste do item (b) para obter um teste com nível de significância de 0,05? Que impacto essa mudança teria na probabilidade de erro do item (c)?
- e. Considere a estatística de teste padronizada  $Z = (\bar{X} - 1300)/(\sigma/\sqrt{n}) = (\bar{X} - 1300)/13,42$ . Quais são os valores de  $Z$  correspondentes à região de rejeição do item (b)?
11. A calibração de uma balança será verificada pela pesagem de um espécime de teste de 10 kg 25 vezes. Suponha que os resultados das diferentes pesagens sejam independentes entre si e que o peso em cada tentativa seja normalmente distribuído com  $\sigma = 0,200$  kg. Seja  $\mu$  a leitura do peso médio real da balança.
- a. Quais hipóteses devem ser testadas?
- b. Suponha que a balança seja recalibrada se  $\bar{x} \geq 10,1032$  ou  $\bar{x} \leq 9,8968$ . Qual é a probabilidade de a recalibração ser realizada quando for realmente desnecessária?
- c. Qual é a probabilidade de a recalibração ser considerada desnecessária quando na verdade  $\mu = 10,1$ ? E quando  $\mu = 9,8$ ?
- d. Seja  $z = (\bar{x} - 10)/(\sigma/\sqrt{n})$ . Para que valor de  $c$  a região de rejeição do item (b) é equivalente à região "bicaudal"  $z \geq c$  ou  $z \leq -c$ ?
- e. Se o tamanho da amostra fosse somente 10 em vez de 25, como o procedimento do item (d) deve ser alterado de modo que  $\alpha = 0,05$ ?
- f. Usando o teste do item (e), que conclusão você tiraria dos dados da amostra a seguir:
- |       |        |        |        |       |
|-------|--------|--------|--------|-------|
| 9,981 | 10,006 | 9,857  | 10,107 | 9,888 |
| 9,728 | 10,439 | 10,214 | 10,190 | 9,793 |
- g. Reescreva o procedimento de teste da parte (b) em termos da estatística de teste padronizada  $Z = (\bar{X} - 10)/(\sigma/\sqrt{n})$ .
12. Foi proposto um novo modelo para o sistema de frenagem em um determinado tipo de carro. Para o sistema atual, a distância de frenagem média real em 40 mph sob condições específicas vale 120 pés. Foi proposto que o novo modelo fosse implementado somente se os dados da amostra indicassem fortemente uma redução da distância de frenagem média real para o novo modelo.
- a. Defina o parâmetro de interesse e determine as hipóteses relevantes.
- b. Suponha que a distância de frenagem do novo sistema seja normalmente distribuída com  $\sigma = 10$ . Seja  $\bar{X}$  a distância de frenagem média amostral de uma amostra aleatória de 36 observações. Qual das regiões de rejeição a seguir é apropriada:  $R_1 = \{\bar{x}: \bar{x} \geq 124,80\}$ ,  $R_2 = \{\bar{x}: \bar{x} \leq 115,20\}$ ,  $R_3 = \{\bar{x}: \text{qualquer } \bar{x} \geq 125,13 \text{ ou } \bar{x} \leq 114,87\}$ ?
- c. Qual é o nível de significância para a região apropriada do item (b)? Como você mudaria a região para obter um teste com  $\alpha = 0,001$ ?
- d. Qual é a probabilidade de o novo modelo não ser implementado quando sua distância de frenagem média real for realmente de 115 pés e a região apropriada do item (b) for usada?
- e. Seja  $Z = (\bar{X} - 120)/(\sigma/\sqrt{n})$ . Qual é o nível de significância para a região de rejeição  $\{z: z \leq -2,33\}$ ? Para a região  $\{z: z \leq -2,88\}$ ?
13. Seja  $X_1, \dots, X_n$  uma amostra aleatória de uma distribuição de uma população normal com um valor conhecido de  $\sigma$ .
- a. Para testar as hipóteses  $H_0: \mu = \mu_0$  versus  $H_a: \mu = \mu_0$  (onde  $\mu_0$  é um número fixo), mostre que o teste com estatística de teste  $\bar{X}$  e região de rejeição  $\bar{x} \geq \mu_0 + 2,33\sigma/\sqrt{n}$  possui nível de significância 0,01.
- b. Suponha que o procedimento do item (a) seja usado para testar  $H_0: \mu \leq \mu_0$  versus  $H_a: \mu > \mu_0$ . Se  $\mu_0 = 100$ ,  $n = 25$ , e  $\sigma = 5$ , qual é a probabilidade de se cometer um erro tipo I quando  $\mu = 99$ ? Quando  $\mu = 98$ ? Em geral, o que se pode dizer sobre a probabilidade de um erro tipo I quando o valor real de  $\mu$  for menor que  $\mu_0$ ? Prove sua afirmação.
14. Reconsidere a situação do Exercício 11 e suponha que a região de rejeição seja  $\{\bar{x}: \bar{x} \geq 10,1004 \text{ ou } \bar{x} \leq 9,8940\} = \{z: z \geq 2,51 \text{ ou } z \leq -2,65\}$ .
- a. Qual é o  $\alpha$  para esse procedimento?
- b. Qual o  $\beta$  quando  $\mu = 10,1$ ? Quando  $\mu = 9,9$ ? Isso é desejável?

## 8.2 Testes Quanto à Média de uma População

A discussão geral no Capítulo 7 sobre intervalos de confiança da média de uma população  $\mu$  focalizou três casos diferentes. Agora, desenvolvemos procedimentos de teste para esses três casos.

### Caso 1: População normal com $\sigma$ conhecido

Embora a suposição de que o valor de  $\sigma$  é conhecido seja raramente encontrada na prática, esse caso fornece um bom ponto de partida por causa da facilidade com que os procedimentos gerais e suas propriedades podem ser desenvolvidos. A hipótese nula nos três casos dirá que  $\mu$  possui um valor numérico específico, o *valor nulo*, que denotaremos por  $\mu_0$ . Seja  $X_1, \dots, X_n$  uma amostra aleatória de tamanho  $n$  da população normal. Então, a média amostral  $\bar{X}$  possui uma distribuição normal com valor esperado  $\mu_{\bar{X}} = \mu$  e desvio padrão  $\sigma_{\bar{X}} = \sigma/\sqrt{n}$ . Quando  $H_0$  é verdadeira,  $\mu_{\bar{X}} = \mu_0$ . Considere agora a estatística  $Z$  obtida pela padronização de  $\bar{X}$  sob a suposição de  $H_0$  ser verdadeira:

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

A substituição da média amostral calculada  $\bar{x}$  fornece  $z$ , a distância entre  $\bar{x}$  e  $\mu_0$  expressa em “unidades de desvio padrão”. Por exemplo, se a hipótese nula for  $H_0: \mu = 100$ ,  $\sigma_{\bar{X}} = \sigma/\sqrt{n} = 10/\sqrt{25} = 2,0$  e  $\bar{x} = 103$ , então, o valor da estatística de teste é  $z = (103 - 100)/2,0 = 1,5$ . Isto é, o valor observado de  $\bar{x}$  está 1,5 desvio padrão (de  $\bar{X}$ ) acima do que esperamos quando  $H_0$  é verdadeira. A estatística  $Z$  é uma medida natural da distância entre  $\bar{X}$ , o estimador de  $\mu$ , e seu valor esperado quando  $H_0$  é verdadeira. Se essa distância for muito grande em uma direção consistente com  $H_a$ , a hipótese nula deve ser rejeitada.

Suponha primeiro que a hipótese alternativa tenha a forma  $H_a: \mu > \mu_0$ . Então, um valor de  $\bar{x}$  menor que  $\mu_0$  certamente não oferece apoio a  $H_a$ . Esse  $\bar{x}$  corresponde a um valor negativo de  $z$  (uma vez que  $\bar{x} - \mu_0$  é negativo e o divisor  $\sigma/\sqrt{n}$ , positivo). De maneira semelhante, um valor de  $\bar{x}$  que excede  $\mu_0$  somente por um valor pequeno (correspondente a  $z$ , que é positivo, porém pequeno) não sugere que  $H_0$  deva ser rejeitada em favor de  $H_a$ . A rejeição de  $H_0$  é apropriada somente quando  $\bar{x}$  excede consideravelmente  $\mu_0$ , isto é, quando o valor de  $z$  é positivo e grande. Resumindo, a região de rejeição apropriada, com base na estatística de teste  $Z$  em vez de  $\bar{X}$ , possui a forma  $z \geq c$ .

Conforme discutido na Seção 8.1, o valor de corte  $c$  deve ser escolhido para controlar a probabilidade de um erro tipo I no nível desejado de  $\alpha$ . Isso é feito com facilidade porque a distribuição da estatística de teste  $Z$  quando  $H_0$  é verdadeira é a distribuição normal padronizada (é por isso que  $\mu_0$  foi subtraído na padronização). O corte requerido  $c$  é o valor crítico de  $z$  que inclui a área da cauda superior  $\alpha$  sob a curva normal padronizada. Como exemplo, seja  $c = 1,645$  o valor que inclui a área da cauda 0,05 ( $z_{0,05} = 1,645$ ). Então,

$$\begin{aligned} \alpha &= P(\text{erro tipo I}) = P(H_0 \text{ é rejeitada quando } H_0 \text{ é verdadeira}) \\ &= P(Z \geq 1,645 \text{ quando } Z \sim N(0, 1)) = 1 - \Phi(1,645) = 0,05 \end{aligned}$$

Sob forma mais geral, a região de rejeição  $z \geq z_\alpha$  possui probabilidade  $\alpha$  de erro tipo I. O procedimento de teste é de *cauda superior* porque a região de rejeição consiste somente em valores grandes da estatística de teste.

Raciocínio análogo para a hipótese alternativa  $H_a: \mu < \mu_0$  sugere uma região de rejeição da forma  $z \leq c$ , em que  $c$  é um número negativo escolhido adequadamente ( $\bar{x}$  está muito abaixo de  $\mu_0$  se, e somente se,  $z$  for negativo). Em virtude de  $Z$  ter distribuição normal padronizada quando  $H_0$  é verdadeira, assumir  $c = -z_\alpha$  produz  $P(\text{erro tipo I}) = \alpha$ . Esse é um teste de *cauda inferior*. Por exemplo,  $z_{0,10} = 1,28$  implica que a região de rejeição  $z \leq -1,28$  especifica um teste com nível de significância 0,10.

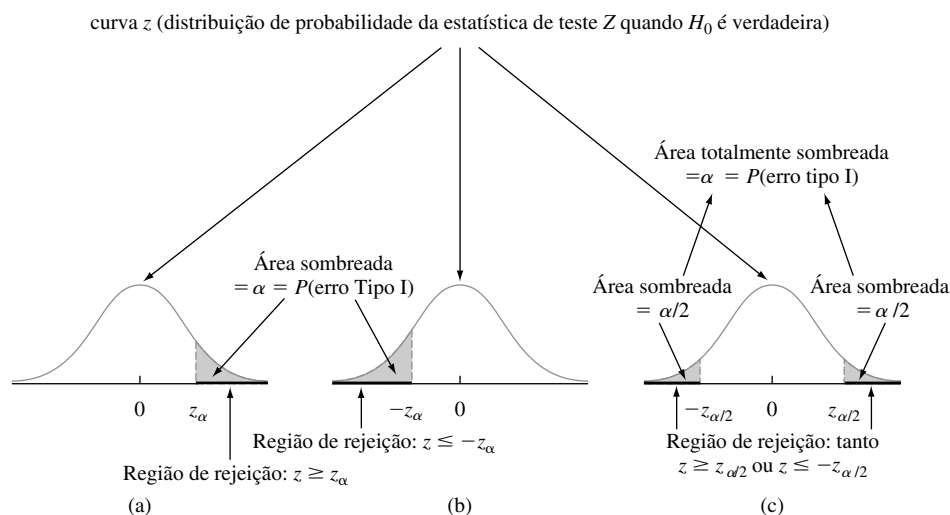
Finalmente, quando a hipótese alternativa for  $H_a: \mu \neq \mu_0$ ,  $H_0$  deve ser rejeitada se  $\bar{x}$  estiver muito longe de qualquer lado de  $\mu_0$ . O que equivale a rejeitar  $H_0$  se  $z \geq c$  ou se  $z \leq -c$ . Suponha que desejemos  $\alpha = 0,05$ . Então,

$$\begin{aligned} 0,05 &= P(Z \geq c \text{ ou } Z \leq -c \text{ quando } Z \text{ possui uma distribuição normal padronizada}) \\ &= \Phi(-c) + 1 - \Phi(c) = 2[1 - \Phi(c)] \end{aligned}$$

Dessa forma,  $c$  é tal que  $1 - \Phi(c)$ , a área sob a curva normal padronizada à direita de  $c$  é 0,025 (e não 0,05!). A partir da Seção 4.3 ou da Tabela A.3 do Apêndice,  $c = 1,96$ , e a região de rejeição é  $z \geq 1,96$  ou  $z \leq -1,96$ . Para qualquer  $\alpha$ , a região de rejeição *bicaudal*  $z \geq z_{\alpha/2}$  ou  $z \leq -z_{\alpha/2}$  possui probabilidade  $\alpha$  de erro tipo I (uma vez que a área  $\alpha/2$  está incluída sob cada uma das duas caudas da curva  $z$ ). Novamente, o motivo principal para usar a estatística de teste padronizada  $Z$  é que, em virtude de  $Z$  ter distribuição conhecida quando  $H_0$  é verdadeira (normal padronizada), uma região de rejeição com probabilidade de erro desejado tipo I é facilmente obtida, usando-se um valor crítico apropriado.

O procedimento de teste para o caso I está resumido no quadro a seguir, e as regiões de rejeição correspondentes estão ilustradas na Figura 8.2.

Hipótese nula: $H_0: \mu = \mu_0$	
Valor da estatística de teste: $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$	
<b>Hipótese alternativa</b>	<b>Região de rejeição para o teste nível <math>\alpha</math></b>
$H_a: \mu > \mu_0$	$z \geq z_{\alpha}$ (teste de cauda superior)
$H_a: \mu < \mu_0$	$z \leq -z_{\alpha}$ (teste de cauda inferior)
$H_a: \mu \neq \mu_0$	tanto $z \geq z_{\alpha/2}$ ou $z \leq -z_{\alpha/2}$ (teste bicaudal)



**Figura 8.2** Regiões de rejeição para os testes  $z$ : (a) teste de cauda superior; (b) teste de cauda inferior; (c) teste bicaudal

O uso da seguinte seqüência de passos é recomendado ao testar as hipóteses de um parâmetro.

1. Identifique o parâmetro de interesse e descreva-o no contexto do problema.
2. Determine o valor nulo e a hipótese nula.
3. Especifique a hipótese alternativa apropriada.
4. Dê a fórmula para o valor calculado da estatística de teste (substituindo o valor nulo e os valores conhecidos de quaisquer outros parâmetros, mas *não* os parâmetros com base na amostra).
5. Especifique a região de rejeição para o nível de significância selecionado  $\alpha$ .
6. Calcule quaisquer quantidades necessárias da amostra, substitua na fórmula para o valor da estatística de teste e calcule esse valor.
7. Decida se  $H_0$  deve ser rejeitada e enuncie a conclusão no contexto do problema.

A formulação das hipóteses (Passos 2 e 3) deve ser feita antes de examinar os dados.

**Exemplo 8.6**

Um fabricante de sistemas anti-incêndio por aspersão, usados para proteção contra incêndio em edifícios de escritórios, alega que a temperatura média real de ativação do sistema é 130°. Uma amostra de  $n = 9$  sistemas, quando testados, produz uma temperatura média amostral de ativação de 131,08°F. Se a distribuição dos tempos de ativação fosse normal com desvio padrão de 1,5°F, os dados contradiriam a alegação do fabricante com nível de significância  $\alpha = 0,01$ ?

1. Parâmetro de interesse:  $\mu =$  temperatura média real de ativação.
2. Hipótese nula:  $H_0: \mu = 130$  (valor nulo =  $\mu_0 = 130$ ).
3. Hipótese alternativa:  $H_a: \mu \neq 130$  (um desvio do valor alegado em *qualquer* direção é preocupante).
4. Valor da estatística de teste:

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{x} - 130}{1,5/\sqrt{n}}$$

5. Região de rejeição: a forma de  $H_a$  implica o uso de um teste bicaudal com região de rejeição tanto  $z \geq z_{0,005}$  ou  $z \leq -z_{0,005}$ . A partir da Seção 4.3 ou da Tabela A.3 do Apêndice,  $z_{0,005} = 2,58$  então, rejeitamos  $H_0$  se  $z \geq 2,58$  ou  $z \leq -2,58$ .
6. Substituindo  $n = 9$  e  $\bar{x} = 131,08$ ,

$$z = \frac{131,08 - 130}{1,5/\sqrt{9}} = \frac{1,08}{0,5} = 2,16$$

Isto é, a média amostral observada estaria um pouco mais de 2 desvios padrão acima do esperado, se  $H_0$  fosse verdadeira.

7. O valor calculado  $z = 2,16$  não cai na região de rejeição ( $-2,58 < 2,16 < 2,58$ ), de modo que  $H_0$  não pode ser rejeitada com nível de significância 0,01. Os dados não dão forte apoio à alegação de que a média real difere do valor projetado de 130. ■

**$\beta$  e a Determinação do Tamanho da Amostra** Os testes  $z$  para o caso I estão entre poucos na estatística para os quais existem fórmulas simples disponíveis para  $\beta$ , a probabilidade de um erro tipo II. Considere primeiro o teste de cauda superior com região de rejeição  $z \geq z_\alpha$ . Isso equivale a  $\bar{x} \geq \mu_0 + z_\alpha \cdot \sigma/\sqrt{n}$ , de modo que  $H_0$  não será rejeitada se  $\bar{x} < \mu_0 + z_\alpha \cdot \sigma/\sqrt{n}$ . Agora, seja  $\mu'$  um valor específico de  $\mu$  que excede o valor nulo  $\mu_0$ . Então,

$$\begin{aligned} \beta(\mu') &= P(H_0 \text{ não é rejeitada quando } \mu = \mu') \\ &= P(\bar{X} < \mu_0 + z_\alpha \cdot \sigma/\sqrt{n} \text{ quando } \mu = \mu') \\ &= P\left(\frac{\bar{X} - \mu'}{\sigma/\sqrt{n}} < z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}} \text{ quando } \mu = \mu'\right) \\ &= \Phi\left(z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right) \end{aligned}$$

À medida que  $\mu'$  aumenta,  $\mu_0 - \mu'$  fica mais negativo, de modo que  $\beta(\mu')$  será pequeno quando  $\mu'$  exceder muito  $\mu_0$  (pois o valor pelo qual  $\Phi$  é avaliado será, então, negativo). As probabilidades de erro dos testes de cauda inferior e bicaudal são deduzidas de uma forma análoga.

Se  $\sigma$  for grande, a probabilidade de um erro tipo II pode ser grande para um valor alternativo  $\mu'$  que seja de interesse específico a um investigador. Suponha que determinemos  $\alpha$  e especifiquemos  $\beta$  para tal valor alternativo. No exemplo do extintor, os administradores da empresa podem considerar  $\mu' = 132$  como um desvio bastante substancial de  $H_0: \mu = 130$  e, portanto, desejarem que  $\beta(132) = 0,10$  além de  $\alpha = 0,01$ . De modo

genérico, considere as duas restrições  $P(\text{erro tipo I}) = \alpha$  e  $\beta(\mu') = \beta$  para  $\alpha$ ,  $\mu'$  e  $\beta$  especificados. Então, para um teste de cauda superior, o tamanho da amostra  $n$  deve ser escolhido para satisfazer

$$\Phi\left(z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right) = \beta$$

Isso implica que

$$-z_\beta = \begin{matrix} \text{valor crítico de } z \text{ que} \\ \text{inclui a área da cauda inferior } \beta \end{matrix} = z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}$$

É fácil resolver essa equação para o  $n$  desejado. Um argumento paralelo produz o tamanho necessário da amostra para os testes de cauda inferior e bicaudal, conforme resumido no quadro a seguir.

Hipótese alternativa	Probabilidade de Erro Tipo II $\beta(\mu')$ para um Teste Nível $\alpha$
$H_a: \mu > \mu_0$	$\Phi\left(z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)$
$H_a: \mu < \mu_0$	$1 - \Phi\left(-z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)$
$H_a: \mu \neq \mu_0$	$\Phi\left(z_{\alpha/2} + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right) - \Phi\left(-z_{\alpha/2} + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)$

onde  $\Phi(z)$  = a fdc normal padronizada.  
 O tamanho da amostra  $n$  para o qual um teste nível  $\alpha$  também possui  $\beta(\mu') = \beta$  com o valor alternativo  $\mu'$  é

$$n = \begin{cases} \left[ \frac{\sigma(z_\alpha + z_\beta)}{\mu_0 - \mu'} \right]^2 & \text{para teste monocaudal} \\ & \text{(superior ou inferior)} \\ \left[ \frac{\sigma(z_{\alpha/2} + z_\beta)}{\mu_0 - \mu'} \right]^2 & \text{para teste bicaudal} \\ & \text{(uma solução aproximada)} \end{cases}$$

**Exemplo 8.7**

Seja  $\mu$  a vida útil média real de certo tipo de pneu. Considere o teste de  $H_0: \mu = 30.000$  versus  $H_a: \mu > 30.000$  com base em um tamanho de amostra  $n = 16$  da distribuição de uma população normal com  $\sigma = 1500$ . Um teste com  $\alpha = 0,01$  requer que  $z_\alpha = z_{0,01} = 2,33$ . A probabilidade de cometer um erro tipo II quando  $\mu = 31.000$  é

$$\beta(31.000) = \Phi\left(2,33 + \frac{30.000 - 31.000}{1500/\sqrt{16}}\right) = \Phi(-0,34) = 0,3669$$

Uma vez que  $z_{0,1} = 1,28$ , um teste ao nível de 0,01 de  $\beta(31.000) = 0,1$  requer

$$n = \left[ \frac{1500(2,33 + 1,28)}{30.000 - 31.000} \right]^2 = (-5,42)^2 = 29,32$$

O tamanho da amostra deve ser um inteiro, de modo que deve ser usado  $n = 30$  pneus. ■

### Caso II: Testes de amostras grandes

Quando o tamanho da amostra é grande, os testes de  $z$  para o caso I são facilmente modificados para produzir procedimentos de teste válidos, sem precisar que a distribuição da população seja normal ou que  $\sigma$  seja conhecido. O resultado principal foi usado no Capítulo 7 para demonstrar os intervalos de confiança de amostra grande: um  $n$  grande implica que o desvio padrão da amostra  $s$  estará próximo de  $\sigma$  para a maioria das amostras, de modo que a variável padronizada

$$Z = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

tenha distribuição *aproximadamente* normal padronizada. A substituição do valor nulo  $\mu_0$  no lugar de  $\mu$  produz a estatística de teste

$$Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

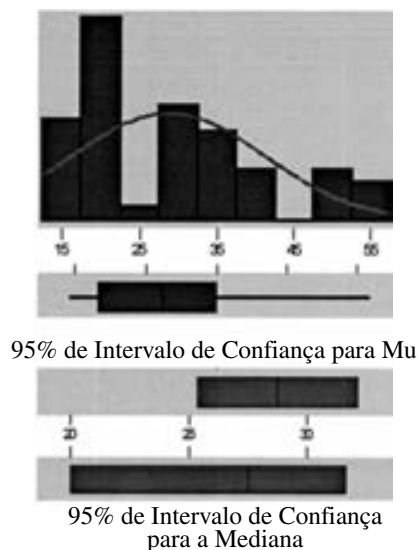
que tem distribuição aproximadamente normal padronizada quando  $H_0$  for verdadeira. O uso de regiões de rejeição dadas anteriormente para o caso I (por exemplo,  $z \geq z_\alpha$  quando a hipótese alternativa é  $H_a: \mu > \mu_0$ ), então, resulta em procedimentos de teste para os quais o nível de significância é aproximadamente (e não exatamente)  $\alpha$ . A regra prática  $n > 40$  novamente será usada para caracterizar o tamanho grande da amostra.

#### Exemplo 8.8

Um cone de penetração dinâmico (DCP) é usado para medir a resistência do material à penetração (mm/fluxo), à medida que um cone é levado ao pavimento ou à base. Suponha que, para uma aplicação específica, é necessário que o valor médio real do DCP de certo tipo de pavimento seja menor que 30. O pavimento não será usado a menos que haja evidência conclusiva de que a especificação foi satisfeita. Vamos determinar e testar as hipóteses apropriadas usando os seguintes dados (“Probabilistic Model for the Analysis of Dynamic Cone Penetrometer Test Values in Pavement Structure Evaluation,” *J. of Testing and Evaluation*, 1999: 7-14):

14,1	14,5	15,5	16,0	16,0	16,7	16,9	17,1	17,5	17,8
17,8	18,1	18,2	18,3	18,3	19,0	19,2	19,4	20,0	20,0
20,8	20,8	21,0	21,5	23,5	27,5	27,5	28,0	28,3	30,0
30,0	31,6	31,7	31,7	32,5	33,5	33,9	35,0	35,0	35,0
36,7	40,0	40,0	41,3	41,7	47,5	50,0	51,0	51,8	54,4
55,0	57,0								

#### Estatísticas Descritivas



#### Variável PCD

Um teste de normalidade de Anderson-Darling

A-Quadrado	1,902
Valor P	0,000
Média	28,7615
Saldo Devedor	12,2647
Varição	150,423
Parte central da amostra	0,8086264
Kutosis	-3,9E-01
N	52
Mínimo	14,1000
Primeiro Quartil	18,2250
Mediana	27,5000
Terceiro Quartil	35,0000
Máximo	57,0000
95% de intervalo de confiança para Mu	25,3470 32,1761
95% de intervalo de confiança para Sigma	10,2784 15,2098
95% de intervalo de confiança para a Média	20,0000 31,7000

Figura 8.3 Resumo descritivo do MINITAB para os dados do DCP do Exemplo 8.8



A Figura 8.3 mostra um resumo descritivo obtido do MINITAB. O DCP da média amostral é menor que 30. Entretanto, existe uma quantidade substancial de variação nos dados (coeficiente de variação amostral =  $s/\bar{x} = 0,4265$ ), de modo que o fato de a média ser menor que a especificação de corte do projeto possa ser apenas uma consequência da variabilidade da amostragem. Observe que o histograma não se assemelha a uma curva normal (e um gráfico de probabilidade normal não exhibe padrão linear), mas os testes de  $z$  para amostra grande não requerem uma população com distribuição normal.

1.  $\mu$  = valor médio real do DCP
2.  $H_0: \mu = 30$
3.  $H_a: \mu < 30$  (de modo que o pavimento não será usado a menos que a hipótese nula seja rejeitada)
4.  $z = \frac{\bar{x} - 30}{s/\sqrt{n}}$
5. Usando um teste com nível de significância 0,05,  $H_0$  será rejeitada se  $z \leq -1,645$  (teste de cauda inferior).
6. Com  $n = 52$ ,  $\bar{x} = 28,76$ , e  $s = 12,2647$ ,

$$z = \frac{28,76 - 30}{12,2647/\sqrt{52}} = \frac{-1,24}{1,701} = -0,73$$

7. Uma vez que  $-0,73 > -1,645$ ,  $H_0$  não pode ser rejeitada. Não temos evidências suficientes para concluir que  $\mu < 30$ , portanto, o uso do pavimento não é justificado. ■

A determinação de  $\beta$  e do tamanho da amostra necessário para esses testes de amostra grande pode ser feita com base na especificação de um valor plausível de  $\sigma$  e no uso das fórmulas do caso I (mesmo se  $s$  for usado no teste) ou no uso das curvas a serem introduzidas em breve, relativas ao caso III.

### Caso III: Uma população com distribuição normal

Quando  $n$  é pequeno, o Teorema do Limite Central (TLC) não pode mais ser usado para justificar um teste de amostra grande. Encontramos essa mesma dificuldade ao obtermos um intervalo de confiança (IC) da amostra pequena para  $\mu$ , no Capítulo 7. Nossa abordagem aqui será igual à utilizada lá: assumiremos que a distribuição da população é pelo menos aproximadamente normal e descreveremos procedimentos de teste cuja validade apóia-se nessa suposição. Se um investigador tem bons motivos para acreditar que a distribuição da população é não-normal, um teste de distribuição livre do Capítulo 15 pode ser utilizado. De maneira alternativa, um estatístico pode ser consultado com relação aos procedimentos válidos para famílias específicas de distribuições de população que não sejam a família normal ou um procedimento de *bootstrap* pode ser desenvolvido.

O resultado principal no qual testes da média de uma população normal foram usados no Capítulo 7 para derivar o IC de  $t$  de uma amostra: Se  $X_1, X_2, \dots, X_n$  for a amostra aleatória de uma distribuição normal, a variável padronizada

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

possui distribuição  $t$  com  $n - 1$  graus de liberdade (gl). Considere o teste de  $H_0: \mu = \mu_0$  contra  $H_a: \mu > \mu_0$  usando a estatística de teste  $T = (\bar{X} - \mu_0)/(S/\sqrt{n})$ . Isto é, a estatística de teste resulta da padronização de  $\bar{X}$  sob a suposição de que  $H_0$  é verdadeira (usando  $S/\sqrt{n}$ , o desvio padrão estimado de  $\bar{X}$ , em vez de  $\sigma/\sqrt{n}$ ). Quando  $H_0$  é verdadeira, a estatística de teste possui uma distribuição  $t$  com  $n - 1$  gl. Conhecer a distribuição da estatística de teste quando  $H_0$  é verdadeira (a “distribuição nula”) permite-nos construir uma região de rejeição para a qual a probabilidade de erro tipo I é controlada no nível desejado. Em particular, o uso do valor crítico  $t_{\alpha, n-1}$  da cauda superior  $t$  para especificar a região de rejeição  $t \geq t_{\alpha, n-1}$  implica

$$\begin{aligned} P(\text{erro tipo I}) &= P(H_0 \text{ é rejeitada quando é verdadeira}) \\ &= P(T \geq t_{\alpha, n-1} \text{ quando } T \text{ possui distribuição } t \text{ com } n - 1 \text{ gl}) \\ &= \alpha \end{aligned}$$

A estatística de teste aqui é realmente a mesma que no caso da amostra grande, mas é denominada  $T$ , para enfatizar que sua distribuição nula é uma distribuição  $t$  com  $n - 1$  gl em vez da distribuição normal padronizada ( $z$ ). A região de rejeição para o teste  $t$  difere da do teste de  $z$  somente no que diz respeito ao valor  $t$  crítico de  $t_{\alpha, n-1}$  substituir o valor  $z$  crítico de  $z_{\alpha}$ . Comentários semelhantes referem-se às alternativas para as quais um teste de cauda inferior ou bicaudal seja apropriado.

**O Teste  $t$  de uma Amostra**

Hipótese nula:  $H_0: \mu = \mu_0$

Valor da estatística de teste:  $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$

**Hipótese alternativa**

$H_a: \mu > \mu_0$

$H_a: \mu < \mu_0$

$H_a: \mu \neq \mu_0$

**Região de rejeição para o teste nível  $\alpha$**

$t \geq t_{\alpha, n-1}$  (com cauda superior)

$t \leq -t_{\alpha, n-1}$  (com cauda inferior)

tanto  $t \geq t_{\alpha/2, n-1}$  ou  $t \leq -t_{\alpha/2, n-1}$  (bicaudal)

**Exemplo 8.9**

Um local de trabalho seguro e bem-projetado pode contribuir muito para aumentar a produtividade. É especialmente importante que os funcionários não sejam solicitados a realizar tarefas – como levantamento de peso – que excedam suas capacidades. Os dados a seguir sobre o peso máximo de levantamento (MAWL, em kg) para uma frequência de quatro levantamentos/min. foram relatados no artigo “The Effects of Speed, Frequency, and Load on Measured Hand Forces for a Floor-to-Knuckle Lifting Task” (*Ergonomics*, 1992, p. 833-843); as pessoas foram selecionadas aleatoriamente de uma população de homens saudáveis entre 18-30 anos. Assumindo que MAWL é normalmente distribuído, os dados a seguir sugerem que o valor médio de MAWL da população excede 25?

25,8    36,6    26,3    21,8    27,2

Vamos fazer um teste usando um nível de significância de 0,05.

1.  $\mu$  = valor médio de MAWL da população
2.  $H_0: \mu = 25$
3.  $H_a: \mu > 25$
4.  $t = \frac{\bar{x} - 25}{s/\sqrt{n}}$
5. Rejeite  $H_0$  se  $t \geq t_{\alpha, n-1} = t_{0,05,4} = 2,132$ .
6.  $\sum x_i = 137,7$  e  $\sum x_i^2 = 3911,97$ , do qual  $\bar{x} = 27,54$ ,  $s = 5,47$ , e

$$t = \frac{27,54 - 25}{5,47/\sqrt{5}} = \frac{2,54}{2,45} = 1,04$$

O resultado do MINITAB a seguir de um teste  $t$  para uma amostra possui os mesmos valores calculados (o valor- $P$  é discutido na Seção 8.4).

```
Test of mu = 25,00 vs mu > 25,00
Variable N Mean StDev SE Mean T P-Value
mawl 5 27,54 5,47 2,45 1,04 0,18
```

7. Uma vez que 1,04 não cai na região de rejeição ( $1,04 < 2,132$ ),  $H_0$  não pode ser rejeitada no nível de significância 0,05. Ainda é plausível que  $\mu$  seja (no máximo) 25. ■

**$\beta$  e a Determinação do Tamanho da Amostra** O cálculo de  $\beta$  no valor alternativo  $\mu'$  no caso I foi feito para expressar a região de rejeição em termos de  $\bar{x}$  (por exemplo,  $\bar{x} \geq \mu_0 + z_\alpha \cdot \sigma/\sqrt{n}$ ) e então subtraindo  $\mu'$  para padronizar corretamente. Uma abordagem equivalente envolve a observação de que, quando  $\mu = \mu'$ , a estatística de teste  $Z = (\bar{X} - \mu_0)/(\sigma/\sqrt{n})$  ainda possui distribuição normal com variância 1, mas agora o valor médio de  $Z$  é dado por  $(\mu' - \mu_0)/(\sigma/\sqrt{n})$ . Isto é, quando  $\mu = \mu'$ , a estatística de teste ainda possui distribuição normal, embora não seja a distribuição normal padronizada. Por causa disso,  $\beta(\mu')$  é a área sob a curva normal correspondente ao valor médio  $(\mu' - \mu_0)/(\sigma/\sqrt{n})$  e variância 1. Ambos,  $\alpha$  e  $\beta$ , envolvem o trabalho com variáveis normalmente distribuídas.

O cálculo de  $\beta(\mu')$  para o teste  $t$  é muito menos direto. Isso porque a distribuição da estatística de teste  $T = (\bar{X} - \mu_0)/(S/\sqrt{n})$  é bastante complicada quando  $H_0$  é falsa e  $H_a$  verdadeira. Dessa forma, para um teste de cauda superior, determinar

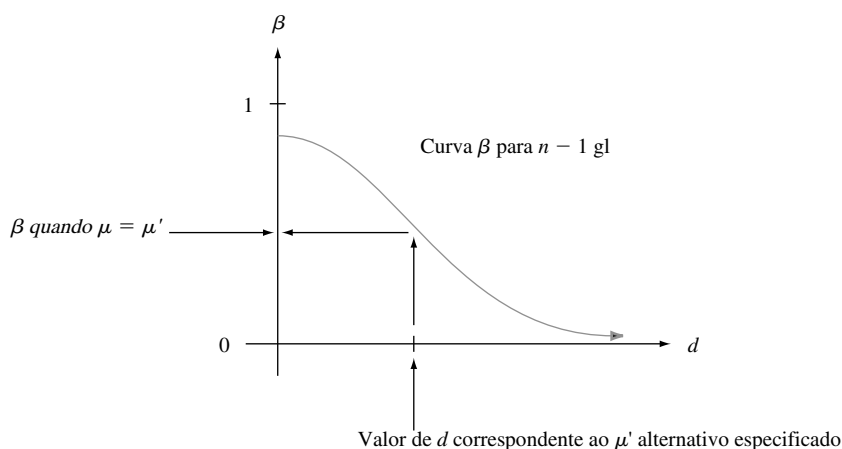
$$\beta(\mu') = P(T < t_{\alpha, n-1} \text{ quando } \mu = \mu' \text{ em vez de } \mu_0)$$

envolve a integração de uma função de densidade bastante desagradável. Isso deve ser feito numericamente, mas felizmente foi feito pelos pesquisadores para os testes  $t$  monocaudal e bicaudal. Os resultados são resumidos em gráficos de  $\beta$  que aparecem na Tabela A.17 do Apêndice. Existem quatro conjuntos de gráficos correspondentes a testes monocaudais nos níveis 0,05 e 0,01 e a testes bicaudais nos mesmos níveis.

Para compreender como esses gráficos são utilizados, observe primeiro que  $\beta$  e o tamanho da amostra necessário  $n$  no caso I são funções não somente da diferença absoluta  $|\mu_0 - \mu'|$  como também de  $d = |\mu_0 - \mu'|/\sigma$ . Suponha, por exemplo, que  $|\mu_0 - \mu'| = 10$ . Esse desvio de  $H_0$  será muito mais fácil de ser detectado ( $\beta$  menor) quando  $\sigma = 2$ , que no caso  $\mu_0$  e  $\mu'$  estão distantes de 5 desvios padrão da população, do que quando  $\sigma = 10$ . O fato de  $\beta$  para o teste  $t$  depender de  $d$  em vez apenas de  $|\mu_0 - \mu'|$  é desastroso, uma vez que, para usar os gráficos, deve-se ter alguma idéia do valor real de  $\sigma$ . Uma estimativa conservadora (grande) de  $\sigma$  produzirá um valor conservador (grande) de  $\beta(\mu')$  e uma estimativa conservadora do tamanho da amostra necessário para  $\alpha$  e  $\beta(\mu')$  recomendados.

Uma vez que  $\mu'$  alternativo e o valor de  $\sigma$  sejam selecionados,  $d$  é calculado e seu valor estabelecido no eixo horizontal do conjunto de curvas. O valor de  $\beta$  é a altura da curva  $n - 1$  gl acima do valor de  $d$  (a interpolação visual é necessária se  $n - 1$  não for um valor para o qual a curva correspondente aparece), conforme ilustrado na Figura 8.4.

Em vez de fixar  $n$  (isto é,  $n - 1$ , e portanto a curva específica de onde  $\beta$  é lido), pode-se estabelecer que  $\alpha$  (0,05 ou 0,01 aqui) e um valor de  $\beta$  para  $\mu'$  e  $\sigma$  escolhidos. Após calcular  $d$ , o ponto  $(d, \beta)$  é focado no conjunto de gráficos. A curva abaixo e mais próxima desse ponto fornece  $n - 1$  e portanto  $n$  (novamente, a interpolação geralmente é necessária).



**Figura 8.4** Uma curva  $\beta$  típica para o teste  $t$

**Exemplo 8.10**

Supõe-se que a queda de tensão média real do coletor ao emissor de certo tipo de transistores bipolares de porta isolada é no máximo de 2,5 volts. Um investigador seleciona uma amostra de  $n = 10$  transistores e utiliza as tensões resultantes como base para testar  $H_0: \mu = 2,5$  versus  $H_a: \mu > 2,5$  usando um teste  $t$  com nível de significância  $\alpha = 0,05$ . Se o desvio padrão da distribuição da tensão for  $\sigma = 0,100$ , qual é a probabilidade de essa  $H_0$  não ser rejeitada quando na verdade  $\mu = 2,6$ ? Com  $d = |2,5 - 2,6|/0,100 = 1,0$ , o ponto na curva  $\beta$  em 9 gl de um teste monocaudal com  $\alpha = 0,05$  acima de 1,0 possui altura de aproximadamente 0,1, de modo que  $\beta \approx 0,1$ . O investigador pode considerar que esse valor de  $\beta$  é muito grande para um desvio substancial de  $H_0$  e pode querer  $\beta = 0,05$  para esse valor alternativo de  $\mu$ . Uma vez que  $d = 1,0$ , o ponto  $(d, \beta) = (1,0, 0,05)$  deve ser localizado. Esse ponto está muito próximo da curva de 14 gl, de modo que usar  $n = 15$  levará a  $\alpha = 0,05$  e  $\beta = 0,05$  quando o valor de  $\mu$  for 2,6 e  $\sigma = 0,10$ . Um valor maior de  $\sigma$  daria um  $\beta$  maior para essa alternativa, e um valor alternativo de  $\mu$  mais próximo de 2,5 também resultaria em um aumento no valor de  $\beta$ . ■

A maioria dos pacotes estatísticos de computador usados também calculará as probabilidades de erro tipo II e determinará os tamanhos de amostra necessários. Como exemplo, usamos o MINITAB para realizar os cálculos do Exemplo 8.10. Seus cálculos baseiam-se na **potência**, que é simplesmente  $1 - \beta$ . Queremos que  $\beta$  seja pequeno, o que equivale a querer que a potência do teste seja grande. Por exemplo,  $\beta = 0,05$  corresponde a um valor de 0,95 para potência. A seguir é mostrada a saída do MINITAB resultante.

```

Power and Sample Size
Testing mean = null (versus > null)
Calculating power for mean = null + 0,1
Alpha = 0,05   Sigma = 0,1

Sample
Size      Power
10        0,8975

Power and Sample Size
1-Sample t Test
Testing mean = null (versus > null)
Calculating power for mean = null + 0,1
Alpha = 0,05   Sigma = 0,1

Sample  Target  Actual
Size    Power   Power
13      0,9500   0,9597
    
```

Observe, a partir da segunda parte da saída, que o tamanho da amostra necessário para obter uma potência de 0,95 ( $\beta = 0,05$ ) de um teste de cauda superior, com  $\alpha = 0,05$  quando  $\sigma = 0,1$  e  $\mu'$  é 0,1 maior que  $\mu_0$ , é somente  $n = 13$ , enquanto, em nossa aproximação da curva,  $\beta$  forneceu 15. Quando disponível, esse tipo de software é mais confiável que as curvas.

---

**Exercícios** | Seção 8.2 (15–34)

- 15. Seja a estatística de teste  $z$  normalmente distribuído quando  $H_0$  for verdadeira. Forneça o nível de significância para cada uma das situações seguintes:
  - a.  $H_a: \mu > \mu_0$ , região de rejeição  $z \geq 1,88$
  - b.  $H_a: \mu < \mu_0$ , região de rejeição  $z \leq -2,75$
  - c.  $H_a: \mu \neq \mu_0$ , região de rejeição  $z \geq 2,88$  ou  $z \leq -2,88$
- 16. Seja a estatística de teste  $T$  uma distribuição  $t$  quando  $H_0$  for verdadeira. Forneça o nível de significância de cada situação a seguir:
  - a.  $H_a: \mu > \mu_0$ ,  $df = 15$ , região de rejeição  $t \geq 3,733$
  - b.  $H_a: \mu < \mu_0$ ,  $n = 24$ , região de rejeição  $t \leq -2,500$
  - c.  $H_a: \mu \neq \mu_0$ ,  $n = 31$ , região de rejeição  $t \geq 1,697$  ou  $t \leq -1,697$
- 17. Responda às questões a seguir referentes ao problema do pneu no Exemplo 8.7.
  - a. Se  $\bar{x} = 20,960$  e um teste nível  $\alpha = 0,01$  for usado, qual é a decisão?
  - b. Se um teste nível 0,01 for usado, qual é  $\beta(20.500)$ ?
  - c. Se um teste nível 0,01 for usado e também for necessário que  $\beta(20.500) = 0,05$ , que tamanho da amostra  $n$  é necessária?
  - d. Se  $\bar{x} = 20.960$ , qual é o menor  $\alpha$  em que  $H_0$  pode ser rejeitada (com base em  $n = 16$ )?
- 18. Reconsidere a situação de secagem da pintura do Exemplo 8.2, em que o tempo de secagem de um espécime de teste é normalmente distribuído com  $\sigma = 9$ .

As hipóteses  $H_0: \mu = 75$  versus  $H_a: \mu < 75$  serão testadas usando uma amostra aleatória de  $n = 25$  observações.

- a. Quantos desvios padrão (de  $\bar{X}$ ) abaixo do valor nulo são  $\bar{x} = 72,3$ ?
  - b. Se  $\bar{x} = 72,3$ , qual é a conclusão usando  $\alpha = 0,01$ ?
  - c. Qual é o  $\alpha$  para o procedimento de teste que rejeita  $H_0$  quando  $z \leq -2,88$ ?
  - d. Para o procedimento de teste da parte (c), qual é  $\beta(70)$ ?
  - e. Se o procedimento de teste da parte (c) for usado, que  $n$  é necessário para garantir que  $\beta(70) = 0,01$ ?
  - f. Se um teste nível 0,01 for usado com  $n = 100$ , qual é a probabilidade de um erro tipo I quando  $\mu = 76$ ?
19. O ponto de fundição de cada uma das 16 amostras de certa marca de óleo vegetal hidrogenado foi determinado, resultando em  $\bar{x} = 94,32$ . Assuma que a distribuição do ponto de fundição seja normal com  $\sigma = 1,20$ .
- a. Teste  $H_0: \mu = 95$  versus  $H_a: \mu \neq 95$  usando um teste nível 0,01 bicaudal.
  - b. Se um teste nível 0,01 for usado, qual é  $\beta(94)$ , ou seja, a probabilidade de um erro tipo II quando  $\mu = 94$ ?
  - c. Que valor de  $n$  é necessário para garantir que  $\beta(94) = 0,1$  quando  $\alpha = 0,01$ ?
20. Certo tipo de lâmpadas é anunciado como tendo vida útil média de 750 horas. O preço dessas lâmpadas é bastante favorável, de modo que um cliente em potencial decidiu prosseguir com o pedido de compra, a menos que possa ser conclusivamente demonstrado que a vida útil média real seja menor que a anunciada. Uma amostra aleatória de 50 lâmpadas foi selecionada, a vida útil de cada lâmpada determinada e as hipóteses apropriadas, testadas usando o MINITAB, resultando na saída a seguir.

Variável	N	Média	Des Pad	MédiaSE	Z	Valor de P
vida útil	50	738,44	38,20	5,40	-2,14	0,016

Que conclusão seria apropriada para um nível de significância de 0,05? Um nível de significância de 0,01? Que nível de significância e conclusão você recomendaria?

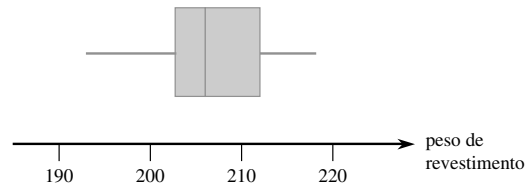
21. Supõe-se que o diâmetro médio real de certo tipo de rolamentos de esferas é 0,5 in. Um teste  $t$  de uma amostra será realizado para ver se esse é o caso. Que conclusão é apropriada em cada uma das situações a seguir?
- a.  $n = 13, t = 1,6, \alpha = 0,05$
  - b.  $n = 13, t = -1,6, \alpha = 0,05$
  - c.  $n = 25, t = -2,6, \alpha = 0,01$
  - d.  $n = 25, t = -3,9$
22. O artigo "The Foreman's View of Quality Control" (*Quality Engr.*, 1990, p. 257-280) descreveu uma investigação a respeito dos pesos do revestimento de canos grandes que resultam de um processo de revestimento galvanizado. Os padrões de produção exigem peso mé-

dio real de 200 lb por cano. O resumo descritivo e o *boxplot* a seguir são do MINITAB.

Variável	N	Média	Mediana	MédiaAp	Z	Des Pad
ctg wt	30	206,73	206,00	206,81	6,35	1,16

Variável	Mín	Máx	Q1	Q3
ctg wt	193,00	218,00	202,75	212,00



- a. O que o *boxplot* sugere sobre o *status* da especificação do peso de revestimento médio real?
  - b. Um gráfico de probabilidade normal dos dados foi bastante direto. Use a saída descritiva para testar as hipóteses apropriadas.
23. O Exercício 36 do Capítulo 1 forneceu  $n = 26$  observações sobre o tempo de fuga (s) para trabalhadores de uma plataforma de petróleo em um exercício simulado, no qual a média e o desvio padrão amostrais são 370,69 e 24,36, respectivamente. Suponha que os investigadores acreditaram, *a priori*, que o tempo de fuga médio real seria no máximo 6 min. Os dados contradizem esse ponto de vista anterior? Assumindo a normalidade, teste as hipóteses apropriadas usando um nível de significância de 0,05.
24. Reconsidere as observações da amostra sobre a viscosidade estabilizada dos espécimes de asfalto introduzidos no Exercício 46 do Capítulo 1 (2781, 2900, 3013, 2856 e 2888). Suponha que, para uma aplicação específica, seja necessário que a viscosidade média real seja 3000. Esse requisito parece ter sido satisfeito? Determine e teste as hipóteses apropriadas.
25. A porcentagem desejada de  $\text{SiO}_2$  em um tipo de cimento aluminoso é 5,5. Para testar se a porcentagem média real é 5,5 de uma instalação de produção específica, 16 amostras obtidas independentemente são analisadas. Suponha que a porcentagem de  $\text{SiO}_2$  em uma amostra seja normalmente distribuída com  $\sigma = 0,3$  e que  $\bar{x} = 5,25$ .
- a. Isso indica conclusivamente que a porcentagem média real é diferente de 5,5? Faça a análise usando a seqüência de passos sugerida no texto.
  - b. Se a porcentagem média real é  $\mu = 5,6$  e um teste nível  $\alpha = 0,01$  com base em  $n = 16$  for usado, qual é a probabilidade de detectar esse desvio de  $H_0$ ?
  - c. Que valor de  $n$  é necessário para satisfazer  $\alpha = 0,01$  e  $\beta(5,6) = 0,01$ ?
26. Para obter informações sobre as propriedades de resistência à corrosão de um determinado tipo de conduíte de aço, 45 espécimes são enterrados no solo por um período de 2 anos. A penetração máxima (em mils) de cada espécime é então medida, produzindo uma

penetração média amostral de  $\bar{x} = 52,7$  e um desvio padrão da amostra de  $s = 4,8$ . Os conduítes foram fabricados com a especificação de a penetração média real ser de no máximo 50 milis. Eles serão usados, a menos que seja demonstrado de forma conclusiva que a especificação não foi satisfeita. A que conclusão você chegaria?

27. O teste de impacto Charpy de entalhe em V é a base para estudar muitos critérios de dureza do material. Esse teste foi aplicado em 42 espécimes de uma liga específica a 110°F. A quantidade média amostral de expansão transversal foi calculada como sendo 73,1 milis, e o desvio padrão da amostra,  $s = 5,9$  milis. Para ser adequado para uma aplicação específica, a quantidade média real de expansão deve ser menor que 75 milis. A liga não será usada, a menos que a amostra forneça forte evidência de que esse critério foi satisfeito. Teste as hipóteses relevantes, usando  $\alpha = 0,01$  para decidir se a liga é adequada.
28. Cirurgias menores em cavalos com condições de campo requerem um anestésico de curto prazo confiável que produza bom relaxamento muscular, alterações respiratória e cardiovascular mínimas e uma recuperação rápida e fácil com efeitos subseqüentes mínimos, de modo que os cavalos possam ser deixados sozinhos. O artigo “A Field Trial of Ketamine Anesthesia in the Horse” (*Equine Vet. J.*, 1984, p. 176-179) relata que, para uma amostra de  $n = 73$  cavalos, à qual a cetamina foi administrada sob certas condições, o tempo de repouso (deitado) lateral médio da amostra foi de 18,86 min. e o desvio padrão, de 8,6 min. Esses dados sugerem que o tempo de repouso lateral médio real sob tais condições é menor que 20 min.? Teste as hipóteses apropriadas no nível de significância 0,10.
29. A quantidade de desgaste do eixo (0,0001 in.) após uma milhagem fixa foi determinada para cada um dos  $n = 8$  motores a combustão interna que têm fio de cobre como um material do mancal, resultando em  $\bar{x} = 3,72$  e  $s = 1,25$ .
- Assumindo que a distribuição do desgaste do eixo é normal com média  $\mu$ , use o teste  $t$  no nível 0,05 para testar  $H_0: \mu = 3,50$  versus  $H_a: \mu > 3,50$ .
  - Usando  $\sigma = 1,25$ , qual é a probabilidade de erro tipo II  $\beta(\mu')$  do teste para a alternativa  $\mu' = 4,00$ ?
30. A tolerância de dieta diária de zinco recomendada entre homens com mais de 50 anos é 15 mg/dia. O artigo “Nutrient Intakes and Dietary Patterns of Older Americans: A National Study” (*J. Gerontology*, 1992, p. M145-150) relata os seguintes dados resumidos sobre ingestão para uma amostra de homens entre 65-74 anos:  $n = 115$ ,  $\bar{x} = 11,3$  e  $s = 6,43$ . Esses dados indicam que a ingestão de zinco diária média na população dos homens entre 65-74 anos cai abaixo da tolerância permitida?
31. Em um experimento designado para medir o tempo adequado de exposição visual de um inspetor, a fim de realizar uma atividade com luz reduzida, o tempo médio da amostra para  $n = 9$  inspetores foi 6,32 s e o desvio padrão da amostra 1,65 s. Assumiu-se anteriormente que o tempo de adaptação médio foi de pelo menos 7 s. Assumindo que o tempo de adaptação seja normalmente distribuído, os dados contradizem o ponto de vista anterior? Use o teste  $t$  com  $\alpha = 0,1$ .
32. Uma amostra de determinado tipo de 12 detectores de radônio de um determinado tipo foi selecionada, e cada um foi exposto a 100 pCi/L de radônio. As leituras resultantes foram as seguintes:
- |       |       |      |       |       |      |
|-------|-------|------|-------|-------|------|
| 105,6 | 90,9  | 91,2 | 96,9  | 96,5  | 91,3 |
| 100,1 | 105,0 | 99,6 | 107,7 | 103,3 | 92,4 |
- Esses dados sugerem que a leitura média da população sob essas condições é diferente de 100? Determine e teste as hipóteses apropriadas, usando  $\alpha = 0,05$ .
  - Suponha que antes do experimento um valor de  $\sigma = 7,5$  foi assumido. Quantas determinações, então, teriam sido apropriadas para obter  $\beta = 0,10$  para a alternativa  $\mu = 95$ ?
33. Mostre que, para qualquer  $\Delta > 0$ , quando a distribuição da população é normal e o  $\sigma$  é conhecido, o teste bicaudal satisfaz  $\beta(\mu_0 - \Delta) = \beta(\mu_0 + \Delta)$ , de modo que  $\beta(\mu')$  seja simétrico em relação a  $\mu_0$ .
34. Para um valor alternativo fixo  $\mu'$ , mostre que  $\beta(\mu') \rightarrow 0$  quando  $n \rightarrow \infty$  para um teste  $z$  monocaudal ou bicaudal, no caso de uma distribuição de população normal com  $\sigma$  conhecido.

### 8.3 Testes com Relação a uma Proporção da População

Seja  $p$  a proporção de indivíduos ou objetos em uma população que possui uma propriedade especificada (por exemplo, carros com transmissões manuais ou fumantes que utilizam cigarro de filtro). Se uma pessoa ou um objeto com a propriedade for considerada um sucesso ( $S$ ), então  $p$  é a proporção da população de sucessos. Os testes relacionados a  $p$  serão feitos com base em uma amostra aleatória de tamanho  $n$  da população. Contudo que  $n$  seja pequeno em relação ao tamanho da população,  $X$  (o número de  $S$ s na amostra) possui (aproximadamente) uma distribuição binomial. Além disso, se o próprio  $n$  for grande, ambos  $X$  e o estimador  $\hat{p} = X/n$  têm

aproximadamente uma distribuição normal. Consideramos primeiro os testes de amostra grande com base nesse último fato, e então mudamos para o caso de amostra pequena que utiliza diretamente a distribuição binomial.

## Testes de Amostra Grande

Os testes de amostra grande relacionados a  $p$  são um caso especial dos procedimentos de amostra grande mais gerais para um parâmetro  $\theta$ . Seja  $\hat{\theta}$  um estimador de  $\theta$  (pelo menos aproximadamente) não-tendencioso e que possua aproximadamente distribuição normal. A hipótese nula tem a forma  $H_0: \theta = \theta_0$ , onde  $\theta_0$  representa um número (o valor nulo) apropriado para o contexto do problema. Suponha que, quando  $H_0$  for verdadeira, o desvio padrão de  $\hat{\theta}$ ,  $\sigma_{\hat{\theta}}$ , não envolva nenhum parâmetro desconhecido. Por exemplo, se  $\theta = \mu$  e  $\hat{\theta} = \bar{X}$ ,  $\sigma_{\hat{\theta}} = \sigma_{\bar{X}} = \sigma/\sqrt{n}$ , que não envolve parâmetros desconhecidos somente se o valor de  $\sigma$  for conhecido. Uma estatística de teste de amostra grande resulta da padronização de  $\hat{\theta}$  sob a suposição de que  $H_0$  é verdadeira (de modo que  $E(\hat{\theta}) = \theta_0$ ):

$$\text{Estatística de teste: } Z = \frac{\hat{\theta} - \theta_0}{\sigma_{\hat{\theta}}}$$

Se a hipótese alternativa é  $H_a: \theta > \theta_0$ , um teste de cauda superior, cujo nível de significância é aproximadamente  $\alpha$ , é especificado pela região de rejeição  $z \geq z_{\alpha}$ . As outras duas alternativas,  $H_a: \theta < \theta_0$  e  $H_a: \theta \neq \theta_0$ , são testadas usando um teste  $z$  de cauda inferior e um teste  $z$  bicaudal, respectivamente.

No caso  $\theta = p$ ,  $\sigma_{\hat{\theta}}$  não envolverá quaisquer parâmetros desconhecidos quando  $H_0$  for verdadeira, mas isso é atípico. Quando  $\sigma_{\hat{\theta}}$  envolve parâmetros desconhecidos, geralmente é possível usar um desvio padrão estimado  $S_{\hat{\theta}}$  no lugar de  $\sigma_{\hat{\theta}}$  e ainda ter  $Z$  aproximadamente com distribuição normal quando  $H_0$  for verdadeira (pois, quando  $n$  é grande,  $s_{\hat{\theta}} \approx \sigma_{\hat{\theta}}$  para a maioria das amostras). O teste de amostra grande da seção anterior fornece um exemplo disso: em virtude de  $\sigma$  ser geralmente desconhecido, usamos  $s_{\hat{\theta}} = s_{\bar{X}} = s/\sqrt{n}$  no lugar de  $\sigma/\sqrt{n}$  no denominador de  $z$ .

O estimador  $\hat{p} = X/n$  é não-tendencioso ( $E(\hat{p}) = p$ ), possui aproximadamente distribuição normal e seu desvio padrão é  $\sigma_{\hat{p}} = \sqrt{p(1-p)/n}$ . Esses fatos foram usados na Seção 7.2 para obter um intervalo de confiança de  $p$ . Quando  $H_0$  é verdadeira,  $E(\hat{p}) = p_0$  e  $\sigma_{\hat{p}} = \sqrt{p_0(1-p_0)/n}$ , então  $\sigma_{\hat{p}}$  não envolve quaisquer parâmetros desconhecidos. Segue, então, que, quando  $n$  é grande e  $H_0$  verdadeira, a estatística de teste

$$Z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$$

possui aproximadamente uma distribuição normal padronizada. Se a hipótese alternativa for  $H_a: p > p_0$  e a região de rejeição de cauda superior  $z \geq z_{\alpha}$  for usada, então

$$\begin{aligned} P(\text{erro tipo I}) &= P(H_0 \text{ é rejeitada quando é verdadeira}) \\ &= P(Z \geq z_{\alpha} \text{ quando } Z \text{ possui aproximadamente uma distribuição normal} \\ &\quad \text{padronizada}) \approx \alpha \end{aligned}$$

Dessa forma, o nível de significância desejado  $\alpha$  é obtido usando o valor crítico que inclui a área  $\alpha$  na cauda superior da curva  $z$ . As regiões de rejeição das outras duas hipóteses alternativas de cauda inferior para  $H_a: p < p_0$  e bicaudal para  $H_a: p \neq p_0$  são justificadas de forma análoga.

Hipótese nula:  $H_0: p = p_0$

Valor da estatística de teste:  $z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$

**Hipóteses alternativas**

**Região rejeitada**

$H_a: p > p_0$

$z \geq z_{\alpha}$  (de cauda superior)

$H_a: p < p_0$

$z \leq -z_{\alpha}$  (de cauda inferior)

$H_a: p \neq p_0$

tanto  $z \geq z_{\alpha/2}$  como  $z \leq -z_{\alpha/2}$  (bicaudal)

Esses procedimentos de teste são válidos desde que  $np_0 \geq 10$  e  $n(1-p_0) \geq 10$ .

**Exemplo 8.11**

Informações recentes sugerem que a obesidade é um problema crescente nos Estados Unidos entre todos os grupos de idade. A Associated Press (9 de outubro de 2002) relatou que 1276 pessoas em uma amostra de 4115 adultos foram consideradas obesas (índice de massa corporal que excede 30; esse índice é uma medida do peso com relação à altura). Uma pesquisa de 1998, com base na avaliação das próprias pessoas, revelou que 20% dos americanos adultos se consideravam obesos. Os dados mais recentes sugerem que a proporção real de adultos obesos é mais de 1,5 vez a porcentagem da pesquisa de auto-avaliação? Vamos fazer um teste de hipóteses, usando um nível de significância de 0,10.

1.  $p$  = proporção de todos os adultos americanos obesos.
2. Dizer que a porcentagem atual de 1,5 vez a porcentagem de auto-avaliação equivale à afirmação de que a porcentagem atual é de 30%, da qual temos a hipótese nula como  $H_0: p = 0,30$ .
3. A frase “mais que” na descrição do problema implica a hipótese alternativa ser  $H_a: p > 0,30$ .
4. Uma vez que  $np_0 = 4115(0,3) > 10$  e  $np_0 = 4115(0,7) > 10$ , o teste  $z$  de amostra grande certamente pode ser usado. O valor da estatística de teste é

$$z = (\hat{p} - 0,3) / \sqrt{(0,3)(0,7)/n}$$

5. A forma de  $H_a$  implica que um teste de cauda superior é apropriado: rejeite  $H_0$  se  $z > z_{0,10} = 1,28$ .
6.  $\hat{p} = 1276/4115 = 0,310$ , do qual  $z = (0,310 - 0,3) / \sqrt{(0,3)(0,7)/4115} = 0,010/0,0071 = 1,40$ .
7. Uma vez que 1,40 excede o valor crítico 1,28,  $z$  permanece na região de rejeição, o que justifica a rejeição da hipótese nula. Usando um nível de significância de 0,10, parece que mais de 30% dos adultos americanos são obesos. ■

**$\beta$  e a Determinação do Tamanho da Amostra** Quando  $H_0$  é verdadeira, a estatística de teste  $Z$  possui aproximadamente distribuição normal padronizada. Agora, suponha que  $H_0$  não seja verdadeira e que  $p = p'$ . Então,  $Z$  ainda possui aproximadamente uma distribuição normal (porque é uma função linear de  $\hat{p}$ ), mas seu valor médio e variância não são mais 0 e 1, respectivamente. Em vez disso,

$$E(Z) = \frac{p' - p_0}{\sqrt{p_0(1 - p_0)/n}} \quad V(Z) = \frac{p'(1 - p')/n}{p_0(1 - p_0)/n}$$

a probabilidade de um erro tipo II para o teste de cauda superior é  $\beta(p') = P(Z < z_\alpha \text{ quando } p = p')$ . Isso pode ser calculado usando a média e a variância fornecidas para padronizar e depois aplicar a fdc normal padronizada. Além disso, se for desejado que o teste nível  $\alpha$  também tenha  $\beta(p') = \beta$  para um valor especificado de  $\beta$ , a equação pode ser resolvida para o  $n$  necessário, como na Seção 8.2. As expressões gerais para  $\beta(p')$  e  $n$  são dadas no quadro a seguir.

Hipótese alternativa	$\beta(p')$
$H_a: p > p_0$	$\Phi\left[\frac{p_0 - p' + z_\alpha \sqrt{p_0(1 - p_0)/n}}{\sqrt{p'(1 - p')/n}}\right]$
$H_a: p < p_0$	$1 - \Phi\left[\frac{p_0 - p' - z_\alpha \sqrt{p_0(1 - p_0)/n}}{\sqrt{p'(1 - p')/n}}\right]$
$H_a: p \neq p_0$	$\Phi\left[\frac{p_0 - p' + z_{\alpha/2} \sqrt{p_0(1 - p_0)/n}}{\sqrt{p'(1 - p')/n}}\right]$ $- \Phi\left[\frac{p_0 - p' - z_{\alpha/2} \sqrt{p_0(1 - p_0)/n}}{\sqrt{p'(1 - p')/n}}\right]$



O tamanho da amostra  $n$  para o qual o teste nível  $\alpha$  também satisfaz  $\beta(p') = \beta$  é

$$n = \begin{cases} \left[ \frac{z_\alpha \sqrt{p_0(1-p_0)} + z_\beta \sqrt{p'(1-p')}}{p' - p_0} \right]^2 & \text{teste monocaudal} \\ \left[ \frac{z_{\alpha/2} \sqrt{p_0(1-p_0)} + z_\beta \sqrt{p'(1-p')}}{p' - p_0} \right]^2 & \text{teste bicaudal (solução aproximada)} \end{cases}$$

**Exemplo 8.12**

Um serviço de entrega informa que no mínimo 90% dos pacotes deixados em seu escritório por volta das 9 da manhã para distribuição na mesma cidade são entregues até meio-dia desse dia. Seja  $p$  a proporção real de tais pacotes entregues como anunciado e considere as hipóteses  $H_0: p = 0,9$  versus  $H_a: p < 0,9$ . Se somente 80% dos pacotes são entregues conforme anunciado, qual é a probabilidade de um teste nível 0,01 com base em  $n = 225$  pacotes detectar tal desvio de  $H_0$ ? Qual deve ser o tamanho da amostra para garantir que  $\beta(0,8) = 0,01$ ? Com  $\alpha = 0,01$ ,  $p_0 = 0,9$ ,  $p' = 0,8$  e  $n = 225$ ,

$$\begin{aligned} \beta(0,8) &= 1 - \Phi\left(\frac{0,9 - 0,8 - 2,33\sqrt{(0,9)(0,1)/225}}{\sqrt{(0,8)(0,2)/225}}\right) \\ &= 1 - \Phi(2,00) = 0,0228 \end{aligned}$$

Dessa forma, a probabilidade de  $H_0$  ser rejeitada usando o teste quando  $p = 0,8$  é 0,9772 – aproximadamente 98% de todas as amostras resultarão em rejeição correta de  $H_0$ .

Usar  $z_\alpha = z_\beta = 2,33$  na fórmula do tamanho da amostra produz

$$n = \left[ \frac{2,33\sqrt{(0,9)(0,1)} + 2,33\sqrt{(0,8)(0,2)}}{0,8 - 0,9} \right]^2 \approx 266 \quad \blacksquare$$

**Testes de amostra pequena**

Os procedimentos de teste, quando o tamanho da amostra  $n$  é pequeno, baseiam-se diretamente na distribuição binomial e não na aproximação normal. Considere a hipótese alternativa  $H_a: p > p_0$  e novamente seja  $X$  o número de sucessos na amostra. Então,  $X$  é a estatística de teste, e a região de rejeição de cauda superior possui a forma  $x \geq c$ . Quando  $H_0$  é verdadeira,  $X$  possui distribuição binomial com parâmetros  $n$  e  $p_0$ , de modo que

$$\begin{aligned} P(\text{erro tipo I}) &= P(H_0 \text{ é rejeitada quando verdadeira}) \\ &= P(X \geq c \text{ quando } X \sim \text{Bin}(n, p_0)) \\ &= 1 - P(X \leq c - 1 \text{ quando } X \sim \text{Bin}(n, p_0)) \\ &= 1 - B(c - 1; n, p_0) \end{aligned}$$

À medida que o valor crítico  $c$  diminui, mais valores  $x$  são incluídos na região de rejeição e  $P(\text{erro tipo I})$  aumenta. Em virtude de  $X$  ter distribuição de probabilidades discreta, geralmente não é possível determinar um valor de  $c$  para o qual  $P(\text{erro tipo I})$  é exatamente o nível de significância desejado  $\alpha$  (por exemplo, 0,05 ou 0,01). Em vez disso, a região de rejeição maior da forma  $\{c, c + 1, \dots, n\}$  que satisfaz  $1 - B(c - 1; n, p_0) \leq \alpha$  é usada.

Seja  $p'$  um valor alternativo de  $p$  ( $p' > p_0$ ). Quando  $p = p'$ ,  $X \sim \text{Bin}(n, p')$ , de modo que

$$\begin{aligned} \beta(p') &= P(\text{erro tipo II quando } p = p') \\ &= P(X < c \text{ quando } X \sim \text{Bin}(n, p')) = B(c - 1; n, p') \end{aligned}$$

Isto é,  $\beta(p')$  é o resultado de um cálculo direto da probabilidade binomial. O tamanho da amostra  $n$ , necessário para garantir que um teste nível  $\alpha$  também tenha especificado  $\beta$  em um valor alternativo específico  $p'$ , deve ser determinado por tentativa e erro, usando a fdc binomial.

Os procedimentos de teste para  $H_a: p < p_0$  e para  $H_a: p \neq p_0$  são construídos de forma semelhante. No primeiro caso, a região de rejeição apropriada tem a forma  $x \leq c$  (teste de cauda inferior). O valor crítico  $c$  é o maior número que satisfaz  $B(c; n, p_0) \leq \alpha$ . A região de rejeição, quando a hipótese alternativa é  $H_a: p \neq p_0$  consiste em valores grandes e pequenos de  $x$ .

### Exemplo 8.13

Um fabricante de materiais plásticos desenvolveu um novo tipo de lixeira e propõe vendê-las com uma garantia incondicional de 6 anos. Para ver se isso é economicamente viável, 20 latas-modelo são submetidas a um teste de vida útil acelerado para simular 6 anos de uso. A garantia proposta será modificada somente se os dados da amostra sugerirem fortemente que menos de 90% das lixeiras sobreviveriam ao período de 6 anos. Seja  $p$  a proporção de lixeiras que sobrevivem ao teste acelerado. As hipóteses relevantes são  $H_0: p = 0,9$  versus  $H_a: p < 0,9$ . Uma decisão será tomada com base na estatística de teste  $X$ , o número das que sobrevivem entre as 20 lixeiras. Se o nível de significância desejado for  $\alpha = 0,05$ ,  $c$  deve satisfazer  $B(c; 20, 0,9) \leq 0,05$ . Pela Tabela A.1 do Apêndice,  $B(15; 20, 0,9) = 0,043$ , enquanto  $B(16; 20, 0,9) = 0,133$ . A região de rejeição apropriada é, portanto,  $x \leq 15$ . Se o teste acelerado resultasse em  $x = 14$ ,  $H_0$  seria rejeitada em favor de  $H_a$ , precisando de uma modificação da garantia proposta. A probabilidade de um erro tipo II para o valor alternativo  $p' = 0,8$  é

$$\begin{aligned} \beta(0,8) &= P(H_0 \text{ não é rejeitada quando } X \sim \text{Bin}(20, 0,8)) \\ &= P(X \geq 16 \text{ quando } X \sim \text{Bin}(20, 0,8)) \\ &= 1 - B(15; 20, 0,8) = 1 - 0,370 = 0,630 \end{aligned}$$

isto é, quando  $p = 0,8$ , 63% de todas as amostras que consistem em  $n = 20$  lixeiras resultariam na não-rejeição incorreta de  $H_0$ . Essa probabilidade de erro é alta porque 20 é um tamanho da amostra pequeno e  $p' = 0,8$  está próximo do valor nulo  $p_0 = 0,9$ . ■

---

## Exercícios | Seção 8.3 (35–43)

35. Os registros de DMV do estado indicam que, de todos os veículos que se submeteram ao teste de emissões durante o ano anterior, 70% passaram na primeira tentativa. Uma amostra aleatória de 200 carros testados em um país específico durante o ano atual indica que 124 passaram pelo teste inicial. Isso sugere que a proporção real desse país durante o ano atual é diferente da proporção anterior de todo o estado? Teste as hipóteses relevantes, usando  $\alpha = 0,05$ .
36. Um fabricante de baterias de níquel hidrogênio seleciona aleatoriamente 100 placas de níquel para testar as células, ciclando-as certa quantidade de vezes e determinando que 14 placas foram cobertas de bolhas.
  - a. Isso fornece uma forte evidência de que mais de 10% das placas formam bolhas sob tais circunstâncias? Determine e teste as hipóteses apropriadas, usando um nível de significância de 0,05. Ao chegar à sua conclusão, que tipo de erro você pode ter cometido?
  - b. Se for realmente o caso de 15% das placas formarem bolhas sob essas circunstâncias e se for usado um tamanho de amostra de 100, qual é a probabilidade de a hipótese nula da parte (a) não ser rejeitada pelo teste nível 0,05? Responda a essa pergunta para um tamanho da amostra de 200.
  - c. Quantas placas devem ser testadas para se ter  $\beta(0,15) = 0,10$  para o teste da parte (a)?
37. Uma amostra aleatória de 150 doações recentes em certo banco de sangue revela que 82 eram sangue tipo A. Isso sugere que a porcentagem real de doações tipo A é diferente de 40%, ou seja, a porcentagem da população que tem sangue tipo A? Faça um teste das hipóteses apropriadas, usando um nível de significância de 0,01. Sua conclusão teria sido diferente se fosse usado um nível de significância de 0,05?
38. A biblioteca de uma universidade rotineiramente faz o inventário completo das prateleiras uma vez por ano. Devido às novas regras de armazenamento nas prateleiras instituídas no ano anterior, o bibliotecário-chefe acredita ser possível economizar recursos adiando o inventário. Ele decide selecionar aleatoriamente 1000 livros da coleção da biblioteca para serem procurados de maneira preliminar. Se a evidência indicar fortemente

que a proporção real de livros mal arquivados nas prateleiras ou que não podem ser localizados for menor que 0,02, então, o inventário será adiado.

- a. Dentre os 1000 livros procurados, 15 estavam mal arquivados ou não foram localizados. Teste as hipóteses relevantes e aconselhe o bibliotecário sobre o que fazer (use  $\alpha = 0,05$ ).
  - b. Se a proporção real de livros mal arquivados ou perdidos nas prateleiras for realmente 0,01, qual é a probabilidade de o inventário ser (desnecessariamente) feito?
  - c. Se a proporção real for 0,05, qual é a probabilidade de o inventário ser adiado?
- 39.** O artigo “Statistical Evidence of Discrimination” (*J. Amer Stat. Assoc.*, 1982, p. 773-783) discute o caso do processo jurídico *Swain versus Alabama* (1965), no qual foi alegado haver discriminação contra os negros na seleção dos jurados. Os dados do censo sugeriram que 25% das pessoas elegíveis para jurados eram negras, ainda que uma amostra aleatória de 1050 chamadas para servirem a um possível trabalho havia apenas 177 negros. Usando um teste nível 0,01, esses dados estão fortemente de acordo para uma conclusão de discriminação?
- 40.** Um plano de fidelidade de viagens de executivos foi desenvolvido por uma companhia aérea sob a premissa de que 5% de seus clientes atuais se qualificariam para serem membros. Uma amostra aleatória de 500 clientes resultou em 40 que se qualificariam.
- a. Usando esses dados, teste no nível 0,01 a hipótese nula de que a premissa da companhia está correta contra a alternativa de não estar correta.
  - b. Qual é a probabilidade, quando o teste da parte (a) for usado, de a premissa da companhia ser considerada correta quando, na verdade, 10% dos clientes atuais se qualificam?
- 41.** A cada um, de um grupo de 20 jogadores de tênis intermediários, são dadas duas raquetes, uma com cordas de náilon e outra com cordas de tripa sintética. Após várias semanas jogando com as duas raquetes, cada jogador será solicitado a dizer qual dos dois tipos de corda ele preferiu. Seja  $p$  a proporção de todos os jogadores que preferem a tripa ao náilon, e  $X$  o número de jogadores na amostra que preferem tripa. Em virtude de as cordas de tripa serem mais caras, considere a hipótese nula de que no máximo 50% desses jogadores preferem tripa. Simplificamos o resultado a  $H_0: p = 0,5$ , planejando rejeitar  $H_0$  somente se a evidência da amostra favorecer fortemente as cordas de tripa.
- a. Qual das regiões de rejeição  $\{15, 16, 17, 18, 19, 20\}$ ,  $\{0, 1, 2, 3, 4, 5\}$ , ou  $\{0, 1, 2, 3, 17, 18, 19, 20\}$  é mais apropriada, e por que as outras duas não são adequadas?
  - b. Qual é a probabilidade de um erro tipo I para a região escolhida da parte (a)? A região especifica um teste nível 0,05? O teste nível 0,05 é o melhor?
  - c. Se 60% dos jogadores preferirem tripa, calcule a probabilidade de um erro tipo II, usando a região apropriada da parte (a). Repita, se 80% de todos os jogadores preferirem tripa.
  - d. Se 13 dos 20 jogadores preferirem tripa,  $H_0$  deve ser rejeitada usando um nível de significância de 0,10?
- 42.** Um fabricante de acessórios de encanamento desenvolveu um novo tipo de torneira sem arruela. Seja  $p = P$  (uma torneira desse tipo selecionada aleatoriamente ter um vazamento dentro de 2 anos sob uso normal). O fabricante decidiu continuar com a produção, a menos que seja determinado que  $p$  é muito grande: o valor aceitável limítrofe de  $p$  é especificado como 0,10. O fabricante decide submeter  $n$  dessas torneiras a um teste acelerado (aproximadamente 2 anos de uso normal). Com  $X =$  número entre as  $n$  torneiras que vazam antes da conclusão do teste, a produção começará, a menos que  $X$  observado seja muito grande. Decidiu-se que, se  $p = 0,10$ , a probabilidade de não continuar deve ser de no máximo 0,10, enquanto, se  $p = 0,30$ , a probabilidade de continuar deve ser de no máximo 0,10.  $n = 10$  pode ser usado?  $n = 20$ ?  $n = 25$ ? Qual é a região de rejeição apropriada para o  $n$  escolhido e quais as probabilidades de erro reais quando essa região é usada?
- 43.** Os cientistas acreditam que os robôs terão papel crucial nas fábricas, nas próximas décadas. Suponha que, em um experimento para determinar se o uso de robôs para fazer cabos de computador é viável, um robô tenha sido usado para montar 500 cabos. Os cabos foram examinados e havia 15 com defeito. Se os montadores humanos têm um índice de falha de 0,035 (3,5%), esses dados apóiam a hipótese de que a proporção de defeituosos é menor para os robôs do que para os homens? Use nível de significância de 0,01.

## 8.4 | Valores P

Uma forma de relatar o resultado da análise do teste de hipóteses é simplesmente dizer se a hipótese nula foi rejeitada em um nível de significância especificado. Dessa forma, um investigador pode afirmar que  $H_0$  foi rejeitada no nível de significância 0,05 ou que o uso de um teste nível 0,01 resultou na não-rejeição de  $H_0$ . Esse

tipo de afirmação é um tanto inadequada, pois não informa se o valor calculado da estatística de teste apenas caiu na região de rejeição ou se excedeu o valor crítico por um valor grande. Uma dificuldade adicional é que esse relatório impõe o nível de significância especificado sobre outros tomadores de decisões. Em muitas situações de decisão, os indivíduos podem ter visões diferentes com relação às conseqüências de um erro tipo I ou tipo II. Cada indivíduo, então, gostaria de selecionar o próprio nível de significância – alguns selecionam  $\alpha = 0,05$ , outros 0,01, e assim por diante – e, conseqüentemente, chegam a uma conclusão. Isso pode resultar em alguns indivíduos rejeitando  $H_0$  enquanto outros concluem que os dados não mostram uma contradição forte o suficiente de  $H_0$  que justifique sua rejeição.

**Exemplo 8.14**

O tempo médio real para o alívio inicial da dor do lenitivo mais vendido é de 10 min. Seja  $\mu$  o tempo médio real de alívio de um lenitivo desenvolvido recentemente na empresa, que espera produzi-lo e comercializá-lo somente se oferecer alívio mais rápido que o mais vendido, de modo que deseja testar  $H_0: \mu = 10$  versus  $H_a: \mu < 10$ . Somente se a evidência experimental levar à rejeição de  $H_0$  o novo lenitivo será introduzido. Após considerar a gravidade relativa de cada tipo de erro, um único nível de significância deve ser aceito, para a decisão – tomada nesse nível – rejeitar  $H_0$  e introduzir ou não o lenitivo.

Suponha que o novo lenitivo tenha sido introduzido. A empresa apóia sua alegação de alívio mais rápido, afirmando que, com base em uma análise dos dados experimentais,  $H_0: \mu = 10$  foi rejeitada em favor de  $H_a: \mu < 10$ , usando o nível de significância  $\alpha = 0,10$ . Quaisquer pessoas que considerem a mudança para esse novo lenitivo naturalmente gostariam de chegar às próprias conclusões relacionadas à validade da alegação. As pessoas satisfeitas com o lenitivo mais vendido veriam um erro tipo I (concluindo que o novo produto oferece alívio mais rápido, quando ele realmente não o faz) tão sério que podem querer usar  $\alpha = 0,05, 0,01$ , ou mesmo níveis menores. Infelizmente, a natureza da afirmação da empresa evita que um tomador de decisões individual tire uma conclusão nesse nível. A empresa impôs a própria escolha de nível de significância para os outros. O relatório podia ter sido feito de forma que permitisse flexibilidade a cada indivíduo para tirar uma conclusão com um  $\alpha$  selecionado pessoalmente. ■

Um valor  $P$  transmite muitas informações sobre a força da evidência contra  $H_0$  e permite que um tomador de decisões individual tire uma conclusão em qualquer nível  $\alpha$  especificado. Antes de darmos uma definição geral, considere como a conclusão em um problema de teste de hipóteses depende do nível  $\alpha$  selecionado.

**Exemplo 8.15**

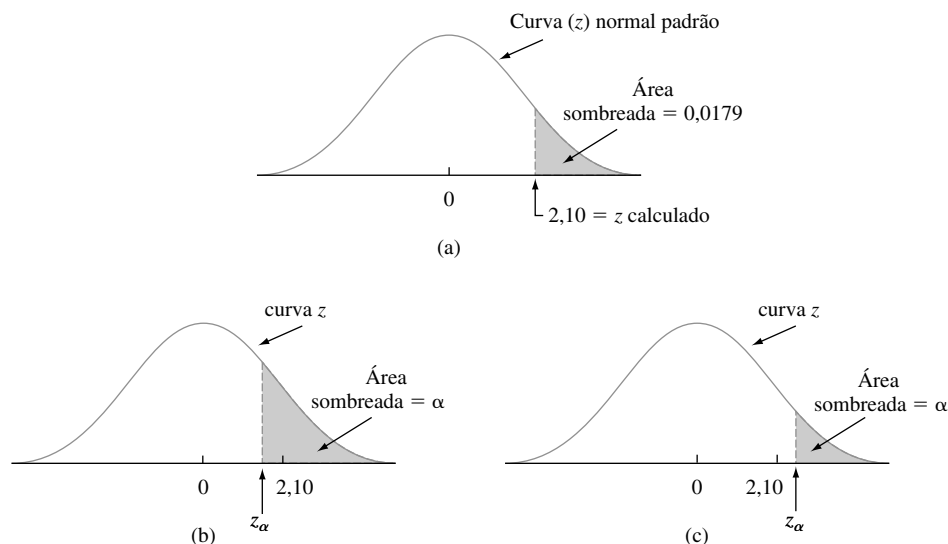
O problema do teor de nicotina discutido no Exemplo 8.5 envolveu o teste de  $H_0: \mu = 1,5$  versus  $H_a: \mu > 1,5$ . Devido à desigualdade em  $H_a$ , a região de rejeição é de cauda superior com  $H_0$  rejeitada se  $z \geq z_\alpha$ . Suponha que  $z = 2,10$ . A tabela a seguir exibe a região de rejeição para cada um dos quatro  $\alpha$ s diferentes, com a conclusão resultante.

Nível de Significância $\alpha$	Região de Rejeição	Conclusão
0,05	$z \geq 1,645$	Rejeite $H_0$
0,025	$z \geq 1,96$	Rejeite $H_0$
0,01	$z \geq 2,33$	Não rejeite $H_0$
0,005	$z \geq 2,58$	Não rejeite $H_0$

Para  $\alpha$  relativamente grande, o valor crítico  $z_\alpha$  não está muito longe na cauda superior; 2,10 excede o valor crítico e  $H_0$  é rejeitada. Entretanto, à medida que  $\alpha$  diminui, o valor crítico aumenta. Para  $\alpha$  pequeno, o valor crítico  $z$  é grande, 2,10 é menor que  $z_\alpha$  e  $H_0$  não é rejeitada.

Lembre-se de que, para um teste  $z$  de cauda superior,  $\alpha$  é apenas a área sob a curva  $z$  à direita do valor crítico de  $z_\alpha$ . Isto é, uma vez que  $\alpha$  é especificado, o valor crítico é escolhido para incluir a área de cauda superior  $\alpha$ . A Tabela A.3 do Apêndice mostra que a área à direita de 2,10 é 0,0179. Usando um  $\alpha$  maior que

0,0179 corresponde a  $z_\alpha < 2,10$ . Um  $\alpha$  menor que 0,0179 precisa usar um valor crítico  $z$  que exceda 2,10. A decisão em um nível específico  $\alpha$  depende, portanto, de como o  $\alpha$  selecionado se compara à área da cauda incluída pelo  $z$  calculado. Isso é ilustrado na Figura 8.5. Observe em particular que 0,0179, a área da cauda incluída, é o menor nível  $\alpha$  em que  $H_0$  seria rejeitada, pois usar qualquer  $\alpha$  menor resulta em um valor crítico  $z$  que excede 2,10, de modo que 2,10 não esteja na região de rejeição.



**Figura 8.5** Relação entre  $\alpha$  e a área da cauda incluída pelo  $z$  calculado: (a) área da cauda incluída pelo  $z$  calculado; (b) quando  $\alpha > 0,0179$ ,  $z_\alpha < 2,10$  e  $H_0$  é rejeitada; (c) quando  $\alpha < 0,0179$ ,  $z_\alpha > 2,10$  e  $H_0$  não é rejeitada

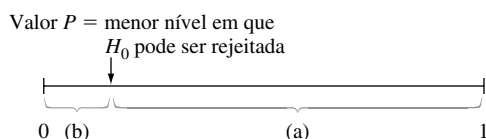
Em geral, suponha que a distribuição de probabilidades de uma estatística de teste quando  $H_0$  é verdadeira tenha sido determinada. Então, para  $\alpha$  especificado, a região de rejeição é determinada, encontrando um valor crítico ou valores que incluam a área da cauda  $\alpha$  (superior, inferior ou bicaudal, qualquer que seja apropriada) sob a curva da distribuição de probabilidades. O  $\alpha$  menor para o qual  $H_0$  seria rejeitada é a área da cauda incluída pelo valor calculado da estatística de teste. Esse  $\alpha$  menor é o valor  $P$ .

#### DEFINIÇÃO

O **valor  $P$**  (ou *nível de significância observado*) é o menor nível de significância em que  $H_0$  seria rejeitada, quando um procedimento de teste especificado é usado em um determinado conjunto de dados. Uma vez que o valor  $P$  tenha sido determinado, a conclusão, em qualquer nível específico  $\alpha$ , resulta da comparação do valor  $P$  a  $\alpha$ :

1. Valor  $P \leq \alpha \Rightarrow$  implica rejeição de  $H_0$  no nível  $\alpha$ .
2. Valor  $P > \alpha \Rightarrow$  implica não-rejeição de  $H_0$  no nível  $\alpha$ .

É comum chamar os dados de *significativos* quando  $H_0$  é rejeitada e de *não-significativos* em caso contrário. O valor  $P$  é então o menor nível em que os dados são significativos. Uma maneira fácil de visualizar a comparação do valor  $P$  com o  $\alpha$  escolhido é fazer um desenho como o da Figura 8.6. O cálculo do valor  $P$  depende do teste ser bicaudal ou de cauda superior ou inferior. Entretanto, uma vez que tenha sido calculada, a comparação com  $\alpha$  não depende do tipo de teste que foi utilizado.



**Figura 8.6** Comparando  $\alpha$  e o valor  $P$ : (a) rejeite  $H_0$  quando  $\alpha$  estiver aqui; (b) não rejeite  $H_0$  quando  $\alpha$  estiver aqui

**Exemplo 8.16** (continuação do Exemplo 8.14)

Suponha que, quando os dados de um experimento envolvendo o novo lenitivo para dor foram analisados, o valor  $P$  para teste de  $H_0: \mu = 10$  versus  $H_a: \mu < 10$  foi calculado como 0,0384. Uma vez que  $\alpha = 0,05$  é maior que o valor  $P$  (0,05 está no intervalo (a) da Figura 8.6),  $H_0$  seria rejeitada por qualquer pessoa que fizesse o teste no nível 0,05. Entretanto, no nível 0,01,  $H_0$  não seria rejeitada, pois 0,01 é menor que o menor nível (0,0384) em que  $H_0$  pode ser rejeitada. ■

Os pacotes de softwares estatísticos mais amplamente usados incluem automaticamente um valor  $P$  quando é feita uma análise do teste de hipóteses. Pode-se tirar uma conclusão diretamente da saída, sem consultar uma tabela de valores críticos.

Uma definição alternativa útil equivalente à dada é a seguinte:

**DEFINIÇÃO**

O **valor  $P$**  é a probabilidade calculada, assumindo-se que  $H_0$  seja verdadeira, de se obter um valor da estatística de teste pelo menos tão contraditório a  $H_0$  quanto o valor que realmente resultou. Quanto menor o valor  $P$ , mais contraditórios são os dados  $H_0$ .

Dessa forma, se  $z = 2,10$  para um teste  $z$  de cauda superior, o valor  $P(Z \geq 2,10, \text{ quando } H_0 \text{ é verdadeira}) = 1 - \Phi(2,10) = 0,0179$ , como anteriormente. Cuidado: o valor  $P$  não é a probabilidade de  $H_0$  ser verdadeira nem probabilidade de erro!

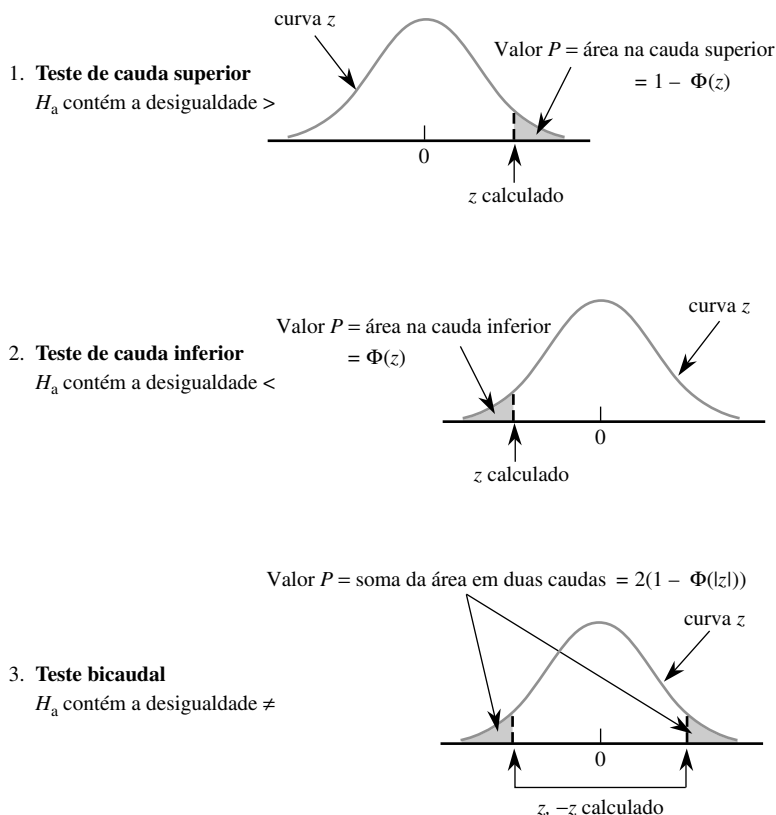
**Valores  $P$  para Testes  $z$**

O valor  $P$  para um teste  $z$  (baseado em uma estatística de teste cuja distribuição, quando  $H_0$  é verdadeira, é pelo menos aproximadamente normal padronizada) é determinado com facilidade a partir das informações na Tabela A.3 do Apêndice. Considere um teste de cauda superior e seja  $z$  o valor calculado da estatística de teste  $Z$ . A hipótese nula é rejeitada se  $z \geq z_\alpha$  e o valor  $P$  é o menor  $\alpha$  para o qual esse é o caso. Uma vez que  $z_\alpha$  aumenta à medida que  $\alpha$  diminui, o valor  $P$  é o valor de  $\alpha$  para o qual  $z = z_\alpha$ . Isto é, o valor  $P$  é apenas a área incluída pelo valor calculado de  $z$  na cauda superior da curva normal padronizada. A área acumulada correspondente é  $\Phi(z)$ , de modo que, nesse caso, o valor  $P = 1 - \Phi(z)$ .

Um argumento análogo para um teste de cauda inferior mostra que o valor  $P$  é a área incluída pelo valor calculado  $z$  na cauda inferior da curva normal padronizada. Deve-se tomar mais cuidado no caso de um teste bicaudal. Suponha primeiro que  $z$  seja positivo. Então, o valor  $P$  é o valor de  $\alpha$  que satisfaz  $z = z_{\alpha/2}$  (isto é,  $z$  calculado = valor crítico de cauda superior). Isso diz que a área incluída na cauda superior é metade do valor  $P$ , de modo que o valor  $P = 2[1 - \Phi(z)]$ . Se  $z$  é negativo, o valor  $P$  é o  $\alpha$  para o qual  $z = -z_{\alpha/2}$ , ou, de modo equivalente,  $-z = z_{\alpha/2}$ , de modo que o valor  $P = 2[1 - \Phi(-z)]$ . Uma vez que  $-z = |z|$  quando  $z$  é negativo, o valor  $P = 2[1 - \Phi(|z|)]$  para  $z$  negativo ou positivo.

$$\text{Valor } P: \quad P = \begin{cases} 1 - \Phi(z) & \text{para teste de cauda superior} \\ \Phi(z) & \text{para teste de cauda inferior} \\ 2[1 - \Phi(|z|)] & \text{para teste bicaudal} \end{cases}$$

Cada um desses é a probabilidade de se obter um valor pelo menos tão extremo quanto o que foi obtido (assumindo  $H_0$  como verdadeira). Os três casos são ilustrados na Figura 8.7.



**Figura 8.7** Determinação do valor  $P$  para o teste  $z$

O próximo exemplo ilustra o uso da abordagem do valor  $P$  ao teste de hipóteses por meio de uma seqüência de passos modificados de nossa seqüência recomendada anteriormente.

### Exemplo 8.17

A espessura-alvo dos *wafers* de silício usados em um tipo de circuito integrado é  $245 \mu\text{m}$ . Uma amostra de 50 *wafers* é obtida, e a espessura de cada uma é determinada, resultando em uma espessura média amostral de  $246,8 \mu\text{m}$  e um desvio padrão da amostra de  $3,60 \mu\text{m}$ . Esses dados sugerem que a espessura média real do *wafers* é diferente do valor-alvo?

1. Parâmetro de interesse:  $\mu = \text{espessura média real do wafer}$
2. Hipótese nula:  $H_0: \mu = 245$
3. Hipótese alternativa:  $H_a: \mu \neq 245$
4. Fórmula do valor da estatística de teste:  $z = \frac{\bar{x} - 245}{s/\sqrt{n}}$
5. Cálculo do valor da estatística de teste:  $z = \frac{246,18 - 245}{3,60/\sqrt{50}} = 2,32$
6. Determinação do valor  $P$ : em virtude de o teste ser bicaudal.

$$\text{Valor } P = 2(1 - \Phi(2,32)) = 0,0204$$

7. Conclusão: Usando um nível de significância de  $0,01$ ,  $H_0$  não seria rejeitada, uma vez que  $0,0204 > 0,01$ . Nesse nível de significância, existe evidência insuficiente para concluir que a espessura média real é diferente do valor-alvo. ■

### Valores $P$ para testes $t$

Assim como o valor  $P$  para um teste  $z$  é a área de uma curva  $z$ , o valor  $P$  para um teste  $t$  será a área de uma curva  $t$ . A Figura 8.8 ilustra os três diferentes casos. O número de gl para o teste  $t$  de uma amostra é  $n - 1$ .

A tabela de valores críticos  $t$  usada anteriormente para intervalos de previsão e confiança não contém informações suficientes sobre qualquer distribuição  $t$  específica para permitir determinação precisa das áreas desejadas. Portanto, incluímos outra tabela  $t$  na Tabela A.8 do Apêndice, que contém a tabulação de áreas da curva  $t$  de cauda superior. Cada coluna diferente da tabela é para um número diferente de gl, e as linhas são para valores calculados da estatística de teste  $t$  que varia de 0,0 a 4,0 em incrementos de 0,1. Por exemplo, o número 0,074 aparece na interseção da linha 1,6 e a coluna de gl 8, de modo que a área sob a curva de gl 8 à direita de 1,6 (área de cauda superior) é 0,074. Em virtude de as curvas  $t$  serem simétricas, 0,074 também é a área sob a curva de gl 8 à esquerda de  $-1,6$  (área da cauda inferior).

Suponha, por exemplo, que um teste de  $H_0: \mu = 100$  versus  $H_a: \mu > 100$  baseie-se na distribuição  $t$  de gl 8. Se o valor calculado da estatística de teste for  $t = 1,6$ , então o valor  $P$  para esse teste de cauda superior é 0,074. Em virtude de 0,074 exceder 0,05, não conseguiríamos rejeitar  $H_0$  em um nível de significância de 0,05. Se a hipótese alternativa for  $H_a: \mu < 100$  e um teste com base em gl 20 produz  $t = -3,2$ , então a Tabela A.8 do Apêndice mostra que o valor  $P$  é a área incluída de cauda inferior 0,002. A hipótese nula pode ser rejeitada no nível 0,05 ou 0,01. Considere o teste de  $H_0: \mu_1 - \mu_2 = 0$  versus  $H_a: \mu_1 - \mu_2 \neq 0$ ; a hipótese nula

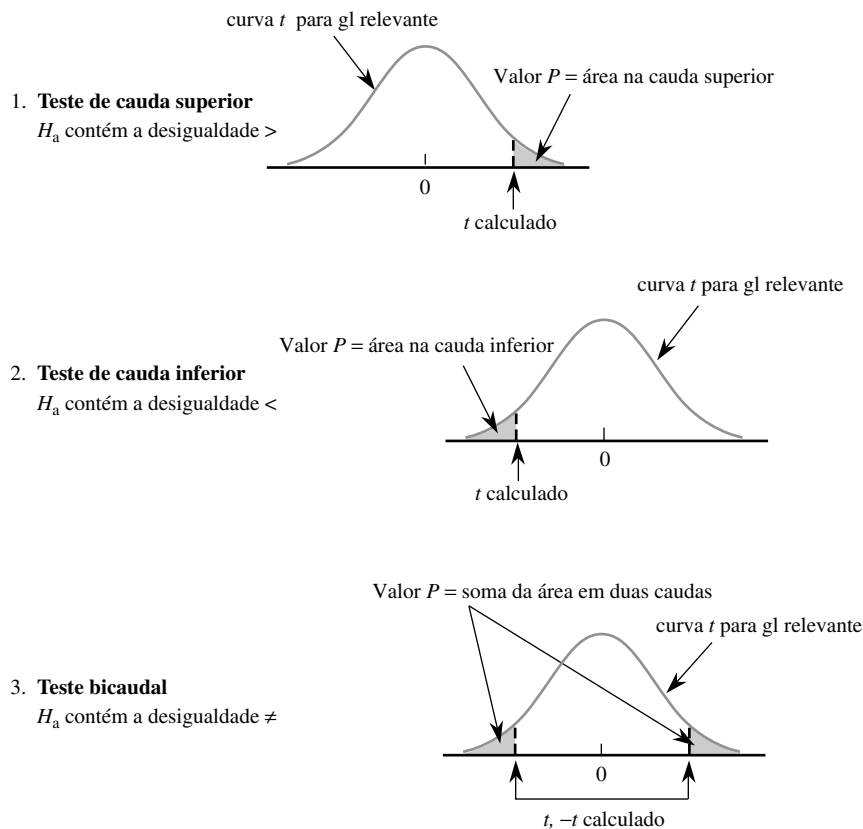


Figura 8.8 Valores  $P$  para testes  $t$

afirma que as médias das duas populações são idênticas, enquanto a hipótese alternativa afirma que são diferentes, sem especificar a direção do desvio de  $H_0$ . Se um teste  $t$  baseia-se em gl 20 e  $t = 3,2$ , então o valor  $P$  para esse teste bicaudal é  $2(0,002) = 0,004$ . Esse também seria o valor  $P$  para  $t = -3,2$ . A área da cauda é dobrada, pois os valores maiores que 3,2 e menores que  $-3,2$  são mais contraditórios para  $H_0$  do que o calculado (valores mais distantes em *qualquer* cauda da curva  $t$ ).



## Exemplo 8.18

No Exemplo 8.9, fizemos um teste de  $H_0: \mu = 25$  versus  $H_a: \mu > 25$  com base em gl 4. O valor calculado de  $t$  foi 1,04. Observando a coluna de gl 4 da Tabela A.8 do Apêndice e a linha 1,0, vemos que a entrada é 0,187, de modo que o valor  $P \approx 0,187$ . Esse valor  $P$  é claramente maior que qualquer nível de significância razoável  $\alpha$  (0,01, 0,05 e mesmo 0,10), assim, não há motivo para rejeitar a hipótese nula. A saída do MINITAB incluída no Exemplo 8.9 possui valor  $P = 0,18$ . Os valores  $P$  dos pacotes de software serão mais precisos que o que resulta da Tabela A.8 do Apêndice, uma vez que os valores de  $t$  em nossa tabela são precisos somente até dígitos decimais. ■

## Exercícios | Seção 8.4 (44–58)

44. Para qual dos valores  $P$  dados a hipótese nula seria rejeitada ao realizar um teste nível 0,05?
- 0,001
  - 0,021
  - 0,078
  - 0,047
  - 0,148
45. São dados pares de valores  $P$  e níveis de significância,  $\alpha$ . Para cada par, afirme se o valor  $P$  observado levaria à rejeição de  $H_0$  no nível de significância dado.
- Valor  $P = 0,084$ ,  $\alpha = 0,05$
  - Valor  $P = 0,003$ ,  $\alpha = 0,001$
  - Valor  $P = 0,498$ ,  $\alpha = 0,05$
  - Valor  $P = 0,084$ ,  $\alpha = 0,10$
  - Valor  $P = 0,039$ ,  $\alpha = 0,01$
  - Valor  $P = 0,218$ ,  $\alpha = 0,10$
46. Seja  $\mu$  o tempo de reação médio a um certo estímulo. Para um teste  $z$  de amostra grande de  $H_0: \mu = 5$  versus  $H_a: \mu > 5$ , determine o valor  $P$  associado a cada um dos valores dados da estatística de teste  $z$ .
- 1,42
  - 0,90
  - 1,96
  - 2,48
  - 0,11
47. Pneus de um determinado tipo recentemente comprados são supostamente enchidos a uma pressão de 30 lb/in<sup>2</sup>. Seja  $\mu$  a pressão média real. Determine o valor  $P$  associado a cada valor da estatística  $z$  dado para testar  $H_0: \mu = 30$  versus  $H_a: \mu \neq 30$ .
- 2,10
  - 1,75
  - 0,55
  - 1,41
  - 5,3
48. Forneça o máximo de informações que puder sobre o valor  $P$  de um teste  $t$  em cada uma das situações a seguir:
- Teste de cauda superior, gl = 8,  $t = 2,0$
  - Teste de cauda inferior, gl = 11,  $t = -2,4$
  - Teste bicaudal, gl = 15,  $t = -1,6$
  - Teste de cauda superior, gl = 19,  $t = -0,4$
  - Teste de cauda superior, gl = 5,  $t = 5,0$
  - Teste bicaudal, gl = 40,  $t = -4,8$
49. A tinta usada para fazer as linhas nas estradas deve refletir luz suficiente para ser claramente visível à noite. Seja  $\mu$  a leitura média real do reflectômetro de um novo tipo de tinta em consideração. Um teste de  $H_0: \mu = 20$  versus  $H_a: \mu > 20$  se baseará em uma amostra aleatória de tamanho  $n$  de uma distribuição de população normal.
- Que conclusão é apropriada em cada uma das situações a seguir?
- $n = 15$ ,  $t = 3,2$ ,  $\alpha = 0,05$
  - $n = 9$ ,  $t = 1,8$ ,  $\alpha = 0,01$
  - $n = 24$ ,  $t = -0,2$
50. Seja  $\mu$  a concentração média real de receptor de soro de todas as gestantes. A média de todas as mulheres é 5,63. O artigo “Serum Transferrin Receptor for the Detection of Iron Deficiency in Pregnancy” (*Amer J. Clinical Nutr*, 1991, p. 1077-1081) relata que o valor  $P > 0,10$  para um teste de  $H_0: \mu = 5,63$  versus  $H_a: \mu \neq 5,63$  com base em  $n = 176$  gestantes. Usando um nível de significância de 0,01, a que conclusão você chegaria?
51. Um fabricante de aspirina enche os frascos por peso em vez de fazê-lo por quantidade. Uma vez que cada frasco deve conter 100 comprimidos, o peso médio por unidade deve ser de 5 *grains*. Cada um dos 100 comprimidos tirados de um lote bastante grande é pesado, resultando em um peso médio amostral por comprimido de 4,87 *grains* e um desvio padrão da amostra de 35 *grains*. Essas informações fornecem forte evidência para concluir que a empresa não está enchendo seus vidros conforme anunciado? Teste as hipóteses apropriadas, usando  $\alpha = 0,01$  calculando primeiro o valor  $P$  e depois comparando-o com o nível de significância especificado.
52. Devido à variabilidade no processo de manufatura, o ponto de escoamento real de uma amostra de aço leve sujeita à força crescente geralmente será diferente do ponto de escoamento teórico. Seja  $p$  a proporção real de amostras que escoam antes de seu ponto de escoamento teórico. Se, com base em uma amostra, é possível concluir que mais de 20% dos espécimes que escoam antes do ponto teórico, o processo de produção deve ser modificado.
- Se 15 de 60 espécimes escoam antes do ponto teórico, qual é o valor  $P$  quando o teste apropriado é usado e o que você recomendaria que a empresa fizesse?
  - Se a percentagem real dos “primeiros escoamentos” é realmente 50% (de modo que o ponto teórico é a mediana da distribuição do escoamento), e for usado um teste nível 0,01, qual é a probabilidade de a empresa concluir que é necessário modificar o processo?

53. Muitos consumidores estão mudando para genéricos como forma de reduzir o custo dos medicamentos prescritos. O artigo “Commercial Information on Drugs: Confusing to the Physician?” (*J. of Drug Issues*, 1988, p. 245-257) fornece os resultados de uma pesquisa de 102 médicos. Somente 47 dos pesquisados sabiam o nome genérico do medicamento à base de metadona. Esse resultado indica forte evidência de que menos da metade dos médicos sabe o nome genérico da metadona? Faça um teste de hipóteses usando nível de significância de 0,01 e utilizando o método do valor  $P$ .

54. Uma amostra aleatória de espécimes de solo foi obtida, e a quantidade de material orgânico (%) no solo foi determinada para cada espécime, resultando nos dados a seguir (de “Engineering Properties of Soil,” *Soil Science*, 1998, p. 93-102):

1,10	5,09	0,97	1,59	4,60	0,32	0,55	1,45
0,14	4,47	1,20	3,50	5,02	4,67	5,22	2,69
3,98	3,17	3,03	2,21	0,69	4,47	3,31	1,17
0,76	1,17	1,57	2,62	1,66	2,05		

Os valores da média amostral, do desvio padrão da amostra e do erro-padrão (estimado) da média são 2,481, 1,616 e 0,295, respectivamente. Esses dados sugerem que a porcentagem média real de material orgânico no solo é diferente de 3%? Faça um teste das hipóteses apropriadas com nível de significância 0,10, determinando primeiro o valor  $P$ . Sua conclusão seria diferente se  $\alpha = 0,05$  fosse usado? (*Observação*: um gráfico de probabilidade normal dos dados mostra um padrão aceitável considerando o tamanho da amostra razoavelmente grande.)

55. Os tempos de ativação inicial de aspersores de uma série de testes com sistemas de extinção de incêndio por aspersão, usando uma espuma de formação de película aquosa, foram (em segundos):

27	41	22	27	23	35	30	33	24	27	28	22	24
----	----	----	----	----	----	----	----	----	----	----	----	----

(consulte “Use of AFFF in Sprinkler Systems,” *Fire Technology*, 1976, p. 5). O sistema foi projetado de modo que o tempo de ativação médio real seja de no máximo 25 s sob tais condições. Os dados contradizem fortemente a validade da especificação desse projeto? Teste as hipóteses relevantes no nível de significância 0,05 usando a abordagem do valor  $P$ .

56. Uma caneta foi projetada de modo que o tempo de escrita médio real sob condições controladas (que envolvem o uso de uma máquina de escrever) seja de no mínimo 10 horas. Uma amostra aleatória de 18 canetas é selecionada, o tempo de escrita de cada uma determinado, e um gráfico de probabilidade normal dos dados resultantes auxilia o uso de um teste  $t$  de uma amostra.

- a. Quais hipóteses devem ser testadas se os investigadores acreditam, *a priori*, que a especificação do projeto foi satisfeita?
- b. Que conclusão é apropriada, se as hipóteses da parte (a) forem testadas,  $t = -2,3$ , e  $\alpha = 0,05$ ?
- c. Que conclusão é apropriada se as hipóteses da parte (a) forem testadas,  $t = -1,8$ , e  $\alpha = 0,01$ ?
- d. O que se deve concluir se as hipóteses da parte (a) forem testadas e  $t = -3,6$ ?

57. Está sendo testada a precisão de um espectrofotômetro usado para medir a concentração de CO [ppm (partes por milhão) por volume] fazendo leituras de um gás fabricado (chamado gás de calibração) em que a concentração de CO é precisamente controlada em 70 ppm. Se as leituras sugerirem que o espectrofotômetro não está funcionando corretamente, ele deve ser recalibrado. Assuma que, se ele for corretamente calibrado, a concentração medida das amostras de gás de calibração é normalmente distribuída. Com base nas seis leituras — 85, 77, 82, 68, 72 e 69 — a recalibração é necessária? Faça um teste das hipóteses relevantes, usando a abordagem do valor  $P$  com  $\alpha = 0,05$ .

58. A condutividade relativa de um dispositivo semicondutor é determinada pela quantidade de impureza “aplicada” no dispositivo durante sua fabricação. Um diodo de silício a ser usado para um determinado fim requer uma voltagem de corte média de 0,60 V e, se ela não for obtida, a quantidade de impureza deve ser ajustada. Uma amostra de diodos foi selecionada, e a voltagem de corte, determinada. A saída do SAS a seguir resultou de uma solicitação para testar as hipóteses apropriadas.

N	Média	Desv Pad	T	Prob> T
15	0,0453333	0,0899100	1,9527887	0,0711

[*Observação*: SAS testa explicitamente  $H_0: \mu = 0$ , de modo a testar  $H_0: \mu = 0,60$ , o valor nulo 0,60 deve ser subtraído de cada  $x_i$ ; a média relatada é, então, a média dos valores de  $(x_i - 0,60)$ . Além disso, o valor  $P$  de SAS é sempre para teste bicaudal.] O que se concluiria para um nível de significância de 0,01? 0,05? 0,10?

## 8.5 | Alguns Comentários sobre como Selecionar um Procedimento de Teste

Uma vez que o investigador decidiu sobre a questão de interesse e o método de se obter dados (o modelo do experimento), a construção de um procedimento de teste apropriado consiste de três passos distintos:

1. Especifique uma estatística de teste (a função dos valores observados que servirão para a tomada de decisões).
2. Decida sobre a forma geral da região de rejeição (rejeitar tipicamente  $H_0$  para valores apropriadamente grandes da estatística de teste, rejeitar para valores apropriadamente pequenos ou rejeitar para valores grandes ou pequenos).
3. Selecione o valor ou valores críticos numéricos específicos que separarão a região de rejeição da região de aceitação (obtendo a distribuição da estatística de teste quando  $H_0$  é verdadeira e, depois, selecionando um nível de significância).

Nos exemplos fornecidos até agora, os Passos 1 e 2 foram realizados de modo particular por meio da intuição. Por exemplo, quando a população subjacente foi considerada normal com média  $\mu$  e  $\sigma$  conhecido, fomos levados de  $\bar{X}$  à estatística de teste padronizada

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

Para testar  $H_0: \mu = \mu_0$  versus  $H_a: \mu > \mu_0$ , a intuição então sugeriu rejeitar  $H_0$  quando  $z$  fosse grande. Finalmente, o valor crítico foi determinado especificando o nível de significância  $\alpha$  e usando o fato de  $Z$  ter uma distribuição normal padronizada quando  $H_0$  é verdadeira. A confiabilidade do teste ao chegar a uma decisão correta pode ser avaliada, estudando-se as probabilidades de erro tipo II.

Os problemas a serem considerados na realização dos Passos 1-3 incluem as questões a seguir:

1. Quais as implicações práticas e as conseqüências de se escolher um nível de significância específico, uma vez que os outros aspectos de um procedimento de teste foram determinados?
2. Existe um princípio geral — que não depende apenas da intuição — que pode ser usado para obter procedimentos de teste bons ou melhores?
3. Quando dois ou mais testes são apropriados em uma determinada situação, como podem ser comparados para escolher qual deles deve ser usado?
4. Se um teste é derivado sob suposições específicas com relação à distribuição ou população que está servindo de amostra, quão bem o procedimento de teste funcionará quando as suposições forem violadas?

## Significância Prática Versus Estatística

Embora o processo de se chegar a uma decisão usando a metodologia de teste clássico de hipóteses envolva a seleção de um nível de significância e a rejeição ou não-rejeição de  $H_0$  nesse nível  $\alpha$ , relatar simplesmente o  $\alpha$  usado e a decisão tomada transmite pouco sobre a informação contida nos dados da amostra. Especialmente quando os resultados de um experimento devem ser comunicados a um grande público, a rejeição de  $H_0$  no nível 0,05 será muito mais convincente se o valor observado da estatística de teste exceder em muito o valor crítico de 5% do que se exceder por pouco esse valor. Foi precisamente o que levou ao conceito do valor  $P$  como uma forma de relatar a significância sem impor um  $\alpha$  específico sobre outros que possam querer tirar as próprias conclusões.

Mesmo se um valor  $P$  for incluído em um resumo de resultados, entretanto, pode haver dificuldade ao interpretar esse valor e ao tomar uma decisão. Isso porque um valor  $P$  pequeno, que geralmente indicaria **significância estatística**, que sugere fortemente a rejeição de  $H_0$  em favor de  $H_a$ , pode ser o resultado de uma amostra de tamanho grande em combinação com um desvio de  $H_0$  que possui pouca **significância prática**. Em muitas situações experimentais, somente desvios de  $H_0$  de grande magnitude seriam de detecção útil, enquanto um pequeno desvio de  $H_0$  teria pouca significância prática.

Considere como exemplo o teste de  $H_0: \mu = 100$  versus  $H_a: \mu > 100$  em que  $\mu$  é a média de uma população normal com  $\sigma = 10$ . Suponha que um valor real de  $\mu = 101$  não representasse um desvio grave de  $H_0$  no sentido de que não rejeitar  $H_0$  quando  $\mu = 101$  seria um erro relativamente barato. Para um tamanho  $n$  de amostra

razoavelmente grande, esse  $\mu$  levaria a um valor de  $\bar{x}$  próximo a 101, de modo que não gostaríamos que essa evidência da amostra sustentasse fortemente a rejeição de  $H_0$ , quando  $\bar{x} = 101$  é observado. Para vários tamanhos de amostra, a Tabela 8.1 registra o valor  $P$  quando  $\bar{x} = 101$  e também a probabilidade de não rejeitar  $H_0$  no nível 0,01 quando  $\mu = 101$ .

A segunda coluna na Tabela 8.1 mostra que, mesmo para tamanhos de amostra moderadamente grandes, o valor  $P$  de  $\bar{x} = 101$  sustenta fortemente a rejeição de  $H_0$ , enquanto o próprio  $\bar{x}$  observado sugere que em termos práticos o valor real de  $\mu$  difere pouco do valor nulo  $\mu_0 = 100$ . A terceira coluna indica que, mesmo quando existe pouca diferença prática entre o  $\mu$  real e o valor nulo, para um nível fixo de significância, um tamanho de amostra grande quase sempre levará à rejeição da hipótese nula naquele nível. Para resumir, *deve-se ser especialmente cuidadoso ao interpretar a evidência quando o tamanho da amostra é grande, uma vez ser quase certo que qualquer desvio pequeno de  $H_0$  será detectado por um teste, ainda que possa ter pouca significância prática.*

**Tabela 8.1** Uma ilustração do efeito do tamanho da amostra sobre os valores  $P$  e  $\beta$

$n$	Valor $P$ quando $\bar{x} = 101$	$\beta(101)$ para Teste de nível 0,01
25	0,3085	0,9664
100	0,1587	0,9082
400	0,0228	0,6293
900	0,0013	0,2514
1600	0,0000335	0,0475
2500	0,000000297	0,0038
10.000	$7,69 \times 10^{-24}$	0,0000

### O Princípio de Razão de Verossimilhança

Sejam  $x_1, x_2, \dots, x_n$  as observações em uma amostra aleatória de tamanho  $n$  de uma distribuição de probabilidades  $f(x; \theta)$ . A distribuição conjunta avaliada nesses valores da amostra é o produto  $f(x_1; \theta) \cdot f(x_2; \theta) \cdot \dots \cdot f(x_n; \theta)$ . Como na discussão da estimativa da máxima verossimilhança, a *função verossimilhança* é essa distribuição conjunta considerada como função de  $\theta$ . Considere o teste de  $H_0: \theta$  está em  $\Omega_0$ , versus  $H_a: \theta$  está em  $\Omega_a$ , em que  $\Omega_0$  e  $\Omega_a$  são disjuntos (por exemplo,  $H_0: \theta \leq 100$  versus  $H_a: \theta > 100$ ). O **princípio de razão de verossimilhança** para a construção do teste continua como segue:

1. Determine o maior valor da verossimilhança para qualquer  $\theta$  em  $\Omega_0$ , (determinando a estimativa da máxima verossimilhança dentro de  $\Omega_0$  e o substituindo na função verossimilhança).
2. Determine o maior valor da verossimilhança para qualquer  $\theta$  em  $\Omega_a$ .
3. Forme a razão

$$\lambda(x_1, \dots, x_n) = \frac{\text{máxima verossimilhança de } \theta \text{ em } \Omega_0}{\text{máxima verossimilhança de } \theta \text{ em } \Omega_a}$$

A razão  $\lambda(x_1, \dots, x_n)$  é chamada de *valor estatístico de razão de verossimilhança*. O procedimento de teste consiste em rejeitar  $H_0$  quando essa razão é pequena. Isto é, uma constante  $k$  é escolhida, e  $H_0$  é rejeitada se  $\lambda(x_1, \dots, x_n) \leq k$ . Dessa forma,  $H_0$  é rejeitada quando o denominador de  $\lambda$  excede muito o numerador, indicando que os dados são muito mais consistentes com  $H_a$  do que com  $H_0$ .

A constante  $k$  é selecionada para produzir a probabilidade de erro tipo I desejada. Geralmente, a desigualdade  $\lambda \leq k$  pode ser manipulada para produzir uma condição equivalente mais simples. Por exemplo, para testar  $H_0: \mu \leq \mu_0$  versus  $H_a: \mu > \mu_0$  no caso de normalidade,  $\lambda \leq k$  é equivalente a  $t \geq c$ . Dessa forma, com  $c = t_{\alpha, n-1}$ , o teste de razão de verossimilhança é o teste  $t$  de uma amostra.

O princípio de razão de verossimilhança também pode ser aplicado quando os  $X_i$ s possuem distribuições diferentes e mesmo quando são dependentes, embora a função verossimilhança possa ser complicada em tais

casos. Muitos dos procedimentos de teste a serem apresentados nos capítulos subsequentes são obtidos a partir do princípio de razão de verossimilhança. Esses testes geralmente são formados para minimizar  $\beta$  entre todos os testes que possuem o  $\alpha$  desejado, portanto são de fato os melhores testes. Para mais detalhes e alguns exemplos, consulte o livro de DeGroot ou o de Hogg e Craig listados na bibliografia do Capítulo 6.

Uma limitação prática sobre o uso do princípio de razão de verossimilhança é que, para construir a estatística de teste dessa razão, a forma da distribuição de probabilidades, da qual a amostra é tirada, deve ser especificada. Para encontrar o teste  $t$  do princípio de razão de verossimilhança, o investigador deve assumir um fdp normal. Se o investigador deseja assumir que a distribuição é simétrica, mas não quer ser específico com relação a sua forma exata (tal como normal, uniforme ou Cauchy), então o princípio falha, pois não há como escrever uma fdp conjunta simultaneamente válida para todas as distribuições simétricas. No Capítulo 15, apresentaremos diversos procedimentos de teste **distribuição livre**, assim chamado porque a probabilidade de erro tipo I é controlada simultaneamente para muitas distribuições subjacentes diferentes. Esses procedimentos são úteis quando o investigador tem conhecimento limitado da distribuição subjacente. Devemos também falar mais sobre as questões 3 e 4 listadas no início desta seção.

---

## Exercícios | Seção 8.5 (59–60)

59. Reconsidere o problema de secagem de pintura discutido no Exemplo 8.2. As hipóteses eram  $H_0: \mu = 75$  versus  $H_a: \mu < 75$ , com  $\sigma$  assumido para ter valor 9,0. Considere o valor alternativo  $\mu = 74$ , que no contexto do problema não seria presumivelmente um desvio praticamente significativo de  $H_0$ .
- Para um teste nível 0,01, calcule  $\beta$  para essa alternativa para os tamanhos de amostra  $n = 100, 900$  e 2500.
  - Se o valor observado de  $\bar{X}$  é  $\bar{x} = 74$ , o que você pode dizer sobre o valor  $P$  resultante quando  $n = 2500$ ? Os dados são estatisticamente significativos em quaisquer dos valores padrão de  $\alpha$ ?
- Você realmente quer usar um tamanho de amostra de 2500 em um teste nível 0,01 (desconsiderando o custo de tal experimento)? Explique.
60. Considere o teste nível 0,01 de amostra grande na Seção 8.3 para testar  $H_0: p = 0,2$  contra  $H_a: p > 0,2$ .
- Para o valor alternativo  $p = 0,21$ , calcule  $\beta(0,21)$  para os tamanhos de amostra  $n = 100, 2500, 10.000, 40.000$  e 90.000.
  - Para  $\hat{p} = x/n = 0,21$ , calcule o valor  $P$  quando  $n = 100, 2500, 10.000$  e 40.000.
  - Na maioria das situações, seria razoável usar um teste nível 0,01 em conjunto com um tamanho de amostra de 40.000? Por quê?

---

## Exercícios Suplementares (61–82)

61. Uma amostra de 50 lentes usadas em óculos produz uma espessura média amostral de 3,05 mm e um desvio padrão da amostra de 0,34 mm. A espessura média real desejada de tais lentes é 3,20 mm. Os dados sugerem fortemente que a espessura média real das lentes é diferente da desejada? Teste usando  $\alpha = 0,05$ .
62. No Exercício 61, suponha que o investigador, antes de coletar os dados, acreditou que o valor de  $\sigma$  era aproximadamente 0,30. Se ele quisesse que a probabilidade de um erro tipo II fosse 0,05 quando  $\mu = 3,00$ , um tamanho de amostra 50 era desnecessariamente grande?
63. Especificou-se que certo tipo de ferro devia conter 0,85 g de silício por 100 g de ferro (85%). O conteúdo de silício de cada um dos 25 espécimes de ferro selecionados aleatoriamente foi determinado, e a saída do MINITAB a seguir resultou de um teste das hipóteses apropriadas.

Variável	N	Média	Desv Pad	SE Média	T	P
sil cont	25	0,8880	0,1807	0,0361	1,05	0,30

- Quais hipóteses foram testadas?
  - A que conclusão se pode chegar para um nível de significância de 0,05, e por quê? Responda à mesma pergunta para um nível de significância de 0,10.
64. Um método de endireitar fios antes de enrolá-los para fazer um rolo é chamado de *roller straightening*. O artigo “The Effect of Roller and Spinner Wire Straightening on Coiling Performance and Wire Properties” (*Springs*, 1987, p. 27-28) relata as propriedades de tensão do fio. Suponha que uma amostra de 16 fios seja selecionada, e cada um deles testado para determinar a resistência à tensão ( $N/mm^2$ ). A média amostral e o desvio padrão resultantes são 2160 e 30, respectivamente.
- A resistência à tensão média dos rolos feitos com *spinner straightening* é 2150  $N/mm^2$ . Quais hipóte-

ses devem ser testadas para determinar se a resistência à tensão média do método do rolo excede 2150?

- b. Assumindo que a distribuição da resistência à tensão é aproximadamente normal, qual estatística de teste você usaria para testar as hipóteses na parte (a)?
- c. Qual é o valor da estatística de teste para esses dados?
- d. Qual é o valor  $P$  para o valor da estatística de teste calculado na parte (c)?
- e. Para um teste nível 0,05, a que conclusão você chegaria?

65. Um novo método para medir os níveis de fósforo no solo é descrito no artigo “A Rapid Method to Determine Total Phosphorus in Soils” (*Soil Sci. Amer J.* 1988, p. 1301-1304). Suponha que uma amostra de 11 espécimes de solo, cada um com um teor real de fósforo de 548 mg/kg, seja analisada usando-se o novo método. A média amostral e o desvio padrão resultantes do nível de fósforo são 587 e 10, respectivamente.

- a. Existe evidência de que o nível médio de fósforo relatado pelo novo método difere significativamente do valor real de 548 mg/kg? Use  $\alpha = 0,05$ .
- b. Quais suposições você deve fazer para que o teste na parte (a) seja apropriado?

66. O artigo “Orchard Floor Management Utilizing Soil-Applied Coal Dust for Frost Protection” (*Agri. and Forest Meteorology*, 1988, p. 71-82) relata os seguintes valores para o fluxo de calor do solo de oito canteiros cobertos com pó de carvão.

34,7 35,4 34,7 37,7 32,5 28,0 18,4 24,9

O fluxo de calor médio do solo dos canteiros cobertos somente com grama é 29,0. Assumindo que a distribuição do fluxo de calor seja aproximadamente normal, os dados sugerem que o pó de carvão é eficaz no aumento do fluxo de calor médio com relação ao da grama? Teste as hipóteses apropriadas usando  $\alpha = 0,05$ .

67. O artigo “Caffeine Knowledge, Attitudes, and Consumption in Adult Women” (*J. of Nutrition Educ.*, 1992, p. 179-184) relata os seguintes dados resumidos sobre o consumo diário de cafeína de uma amostra de mulheres adultas:  $n = 47$ ,  $\bar{x} = 215$  mg,  $s = 235$  mg, e a variação = 5–1176.

- a. Parece plausível que a distribuição da população do consumo diário de cafeína seja normal? É necessário assumir uma distribuição de população normal para testar as hipóteses com relação ao valor do consumo médio da população? Explique seu raciocínio.
- b. Suponha que se tenha acreditado anteriormente que o consumo médio era de no máximo 200 mg. Os dados fornecidos contradizem essa opinião anterior? Teste as hipóteses apropriadas no nível de significância 0,10 e inclua o valor  $P$  em sua análise.

68. A saída a seguir resultou quando o MINITAB foi usado para testar as hipóteses apropriadas em relação ao tempo de ativação médio real, com base nos dados do Exercício 55. Use essas informações para chegar a uma conclusão no nível de significância 0,05 e no nível 0,01.

TESTE de MU = 25,000 VS MU G.T. 25,000

	N	MÉDIA	DESV PAD	MÉDIA SE	T	VALOR P
tempo	13	27,923	5,619	1,559	1,88	0,043

69. Supõe-se que a tensão de quebra média real de certo tipo de isoladores cerâmicos seja de no mínimo 10 psi. Eles serão usados para uma aplicação específica, a menos que os dados da amostra indiquem conclusivamente que essa especificação não foi satisfeita. Um teste de hipóteses que utiliza  $\alpha = 0,01$  será feito com base em uma amostra aleatória de dez isoladores. Assuma que a distribuição da tensão de quebra seja normal com desvio padrão desconhecido.

- a. Se o desvio padrão real é 0,80, qual é a probabilidade de os isoladores serem considerados satisfatórios quando a tensão de quebra média real for realmente de apenas 9,5? Somente 9,0?
- b. Que tamanho de amostra seria necessário para ter uma chance de 75% de detectar que a tensão de quebra média real seja 9,5 quando o desvio padrão real for 0,80?

70. As observações a seguir sobre o tempo de queima residual (segundos) de tiras de pijamas tratados de crianças foram dadas no artigo “An Introduction to Some Precision and Accuracy of Measurement Problems” (*J. of Testing and Eval.*, 1982, p. 132-140). Suponha que um tempo de queima médio real de no máximo 9,75 tenha sido determinado. Os dados sugerem que essa condição não foi satisfeita? Faça um teste apropriado depois de investigar a plausibilidade das suposições que servem de base a seu método de inferência.

9,85 9,93 9,75 9,77 9,67 9,87 9,67

9,94 9,85 9,75 9,83 9,92 9,74 9,99

9,88 9,95 9,95 9,93 9,92 9,89

71. Acredita-se que a incidência de um tipo de defeito do cromossomo na população de homens adultos nos Estados Unidos seja 1 a cada 75. Uma amostra aleatória de 800 pessoas das instituições penais nos Estados Unidos revela que 16 têm tais defeitos. Pode-se concluir que a taxa de incidência desse defeito entre os presos é diferente da taxa presumida para toda a população de homens adultos?

- a. Determine e teste as hipóteses relevantes usando  $\alpha = 0,05$ . Que tipo de erro você pode ter feito ao chegar a uma conclusão?
- b. Que valor  $P$  está associado a esse teste? Com base nesse valor  $P$ ,  $H_0$  pode ser rejeitada no nível de significância 0,20?

72. Em uma investigação da toxina produzida por certa cobra venenosa, um pesquisador preparou 26 frascos diferentes, cada um com 1 g da toxina, e então determinou a quantidade de antitoxina necessária para neutralizar a toxina. A quantidade de antitoxina média amostral necessária foi de 1,89 mg, e o desvio padrão da amostra, 0,42. A pesquisa anterior indicou que a média real da quantidade de neutralizante era de 1,75 mg/g

de toxina. Os novos dados contradizem o valor sugerido pela pesquisa anterior? Teste as hipóteses relevantes usando a abordagem do valor  $P$ . A validade de sua análise depende de qualquer suposição sobre a distribuição da população da quantidade neutralizante? Explique.

73. A resistência à compressão ilimitada média amostral de 45 espécimes de um tipo específico de tijolo foi calculada como sendo 3107 psi, e o desvio padrão da amostra, de 188. A distribuição da resistência à compressão ilimitada pode ser um tanto variada. Os dados indicam fortemente que a resistência à compressão ilimitada média real é menor que o valor-modelo de 3200? Teste, usando  $\alpha = 0,001$ .

74. Com a finalidade de testar a habilidade dos mecânicos de automóveis para identificar problemas simples de motor, um automóvel com um único problema foi levado a 72 oficinas mecânicas diferentes. Somente 42 das 72 mecânicas que examinaram o carro identificaram corretamente o problema. Isso indica fortemente que a proporção real de mecânicas que podem identificar esse problema é menor que 0,75? Calcule o valor  $P$  e tire uma conclusão.

75. Quando  $X_1, X_2, \dots, X_n$  são variáveis independentes de Poisson, cada uma com parâmetro  $\lambda$ , e  $n$  é grande, a média amostral  $\bar{X}$  possui aproximadamente distribuição normal com  $\mu = E(\bar{X}) = \lambda$  e  $\sigma^2 = V(\bar{X}) = \lambda/n$ . Isso implica que

$$Z = \frac{\bar{X} - \lambda}{\sqrt{\lambda/n}}$$

possui aproximadamente distribuição normal padronizada. Para testar  $H_0: \lambda = \lambda_0$ , podemos substituir  $\lambda$  por  $\lambda_0$  na equação de  $Z$  para obter uma estatística de teste. Prefere-se realmente essa estatística à estatística de amostra grande com denominador  $S/\sqrt{n}$  (quando os  $X_i$ s são Poisson), pois é feita explicitamente à suposição de Poisson. Se o número de pedidos de consulta recebido por um estatístico durante uma semana de trabalho de 5 dias tiver distribuição de Poisson, e o número total de pedidos de consulta durante um período de 36 semanas for 160, sugere que o número médio real de pedidos semanais excede 4,0? Teste, usando  $\alpha = 0,02$ .

76. Um fabricante de banheira de água quente anuncia que, com seu equipamento de aquecimento, uma temperatura de 100°F pode ser obtida em no máximo 15 min. Uma amostra aleatória de 32 banheiras é selecionada, e o tempo necessário para se chegar a uma temperatura 100°F é determinado para cada banheira. O tempo médio e o desvio padrão amostrais são 17,5 min. e 2,2 min., respectivamente. Esses dados deixam alguma dúvida na alegação da empresa? Calcule o valor  $P$  e use-o para chegar a uma conclusão no nível 0,05 (assuma que a distribuição do tempo de aquecimento seja aproximadamente normal).

77. O Capítulo 7 apresentou um IC para a variância  $\sigma^2$  de uma distribuição de população normal. O resultado-

chave foi que a va  $\chi^2 = (n-1)S^2/\sigma^2$  possui distribuição qui-quadrado com gl  $n-1$ . Considere a hipótese nula  $H_0: \sigma^2 = \sigma_0^2$  (de forma equivalente,  $\sigma = \sigma_0$ ). Então, quando  $H_0$  é verdadeira, a estatística de teste  $\chi^2 = (n-1)S^2/\sigma_0^2$  possui distribuição qui-quadrado com gl  $n-1$ . Se a alternativa relevante for  $H_a: \sigma^2 > \sigma_0^2$ , rejeitar  $H_0$  se  $(n-1)S^2/\sigma_0^2 \geq \chi_{\alpha, n-1}^2$  fornece um teste com nível de significância  $\alpha$ . Para garantir características razoavelmente uniformes para uma aplicação específica, deseja-se que o desvio padrão real do ponto de amolecimento de um tipo de piche de petróleo seja de no máximo 0,50°C. Os pontos de amolecimento de dez espécimes diferentes foram determinados, produzindo um desvio padrão da amostra de 0,58°C. Isso contradiz fortemente a especificação de uniformidade? Teste as hipóteses apropriadas, usando  $\alpha = 0,01$ .

78. Referindo-se ao Exercício 77, suponha que um pesquisador deseja testar  $H_0: \sigma^2 = 0,04$  versus  $H_a: \sigma^2 < 0,04$  com base em uma amostra de 21 observações. O valor calculado de  $20s^2/0,04$  é 8,58. Coloque limites no valor  $P$  e tire uma conclusão no nível 0,01.

79. Quando a distribuição da população é normal e  $n$  é grande, o desvio padrão da amostra  $S$  possui aproximadamente distribuição normal com  $E(S) \approx \sigma$  e  $V(S) \approx \sigma^2/(2n)$ . Já sabemos que, nesse caso, para qualquer  $n$ ,  $\bar{X}$  é normal com  $E(\bar{X}) = \mu$  e  $V(\bar{X}) = \sigma^2/n$ .

a. Assumindo que a distribuição subjacente seja normal, o que é um estimador aproximadamente não-tendencioso do 99º percentil  $\theta = \mu + 2,33\sigma$ ?

b. Quando os  $X_i$ s são normais, pode-se mostrar que  $\bar{X}$  e  $S$  são v.a.s independentes (uma mede a posição e a outra, a dispersão). Use isso para calcular  $V(\hat{\theta})$  e  $\sigma_{\hat{\theta}}$  para o estimador  $\hat{\theta}$  da parte (a). Qual é o erro-padrão estimado  $\hat{\sigma}_{\hat{\theta}}$ ?

c. Escreva uma estatística de teste para testar  $H_0: \theta = \theta_0$  que possui aproximadamente distribuição normal padronizada quando  $H_0$  é verdadeira. Se o pH do solo é normalmente distribuído em certa região e 64 amostras de solo produzem  $\bar{x} = 6,33$ ,  $s = 0,16$ , isso provoca forte evidência para concluir que no máximo 99% de todas as amostras possíveis teriam um pH menor que 6,75? Teste, usando  $\alpha = 0,01$ .

80. Seja  $X_1, X_2, \dots, X_n$  a amostra aleatória de uma distribuição exponencial com parâmetro  $\lambda$ . Então, pode-se mostrar que  $2\lambda\sum X_i$  possui distribuição qui-quadrado com  $\nu = 2n$  (mostrando primeiro que  $2\lambda X_i$  possui uma distribuição qui-quadrado com  $\nu = 2$ ).

a. Use esse fato para obter uma estatística de teste e a região de rejeição que, juntas, especifiquem um teste nível  $\alpha$  para  $H_0: \mu = \mu_0$ , versus cada uma das três alternativas comumente encontradas. [Sugestão:  $E(X_i) = \mu = 1/\lambda$ , assim,  $\mu = \mu_0$ , é equivalente a  $\lambda = 1/\mu_0$ .]

b. Suponha que dez componentes idênticos, cada um com tempo exponencialmente distribuído até falhar, sejam testados. Os tempos de falha resultantes são:

95 16 11 3 42 71 225 64 87 123

Use o procedimento de teste da parte (a) para decidir se os dados sugerem fortemente que a vida útil média real é menor que o valor de 75 alegado anteriormente.

- 81.** Suponha que a distribuição da população seja normal com  $\sigma$  conhecido. Seja  $\gamma$  tal que  $0 < \gamma < \alpha$ . Para testar  $H_0: \mu = \mu_0$  versus  $H_a: \mu \neq \mu_0$ , considere o teste que rejeita  $H_0$  se tanto  $z \geq z_\gamma$  como  $z \leq -z_{\alpha-\gamma}$ , onde a estatística de teste é  $Z = (\bar{X} - \mu_0)/(\sigma/\sqrt{n})$ .
- Mostre que  $P(\text{erro tipo I}) = \alpha$ .
  - Encontre uma expressão para  $\beta(\mu')$ . (Sugestão: expresse o teste na forma “rejeitar  $H_0$  se  $\bar{x} \geq c_1$  ou  $\leq c_2$ .”)
  - Seja  $\Delta > 0$ . Para quais valores de  $\gamma$  (relativos a  $\alpha$ ) será  $\beta(\mu_0 + \Delta) < \beta(\mu_0 - \Delta)$ ?
- 82.** Após um período de aprendizado, uma organização aplica um exame no qual o indivíduo deve ser aprovado a fim de ser elegível para se tornar sócio. Seja  $p = P(\text{aprovação de principiantes escolhidos aleatoriamente})$ . A organização quer um exame que a maioria, mas não todos, consiga passar, de modo que decide que  $p = 0,90$  é desejável. Para um exame específico, as hipóteses relevantes são  $H_0: p = 0,90$  versus a alternativa  $H_a: p \neq 0,90$ . Suponha que dez pessoas façam o exame, e seja  $X =$  o número de aprovados.
- A região de cauda inferior  $\{0, 1, \dots, 5\}$  especifica um teste nível 0,01?
  - Mostre que, apesar de  $H_a$  ser bicaudal, nenhum teste bicaudal é teste nível 0,01.
  - Desenhe um gráfico de  $\beta(p')$  em função de  $p'$  para o teste. Isso é desejável?

---

## Bibliografia

Veja as bibliografias do final dos capítulos 6 e 7.



## Inferências Baseadas em Duas Amostras

### Introdução

Os Capítulos 7 e 8 apresentaram Intervalos de Confiança (ICs) e procedimentos de teste de hipóteses para uma única média  $\mu$ , uma única proporção  $p$  e uma única variância  $\sigma^2$ . Aqui, estendemos esses métodos a situações que envolvem as médias, proporções e variâncias de duas distribuições de populações diferentes. Por exemplo: seja  $\mu_1$  a dureza de Rockwell média real de espécimes de aço tratados à base de calor e  $\mu_2$  a dureza média real de espécimes tratados a frio. Nesse caso, um investigador pode usar amostras de observações de dureza de cada tipo de aço como base de cálculo de uma estimativa do intervalo de  $\mu_1 - \mu_2$ , a diferença entre as duas durezas médias reais. Expondo outro exemplo, seja  $p_1$  a proporção real de células de níquel-cádmio produzidas sob condições operacionais atuais que estão com defeito devido a curto-circuitos internos, e por  $p_2$  a proporção real de células com curto-circuitos internos produzidas sob condições operacionais modificadas. Se o fundamento lógico das condições modificadas é reduzir a proporção de células com defeito, um engenheiro de qualidade gostaria de usar as informações da amostra para testar a hipótese nula  $H_0: p_1 - p_2 = 0$  (isto é,  $p_1 = p_2$ ) versus a hipótese alternativa  $H_a: p_1 - p_2 > 0$  (isto é,  $p_1 > p_2$ ).

### 9.1 Testes z e Intervalos de Confiança para Diferença entre as Médias de duas Populações

As inferências discutidas nesta seção dizem respeito a uma diferença  $\mu_1 - \mu_2$  entre as médias de duas distribuições de populações diferentes. Um investigador pode, por exemplo, testar as hipóteses sobre a diferença entre as tensões de quebra médias reais de dois tipos distintos de placa ondulada de fibra. Uma hipótese afirmaria

que  $\mu_1 - \mu_2 = 0$ , isto é, que  $\mu_1 = \mu_2$ . De forma alternativa, pode ser apropriado estimar  $\mu_1 - \mu_2$ , calculando um IC de 95%. Tais inferências baseiam-se em uma amostra de observações de tensão de cada tipo de placa de fibra.

#### Hipóteses Básicas

1.  $X_1, X_2, \dots, X_m$  é uma amostra aleatória de uma população com média  $\mu_1$  e variância  $\sigma_1^2$ .
2.  $Y_1, Y_2, \dots, Y_n$  é uma amostra aleatória de uma população com média  $\mu_2$  e variância  $\sigma_2^2$ .
3. As amostras  $X$  e  $Y$  são independentes entre si.

O estimador natural de  $\mu_1 - \mu_2$  é  $\bar{X} - \bar{Y}$ , a diferença entre as médias amostrais correspondentes. A estatística de teste resulta da padronização desse estimador, de modo que precisamos de expressões para o valor esperado e o desvio padrão de  $\bar{X} - \bar{Y}$ .

#### PROPOSIÇÃO

O valor esperado de  $\bar{X} - \bar{Y}$  é  $\mu_1 - \mu_2$ , de modo que  $\bar{X} - \bar{Y}$  é um estimador não-tendencioso de  $\mu_1 - \mu_2$ . O desvio padrão de  $\bar{X} - \bar{Y}$  é

$$\sigma_{\bar{X} - \bar{Y}} = \sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}$$

**Demonstração** Esses resultados dependem das regras do valor esperado e da variância apresentadas no Capítulo 5. Uma vez que o valor esperado de uma diferença é a diferença de valores esperados,

$$E(\bar{X} - \bar{Y}) = E(\bar{X}) - E(\bar{Y}) = \mu_1 - \mu_2$$

Em virtude das amostras  $X$  e  $Y$  serem independentes,  $\bar{X}$  e  $\bar{Y}$  são quantidades independentes, de modo que a variância da diferença é a soma de  $V(\bar{X})$  e  $V(\bar{Y})$ :

$$V(\bar{X} - \bar{Y}) = V(\bar{X}) + V(\bar{Y}) = \frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}$$

O desvio padrão de  $\bar{X} - \bar{Y}$  é a raiz quadrada dessa expressão. ■

Se pensarmos em  $\mu_1 - \mu_2$  como um parâmetro  $\theta$ , então, seu estimador é  $\hat{\theta} = \bar{X} - \bar{Y}$  com desvio padrão  $\sigma_{\hat{\theta}}$  dado pela proposição. Quando  $\sigma_1^2$  e  $\sigma_2^2$  possuem valores conhecidos, a estatística de teste terá a forma  $(\hat{\theta} - \text{valor nulo})/(\sigma_{\hat{\theta}})$ ; essa forma foi usada em vários problemas de uma amostra no capítulo anterior. Quando  $\sigma_1^2$  e  $\sigma_2^2$  são desconhecidos, as variâncias amostrais devem ser usadas para estimar  $\sigma_{\hat{\theta}}$ .

## Procedimentos de Teste para Populações Normais com Variâncias Conhecidas

Nos capítulos 7 e 8, os primeiros ICs e procedimentos de teste para uma média da população  $\mu$  foram determinados com base na suposição de que a distribuição da população era normal com o valor da variância da população  $\sigma^2$  conhecida pelo investigador. De forma semelhante, assumimos primeiro que *ambas* as distribuições das populações são normais e que os valores de *ambos*  $\sigma_1^2$  e  $\sigma_2^2$  são conhecidos. As situações em que uma ou duas dessas hipóteses podem ser dispensadas serão apresentadas em breve.

Em virtude de as distribuições das populações serem normais,  $\bar{X}$  e  $\bar{Y}$  possuem distribuições normais. Essa afirmativa implica que  $\bar{X} - \bar{Y}$  seja normalmente distribuído, com valor esperado  $\mu_1 - \mu_2$ , e desvio padrão  $\sigma_{\bar{X}-\bar{Y}}$  dados na proposição precedente. Padronizar  $\bar{X} - \bar{Y}$  fornece a variável normal padronizada

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \quad (9.1)$$

Em um problema de teste de hipóteses, a hipótese nula afirmará que  $\mu_1 - \mu_2$  possui um valor especificado. Representando esse valor nulo por  $\Delta_0$ , a hipótese nula será  $H_0: \mu_1 - \mu_2 = \Delta_0$ . Geralmente,  $\Delta_0 = 0$ , caso em que  $H_0$  diz que  $\mu_1 = \mu_2$ . Uma estatística de teste resulta da substituição de  $\mu_1 - \mu_2$  na Expressão (9.1) pelo valor nulo  $\Delta_0$ . Em virtude de a estatística de teste  $Z$  ser obtida pela padronização de  $\bar{X} - \bar{Y}$  sob a suposição de que  $H_0$  é verdadeira, ela possui uma distribuição normal padronizada nesse caso. Considere a hipótese alternativa  $H_a: \mu_1 - \mu_2 > \Delta_0$ . Um valor  $\bar{x} - \bar{y}$  que excede consideravelmente  $\Delta_0$  (o valor esperado de  $\bar{X} - \bar{Y}$  quando  $H_0$  é verdadeira) fornece evidência contra  $H_0$  e a favor de  $H_a$ . Tal valor de  $\bar{x} - \bar{y}$  corresponde a um valor grande e positivo de  $z$ . Dessa forma,  $H_0$  deve ser rejeitada em favor de  $H_a$ , se  $z$  for maior ou igual a um valor crítico escolhido de forma adequada. Em virtude de a estatística de teste  $Z$  ter distribuição normal padronizada quando  $H_0$  é verdadeira, a região de rejeição de cauda superior  $z \geq z_\alpha$  fornece um teste com nível de significância (probabilidade de erro tipo I)  $\alpha$ . As regiões de rejeição de  $H_a: \mu_1 - \mu_2 < \Delta_0$  e  $H_a: \mu_1 - \mu_2 \neq \Delta_0$  que produzem testes com nível de significância desejado  $\alpha$  são de cauda inferior e bicaudal, respectivamente.

Hipótese nula:  $H_0: \mu_1 - \mu_2 = \Delta_0$

Valor da estatística de teste:  $z = \frac{\bar{x} - \bar{y} - \Delta_0}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$

**Hipótese Alternativa**

$H_a: \mu_1 - \mu_2 > \Delta_0$

$H_a: \mu_1 - \mu_2 < \Delta_0$

$H_a: \mu_1 - \mu_2 \neq \Delta_0$

**Região de Rejeição para o Teste Nível  $\alpha$**

$z \geq z_\alpha$  (com cauda superior)

$z \leq -z_\alpha$  (com cauda inferior)

tanto  $z \geq z_{\alpha/2}$  como  $z \leq -z_{\alpha/2}$  (bicaudal)

Em virtude de serem testes  $z$ , um valor  $P$  é calculado como se fosse para os testes  $z$  no Capítulo 8 (por exemplo, valor  $P = 1 - \Phi(z)$  para teste de cauda superior).

### Exemplo 9.1

A análise de uma amostra aleatória que consiste em  $m = 20$  espécimes de aço tratado a frio para determinar a resistência resultou em uma resistência média amostral de  $\bar{x} = 29,8$  ksi. Uma segunda amostra aleatória de  $n = 25$  espécimes de aço galvanizado de dois lados forneceu uma resistência média amostral de  $\bar{y} = 34,7$  ksi. Assumindo que as duas distribuições de resistência de rendimento sejam normais com  $\sigma_1 = 4,0$  e  $\sigma_2 = 5,0$  (sugeridas por um gráfico no artigo "Zinc-Coated Sheet Steel: An Overview," *Automotive Engr.*, Dec. 1984, p. 39-43), os dados indicam que as resistências de rendimento médias reais correspondentes  $\mu_1$  e  $\mu_2$  são diferentes? Vamos fazer um teste no nível de significância  $\alpha = 0,01$ .

1. O parâmetro de interesse é  $\mu_1 - \mu_2$ , a diferença entre as resistências médias reais dos dois tipos de aço.
2. A hipótese nula é  $H_0: \mu_1 - \mu_2 = 0$ .
3. A hipótese alternativa é  $H_a: \mu_1 - \mu_2 \neq 0$ ; se  $H_a$  é verdadeira, então,  $\mu_1$  e  $\mu_2$  são diferentes.

4. Com  $\Delta_0 = 0$ , o valor da estatística de teste é

$$z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$$

5. A desigualdade em  $H_a$  implica que o teste é bicaudal. Para  $\alpha = 0,01$ ,  $\alpha/2 = 0,005$  e  $z_{\alpha/2} = z_{0,005} = 2,58$ .  $H_0$  será rejeitada se  $z \geq 2,58$  ou se  $z \leq -2,58$ .

6. Substituindo  $m = 20$ ,  $\bar{x} = 29,8$ ,  $\sigma_1^2 = 16,0$ ,  $n = 25$ ,  $\bar{y} = 34,7$ , e  $\sigma_2^2 = 25,0$  na fórmula por  $z$  produz

$$z = \frac{29,8 - 34,7}{\sqrt{\frac{16,0}{20} + \frac{25,0}{25}}} = \frac{-4,90}{1,34} = -3,66$$

Isto é, o valor observado de  $\bar{x} - \bar{y}$  é mais do que 3 desvios padrão abaixo do que se esperaria, se  $H_0$  fosse verdadeira.

7. Uma vez que  $-3,66 < -2,58$ ,  $z$  cai na cauda inferior da região de rejeição.  $H_0$  é, portanto, rejeitada no nível 0,01 em favor da conclusão de que  $\mu_1 \neq \mu_2$ . Os dados da amostra sugerem fortemente que a resistência de rendimento média real do aço tratado a frio é diferente da do aço galvanizado. O valor  $P$  para esse teste bicaudal é  $2(1 - \Phi(3,66)) \approx 2(1 - 1) = 0$ , de modo que  $H_0$  deve ser rejeitada em *qualquer* nível de significância razoável. ■

## Usando uma Comparação para Identificar a Causalidade

Os investigadores geralmente se interessam em comparar os efeitos de dois tratamentos diferentes de uma resposta ou a resposta após o tratamento com a resposta sem tratamento (tratamento x controle). Se os indivíduos ou objetos a serem usados na comparação não são designados pelos investigadores às duas condições diferentes, diz-se que o estudo é **observacional**. A dificuldade em tirar conclusões com base em um estudo observacional é que, embora a análise estatística possa indicar uma diferença significativa na resposta entre os dois grupos, a diferença pode ser resultado de alguns fatores subjacentes que não foram controlados em vez de qualquer diferença nos tratamentos.

### Exemplo 9.2

Uma carta no *Journal of the American Medical Association* (19 de maio de 1978) relata que, dos 215 médicos do sexo masculino que se formaram em Harvard e morreram entre novembro de 1974 e outubro de 1977, os 125 em clínica em tempo integral viveram em média 48,9 anos depois da graduação, enquanto os 90 com vínculo acadêmico viveram em média 43,2 anos depois da graduação. Os dados sugerem que o tempo de vida médio depois da graduação dos médicos em clínica em tempo integral excede o tempo de vida médio daqueles que possuem vínculo acadêmico (se for sim, os estudantes de medicina que dizem estar "morrendo para obter um vínculo acadêmico" podem estar mais próximos da verdade do que possam imaginar; em outras palavras, "publicar ou perecer" realmente é "publicar e perecer")?

Seja  $\mu_1$  o número médio real de anos vividos após a graduação de médicos em clínica em tempo integral, e por  $\mu_2$ , a mesma quantidade de médicos com vínculos acadêmicos. Assuma que os 125 e os 90 médicos sejam amostras aleatórias das populações 1 e 2, respectivamente (que pode não ser razoável se houver motivo para acreditar que os graduados em Harvard têm características especiais que os diferenciam de todos os outros médicos – nesse caso, as inferências seriam restritas apenas às "populações de Harvard"). A carta de onde os dados foram retirados não forneceu nenhuma informação sobre as variâncias, de modo que, para ilustração, assumamos que  $\sigma_1 = 14,6$  e  $\sigma_2 = 14,4$ . As hipóteses são  $H_0: \mu_1 - \mu_2 = 0$  versus  $H_a: \mu_1 - \mu_2 > 0$ , assim,  $\Delta_0$  é zero. O valor calculado da estatística de teste é

$$z = \frac{48,9 - 43,2}{\sqrt{\frac{(14,6)^2}{125} + \frac{(14,4)^2}{90}}} = \frac{5,70}{\sqrt{1,70 + 2,30}} = 2,85$$

O valor  $P$  de um teste de cauda superior é  $1 - \Phi(2,85) = 0,0022$ . No nível de significância 0,01,  $H_0$  é rejeitada (pois  $\alpha > \text{valor } P$ ) em favor da conclusão de  $\mu_1 - \mu_2 > 0$  ( $\mu_1 > \mu_2$ ). Isso é consistente com as informações relatadas na carta.

Esses dados resultaram a partir de um estudo observacional **retrospectivo**; o investigador não começou selecionando uma amostra de médicos e designando alguns ao tratamento de "vinculação acadêmica" e os outros ao tratamento de "clínicas em tempo integral", mas, ao contrário, identificou os membros dos dois grupos, procurando por registros passados (por meio de obituários!). Pode o resultado estatisticamente significativo aqui ser realmente atribuído a uma diferença do tipo de prática médica após a graduação ou existe algum outro fator subjacente (por exemplo, idade quando se formou, regimes de exercício etc.) que também possa oferecer uma explicação plausível para a diferença? Os estudos observacionais foram utilizados para discutir uma ligação causal entre fumar e o câncer do pulmão. Existem muitos estudos que mostram que a incidência de câncer do pulmão é significativamente maior entre os fumantes que entre os não-fumantes. Entretanto, os indivíduos decidiram se tornar fumantes muito antes de os investigadores entrarem em cena, e os fatores na tomada dessa decisão podem ter exercido um papel causal na contração de câncer do pulmão. ■

Um **experimento controlado aleatoriamente** surge quando os investigadores submetem as pessoas aos dois tratamentos de forma aleatória. Quando a significância estatística é observada em tal experimento, o investigador e outras partes interessadas terão mais confiança na conclusão de que a diferença na resposta foi provocada por uma diferença nos tratamentos. Um exemplo bastante famoso desse tipo de experimento e conclusão é o de Salk, da vacina contra a poliomielite, descrito na Seção 9.4. Essas questões são discutidas com mais detalhes nos livros (não-matemáticos) de Moore e de Freedman et al., listados nas referências do Capítulo 1.

## $\beta$ e a Escolha do Tamanho da Amostra

A probabilidade de um erro tipo II é facilmente calculada quando as distribuições das populações são normais com valores conhecidos de  $\sigma_1$  e  $\sigma_2$ . Considere o caso em que a hipótese alternativa é  $H_a: \mu_1 - \mu_2 > \Delta_0$ . Seja  $\Delta'$  um valor de  $\mu_1 - \mu_2$  que excede  $\Delta_0$  (um valor para o qual  $H_0$  é falsa). A região de rejeição de cauda superior  $z \geq z_\alpha$  pode ser expressa novamente na forma  $\bar{x} - \bar{y} \geq \Delta_0 + z_\alpha \sigma_{\bar{x}-\bar{y}}$ . Dessa maneira, a probabilidade de um erro tipo II quando  $\mu_1 - \mu_2 = \Delta'$  é

$$\begin{aligned}\beta(\Delta') &= P(\text{não rejeitar } H_0 \text{ quando } \mu_1 - \mu_2 = \Delta') \\ &= P(\bar{X} - \bar{Y} < \Delta_0 + z_\alpha \sigma_{\bar{X}-\bar{Y}} \text{ quando } \mu_1 - \mu_2 = \Delta')\end{aligned}$$

Quando  $\mu_1 - \mu_2 = \Delta'$ ,  $\bar{X} - \bar{Y}$  é normalmente distribuído com valor médio  $\Delta'$  e desvio padrão  $\sigma_{\bar{X}-\bar{Y}}$  (o mesmo desvio padrão quando  $H_0$  é verdadeira), usar esses valores para padronizar a desigualdade em parênteses fornece  $\beta$ .

**Hipótese alternativa**

$\beta(\Delta') = P(\text{erro tipo II quando } \mu_1 - \mu_2 = \Delta')$

$$H_a: \mu_1 - \mu_2 > \Delta_0$$

$$\Phi\left(z_\alpha - \frac{\Delta' - \Delta_0}{\sigma}\right)$$

$$H_a: \mu_1 - \mu_2 < \Delta_0$$

$$1 - \Phi\left(-z_\alpha - \frac{\Delta' - \Delta_0}{\sigma}\right)$$

$$H_a: \mu_1 - \mu_2 \neq \Delta_0$$

$$\Phi\left(z_{\alpha/2} - \frac{\Delta' - \Delta_0}{\sigma}\right) - \Phi\left(-z_{\alpha/2} - \frac{\Delta' - \Delta_0}{\sigma}\right)$$

$$\text{onde } \sigma = \sigma_{\bar{X}-\bar{Y}} = \sqrt{(\sigma_1^2/m) + (\sigma_2^2/n)}$$

**Exemplo 9.3** (continuação do Exemplo 9.1)

Suponha que, quando  $\mu_1$  e  $\mu_2$  (as resistências de rendimento médias reais dos dois tipos de aço) diferem por 5, a probabilidade de detectar tal desvio de  $H_0$  deve ser 0,90. Um teste nível 0,01 com tamanhos das amostras  $m = 20$  e  $n = 25$  satisfaz essa condição? O valor de  $\sigma$  para esses tamanhos das amostras (o denominador de  $z$ ) foi calculado anteriormente como sendo 1,34. A probabilidade de um erro tipo II para o teste nível 0,01 bicaudal quando  $\mu_1 - \mu_2 = \Delta' = 5$  é

$$\begin{aligned}\beta(5) &= \Phi\left(2,58 - \frac{5 - 0}{1,34}\right) - \Phi\left(-2,58 - \frac{5 - 0}{1,34}\right) \\ &= \Phi(-1,15) - \Phi(-6,31) = 0,1251\end{aligned}$$

É fácil verificar que  $\beta(-5) = 0,1251$  (porque a região de rejeição é simétrica). Dessa forma, a probabilidade de detectar tal desvio é  $1 - \beta(5) = 0,8749$ . Em virtude de esse valor ser menor que 0,9, devem ser usadas amostras de tamanhos maiores. ■

Como no Capítulo 8, os tamanhos das amostras  $m$  e  $n$  podem ser determinados de modo que satisfaçam  $P(\text{erro tipo I}) = \text{um } \alpha \text{ especificado}$  e  $P(\text{erro tipo II quando } \mu_1 - \mu_2 = \Delta') = \text{um } \beta \text{ especificado}$ . Para um teste de cauda superior, igualar a expressão anterior para  $\beta(\Delta')$  ao valor especificado de  $\beta$  resulta

$$\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n} = \frac{(\Delta' - \Delta_0)^2}{(z_\alpha + z_\beta)^2}$$

Quando os dois tamanhos das amostras são iguais, essa equação produz

$$m = n = \frac{(\sigma_1^2 + \sigma_2^2)(z_\alpha + z_\beta)^2}{(\Delta' - \Delta_0)^2}$$

Essas expressões também estão corretas para um teste de cauda inferior, enquanto  $\alpha$  é substituído por  $\alpha/2$  para um teste bicaudal.

**Testes de Amostras Grandes**

As hipóteses de distribuições normais das populações e valores conhecidos de  $\sigma_1$  e  $\sigma_2$  são desnecessárias quando ambos os tamanhos das amostras são grandes. Nesse caso, o Teorema do Limite Central garante que  $\bar{X} - \bar{Y}$  possui aproximadamente distribuição normal, independentemente das distribuições das populações subjacentes. Além disso, usar  $S_1^2$  e  $S_2^2$  no lugar de  $\sigma_1^2$  e  $\sigma_2^2$ , na Expressão (9.1), fornece uma variável cuja distribuição é aproximadamente normal padronizada:

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}}$$

Uma estatística de teste de amostra grande resulta da substituição de  $\mu_1 - \mu_2$  por  $\Delta_0$ , o valor esperado de  $\bar{X} - \bar{Y}$  quando  $H_0$  é verdadeira. Essa estatística  $Z$ , então, possui aproximadamente uma distribuição normal padronizada quando  $H_0$  é verdadeira, de modo que os testes nível  $\alpha$  são obtidos usando-se valores críticos  $z$  exatamente como antes.

O uso do valor da estatística de teste

$$z = \frac{\bar{x} - \bar{y} - \Delta_0}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}$$

com as regiões de rejeição de cauda superior, inferior e bicaudal definidas anteriormente, com base em valores críticos  $z$ , fornece testes de amostra grande cujos níveis de significância são aproximadamente  $\alpha$ . Esses testes geralmente são apropriados se  $m > 40$  e  $n > 40$ . Um valor  $P$  é calculado exatamente como se fosse para nossos testes  $z$  anteriores.

### Exemplo 9.4

Ao selecionar um concreto de enxofre para a construção de estradas em regiões que sofrem fortes congelamentos, é importante que o concreto escolhido tenha baixo valor de condutividade térmica para minimizar os danos subsequentes devido às variações de temperatura. Suponha que dois tipos de concreto, um agregado de boa qualidade e um agregado de baixa qualidade, sejam considerados para uma determinada estrada. A Tabela 9.1 resume os dados de um experimento realizado para comparar os dois tipos de concreto. Essas informações sugerem que a condutividade média real do concreto de boa qualidade excede a do concreto de baixa qualidade? Vamos usar um teste com  $\alpha = 0,01$ .

**Tabela 9.1** Dados para o Exemplo 9.4

Tipo	Tamanho	Condutividade Média Amostral	DP Amostral
Bom	42	0,486	0,187
De baixa	42	0,359	0,158

Sejam  $\mu_1$  e  $\mu_2$  a condutividade térmica média real do concreto de agregado de boa qualidade e de baixa qualidade, respectivamente. As duas hipóteses são  $H_0: \mu_1 - \mu_2 = 0$  versus  $H_a: \mu_1 - \mu_2 > 0$ .  $H_0$  será rejeitada se  $z \geq z_{0,01} = 2,33$ . Calculamos

$$z = \frac{0,486 - 0,359}{\sqrt{\frac{(0,187)^2}{42} + \frac{(0,158)^2}{42}}} = \frac{0,127}{0,0378} = 3,36$$

Uma vez que  $3,36 \geq 2,33$ ,  $H_0$  é rejeitada no nível de significância 0,01. De forma alternativa, o valor  $P$  de um teste  $z$  de cauda superior é

$$\text{Valor } P = 1 - \Phi(z) = 1 - \Phi(3,36) = 0,0004$$

$H_0$  deve ser rejeitada não somente para um teste com  $\alpha = 0,01$ , como também para  $\alpha = 0,001$  ou qualquer outro  $\alpha$  que exceda 0,0004. Os dados mantêm fortemente a conclusão de que a condutividade térmica média real do concreto de boa qualidade excede a do concreto de baixa qualidade. ■

### Intervalos de Confiança de $\mu_1 - \mu_2$

Quando as distribuições das populações são normais, padronizar  $\bar{X} - \bar{Y}$  fornece uma variável aleatória  $Z$  com uma distribuição normal padronizada. Uma vez que a área sob a curva  $z$  entre  $-z_{\alpha/2}$  e  $z_{\alpha/2}$  é  $1 - \alpha$ , segue que

$$P\left(-z_{\alpha/2} < \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} < z_{\alpha/2}\right) = 1 - \alpha$$

A manipulação das desigualdades dentro dos parênteses para isolar  $\mu_1 - \mu_2$  produz

$$P\left(\bar{X} - \bar{Y} - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}} < \mu_1 - \mu_2 < \bar{X} - \bar{Y} + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}\right) = 1 - \alpha$$

Isso implica que um IC de  $100(1 - \alpha)\%$  para  $\mu_1 - \mu_2$  possui limite inferior  $\bar{x} - \bar{y} - z_{\alpha/2} \cdot \sigma_{\bar{x}-\bar{y}}$  e limite superior  $\bar{x} - \bar{y} + z_{\alpha/2} \cdot \sigma_{\bar{x}-\bar{y}}$ , onde  $\sigma_{\bar{x}-\bar{y}}$  é a raiz quadrada da expressão. Esse intervalo é um caso especial da fórmula geral  $\hat{\theta} \pm z_{\alpha/2} \cdot \sigma_{\hat{\theta}}$ .

Se  $m$  e  $n$  são grandes, o TLC implica que esse intervalo é válido, mesmo sem a suposição de populações normais; nesse caso, o nível de confiança é aproximadamente  $100(1 - \alpha)\%$ . Além disso, o uso das variâncias amostrais  $S_1^2$  e  $S_2^2$  na variável padronizada  $Z$  produz um intervalo válido no qual  $s_1^2$  e  $s_2^2$  substituem  $\sigma_1^2$  e  $\sigma_2^2$ .

Contanto que  $m$  e  $n$  sejam grandes, um IC para  $\mu_1 - \mu_2$ , com um nível de confiança de aproximadamente  $100(1 - \alpha)\%$  é

$$\bar{x} - \bar{y} \pm z_{\alpha/2} \sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}$$

onde  $-$  fornece o limite inferior e  $+$  o limite superior do intervalo. Um limite de confiança superior ou inferior também pode ser calculado, mantendo-se o sinal apropriado ( $+$  ou  $-$ ) e substituindo-se  $z_{\alpha/2}$  por  $z_\alpha$ .

Nossa regra prática padrão para caracterizar os tamanhos das amostras como grandes é  $m > 40$  e  $n > 40$ .

### Exemplo 9.5

Um experimento realizado para estudar várias características de pinos de ferro resultou em 78 observações sobre a resistência de corte (kip) de pinos de 3/8-pol. de diâmetro e 88 observações sobre a resistência de pinos de 1/2-pol. de diâmetro. As quantidades resumidas do Minitab são apresentadas a seguir, e um *boxplot* comparativo é apresentado na Figura 9.1. Os tamanhos das amostras, as médias amostrais e os desvios padrão amostrais correspondem aos valores dados no artigo "Ultimate Load Capacities of Expansion Anchor Bolts" (*J. Energy Engr.*, 1993, p. 139-158). Os resumos sugerem que a principal diferença entre as duas amostras é o local onde estão centralizadas.

Variável	N	Média	Mediana	MédiaAp	DesvPad	MédiaSE
diâm 3/8	78	4,250	4,230	4,238	1,300	0,147
Variável	Min	Máx	Q1	Q3		
diâm 3/8	1,634	7,327	3,389	5,075		
Variável	N	Média	Mediana	MédiaAP	DesvPad	SEMMean
diâm 1/2	88	7,140	7,113	7,150	1,680	0,179
Variável	Min	Máx	Q1	Q3		
diâm 1/2	2,450	11,343	5,965	8,447		

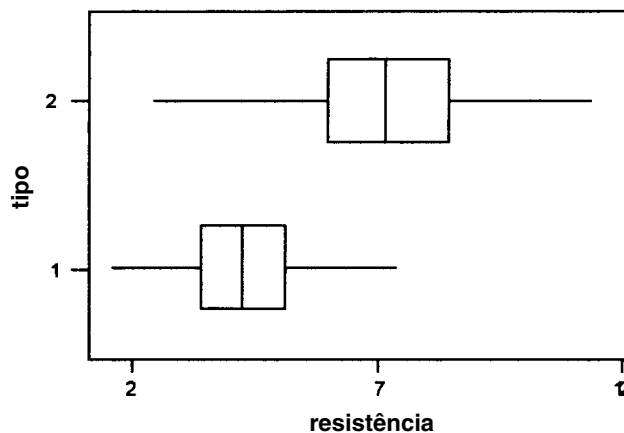


Figura 9.1 Um *boxplot* comparativo dos dados de resistência de corte



Vamos calcular agora um intervalo de confiança para a diferença entre a resistência de corte média real de pinos de 3/8-pol. ( $\mu_1$ ) e a resistência de corte média real de pinos de 1/2-pol. ( $\mu_2$ ) usando um nível de confiança de 95%:

$$\begin{aligned} 4,25 - 7,14 \pm (1,96) \sqrt{\frac{(1,30)^2}{78} + \frac{(1,68)^2}{88}} &= -2,89 \pm (1,96)(0,2318) \\ &= -2,89 \pm 0,45 = (-3,34, -2,44) \end{aligned}$$

Isto é, com confiança de 95%,  $-3,34 < \mu_1 - \mu_2 < -2,44$ . Podemos, portanto, ficar altamente confiantes de que a resistência de corte média real dos pinos de 1/2-pol. excede a dos pinos de 3/8-pol. entre 2,44 kip e 3,34 kip. Observe que, se mudarmos a nomenclatura de modo que  $\mu_1$  se refira aos pinos de 1/2-pol. e  $\mu_2$  aos pinos de 3/8-pol., o intervalo de confiança está agora centralizado em  $+2,89$  e o valor  $0,45$  ainda é subtraído e somado para obter os limites de confiança. O intervalo resultante é  $(2,44, 3,34)$ , e a interpretação é idêntica à do intervalo calculado anteriormente. ■

Se as variâncias  $\sigma_1^2$  e  $\sigma_2^2$  são, pelo menos, conhecidas aproximadamente e o investigador utiliza tamanhos de amostras iguais, então, o tamanho da amostra  $n$  de cada amostra que produz  $100(1 - \alpha)\%$  de um intervalo de largura  $w$  é

$$n = \frac{4z_{\alpha/2}^2(\sigma_1^2 + \sigma_2^2)}{w^2}$$

que geralmente terá de ser arredondado para um número inteiro.

## Exercícios | Seção 9.1 (1–16)

- Um artigo do *Consumer Reports* de novembro de 1983 comparou vários tipos de pilhas. A vida útil média das pilhas alcalinas AA Duracell e das pilhas alcalinas AA Eveready Energizer foi dada como 4,1 horas e 4,5 horas, respectivamente. Suponha que essas sejam a vida útil média da população.
  - Seja  $\bar{X}$  a vida útil média amostral de 100 pilhas Duracell e  $\bar{Y}$  a vida útil da média amostra de 100 pilhas Eveready. Qual é o valor médio de  $\bar{X} - \bar{Y}$  (isto é, onde a distribuição de  $\bar{X} - \bar{Y}$  está centralizada)? Como a sua resposta depende dos tamanhos especificados das amostras?
  - Suponha que os desvios padrão da vida útil da população sejam 1,8 hora para as pilhas Duracell e 2,0 horas para as pilhas Eveready. Com os tamanhos amostrais dados na parte (a), qual é a variância da estatística  $\bar{X} - \bar{Y}$  e qual é o seu desvio padrão?
  - Para os tamanhos amostrais dados na parte (a), faça um desenho da curva de distribuição aproximada de  $\bar{X} - \bar{Y}$  (inclua uma escala de medição no eixo horizontal). O formato da curva seria necessariamente o mesmo para os tamanhos amostrais de 10 pilhas de cada tipo? Explique.
- Sejam  $\mu_1$  e  $\mu_2$  as vidas médias reais de duas marcas concorrentes de pneus radiais de tamanho P205/65R15. Teste  $H_0: \mu_1 - \mu_2 = 0$  versus  $H_a: \mu_1 - \mu_2 \neq 0$  no nível 0,05, usando os seguintes dados:  $m = 45$ ,  $\bar{x} = 42.500$ ,  $s_1 = 2200$ ,  $n = 45$ ,  $\bar{y} = 40.400$  e  $s_2 = 1900$ .
  - Seja  $\mu_1$ , a vida real média de uma marca *premium* de pneu radial P205/65R15 e  $\mu_2$  a vida média real de uma marca econômica do mesmo tamanho. Teste  $H_0: \mu_1 - \mu_2 = 5000$  versus  $H_a: \mu_1 - \mu_2 > 5000$  no nível 0,01, usando os seguintes dados:  $m = 45$ ,  $\bar{x} = 42.500$ ,  $s_1 = 2200$ ,  $n = 45$ ,  $\bar{y} = 36.800$  e  $s_2 = 1500$ .
- Use os dados do Exercício 2 para calcular um IC de 95% para  $\mu_1 - \mu_2$ . O intervalo resultante sugere que  $\mu_1 - \mu_2$  foi estimado com precisão?
  - Use os dados do Exercício 3 para calcular o limite de confiança superior de 95% para  $\mu_1 - \mu_2$ .
- As pessoas que têm a síndrome de Reynaud estão propensas a sofrer uma diminuição repentina da circulação sanguínea nos dedos dos pés e das mãos. Em um experimento para estudar a extensão dessa deficiência, cada pessoa mergulhou o dedo indicador na água, e a emissão de calor resultante ( $\text{cal/cm}^2/\text{min}$ ) foi medida. Para  $m = 10$  pessoas com a síndrome, a saída de calor média foi  $\bar{x} = 0,64$ , e para  $n = 10$  que não sofrem da síndrome, a emissão média foi 2,05. Sejam  $\mu_1$  e  $\mu_2$  as emissões de calor médias reais para os dois tipos de pessoas. Assuma que as duas distribuições de emissão de calor sejam normais com  $\sigma_1 = 0,2$  e  $\sigma_2 = 0,4$ .
  - Considere o teste de  $H_0: \mu_1 - \mu_2 = -1,0$  versus  $H_a: \mu_1 - \mu_2 < -1,0$  no nível 0,01. Descreva em palavras o que diz  $H_a$  e faça o teste.
  - Calcule o valor  $P$  para o valor de  $Z$  obtido na parte (a).
  - Qual é a probabilidade de um erro tipo II quando a diferença real entre  $\mu_1$  e  $\mu_2$  for  $\mu_1 - \mu_2 = -1,2$ ?

- d. Assumindo que  $m = n$ , quais tamanhos das amostras são necessários para garantir que  $\beta = 0,1$  quando  $\mu_1 - \mu_2 = -1,2$ ?
6. Um experimento para comparar a resistência de contato à tensão da argamassa modificada de látex de polímero (argamassa de cimento Portland à qual foram acrescentadas emulsões de látex de polímero durante a mistura) à resistência da argamassa inalterada resultou em  $\bar{x} = 18,12$  kgf/cm<sup>2</sup> para a argamassa modificada ( $m = 40$ ) e  $\bar{y} = 16,87$  kgf/cm<sup>2</sup> para a argamassa inalterada ( $n = 32$ ). Sejam  $\mu_1$  e  $\mu_2$  as resistências de contato à tensão das argamassas modificadas e inalteradas, respectivamente. Assuma que as distribuições da resistência de contato sejam normais.
- a. Assumindo que  $\sigma_1 = 1,6$  e  $\sigma_2 = 1,4$ , teste  $H_0: \mu_1 - \mu_2 = 0$  versus  $H_a: \mu_1 - \mu_2 > 0$  no nível 0,01.
- b. Calcule a probabilidade de um erro tipo II para o teste da parte (a) quando  $\mu_1 - \mu_2 = 1$ .
- c. Suponha que o investigador decidiu usar um teste nível 0,05 e quis que  $\beta = 0,10$  quando  $\mu_1 - \mu_2 = 1$ . Se  $m = 40$ , que valor de  $n$  é necessário?
- d. Como a análise e a conclusão da parte (a) mudariam, se  $\sigma_1$  e  $\sigma_2$  fossem desconhecidas, mas  $s_1 = 1,6$  e  $s_2 = 1,4$ ?
7. Os alunos universitários do sexo masculino se cansam com mais facilidade que suas colegas? Essa pergunta foi avaliada no artigo "Boredom in Young Adults – Gender and Cultural Comparisons" (*J. of CrossCultural Psych.*, 1991, p. 209-223). Os autores administraram uma escala chamada Escala de Propensão ao Cansaço a 97 homens e 148 mulheres de uma universidade dos Estados Unidos. Os dados a seguir apóiam a hipótese da pesquisa de que o Grau de Propensão ao Cansaço médio é maior para os homens que para as mulheres? Teste as hipóteses apropriadas, usando um nível de significância de 0,05.

Sexo	Tamanho Amostral	Média Amostral	DP Amostral
Homens	97	10,40	4,83
Mulheres	148	9,26	4,68

8. Testes de resistência à tensão foram realizados em dois graus diferentes de haste de arame ("Fluidized Bed Patenting of Wire Rods." *Wire J.*, June 1977, p. 56-61), resultando nos dados a seguir.

Graus	Tamanho Amostral	Média Amostral (kg/mm <sup>2</sup> )	DP Amostral
SI 1064	$m = 129$	$\bar{x} = 107,6$	$s_1 = 1,3$
SI 1078	$n = 129$	$\bar{y} = 123,6$	$s_2 = 2,0$

- a. Os dados fornecem evidências convincentes para concluir que a resistência média real do grau 1078 excede a do grau 1064 por mais de 10 kg/mm<sup>2</sup>? Teste

as hipóteses apropriadas, usando a abordagem do valor  $P$ .

- b. Calcule a diferença entre as resistências médias reais dos dois graus, de forma que forneça informações sobre precisão e confiabilidade.
9. O artigo "Evaluation of a Ventilation Strategy to Prevent Barotrauma in Patients at High Risk for Acute Respiratory Distress Syndrome" (*New Engl. J. of Med.*, 1998, p. 355-358) relatou um experimento no qual 120 pacientes com características clínicas semelhantes foram divididos aleatoriamente em um grupo de controle e em um grupo de tratamento, cada um com 60 pacientes. A média de permanência na UTI (dias) da amostra e o desvio padrão da amostra para o grupo de tratamento foram, respectivamente, 19,9 e 39,1, enquanto os valores para o grupo de controle foram 13,7 e 15,8.
- a. Calcule uma estimativa pontual da diferença entre a permanência média real na UTI dos grupos de tratamento e de controle. Essa estimativa sugere que existe uma diferença significativa entre as permanências médias reais sob as duas condições?
- b. Responda à pergunta proposta na parte (a) fazendo um teste de hipótese formal. O resultado é diferente do que você supôs na parte (a)?
- c. Parece que a permanência na UTI dos pacientes que receberam o tratamento de ventilação é normalmente distribuída? Explique o seu raciocínio.
- d. Estime a duração da permanência média real dos pacientes que receberam o tratamento de ventilação, de modo que transmita informações sobre precisão e confiabilidade.
10. Foi realizado um experimento para comparar a dureza de fratura do 18 Ni de alta pureza com o aço de pureza comercial do mesmo tipo (*Corrosion Science*, 1971, p. 723-736). Para  $m = 32$  espécimes, a dureza média amostral foi  $\bar{x} = 65,6$  para o aço de alta pureza, enquanto para  $n = 38$  espécimes de aço comercial,  $\bar{y} = 59,8$ . Porque o aço de alta pureza é mais caro, seu uso para determinada aplicação pode ser justificado somente se sua dureza de fratura exceder a do aço de pureza comercial por mais de 5. Suponha que ambas as distribuições de dureza sejam normais.
- a. Assumindo que  $\sigma_1 = 1,2$  e  $\sigma_2 = 1,1$ , teste as hipóteses relevantes, usando  $\alpha = 0,001$ .
- b. Calcule  $\beta$  para o teste realizado na parte (a) quando  $\mu_1 - \mu_2 = 6$ .
11. O nível de chumbo no sangue foi determinado para uma amostra de 152 trabalhadores que lidam com resíduos perigosos com idade entre 20-30 anos e também para uma amostra de 86 trabalhadoras, resultando em uma média  $\pm$  erro-padrão de  $5,5 \pm 0,3$  para os homens e  $3,8 \pm 0,2$  para as mulheres ("Temporal Changes in Blood Lead Levels of Hazardous Waste Workers in New Jersey, 1984-1987," *Environ. Monitoring and Assessment*, 1993 p. 99-107). Calcule uma estimativa da diferença entre os níveis médios reais de chumbo no sangue dos operários e das operárias, de modo que forneça informações sobre precisão e confiabilidade.

12. A tabela a seguir fornece dados resumidos sobre a resistência à compressão ao cubo ( $N/mm^2$ ) de espécimes de concreto feitos com uma mistura pulverizada de combustível pulverizada ("A Study of Twenty-Five-Year-Old Pulverized Fuel Ash Concrete Used in Foundation Structures," *Proc. Inst. Civ. Engrs.*, Mar. 1985, p. 149-165):

Idade (dias)	Tamanho Amostral	Média Amostral	DP Amostral
7	68	26,99	4,89
28	74	35,76	6,43

Calcule e interprete um IC de 99% da diferença entre a resistência média real de 7 dias e a resistência média real de 28 dias.

13. Um engenheiro mecânico quer comparar as propriedades de resistência de vigas de aço com as de vigas semelhantes feitas com uma liga específica. O mesmo número de vigas,  $n$ , de cada tipo será testado. Cada viga será colocada horizontalmente com um apoio em cada extremidade, uma força de 2500 lb será aplicada no centro e a deflexão será medida. A partir de experiência anterior com tais vigas, o engenheiro quer assumir que o desvio padrão real de desvio para ambos os tipos de viga é 0,05 in. Em virtude de a liga ser mais cara, o engenheiro quer testar no nível 0,01 se ela tem deflexão média menor que a viga de aço. Que valor de  $n$  é apropriado, se a probabilidade de erro tipo II desejada for 0,05 quando a diferença da deflexão média real favorecer a liga em 0,04 in.?
14. O nível da atividade de oxidase de monoamina (MAO) nas plaquetas sanguíneas (nm/mg proteína/h) foi determinado para cada indivíduo em uma amostra de 43 esquizofrênicos crônicos, resultando em  $\bar{x} = 2,69$  e  $s_1 = 2,30$ , bem como de 45 pessoas normais, resultando em

$\bar{y} = 6,35$  e  $s_2 = 4,03$ . Esses dados sugerem fortemente que a média real de atividade de MAO das pessoas normais é mais que duas vezes o nível da atividade dos esquizofrênicos? Derive um procedimento de teste e faça o teste usando  $\alpha = 0,01$ . [Sugestão:  $H_0$  e  $H_a$  aqui possuem uma forma diferente dos três casos-padrão. Sejam  $\mu_1$  e  $\mu_2$  a atividade de MAO média real dos esquizofrênicos bem como das pessoas normais, respectivamente, e considere o parâmetro  $\theta = 2\mu_1 - \mu_2$ . Escreva  $H_0$  e  $H_a$  em termos de  $\theta$ , calcule  $\theta$  e deduza  $\hat{\sigma}_\theta$  ("Reduced Monoamine Oxidase Activity in Blood Platelets from Schizophrenic Patients," *Nature*, July 28, 1972, p. 225-226).]

15. a. Mostre para o teste de cauda superior com  $\sigma_1$  e  $\sigma_2$  conhecidos que, à medida que  $m$  ou  $n$  aumenta,  $\beta$  diminui quando  $\mu_1 - \mu_2 > \Delta_0$ .  
b. Para o caso de tamanhos das amostras iguais ( $m = n$ ) e  $\alpha$  fixo, o que acontece com o tamanho da amostra necessário  $n$  quando  $\beta$  é reduzido, onde  $\beta$  é a probabilidade de erro tipo II desejada em uma alternativa fixa?
16. Para decidir se dois tipos diferentes de aço possuem os mesmos valores de dureza de fratura média real,  $n$  espécimes de cada tipo são testados, produzindo os seguintes resultados:

Tipo	Média amostral	DP amostral
1	60,1	1,0
2	59,9	1,0

Calcule o valor  $P$  para o teste apropriado de duas amostras, assumindo que os dados se basearam em  $n = 100$ . Então, repita o cálculo para  $n = 400$ . O valor  $P$  pequeno para  $n = 400$  é indicativo de uma diferença que possui significância prática? Você ficaria satisfeito apenas com um relatório do valor  $P$ ? Comente em poucas palavras.

## 9.2 Teste $t$ e Intervalo de Confiança de Duas Amostras

Na prática, sempre é virtual o caso de os valores das variâncias da população serem desconhecidos. Na seção anterior, ilustramos, para tamanhos das amostras grandes, o uso de um procedimento de teste e IC em que as variâncias amostrais foram utilizadas no lugar das variâncias da população. Na verdade, para amostras grandes, o TLC nos permite usar esses métodos mesmo quando as duas populações de interesse não são normais.

Existem muitos problemas, entretanto, em que pelo menos um tamanho de amostra é pequeno e as variâncias da população possuem valores desconhecidos. Na ausência do TLC, continuamos fazendo hipóteses específicas sobre as distribuições das populações subjacentes. O uso de procedimentos inferenciais que resultam dessas hipóteses é, então, restrito às situações em que as hipóteses são, no mínimo, aproximadamente satisfeitas.

**HIPÓTESES**

Ambas as populações são normais, de modo que  $X_1, X_2, \dots, X_m$ , é uma amostra aleatória de uma distribuição normal, assim como  $Y_1, \dots, Y_n$  (com os  $X$ s e  $Y$ s independentes entre si). A plausibilidade dessas hipóteses pode ser considerada, construindo-se um gráfico de probabilidade normal dos  $x_i$ s e outro dos  $y_i$ s.

A estatística de teste e a fórmula do intervalo de confiança baseiam-se na mesma variável padronizada desenvolvida na Seção 9.1, mas a distribuição relevante é agora  $t$  em vez de  $z$ .

**TEOREMA**

Quando as distribuições das populações são normais, a variável padronizada

$$T = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}} \tag{9.2}$$

possui aproximadamente uma distribuição  $t$  com gl  $\nu$  calculado a partir dos dados por

$$\nu = \frac{\left(\frac{s_1^2}{m} + \frac{s_2^2}{n}\right)^2}{\frac{(s_1^2/m)^2}{m-1} + \frac{(s_2^2/n)^2}{n-1}} = \frac{[(se_1)^2 + (se_2)^2]^2}{\frac{(se_1)^4}{m-1} + \frac{(se_2)^4}{n-1}}$$

onde

$$se_1 = \frac{s_1}{\sqrt{m}}, se_2 = \frac{s_2}{\sqrt{n}}$$

(arredonde  $\nu$  para o número inteiro mais próximo).

Manipular  $T$  em uma fórmula de probabilidades para isolar  $\mu_1 - \mu_2$  fornece um IC, enquanto uma estatística de teste resulta da substituição de  $\mu_1 - \mu_2$  pelo valor nulo  $\Delta_0$ .

**Procedimentos  $t$  de Duas Amostras**

O **intervalo de confiança  $t$  de duas amostras para  $\mu_1 - \mu_2$**  com nível de confiança  $100(1 - \alpha) \%$  é então

$$\bar{x} - \bar{y} \pm t_{\alpha/2, \nu} \sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}$$

Um limite de confiança monocaudal pode ser calculado conforme descrito anteriormente.

O **teste  $t$  de duas amostras** para testar  $H_0: \mu_1 - \mu_2 = \Delta_0$  é como segue:

$$\text{Valor da estatística de teste: } t = \frac{\bar{x} - \bar{y} - \Delta_0}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}$$

**Hipótese alternativa**

$H_a: \mu_1 - \mu_2 > \Delta_0$

$H_a: \mu_1 - \mu_2 < \Delta_0$

$H_a: \mu_1 - \mu_2 \neq \Delta_0$

**Região de Rejeição para o Teste Nível  $\alpha$  Aproximado**

$t \geq t_{\alpha, \nu}$  (com cauda superior)

$t \leq -t_{\alpha, \nu}$  (com cauda inferior)

$H_a: \mu_1 - \mu_2 \neq \Delta_0$  tanto  $t \geq t_{\alpha/2, \nu}$  como  $t \leq -t_{\alpha/2, \nu}$  (bicaudal)

Um valor  $P$  pode ser calculado conforme descrito na Seção 8.4 para o teste  $t$  de uma amostra.

## Exemplo 9.6

O volume nulo em um tecido afeta as propriedades de conforto, flamabilidade e isolamento. A permeabilidade de um tecido refere-se à acessibilidade do espaço vazio ao fluxo de um gás ou líquido. O artigo "The Relationship Between Porosity and Air Permeability of Woven Textile Fabrics" (*J. of Testing and Eval.*, 1997, p. 108-114) forneceu informações resumidas sobre a permeabilidade de ar ( $\text{cm}^3/\text{cm}^2/\text{s}$ ) de vários tipos diferentes de tecido. Considere os dados a seguir sobre dois tipos diferentes de tecido liso:

Tecido Amostral	Tamanho Amostral	Média Amostral	Desvio padrão Amostral
Algodão	10	51,71	0,79
Triacetato	10	136,14	3,59

Assumindo que as distribuições de porosidade de ambos os tipos de tecido sejam normais, vamos calcular um intervalo de confiança para a diferença entre a porosidade média real do tecido de algodão e a do tecido de acetato, usando um nível de confiança de 95%. Antes de o valor crítico  $t$  apropriado poder ser selecionado, o  $gl$  deve ser determinado:

$$df = \frac{\left(\frac{0,6241}{10} + \frac{12,8881}{10}\right)^2}{\frac{(0,6241/10)^2}{9} + \frac{(12,8881/10)^2}{9}} = \frac{1,8258}{0,1850} = 9,87$$

Dessa forma, usamos  $\nu = 9$ ; a Tabela A.5 do Apêndice fornece  $t_{0,025,9} = 2,262$ . O intervalo resultante é

$$51,71 - 136,14 \pm (2,262) \sqrt{\frac{0,6241}{10} + \frac{12,8881}{10}} = -84,43 \pm 2,63 \\ = (-87,06, -81,80)$$

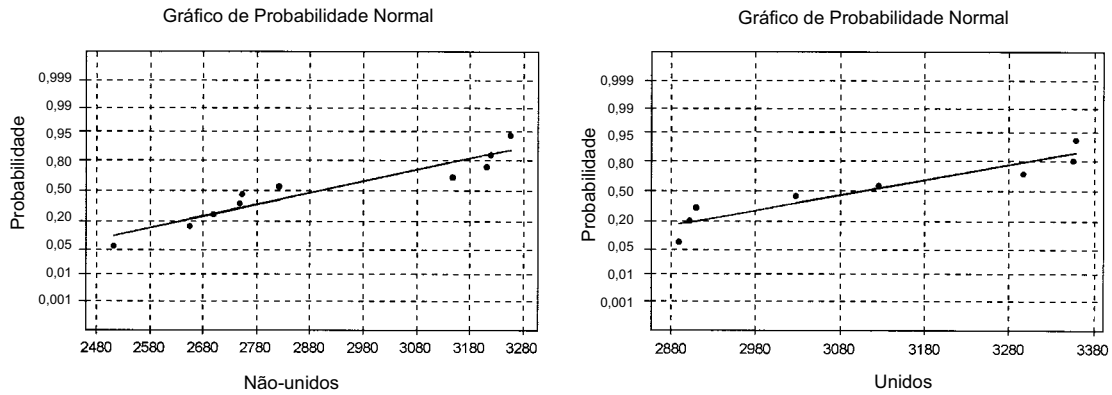
Com alto grau de confiança, podemos dizer que a porosidade média real dos espécimes de tecido de triacetato excede a dos espécimes de algodão entre 81,80 e 87,06  $\text{cm}^3/\text{cm}^2/\text{s}$ . ■

## Exemplo 9.7

A deterioração de muitas redes de oleodutos municipais pelo país é uma preocupação crescente. Uma tecnologia proposta para a reabilitação dos oleodutos utiliza um revestimento flexível que passa pelo tubo existente. O artigo "Effect of Welding on a High-Density Polyethylene Liner" (*J. of Materials in Civil Engr.*, 1996, p. 94-100) relatou os dados a seguir sobre a resistência à tensão (psi) dos espécimes de revestimento quando certo processo de fusão foi utilizado e quando esse processo não foi usado.

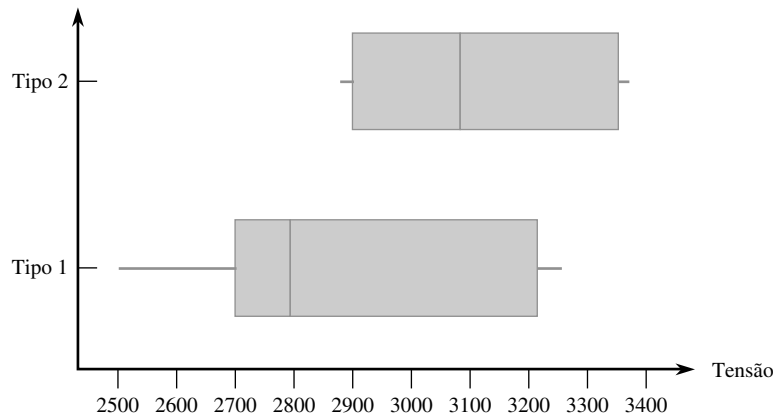
<i>Não-fundido</i>	2748	2700	2655	2822	2511			
	3149	3257	3213	3220	2753			
	$m = 10$	$\bar{x} = 2902,8$		$s_1 = 277,3$				
<i>Fundido</i>	3027	3356	3359	3297	3125	2910	2889	2902
	$n = 8$	$\bar{y} = 3108,1$		$s_2 = 205,9$				

A Figura 9.2 mostra gráficos de probabilidade normal obtidos do MINITAB. O padrão linear em cada gráfico sustenta a suposição de que as distribuições de resistência à tensão sob as duas condições são normais.



**Figura 9.2** Gráficos de probabilidade normal do MINITAB para os dados de resistência à tensão

Os autores do artigo afirmaram que o processo de fusão aumentou a resistência à tensão média. A mensagem do *boxplot* comparativo da Figura 9.3 não é muito clara. Vamos fazer um teste de hipóteses para ver se os dados sustentam essa conclusão.



**Figura 9.3** Um *boxplot* comparativo dos dados de resistência à tensão

1. Sejam  $\mu_1$  a resistência à tensão média real de espécimes, quando o tratamento sem união é usado, e por  $\mu_2$  a resistência à tensão média real quando o tratamento de união é usado.
2.  $H_0: \mu_1 - \mu_2 = 0$  (nenhuma diferença nas resistências à tensão médias reais dos dois tratamentos)
3.  $H_a: \mu_1 - \mu_2 < 0$  (a resistência à tensão média real do tratamento não-fundido é menor que a do tratamento de união, de modo que a conclusão dos investigadores está correta)
4. O valor nulo é  $\Delta_0 = 0$ , assim, a estatística de teste é

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}$$

5. Calculamos agora o valor da estatística de teste e o gl para o teste:

$$t = \frac{2902,8 - 3108,1}{\sqrt{\frac{(277,3)^2}{10} + \frac{(205,9)^2}{8}}} = \frac{-205,3}{113,97} = -1,8$$

Usando  $s_1^2/m = 7689,529$  e  $s_2^2/n = 5299,351$ ,

$$\nu = \frac{(7689,529 + 5299,351)^2}{(7689,529)^2/9 + (5299,351)^2/7} = \frac{168.711.003,7}{10.581.747,35} = 15,94$$

assim, o teste se baseará no gl 15.

6. A Tabela A.8 do Apêndice mostra que a área sob a curva  $t$  do gl 15 à direita de 1,8 é 0,046, assim, o valor  $P$  para um teste de cauda inferior também é 0,046. A saída do MINITAB a seguir resume todos os cálculos:

```

Duas amostras T para não-fundido vs fundido

          N      Média      DesvPadrão      Média SE
não-fundido  10      2903          277          88
fundido      8       3108          206          73

IC de 95% para mu não-fundido - mu fundido: (-488, 38)
mu não-fundido do Teste T = mu fundido (vs<): T = -1,80 P = 0,046 DF = 15

```

7. Usando um nível de significância de 0,05, podemos rejeitar apenas a hipótese nula em favor da hipótese alternativa, confirmando a conclusão do artigo. Entretanto, uma pessoa que necessite de mais evidências pode selecionar  $\alpha = 0,01$ , um nível para o qual  $H_0$  não pode ser rejeitado.

Se a pergunta apresentada questionasse se a fusão aumentou a resistência média real por mais de 100 psi, então as hipóteses relevantes teriam sido  $H_0: \mu_1 - \mu_2 = -100$  versus  $H_a: \mu_1 - \mu_2 < -100$ ; isto é, o valor nulo teria sido  $\Delta_0 = -100$ . ■

## Procedimentos Combinados $t$

As alternativas para os procedimentos  $t$  de duas amostras descritos anteriormente assumem não somente que as duas distribuições das populações são normais, como também que possuem variâncias iguais ( $\sigma_1^2 = \sigma_2^2$ ). Isto é, as duas curvas da distribuição de população são consideradas normais com dispersões iguais, a única diferença possível entre elas é o local onde estão centralizadas.

Seja  $\sigma^2$  a variância da população comum. Então, a padronização de  $\bar{X} - \bar{Y}$  fornece

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma^2}{m} + \frac{\sigma^2}{n}}} = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\sigma^2 \left( \frac{1}{m} + \frac{1}{n} \right)}}$$

que possui uma distribuição normal padronizada. Antes de essa variável poder ser usada como base para fazer inferências sobre  $\mu_1 - \mu_2$ , a variância comum deve ser calculada a partir dos dados da amostra. Um estimador de  $\sigma^2$  é  $S_1^2$ , a variância das  $m$  observações na primeira amostra, e o outro é  $S_2^2$ , a variância da segunda amostra. Intuitivamente, um estimador melhor que a variância amostral individual resulta da combinação das duas variâncias amostrais. A primeira idéia pode ser usar  $(S_1^2 + S_2^2)/2$ , a média comum das duas variâncias amostrais. Entretanto, se  $m > n$ , então, a primeira amostra contém mais informações sobre  $\sigma^2$  que a segunda e um comentário análogo é aplicado se  $m < n$ . A média ponderada a seguir das duas variâncias amostrais, chamada **estimador combinado**  $\sigma^2$ , ajusta-se a qualquer diferença entre os dois tamanhos das amostras:

$$S_p^2 = \frac{m-1}{m+n-2} \cdot S_1^2 + \frac{n-1}{m+n-2} \cdot S_2^2$$

A primeira amostra contribui com  $m-1$  graus de liberdade para a estimativa de  $\sigma^2$  e a segunda amostra contribui com  $n-1$  gl para um total de  $m+n-2$  gl. A teoria estatística diz que, se  $S_p^2$  substitui  $\sigma^2$  na expressão para  $Z$ , a variável padronizada resultante possui uma distribuição  $t$  com base em  $m+n-2$  gl. Da mesma forma que as variáveis padronizadas anteriores foram usadas como base para determinar intervalos de confiança e procedimentos de teste, essa variável  $t$  leva imediatamente ao intervalo de confiança  $t$ , combinado para estimar  $\mu_1 - \mu_2$  e o teste combinado  $t$  para testar as hipóteses sobre uma diferença entre as médias.

No passado, muitos estatísticos recomendavam esses procedimentos combinados  $t$  durante os procedimentos  $t$  de duas amostras. O teste combinado  $t$ , por exemplo, pode ser derivado do princípio de razão de verossimilhança, enquanto o teste  $t$  de duas amostras não é um teste de razão de verossimilhança. Além disso, o nível de significância para o teste  $t$  combinado é exato, enquanto é somente aproximado para o teste  $t$  de duas amostras. Entretanto, uma pesquisa recente mostrou que, embora o teste combinado  $t$  supere um pouco o teste  $t$  de duas amostras ( $\beta$ s menores para o mesmo  $\alpha$ ) quando  $\sigma_1^2 = \sigma_2^2$ , o teste anterior pode facilmente levar a conclusões

errôneas, se aplicado quando as variâncias são diferentes. Comentários análogos aplicam-se ao comportamento dos dois intervalos de confiança, isto é, os procedimentos combinados  $t$  não resistem às violações da suposição de variância igual.

Foi sugerido que se poderia fazer um teste preliminar de  $H_0: \sigma_1^2 = \sigma_2^2$  e usar um procedimento combinado  $t$ , se essa hipótese nula não fosse rejeitada. Infelizmente, o “teste  $F$ ” comum de variâncias iguais (Seção 9.5) é bastante sensível à suposição de distribuições normais das populações, muito mais que os procedimentos  $t$ . Recomendamos, portanto, a abordagem conservadora do uso de procedimentos  $t$  de duas amostras, a menos que realmente haja forte evidência de se fazer o contrário, particularmente quando os dois tamanhos das amostras são diferentes.

## Probabilidades de Erro Tipo II

Determinar as probabilidades de erro tipo II (ou, de forma equivalente, a força  $= 1 - \beta$ ) para o teste  $t$  de duas amostras é complicado. Parece não haver qualquer forma simples de usar as curvas  $\beta$  da Tabela A.17 do Apêndice. A versão mais recente do MINITAB (Versão 13) calculará a força para o teste combinado  $t$ , mas não para o teste  $t$  de duas amostras. Entretanto, a *homepage* do Departamento de Estatística da UCLA (<http://www.stat.ucla.edu>) permite acessar a calculadora de uma força que o fará. Por exemplo: especificamos  $m = 10$ ,  $n = 8$ ,  $\sigma_1 = 300$ ,  $\sigma_2 = 225$  (esses são os tamanhos das amostras para o Exemplo 9.7, cujos desvios padrão da amostra são um pouco menores que os valores de  $\sigma_1$  e  $\sigma_2$ ) e exigidos para a força de um teste nível 0,05 bicaudal de  $H_0: \mu_1 - \mu_2 = 0$  quando  $\mu_1 - \mu_2 = 100, 250$  e  $500$ . Os valores resultantes da força foram 0,1089, 0,4609 e 0,9635 (correspondentes a  $\beta = 0,89, 0,54$  e  $0,04$ ), respectivamente. Em geral,  $\beta$  diminuirá à medida que os tamanhos das amostras aumentam, quando  $\alpha$  aumenta e quando  $\mu_1 - \mu_2$  se distancia de 0. O *software* também calculará os tamanhos das amostras necessários, a fim de obter um valor especificado da força para um valor específico de  $\mu_1 - \mu_2$ .

---

### Exercícios | Seção 9.2 (17–35)

17. Determine a quantidade de graus de liberdade para o teste  $t$  de duas amostras ou IC em cada uma das situações a seguir:
  - a.  $m = 10$ ,  $n = 10$ ,  $s_1 = 5,0$ ,  $s_2 = 6,0$
  - b.  $m = 10$ ,  $n = 15$ ,  $s_1 = 5,0$ ,  $s_2 = 6,0$
  - c.  $m = 10$ ,  $n = 15$ ,  $s_1 = 2,0$ ,  $s_2 = 6,0$
  - d.  $m = 12$ ,  $n = 24$ ,  $s_1 = 5,0$ ,  $s_2 = 6,0$
18. Sejam  $\mu_1$  e  $\mu_2$  as densidades médias reais de dois tipos diferentes de tijolo. Assumindo a normalidade das duas distribuições de densidade, teste  $H_0: \mu_1 - \mu_2 = -10$  versus  $H_a: \mu_1 - \mu_2 \neq 0$ , usando os dados a seguir:  $m = 6$ ,  $\bar{x} = 22,73$ ,  $s_1 = 0,164$ ,  $n = 5$ ,  $\bar{y} = 21,95$  e  $s_2 = 0,240$ .
19. Suponha que  $\mu_1$  e  $\mu_2$  sejam distâncias de parada médias reais em 50 mph de certo tipo de carros equipados com dois tipos diferentes de sistemas de frenagem. Use o teste  $t$  de duas amostras no nível de significância 0,01 para testar  $H_0: \mu_1 - \mu_2 = -10$  versus  $H_a: \mu_1 - \mu_2 < -10$  para os dados a seguir:  $m = 6$ ,  $\bar{x} = 115,7$ ,  $s_1 = 5,03$ ,  $n = 6$ ,  $\bar{y} = 129,3$  e  $s_2 = 5,38$ .
20. Use os dados do Exercício 19 para calcular um IC de 95% para a diferença entre a distância de parada média real dos carros equipados com o sistema 1 e dos carros equipados com o sistema 2. O intervalo sugere que informações precisas sobre o valor dessa diferença estão disponíveis?
21. As técnicas quantitativas não-invasivas são necessárias para avaliar rotineiramente os sintomas das neuropatias periféricas, como Síndrome do Túnel Carpal (STC). O artigo "A Gap Detection Tactility Test for Sensory Deficits Associated with Carpal Tunnel Syndrome" (*Ergonomics*, 1995, p. 2588-2601) relatou um teste que utilizou a percepção de um vão minúsculo em uma superfície plana diferente, investigando-se com um dedo. Esse teste assemelha-se de forma funcional a muitas atividades palpáveis relacionadas ao trabalho, como a detecção de arranhões ou defeitos na superfície. Quando não se pode fazer a sondagem com o dedo, o limiar da detecção do vão médio da amostra para  $m = 8$  pessoas normais foi 1,71 mm, e o desvio padrão da amostra foi 0,53; para  $n = 10$  pessoas com STC, a média e o desvio padrão amostrais foram 2,53 e 0,87, respectivamente. Esses dados sugerem que o limiar da detecção do vão médio real para as pessoas com STC excede o das pessoas normais? Determine e teste as hipóteses relevantes, usando um nível de significância de 0,01.
22. O teste de corte inclinado é amplamente aceito para avaliar a união de materiais de reparo resinosos ao concreto; ele utiliza espécimes cilíndricos feitos de duas metades idênticas unidas a 30°. O artigo "Testing the Bond Between Repair Materials and Concrete Substrate" (*ACI Materials J.*, 1996, p. 553-558) relatou que,



para 12 espécimes preparados usando escova de arame, a resistência de corte média amostral ( $N/mm^2$ ) e o desvio padrão da amostra eram 19,20 e 1,58, respectivamente, enquanto, para 12 espécimes delineados manualmente, os valores correspondentes eram 23,13 e 4,01. A resistência média real parece ser diferente para os dois métodos de preparação de superfície? Determine e teste as hipóteses relevantes, usando um nível de significância de 0,05. O que você supõe com relação às distribuições da resistência de corte?

23. Os entrelinhamentos por fusão estão sendo usados com frequência crescente para suportar os tecidos externos e melhorar o formato e a estampa de várias peças de roupa. O artigo "Compatibility of Outer and Fusible Interlining Fabrics in Tailored Garments" (*Textile Res. J.*, 1997, p. 137-142) forneceu os dados a seguir sobre a extensão (%) em 100 gm/cm de espécimes de tecido de alta qualidade (A) e de baixa qualidade (P).

H	1,2	0,9	0,7	1,0	1,7	1,7	1,1	0,9	1,7
	1,9	1,3	2,1	1,6	1,8	1,4	1,3	1,9	1,6
	0,8	2,0	1,7	1,6	2,3	2,0			
P	1,6	1,5	1,1	2,1	1,5	1,3	1,0	2,6	

- a. Construa gráficos de probabilidade normal para verificar a plausibilidade de ambas as amostras selecionadas das distribuições normais das populações.
- b. Construa um *boxplot* comparativo. Ele sugere que há diferença entre a extensibilidade média real dos espécimes de tecido de alta qualidade e a dos espécimes de baixa qualidade?
- c. A média amostral e o desvio padrão da amostra de alta qualidade são 1,508 e 0,444, respectivamente, e os da amostra de baixa qualidade são 1,588 e 0,530. Use o teste *t* de duas amostras para decidir se a extensibilidade média real é diferente para os dois tipos de tecido.

24. A consistência de um pedaço de fruta é um importante indicador de sua maturidade. A consistência de Magness-Taylor ( $N$ ) foi determinada para uma amostra de 20 maçãs douradas com um prazo de validade de zero dia, resultando em uma média amostral de 8,74 e um desvio padrão da amostra de 0,66, e outra amostra de 20 maçãs com um prazo de validade de 20 dias, com uma média e desvio padrão amostrais de 4,96 e 0,39, respectivamente. Calcule intervalo de confiança para a diferença entre a consistência média real das maçãs de zero dia e a consistência média real das maçãs de 20 dias, usando um nível de confiança de 95% e interprete o intervalo.

25. Dor leve nas costas (DLC) é um problema de saúde sério em muitos ambientes industriais. O artigo "Isodynamic Evaluation of Trunk Muscles and Low-Back Pain Among Workers in a Steel Factory" (*Ergonomics*, 1995, p. 2107-2117) relatou os dados resumidos a seguir sobre o intervalo de movimentos laterais (graus) para uma amostra de trabalhadores sem um histórico de DLC e para outra amostra com histórico dessa doença.

Condição	Tamanho Amostral	Média Amostral	DP Amostral
Sem DLC	28	91,5	5,5
dlc	31	88,3	7,8

Calcule um intervalo de confiança de 90% da diferença entre a extensão média de movimentos laterais da população para as duas condições. O intervalo sugere que o movimento lateral médio da população é diferente para as duas condições? A mensagem é diferente se usarmos um nível de confiança de 95%?

26. O artigo "The Influence of Corrosion Inhibitor and Surface Abrasion on the Failure of AluminumWired Twiston Connections" (*IEEE Trans. on Components, Hybrids, and Manuf. Tech.*, 1984, p. 20-25) relatou dados sobre as medições de queda de potencial de uma amostra de conectores reforçados com arame de liga de alumínio e de outra amostra reforçada com arame de alumínio EC. O relatório SAS a seguir sugere que a queda de potencial médio real das conexões de liga (tipo 1) é maior que a das conexões de EC (conforme afirmado no artigo)? Faça o teste apropriado, usando um nível de significância de 0,01. Ao chegar a sua conclusão, que tipo de erro você pode ter cometido? (Observação: O SAS descreve o valor *P* para um teste bicaudal.)

Tipo	N	Média	Des Pad	Erro-Pad
1	20	17,49900000	0,55012821	0,12301241
2	20	16,90000000	0,48998389	0,10956373
		Variâncias	T	DF
		Variável	3,6362	37,5
		Igual	3,6362	38,0
				Prob>  T
				0,0008
				0,0008

27. Acredita-se que o cotovelo de um tenista seja afetado pelo impacto sentido ao atingir a bola. O artigo "Forces on the Hand in the Tennis One-Handed Backhand" (*Intl. J. of Sport Biomechanics*, 1991, p. 282-292) relatou a força ( $N$ ) na mão, logo depois do impacto de um golpe para trás com uma única mão de seis jogadores avançados e oito jogadores intermediários.

Tipo de Jogador	Tamanho Amostral	Média Amostral	DP Amostral
1. Avançado	6	40,3	11,3
2. Intermediário	8	21,4	8,3

Em sua análise dos dados, os autores assumiram que ambas as distribuições de força eram normais. Calcule um IC de 95% para a diferença entre a força média real dos jogadores avançados ( $\mu_1$ ) e a força média real dos jogadores intermediários ( $\mu_2$ ). Seu intervalo fornece forte evidência de que os dois  $\mu$ s são diferentes? Você teria chegado à mesma conclusão calculando um IC para  $\mu_2 - \mu_1$ , (isto é, invertendo as indicações 1 e 2 nos dois tipos de jogadores)? Explique.

28. À medida que a população envelhece, há uma preocupação crescente com relação às lesões relacionadas a acidentes na velhice. O artigo "Age and Gender Differences in Single-Step Recovery from a Forward Fall" (*J. of Gerontology*, 1999, p. M44-M50) relatou um experimento em que o ângulo de inclinação máximo – o mais afastado que uma pessoa é capaz de se inclinar e ainda voltar ao estado normal em uma etapa – foi determinado para uma amostra de mulheres jovens (21-29 anos) e uma amostra de mulheres idosas (67-81 anos). As observações a seguir são consistentes com os dados resumidos fornecidos no artigo:

MJ: 29, 34, 33, 27, 28, 32, 31, 34, 32, 27

MI: 18, 15, 23, 13, 12

Os dados sugerem que o ângulo de inclinação máximo médio real das mulheres idosas é mais que 10 graus menor que o das mulheres jovens? Determine e teste as hipóteses relevantes no nível de significância 0,10, obtendo um valor  $P$ .

29. O artigo "Effect of Internal Gas Pressure on the Compression Strength of Beverage Cans and Plastic Bottles" (*J. Testing and Evaluation*, 1993, p. 129-131) inclui os dados a seguir sobre a resistência à compressão (lb) de uma amostra de latas de alumínio de 12 oz cheias de suco de morango e de outra amostra com refrigerante. Os dados sugerem que a carbonação extra de refrigerante resulta em uma resistência à compressão média maior? Tome como base de sua resposta um valor  $P$ . Que hipóteses são necessárias para sua análise?

Bebidas	Tamanho Amostral	Média Amostral	DP Amostral
Suco de Morango	15	540	21
Refrigerante	15	554	15

30. O artigo "Flexure of Concrete Beams Reinforced with Advanced Composite Orthogrids" (*J. of Aerospace Engi.*, 1997, p. 7-15) forneceu os dados a seguir sobre carga máxima (kN) de dois tipos diferentes de vigas.

Tipo	Tamanho Amostral	Média Amostral	DP Amostral
Grade de fibra de vidro	26	33,4	2,2
Grade de carbono comercial	26	42,8	4,3

- a. Assumindo que as distribuições subjacentes sejam normais, calcule e interprete um IC de 99% para a diferença entre a carga média real das vigas de fibra de vidro e a das vigas de carbono.
- b. O limite superior do intervalo que você calculou na parte (a) fornece um limite de confiança superior de 99% para a diferença entre as duas  $\mu$ s? Se não, cal-

cule esse limite. Ele sugere fortemente que a carga média real das vigas de carbono é maior que a das vigas de fibra de vidro? Explique.

31. Consulte o Exercício 33 na Seção 7.3. O artigo citado também forneceu as seguintes observações sobre o grau de polimerização dos espécimes que possuem viscosidade  $\times$  concentração em um intervalo mais alto:

429	430	430	431	436	437
440	441	445	446	447	

- a. Construa um *boxplot* comparativo para as duas amostras e comente sobre quaisquer características interessantes.
- b. Calcule um intervalo de confiança de 95% para a diferença entre o grau de polimerização médio real do intervalo intermediário e o do intervalo alto. O intervalo sugere que  $\mu_1$  e  $\mu_2$  podem, de fato, ser diferentes? Explique o seu raciocínio.

32. O artigo citado no Exercício 34 na Seção 7.3 forneceu os dados resumidos a seguir sobre os limites de tensão proporcionais dos espécimes construídos, usando dois tipos diferentes de madeira:

Tipo de Madeira	Tamanho Amostral	Média Amostral	DP Amostral
Carvalho vermelho	14	8,48	0,79
Pinheiro de Douglas	10	6,65	1,28

Assumindo que ambas as amostras foram selecionadas de distribuições normais, faça um teste de hipóteses para decidir se o limite de tensão proporcional médio real das junções de carvalho vermelho excede o das junções de pinheiro de Douglas por mais de 1 MPa.

33. A tabela a seguir resume os dados sobre o ganho de peso corporal (g) de uma amostra de animais, aos quais foi dada uma dose de 1 mg/pílula de um esteróide leve e de uma amostra de animais de controle ("The Soft Drug Approach," *CHEMTECH*, 1984, p. 28-38).

Tratamento	Tamanho Amostral	Média Amostral	DP Amostral
Esteróide	8	32,8	2,6
Controle	10	40,5	2,5

Os dados sugerem que o ganho de peso médio real na situação de controle excede o do tratamento com esteróide por mais de 5 g? Determine e teste as hipóteses apropriadas em um nível de significância de 0,01, utilizando o método do valor  $P$ .

34. Considere a variável combinada  $t$

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{m} + \frac{1}{n}}}$$

que possui distribuição  $t$  com  $m + n - 2$  gl quando ambas as distribuições das populações são normais com  $\sigma_1 = \sigma_2$  (veja a subseção Procedimentos combinados  $t$  para obter uma descrição de  $S_p$ ).

- Use esta variável  $t$  para obter a fórmula de um intervalo de confiança combinado  $t$  para  $\mu_1 - \mu_2$ .
- Uma amostra de umidificadores ultrassônicos de uma marca específica foi selecionada, à qual as observações sobre a saída máxima de umidade (oz) em uma câmara controlada eram 14,0, 14,3, 12,2 e 15,1. Uma amostra da segunda marca forneceu os valores 12,1, 13,6, 11,9 e 11,2 ("Multiple Comparisons of Means Using Simultaneous Confidence Intervals," *J.*

*of Quality Technology*, 1989, p. 232-241). Use a fórmula combinada  $t$  da parte (a) para estimar a diferença entre as saídas médias reais das duas marcas com um intervalo de confiança de 95%.

- Estime a diferença entre os dois  $\mu$ s, usando o intervalo  $t$  de duas amostras discutido nesta seção, e compare-o ao intervalo da parte (b).
35. Consulte o Exercício 34. Descreva o teste combinado  $t$  para testar  $H_0: \mu_1 - \mu_2 = \Delta_0$  quando ambas as distribuições das populações são normais  $\sigma_1 = \sigma_2$ . Então, use esse procedimento de teste para testar as hipóteses sugeridas no Exercício 33.

### 9.3 | Análise de Dados Pareados

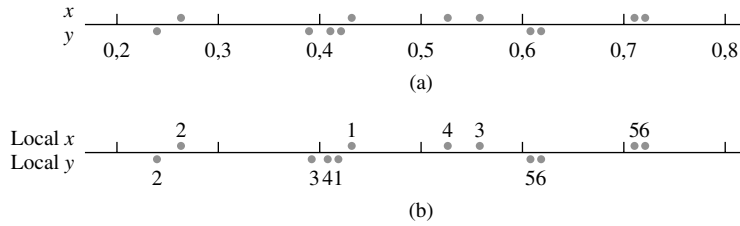
Nas Seções 9.1 e 9.2, consideramos o teste para uma diferença entre duas médias  $\mu_1$  e  $\mu_2$ . Isso foi feito utilizando-se os resultados de uma amostra aleatória  $X_1, X_2, \dots, X_m$  da distribuição com média  $\mu_1$  e uma amostra (dos  $X$ s) completamente independente  $Y_1, \dots, Y_n$  da distribuição com média  $\mu_2$ . Isto é,  $m$  indivíduos foram selecionados da população 1 e  $n$  indivíduos diferentes da população 2, ou a  $m$  indivíduos (ou objetos experimentais) foi dado um tratamento e a outro conjunto de  $n$  indivíduos foi dado outro tratamento. Em contraste, existem várias situações experimentais em que há somente um conjunto de  $n$  indivíduos ou objetos experimentais, e duas observações são feitas sobre cada indivíduo ou objeto, resultando em um pareamento natural de valores.

#### Exemplo 9.8

Um vestígio de metais em água potável afeta o sabor e concentrações extraordinariamente altas podem apresentar risco à saúde. O artigo "Trace Metals of South Indian River" (*Envir. Studies*, 1982, p. 62-66) relata um estudo em que foram selecionados seis locais do rio (seis objetos experimentais) e a concentração de zinco (mg/L) determinada para a água da superfície e para a água mais profunda em cada local. Os seis pares de observações são exibidos na tabela a seguir. Os dados sugerem que a concentração média real na água profunda excede à da água superficial?

	Local					
	1	2	3	4	5	6
Concentração de zinco em água profunda ( $x$ )	0,430	0,266	0,567	0,531	0,707	0,716
Concentração de zinco na água da superfície ( $y$ )	0,415	0,238	0,390	0,410	0,605	0,609
Diferença	0,015	0,028	0,177	0,121	0,102	0,107

A Figura 9.4(a) exhibe o gráfico desses dados. À primeira vista, parece haver pequena diferença entre as amostras  $x$  e  $y$ . De local para local, existe grande quantidade de variabilidade em cada amostra e parece que quaisquer diferenças entre as amostras podem ser atribuídas a essa variabilidade. Entretanto, quando as observações são identificadas pelo local, como na Figura 9.4(b), surge um ponto de vista diferente. Em cada local, a concentração do fundo excede a concentração superficial. Isso é confirmado pelo fato de que todas as  $x - y$  diferenças (concentração da água profunda - concentração da água da superfície) exibidas na linha de baixo da tabela de dados são positivas. Como veremos, uma análise correta desses dados concentra-se nessas diferenças.



**Figura 9.4** Gráfico de dados pareados do Exemplo 9.8: (a) observações não-identificadas pelo local; (b) observações identificadas pelo local

**HIPÓTESES**

Os dados consistem de  $n$  pares selecionados independentemente  $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ , com  $E(X_i) = \mu_1$  e  $E(Y_i) = \mu_2$ . Sejam  $D_1 = X_1 - Y_1, D_2 = X_2 - Y_2, \dots, D_n = X_n - Y_n$ , de modo que os  $D_i$ s sejam as diferenças dentro dos pares. Então, os  $D_i$ s são considerados normalmente distribuídos com valor médio  $\mu_D$  e variância  $\sigma_D^2$  (isso geralmente é uma consequência de os próprios  $X_i$ s e  $Y_i$ s serem normalmente distribuídos).

Estamos novamente interessados em testar as hipóteses sobre a diferença  $\mu_1 - \mu_2$ . O denominador do teste  $t$  de duas amostras foi obtido aplicando-se primeiro a regra  $V(\bar{X} - \bar{Y}) = V(\bar{X}) + V(\bar{Y})$ . Entretanto, com os dados pareados, as observações  $X$  e  $Y$  dentro de cada par, geralmente, não são independentes. Assim,  $\bar{X}$  e  $\bar{Y}$  não são independentes entre si e a regra não é válida. Devemos, portanto, abandonar o teste  $t$  de duas amostras e procurar um método de análise alternativo.

**Teste t Pareado**

Em virtude de pares diferentes serem independentes, os  $D_i$ s são independentes entre si. Se fizermos  $D = X - Y$ , em que  $X$  e  $Y$  são a primeira e a segunda observações, respectivamente, dentro de um par arbitrário, então, a diferença esperada é

$$\mu_D = E(X - Y) = E(X) - E(Y) = \mu_1 - \mu_2$$

(a regra de valores esperados usada aqui é válida, mesmo quando  $X$  e  $Y$  são dependentes). Dessa forma, qualquer hipótese sobre  $\mu_1 - \mu_2$  pode ser expressa como uma hipótese sobre a diferença média  $\mu_D$ . Mas, uma vez que os  $D_i$ s constituem uma amostra aleatória normal (de diferenças) com média  $\mu_D$ , as hipóteses sobre  $\mu_D$  podem ser testadas usando um teste  $t$  de uma amostra. Isto é, para testar as hipóteses sobre  $\mu_1 - \mu_2$  quando os dados são pareados, calcule as diferenças  $D_1, D_2, \dots, D_n$  e faça um teste  $t$  de uma amostra (com base em  $n - 1$  gl) sobre as diferenças.

**Teste t Pareado**

Hipótese nula:  $H_0: \mu_D = \Delta_0$  (onde  $D = X - Y$  é a diferença entre a primeira e a segunda observações dentro de um par e  $\mu_D = \mu_1 - \mu_2$ )

Valor da estatística de teste:  $t = \frac{\bar{d} - \Delta_0}{s_D/\sqrt{n}}$  (onde  $\bar{d}$  e  $s_D$  são a média amostral e o desvio padrão, respectivamente, dos  $d_i$ s)

**Hipótese Alternativa**

**Região de Rejeição para o Teste Nível  $\alpha$**

- $H_a: \mu_D > \Delta_0$   $t \geq t_{\alpha, n-1}$
- $H_a: \mu_D < \Delta_0$   $t \leq -t_{\alpha, n-1}$
- $H_a: \mu_D \neq \Delta_0$  tanto  $t \geq t_{\alpha/2, n-1}$  como  $t \leq -t_{\alpha/2, n-1}$

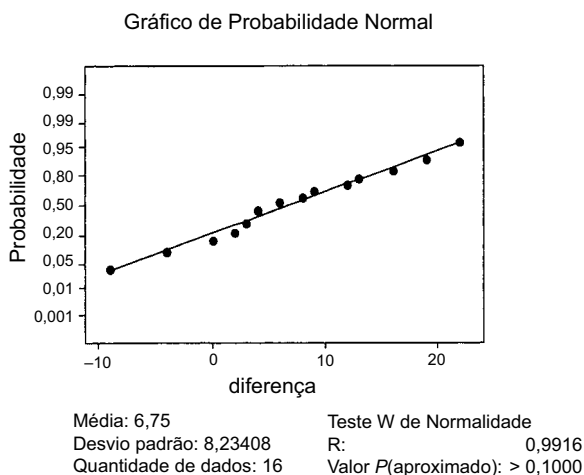
Um valor  $P$  pode ser calculado como foi feito para os testes  $t$  anteriores.

**Exemplo 9.9**

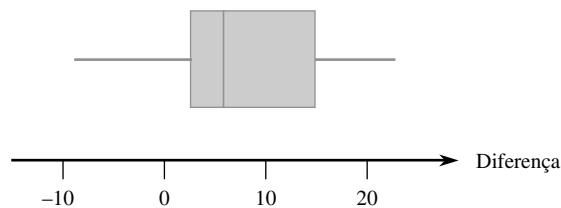
Distúrbios músculo-esqueléticos de pescoço e ombro são muito comuns entre funcionários de escritório que fazem tarefas repetitivas usando unidades visuais de diagrama. O artigo "Upper-Arm Elevation During Office Work" (*Ergonomics*, 1996, p. 1221-1230) relatou um estudo para determinar se condições de trabalho mais variadas teriam qualquer impacto no movimento do braço. Os dados a seguir foram obtidos de uma amostra de  $n = 16$  pessoas. Cada observação é a quantidade de tempo, expressa como proporção de tempo total observado, durante o qual a elevação do braço estava abaixo de  $30^\circ$ . As duas medições de cada pessoa foram obtidas por 18 meses separadamente. Durante esse período, as condições de trabalho foram alteradas e as pessoas puderam se ocupar de uma variedade maior de tarefas. Os dados sugerem que o tempo médio real durante o qual a elevação está abaixo de  $30^\circ$  é diferente após a mudança de tempo daquele tempo médio antes da mudança?

<i>Pessoa</i>	1	2	3	4	5	6	7	8
<i>Antes</i>	81	87	86	82	90	86	96	73
<i>Depois</i>	78	91	78	78	84	67	92	70
<i>Diferença</i>	3	-4	8	4	6	19	4	3
<i>Pessoa</i>	9	10	11	12	13	14	15	16
<i>Antes</i>	74	75	72	80	66	72	56	82
<i>Depois</i>	58	62	70	58	66	60	65	73
<i>Diferença</i>	16	13	2	22	0	12	-9	9

A Figura 9.5 mostra o gráfico de probabilidade normal das 16 diferenças. O padrão no gráfico é bem reto, mantendo a hipótese de normalidade. Um *boxplot* dessas diferenças aparece na Figura 9.6. Ele está localizado consideravelmente à direita de zero, sugerindo que talvez  $\mu_D > 0$  (observe também que 13 das 16 diferenças são positivas e somente duas negativas).



**Figura 9.5** Gráfico de probabilidade normal do MINITAB das diferenças no Exemplo 9.9



**Figura 9.6** *Boxplot* das diferenças no Exemplo 9.9

Vamos usar agora a seqüência de passos recomendada para testar as hipóteses apropriadas.

1. Seja  $\mu_D$  a diferença média real entre o tempo de elevação antes da mudança das condições de trabalho e o tempo depois da mudança.
2.  $H_0: \mu_D = 0$  (não há diferença entre o tempo médio real antes da mudança e o tempo médio real depois da mudança)
3.  $H_a: \mu_D \neq 0$
4.  $t = \frac{\bar{d} - 0}{s_D/\sqrt{n}} = \frac{\bar{d}}{s_D/\sqrt{n}}$
5.  $n = 16$ ,  $\sum d_i = 108$ ,  $\sum d_i^2 = 1746$ , do qual  $\bar{d} = 6,75$ ,  $s_D = 8,234$ , e

$$t = \frac{6,75}{8,234/\sqrt{16}} = 3,28 \approx 3,3$$

6. A Tabela A.8 do Apêndice mostra que a área à direita de 3,3 sob a curva  $t$  com 15 gl é 0,002. A desigualdade em  $H_a$  sugere que um teste bicaudal é apropriado, de modo que o valor  $P$  é aproximadamente  $2(0,002) = 0,004$  (MINITAB fornece 0,0051).
7. Uma vez que  $0,004 < 0,01$ , a hipótese nula pode ser rejeitada no nível de significância 0,05 ou 0,01. Parece que a diferença média real entre os tempos é diferente de zero, isto é, o tempo médio real após a mudança é diferente do tempo antes da mudança. ■

Quando a quantidade de pares é grande, a suposição de distribuição normal de diferença não é necessária. O TLC confirma o teste  $z$  resultante.

### Intervalo de Confiança de $\mu_D$

Da mesma forma que o IC  $t$  da média de uma única população  $\mu$  baseia-se na variável  $t$  de  $T = (\bar{X} - \mu)/(S/\sqrt{n})$ , intervalo de confiança  $t$  de  $\mu_D (= \mu_1 - \mu_2)$  baseia-se no fato de

$$T = \frac{\bar{D} - \mu_D}{S_D/\sqrt{n}}$$

possuir distribuição  $t$  com  $n - 1$  gl. A manipulação dessa variável  $t$ , como nas derivações anteriores dos ICs, produz o seguinte IC  $100(1 - \alpha)\%$ :

O IC  $t$  pareado de  $\mu_D$  é

$$\bar{d} \pm t_{\alpha/2, n-1} \cdot s_D/\sqrt{n}$$

O limite de confiança monocaudal resulta da conservação do sinal relevante e da substituição de  $t_{\alpha/2}$  por  $t_\alpha$ .

Quando  $n$  é pequeno, a validade desse intervalo requer que a distribuição das diferenças seja no mínimo aproximadamente normal. Para  $n$  grande, o TLC garante que o intervalo  $z$  resultante seja válido sem quaisquer restrições na distribuição das diferenças.

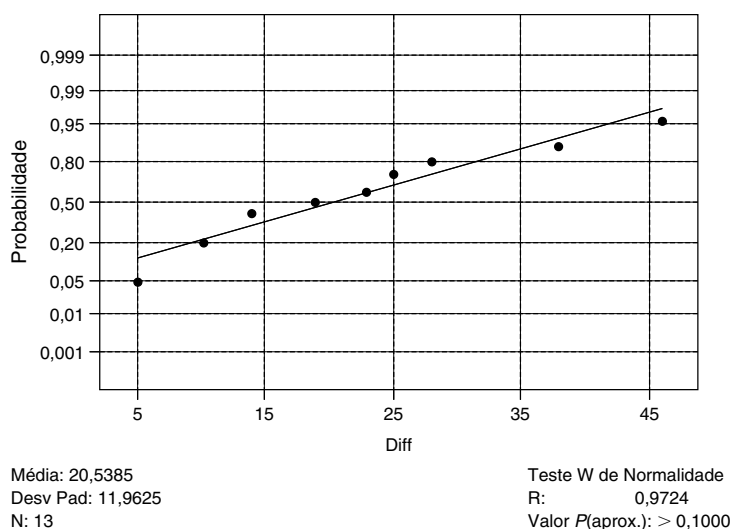
### Exemplo 9.10

Acrescentar imagens médicas computadorizadas a uma base de dados promete fornecer grandes recursos aos médicos. Entretanto, existem outros métodos de se obter tais informações. Assim, a questão da eficiência de acesso precisa ser investigada. O artigo "The Comparative Effectiveness of Conventional and Digital Image Libraries" (*J. of Audiovisual Media in Medicine*, 2001, p. 8-15) relatou um experimento em que 13 profissionais de medicina proficientes em computador foram acompanhados durante a restauração de uma imagem

proveniente de uma biblioteca de *slides* e durante a restauração da mesma imagem da base de dados de um computador conectado à Web.

<i>Pessoa:</i>	1	2	3	4	5	6	7	8	9	10	11	12	13
<i>Slide:</i>	30	35	40	25	20	30	35	62	40	51	25	42	33
<i>Digital:</i>	25	16	15	15	10	20	7	16	15	13	11	19	19
<i>Diferença:</i>	5	19	25	10	10	10	28	46	25	38	14	23	14

Seja  $\mu_D$  a diferença média real entre o tempo de restauração do *slide* (s) e o tempo de restauração digital. Usar o intervalo de confiança  $t$  pareado para calcular  $\mu_D$  requer que a distribuição de diferença seja, no mínimo, aproximadamente normal. O padrão de pontos linear no gráfico de probabilidade normal do MINITAB (Figura 9.7) confirma a suposição de normalidade. (Aparecem somente 9 pontos por causa de empates na diferença).



**Figura 9.7** Gráfico de probabilidade normal das diferenças no Exemplo 9.10

As quantidades resumidas relevantes são  $\sum d_i = 267$ ,  $\sum d_i^2 = 7201$ , das quais  $\bar{d} = 20,5$ ,  $s_D = 11,96$ . O valor crítico  $t$  necessário para um nível de confiança de 95% é  $t_{0,025,12} = 2,179$ , e o IC de 95% é

$$\bar{d} \pm t_{\alpha/2, n-1} \cdot \frac{s_D}{\sqrt{n}} = 20,5 \pm (2,179) \cdot \frac{11,96}{\sqrt{13}} = 20,5 \pm 7,2 = (13,3, 27,7)$$

Assim, podemos ser altamente confiantes (no nível de confiança de 95%) de que  $13,3 < \mu_D < 27,7$ . Esse intervalo de valores plausíveis é bastante amplo, uma consequência do desvio padrão da amostra, grande com relação à média amostral. Um tamanho de amostra muito maior que 13 seria necessário para fazer o cálculo substancialmente com mais precisão. Observe, entretanto, que 0 permanece fora do intervalo, sugerindo que  $\mu_D > 0$ ; isso é confirmado por um teste de hipóteses formal. ■

## Dados Pareados e Procedimentos $t$ de Duas Amostras

Considere o uso do teste  $t$  de duas amostras nos dados pareados. Os numeradores das estatísticas de teste  $t$  de duas amostras e de  $t$  pareado são idênticos, uma vez que  $\bar{d} = \sum d_i/n = [\sum(x_i - y_i)]/n = (\sum x_i)/n - (\sum y_i)/n = \bar{x} - \bar{y}$ . A diferença entre as duas estatísticas deve-se inteiramente aos denominadores. Cada estatística de teste é obtida padronizando-se  $\bar{X} - \bar{Y}$  ( $= \bar{D}$ ), mas na presença de dependência a padronização  $t$  de duas amostras é incorreta. Para ver isso, lembre-se da Seção 5.5 em que

$$V(X \pm Y) = V(X) + V(Y) \pm 2 \text{Cov}(X, Y)$$

Uma vez que a correlação entre  $X$  e  $Y$  é

$$\rho = \text{Corr}(X, Y) = \text{Cov}(X, Y) / [\sqrt{V(X)} \cdot \sqrt{V(Y)}]$$

segue que

$$V(X - Y) = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$$

Aplicando as expressões a  $\bar{X} - \bar{Y}$  produz

$$V(\bar{X} - \bar{Y}) = V(\bar{D}) = V\left(\frac{1}{n}\sum D_i\right) = \frac{V(D_i)}{n} = \frac{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}{n}$$

O teste  $t$  de duas amostras baseia-se na suposição de independência, caso em que  $\rho = 0$ . Mas em muitos experimentos pareados haverá forte dependência *positiva* entre  $X$  e  $Y$  ( $X$  grande associado a  $Y$  grande), de modo que  $\rho$  será positivo e a variância de  $\bar{X} - \bar{Y}$  será menor que  $\sigma_1^2/n + \sigma_2^2/n$ . Assim, *quando há dependência positiva dentro dos pares, o denominador da estatística  $t$  pareada deve ser menor que o de  $t$  do teste de amostras independentes*. Geralmente, o  $t$  de duas amostras estará muito mais próximo de zero que  $t$  pareado, compreendendo consideravelmente a significância dos dados.

De forma semelhante, quando os dados são pareados, o IC de  $t$  pareado geralmente será mais restrito que o IC de  $t$  de duas amostras (incorreto). Isso porque existe tipicamente muito menos variabilidade nas diferenças que nos valores de  $x$  e  $y$ .

## Experimentos Pareados Versus Não-pareados

Em nossos exemplos, os dados pareados resultaram de duas observações sobre a mesma pessoa (Exemplo 9.9) ou objeto experimental (posição no Exemplo 9.8). Mesmo quando isso não pode ser feito, os dados pareados com dependência dentro dos pares podem ser obtidos combinando-se os indivíduos ou objetos em uma ou mais características que se acreditou influenciarem as respostas. Por exemplo: em um experimento médico que visa comparar a eficácia de dois medicamentos para baixar a pressão sanguínea, o orçamento do investigador pode permitir o tratamento de 20 pacientes. Se 10 pacientes são selecionados aleatoriamente para tratamento com o primeiro medicamento, e os outros 10, selecionados independentemente para tratamento com o segundo medicamento, o resultado será um experimento de amostras independentes.

Entretanto, o investigador, sabendo que a pressão sanguínea é influenciada pela idade e pelo peso, pode resolver dispor os pacientes em pares, de modo que, dentro de cada um dos 10 pares resultantes, a idade e o peso fossem aproximadamente iguais (embora pudesse haver diferenças relativamente grandes entre os pares). Então, cada medicamento seria dado a um paciente diferente dentro de cada par para um total de 10 observações sobre cada medicamento.

Sem essa combinação (ou "bloqueio"), um medicamento pode parecer melhor do que o outro apenas porque os pacientes de uma amostra eram mais leves e mais jovens e, assim, mais suscetíveis a uma diminuição da pressão sanguínea que os pacientes mais pesados e mais velhos da segunda amostra. Entretanto, existe um preço a ser pago pelo pareamento – um número menor de graus de liberdade para a análise pareada – assim, devemos perguntar quando se deve preferir um tipo de experimento a outro.

Não há uma resposta direta e precisa para essa pergunta, mas há algumas diretrizes úteis. Se temos de escolher entre dois testes  $t$  válidos (e realizados no mesmo nível de significância  $\alpha$ ), devemos preferir o teste que tenha maior quantidade de graus de liberdade. A razão para isso é que um número maior de graus de liberdade significa  $\beta$  menor para qualquer valor alternativo fixo do parâmetro ou dos parâmetros. Isto é, para uma probabilidade fixa de erro tipo I, a probabilidade de erro tipo II diminuiu pelo incremento nos graus de liberdade.

Entretanto, se as unidades experimentais são bastante heterogêneas nas suas respostas, será difícil detectar diferenças pequenas, porém, significativas, entre os dois tratamentos. Isso foi o que aconteceu essencialmente nos dados definidos no Exemplo 9.8. Para ambos os "tratamentos" (a água do fundo e a da superfície), há uma



grande variabilidade entre os locais, que tende a mascarar as diferenças nos tratamentos dentro dos locais. Se há uma alta correlação positiva dentro das unidades experimentais ou nas pessoas, a variância de  $\bar{D} = \bar{X} - \bar{Y}$  será muito menor que a variância não-pareada. Devido a essa variância reduzida, será mais fácil detectar uma diferença com amostras pareadas que com amostras independentes. Os prós e contras do pareamento podem agora ser resumidos como segue:

1. Se existe grande heterogeneidade entre as unidades experimentais e uma correlação grande nas unidades experimentais ( $\rho$  muito positivo), então, a perda em graus de liberdade será compensada pela precisão elevada associada ao pareamento, de modo que um experimento pareado é preferível a um experimento de amostras independentes.
2. Se as unidades experimentais são relativamente homogêneas e a correlação nos pares não é grande, o ganho na precisão devido ao pareamento passará a ter mais valor pela diminuição nos graus de liberdade, de modo que um experimento de amostras independentes deve ser usado.

Evidentemente, os valores de  $\sigma_1^2$ ,  $\sigma_2^2$ , e  $\rho$  geralmente não serão conhecidos com muita precisão, de modo que um investigador será solicitado a fazer instintivamente um julgamento de como obter a situação 1 ou 2. Em geral, se o número de observações que podem ser obtidas for grande, então uma perda em graus de liberdade (por exemplo, de 40 para 20) não será séria, mas, se o número for pequeno, a perda (digamos, de 16 para 18) devido ao pareamento pode ser séria, caso não seja compensada pela aumentada precisão. Aplicam-se considerações semelhantes ao se escolher entre os dois tipos de experimentos para calcular  $\mu_1 - \mu_2$  com um intervalo de confiança.

**Exercícios** | Seção 9.3 (36–46)

36. Considere os dados a seguir sobre a carga de quebra (kg/25 mm de largura) de vários tecidos em uma condição não-friccionada e uma condição friccionada ("The Effect of Wet Abrasive Wear on the Tensile Properties of Cotton and Polyester-Cotton Fabrics," *J. Testing and Evaluation*, 1993, p. 84-93). Use o teste *t* pareado, conforme fizeram os autores do artigo citado, para testar  $H_0: \mu_D = 0$  versus  $H_a: \mu_D > 0$  no nível de significância 0,01.

	Tecido							
	1	2	3	4	5	6	7	8
Sem	36,4	55,0	51,5	38,7	43,2	48,8	25,6	49,8
Com	28,5	20,0	46,0	34,5	36,5	52,5	26,5	46,5

37. O cromo hexavalente foi identificado como carcinógeno inalável e uma toxina no ar, causando preocupação em várias localidades diferentes. O artigo "Airborne Hexavalent Chromium in Southwestern Ontario" (*J. of Air and Waste Mgmt. Assoc.*, 1997, p. 905-910) forneceu os dados a seguir sobre as concentrações interna e externa (nanogramas/m<sup>3</sup>) de uma amostra de casas selecionadas em uma região.

Casa	1	2	3	4	5	6	7	8	9
Interior	0,07	0,08	0,09	0,12	0,12	0,12	0,13	0,14	0,15
Exterior	0,29	0,68	0,47	0,54	0,97	0,35	0,49	0,84	0,86
Casa	10	11	12	13	14	15	16	17	
Interior	0,15	0,17	0,17	0,18	0,18	0,18	0,18	0,19	
Exterior	0,28	0,32	0,32	1,55	0,66	0,29	0,21	1,02	
Casa	18	19	20	21	22	23	24	25	
Interior	0,20	0,22	0,22	0,23	0,23	0,25	0,26	0,28	
Exterior	1,59	0,90	0,52	0,12	0,54	0,88	0,49	1,24	
Casa	26	27	28	29	30	31	32	33	
Interior	0,28	0,29	0,34	0,39	0,40	0,45	0,54	0,62	
Exterior	0,48	0,27	0,37	1,26	0,70	0,76	0,99	0,36	

- a. Calcule o intervalo de confiança para a diferença média da população entre as concentrações interna e externa, usando nível de confiança de 95% e interprete o intervalo resultante.
  - b. Se a 34ª casa tivesse de ser selecionada aleatoriamente de uma população, entre quais valores você poderia prever a diferença nas concentrações?
38. Espécimes de concreto com razões de altura/diâmetro variáveis cortados em várias posições no cilindro original foram obtidos de uma mistura de concreto de resistência normal e de uma mistura de alta resistência. A

tensão de pico (MPa) foi determinada para cada mistura, resultando nos dados a seguir ("Effect of Length on Compressive Strain Softening of Concrete," *J. of Engr. Mechanics*, 1997, p. 25-35):

Condição de teste	1	2	3	4	5
Normal	42,8	55,6	49,0	48,7	44,1
Alta	90,9	93,1	86,3	90,3	88,5
Condição de teste	6	7	8	9	10
Normal	55,4	50,1	45,7	51,4	43,1
Alta	88,1	93,2	90,8	90,1	92,6
Condição de teste	11	12	13	14	15
Normal	46,8	46,7	47,7	45,8	45,4
Alta	88,2	88,6	91,0	90,0	90,1

- a. Construa um *boxplot* comparativo das tensões de pico para os dois tipos de concreto e comente sobre quaisquer características interessantes.
- b. Calcule a diferença entre as tensões de pico médias reais dos dois tipos de concreto, de forma que forneça informações sobre precisão e confiabilidade. Certifique-se de verificar a plausibilidade de quaisquer hipóteses necessárias em sua análise. Parece plausível que as tensões de pico médias reais dos dois tipos de concreto sejam idênticas? Por que sim ou por que não?

39. Cientistas e engenheiros querem comparar com frequência duas técnicas diferentes para medir ou determinar o valor de uma variável. Em tais situações, o interesse centraliza-se em testar se a diferença média nas medições é zero. O artigo "Evaluation of the Deuterium Dilution Technique Against the Test Weighing Procedure for the Determination of Breast Milk Intake" (*Amer. J. Clinical Nutr.*, 1983, p. 996-1003) relata os dados a seguir sobre a quantidade de leite ingerida por cada uma das 14 crianças selecionadas aleatoriamente.

	Criança			
	1	2	3	4
Método Isotópico	1509	1418	1561	1556
Método teste de pesos	1498	1254	1336	1565
Diferença	11	164	225	-9

	Criança			
	5	6	7	8
Método Isotópico	2169	1760	1098	1198
Método teste de pesos	2000	1318	1410	1129
Diferença	169	442	-312	69

	Criança		
	9	10	11
Método Isotópico	1479	1281	1414
Método teste de pesos	1342	1124	1468
Diferença	137	157	-54

	Criança		
	12	13	14
Método Isotópico	1954	2174	2058
Método teste de pesos	1604	1722	1518
Diferença	350	452	540

- a. É plausível que a distribuição da população de diferenças seja normal?
  - b. Parece que a diferença média real entre os valores de entrada medidos pelos dois métodos é diferente de zero? Determine o valor *P* do teste e use-o para chegar a uma conclusão no nível de significância 0,05.
40. Dois tipos de iscas de peixe, uma feita de tubos de argila vitrificada e outra de blocos de concreto e escova, foram usados durante 16 períodos de tempo diferentes durante 4 anos no lago Tohopekaliga, Flórida ("Two Types of Fish Attractors Compared in Lake Tohopekaliga, Florida", *Trans. Amer. Fisheries Soc.*, 1978, p. 689-695). As observações a seguir referem-se a peixes pescados por dia de pescaria:

	Período							
	1	2	3	4	5	6	7	8
Tubo	6,64	7,89	1,83	0,42	0,85	0,29	0,57	0,63
Escova	9,73	8,21	2,17	0,75	1,61	0,75	0,83	0,56

	Período							
	9	10	11	12	13	14	15	16
Tubo	0,32	0,37	0,00	0,11	4,86	1,80	0,23	0,58
Escova	0,76	0,32	0,48	0,52	5,38	2,33	0,91	0,79

Na média, uma isca parece ser mais eficaz que a outra?

- a. Use o teste *t* pareado com  $\alpha = 0,01$  para testar  $H_0: \mu_D = 0$  versus  $H_a: \mu_D \neq 0$ .
  - b. O que acontece se o teste *t* de duas amostras for usado ( $s_1 = 2,48$  e  $s_2 = 2,91$ )?
41. Em um experimento designado a estudar os efeitos do nível de iluminação no desempenho da tarefa ("Performance of Complex Tasks Under Different Levels of Illumination", *J. Illuminating Eng.*, 1976, p. 235-242), as pessoas foram solicitadas a inserir uma sonda de ponta fina no furo de 10 agulhas em sucessão rápida para um nível de luz baixo com fundo preto e um nível mais alto com fundo branco. Cada valor é o tempo (s) necessário para completar a tarefa.

	Pessoas				
	1	2	3	4	5
Preto	25,85	28,84	32,05	25,74	20,89
Branco	18,23	20,84	22,96	19,68	19,50

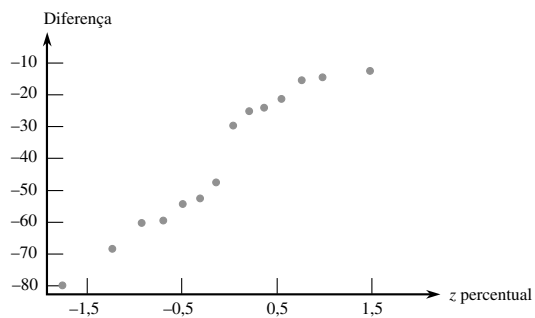
	Pessoas			
	6	7	8	9
Preto	41,05	25,01	24,96	27,47
Branco	24,98	16,61	16,07	24,59

Os dados indicam que o nível mais alto de iluminação produz uma diminuição de mais de 5 s no tempo de conclusão médio real da tarefa? Teste as hipóteses apropriadas, usando a abordagem do valor  $P$ .

42. Estimou-se que, entre 1945 e 1971, aproximadamente 2 milhões de crianças nasceram de mães tratadas com dietilestilbestrol (DES), um estrogênio não-esteróide recomendado para manutenção da gestação. A FDA proibiu esse medicamento em 1971, pois a pesquisa revelou uma ligação com a incidência de câncer cervical. O artigo "Effects of Prenatal Exposure to Diethylstilbestrol (DES) on Hemispheric Laterality and Spatial Ability in Human Males" (*Hormones and Behavior*, 1992, p. 62-75) discutiu um estudo em que 10 homens expostos ao DES e seus irmãos não-expostos se submeteram a vários testes. Esses são os dados resumidos sobre os resultados de um teste de habilidade espacial:  $\bar{x} = 12,6$  (expostos),  $\bar{y} = 13,7$ , e o erro-padrão da diferença média = 0,5. Faça o teste no nível 0,05 para ver se a exposição está associada à habilidade espacial reduzida, obtendo-se o valor  $P$ .
43. A doença de Cushing caracteriza-se pela fraqueza muscular devida à disfunção pituitária ou supra-renal. Para oferecer tratamento eficaz, é importante detectar a doença de Cushing na infância o quanto antes. A idade no início dos sintomas e a idade no diagnóstico de 15 crianças que sofrem da doença foram dadas no artigo "Treatment of Cushing's Disease in Childhood and Adolescence by Transphenoidal Microadenectomy" (*New Engl. J. of Med.*, 1984, p. 889). Aqui estão os valores das diferenças entre a idade no começo dos sintomas e a idade no diagnóstico:

-24 -12 -55 -15 -30 -60 -14 -21  
 -48 -12 -25 -53 -61 -69 -80

- a. O gráfico de probabilidade normal a seguir deixa fortes dúvidas sobre a normalidade aproximada da distribuição de população de diferenças?



- b. Calcule o limite de confiança de 95% inferior para a diferença média da população e interprete o limite resultante.
- c. Suponha que as diferenças (idade no diagnóstico) – (idade no início) tenham sido calculadas. Qual seria um limite de confiança superior de 95% para a diferença média da população correspondente?
44. O Exemplo 7.11 forneceu informações sobre o módulo de elasticidade obtido 1 minuto depois da aplicação de carga em uma determinada configuração. O artigo citado também forneceu os valores do módulo de elasticidade obtido 4 semanas depois da aplicação de carga para os mesmos espécimes de madeira serrada. Os dados são apresentados aqui.

Observação	1 min	4 semanas	Diferença
1	10.490	9.110	1380
2	16.620	13.250	3370
3	17.300	14.720	2580
4	15.480	12.740	2740
5	12.970	10.120	2850
6	17.260	14.570	2690
7	13.400	11.220	2180
8	13.900	11.100	2800
9	13.630	11.420	2210
10	13.260	10.910	2350
11	14.370	12.110	2260
12	11.700	8.620	3080
13	15.470	12.590	2880
14	17.840	15.090	2750
15	14.070	10.550	3520
16	14.760	12.230	2530

Calcule e interprete o limite de confiança superior para a diferença média real entre o módulo de 1 minuto e o módulo de 4 semanas; verifique primeiro a plausibilidade de quaisquer hipóteses necessárias.

45. Consulte o Exercício 41 e calcule a estimativa de intervalo para a diferença entre o tempo médio real da tarefa sob o nível de iluminação alto e o tempo médio real sob o nível baixo.
46. Construa um conjunto de dados pareados para o qual  $t = \infty$ , de modo que os dados sejam altamente significativos quando a análise correta é usada, ainda que  $t$  para o teste  $t$  de duas amostras esteja bem próximo de zero, assim, a análise incorreta produz um resultado insignificante.

## 9.4 Inferências Quanto à Diferença entre Proporções de Populações

Após apresentar os métodos para comparar as médias de duas populações diferentes, agora nos voltamos para a comparação de duas proporções de população. A notação para este problema é uma extensão da notação usada no problema correspondente de uma população. Sejam  $p_1$  e  $p_2$  as proporções de indivíduos nas populações 1 e 2, respectivamente, que possuem uma característica particular. De maneira alternativa, se usarmos o rótulo  $S$  para um indivíduo que possui a característica de interesse (favorece uma proposição específica, leu pelo menos um livro no último mês etc.), então  $p_1$  e  $p_2$  representam as probabilidades de ver o rótulo  $S$  em um indivíduo escolhido aleatoriamente das populações 1 e 2, respectivamente.

Assumiremos a disponibilidade de uma amostra de  $m$  indivíduos da primeira população e  $n$  da segunda. As variáveis  $X$  e  $Y$  representarão o número de indivíduos em cada amostra que possui a característica que define  $p_1$  e  $p_2$ . Desde que os tamanhos da população sejam muito maiores que os tamanhos das amostras, a distribuição de  $X$  pode ser tomada como binomial com parâmetros  $m$  e  $p_1$ , e de forma semelhante,  $Y$  é tomado como uma variável binomial com parâmetros  $n$  e  $p_2$ . Além disso, as amostras são consideradas independentes entre si, de modo que  $X$  e  $Y$  são variáveis independentes.

O estimador óbvio para  $p_1 - p_2$ , a diferença nas proporções da população, é a diferença correspondente nas proporções da amostra  $X/m - Y/n$ . Com  $\hat{p}_1 = X/m$  e  $\hat{p}_2 = Y/n$ , o estimador de  $p_1 - p_2$  pode ser expresso como  $\hat{p}_1 - \hat{p}_2$ .

### PROPOSIÇÃO

Sejam  $X \sim \text{Bin}(m, p_1)$  e  $Y \sim \text{Bin}(n, p_2)$  com  $X$  e  $Y$  variáveis independentes. Então

$$E(\hat{p}_1 - \hat{p}_2) = p_1 - p_2$$

assim  $\hat{p}_1 - \hat{p}_2$  é um estimador não-tendencioso de  $p_1 - p_2$ , e

$$V(\hat{p}_1 - \hat{p}_2) = \frac{p_1 q_1}{m} + \frac{p_2 q_2}{n} \quad (\text{onde } q_i = 1 - p_i) \quad (9.3)$$

**Demonstração** Como  $E(X) = mp_1$  e  $E(Y) = np_2$ ,

$$E\left(\frac{X}{m} - \frac{Y}{n}\right) = \frac{1}{m}E(X) - \frac{1}{n}E(Y) = \frac{1}{m}mp_1 - \frac{1}{n}np_2 = p_1 - p_2$$

Uma vez que  $V(X) = mp_1q_1$ ,  $V(Y) = np_2q_2$ , e  $X$  e  $Y$  são independentes,

$$V\left(\frac{X}{m} - \frac{Y}{n}\right) = V\left(\frac{X}{m}\right) + V\left(\frac{Y}{n}\right) = \frac{1}{m^2}V(X) + \frac{1}{n^2}V(Y) = \frac{p_1 q_1}{m} + \frac{p_2 q_2}{n} \quad \blacksquare$$

Enfocaremos primeiro as situações em que  $m$  e  $n$  são grandes. Então, em virtude de  $\hat{p}_1$  e  $\hat{p}_2$  terem individualmente distribuições aproximadamente normais, o estimador  $\hat{p}_1 - \hat{p}_2$  também possui aproximadamente uma distribuição normal. Padronizar  $\hat{p}_1 - \hat{p}_2$  produz uma variável  $Z$  cuja distribuição é aproximadamente normal padronizada:

$$Z = \frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{m} + \frac{p_2 q_2}{n}}}$$

## Procedimento de Teste de Amostra Grande

Analogamente às hipóteses de  $\mu_1 - \mu_2$ , a hipótese nula mais geral que um investigador pode considerar seria da forma  $H_0: p_1 - p_2 = \Delta_0$ , onde  $\Delta_0$  é novamente um número especificado. Embora para as médias da população o caso  $\Delta_0 \neq 0$  não tenha apresentado dificuldades, para as proporções da população os casos  $\Delta_0 = 0$  e  $\Delta_0 \neq 0$  devem ser considerados separadamente. Uma vez que a maioria absoluta de problemas reais desse tipo compreende  $\Delta_0 = 0$  (isto é, a hipótese nula  $p_1 = p_2$ ), vamos nos concentrar nesse caso. Quando  $H_0: p_1 - p_2 = 0$  é verdadeira, seja  $p$  o valor comum de  $p_1$  e  $p_2$  (e de forma semelhante para  $q$ ). Então, a variável padronizada

$$Z = \frac{\hat{p}_1 - \hat{p}_2 - 0}{\sqrt{pq\left(\frac{1}{m} + \frac{1}{n}\right)}} \quad (9.4)$$

possui aproximadamente distribuição normal padronizada quando  $H_0$  é verdadeira. Entretanto, esse  $Z$  não pode servir como uma estatística de teste, pois o valor de  $p$  é desconhecido,  $H_0$  afirma somente que há um valor comum de  $p$ , mas não diz que valor é esse. Para obter uma estatística de teste que possui aproximadamente uma distribuição normal padronizada quando  $H_0$  é verdadeira (de modo que o uso de um valor crítico  $z$  apropriado especifica um teste nível alfa),  $p$  deve ser estimado a partir dos dados da amostra.

Assumindo, então, que  $p_1 = p_2 = p$ , em vez de amostras separadas de tamanhos  $m$  e  $n$  de duas populações diferentes (duas distribuições binomiais diferentes), temos realmente um único amostra de tamanho  $m + n$  de uma população com proporção  $p$ . Uma vez que o número total de indivíduos nesta amostra combinada que possui a característica de interesse, é  $X + Y$ , o estimador de  $p$  é

$$\hat{p} = \frac{X + Y}{m + n} = \frac{m}{m + n} \hat{p}_1 + \frac{n}{m + n} \hat{p}_2 \quad (9.5)$$

A segunda expressão para  $\hat{p}$  mostra que ele é realmente uma média ponderada de estimadores  $\hat{p}_1$  e  $\hat{p}_2$  obtidos das duas amostras. Se consideramos (9.5) (com  $\hat{q} = 1 - \hat{p}$ ) e substituirmos de volta em (9.4), a estatística resultante possui aproximadamente uma distribuição normal padronizada quando  $H_0$  é verdadeira.

Hipótese nula:  $H_0: p_1 - p_2 = 0$

O valor da estatística de teste (amostras grandes):  $z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{pq\left(\frac{1}{m} + \frac{1}{n}\right)}}$

**Hipótese alternativa**

$H_a: p_1 - p_2 > 0$

$H_a: p_1 - p_2 < 0$

$H_a: p_1 - p_2 \neq 0$

**Região de Rejeição para o Teste Nível  $\alpha$  Aproximado**

$z \geq z_{\alpha}$

$z \leq -z_{\alpha}$

tanto  $z \geq z_{\alpha/2}$  como  $z \leq -z_{\alpha/2}$

Um valor  $P$  é calculado da mesma forma que para os testes  $z$  anteriores.

### Exemplo 9.11

Alguns acusados em processos criminais confessam-se culpados e são sentenciados sem julgamento, enquanto outros, que se declaram inocentes, são subseqüentemente considerados culpados e então sentenciados. Nos últimos anos, estudiosos legais especularam se as sentenças dos que se confessam culpados são diferentes em termos de severidade das sentenças dos que se declaram inocentes e são subseqüentemente julgados culpados. Considere os dados a seguir sobre os réus do município de São Francisco acusados de roubo, todos com

passagem anterior pela prisão ("Does It Pay to Plead Guilty? Differential Sentencing and the Functioning of Criminal Courts", *Law and Society Rev.*, 1981-1982, p. 45-69). Esses dados sugerem que a proporção de todos os acusados nessas circunstâncias que se confessam culpados e são mandados para a prisão difere da proporção dos que são mandados para a prisão depois de se declararem inocentes e serem julgados culpados?

	Declaração	
	Culpado	Inocente
Número julgado como culpado	$m = 191$	$n = 64$
Número sentenciado à prisão	$x = 101$	$y = 56$
Proporção da amostra	$\hat{p}_1 = 0,529$	$\hat{p}_2 = 0,875$

Sejam  $p_1$  e  $p_2$  as duas proporções da população. As hipóteses de interesse são  $H_0: p_1 - p_2 = 0$  versus  $H_a: p_1 - p_2 \neq 0$ . No nível 0,01  $H_0$  deveria ser rejeitada se  $z \geq z_{0,005} = 2,58$  ou se  $z \leq -2,58$ . A estimativa combinada da proporção de sucesso comum é  $\hat{p} = (101 + 56)/(191 + 64) = 0,616$ . O valor da estatística de teste é então

$$z = \frac{0,529 - 0,875}{\sqrt{(0,616)(0,384)\left(\frac{1}{191} + \frac{1}{64}\right)}} = \frac{-0,346}{0,070} = -4,94$$

Uma vez que  $-4,94 \leq -2,58$ ,  $H_0$  deve ser rejeitada.

O valor  $P$  de um teste  $z$  bicaudal é

$$\text{Valor } P = 2[1 - \Phi(|z|)] = 2[1 - \Phi(4,94)] < 2[1 - \Phi(3,49)] = 0,0004$$

Uma tabela normal padronizada mais extensa produz valor  $P \approx 0,0000006$ . Esse valor  $P$  é tão minúsculo que em qualquer nível  $\alpha$  razoável  $H_0$  deve ser rejeitada. Os dados sugerem fortemente que  $p_1 \neq p_2$  e, em particular, que se confessar inicialmente culpado pode ser uma boa estratégia no que diz respeito a escapar da prisão.

O artigo citado também relatou dados sobre acusados em vários outros municípios. Os autores separaram os dados por tipo de crime (roubo ou furto) e por natureza de antecedentes (nenhum; algum, mas sem prisão e com prisão). Em todos os casos, a conclusão foi a mesma; dentre os acusados considerados culpados, aqueles que confessaram ter culpa foram os menos prováveis de receberem sentenças de prisão. ■

### Probabilidades de Erro Tipo II e Tamanhos das Amostras

Aqui, a determinação de  $\beta$  é um pouco mais complicada do que foi para os outros testes de amostra grande. A razão é que o denominador de  $Z$  é uma estimativa do desvio padrão de  $\hat{p}_1 - \hat{p}_2$  assumindo que  $p_1 = p_2 = p$ . Quando  $H_0$  é falsa,  $\hat{p}_1 - \hat{p}_2$  deve ser padronizada novamente, usando-se

$$\sigma_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{p_1q_1}{m} + \frac{p_2q_2}{n}} \tag{9.6}$$

A forma de  $\sigma$  implica que  $\beta$  não é uma função apenas de  $p_1 - p_2$ , então, fazemos sua representação por  $\beta(p_1, p_2)$ .

Hipótese Alternativa	$\beta(p_1, p_2)$
$H_a: p_1 - p_2 > 0$	$\Phi\left[\frac{z_\alpha\sqrt{pq\left(\frac{1}{m} + \frac{1}{n}\right)} - (p_1 - p_2)}{\sigma}\right]$
$H_a: p_1 - p_2 < 0$	$1 - \Phi\left[\frac{-z_\alpha\sqrt{pq\left(\frac{1}{m} + \frac{1}{n}\right)} - (p_1 - p_2)}{\sigma}\right]$

$$H_a: p_1 - p_2 \neq 0$$

$$\Phi \left[ \frac{z_{\alpha/2} \sqrt{pq \left( \frac{1}{m} + \frac{1}{n} \right)} - (p_1 - p_2)}{\sigma} \right]$$

$$- \Phi \left[ \frac{-z_{\alpha/2} \sqrt{pq \left( \frac{1}{m} + \frac{1}{n} \right)} - (p_1 - p_2)}{\sigma} \right]$$

onde  $\bar{p} = (mp_1 + np_2)/(m + n)$ ,  $\bar{q} = (mq_1 + nq_2)/(m + n)$ , e  $\sigma$  é dado por (9.6).

**Demonstração** Para o teste de cauda superior ( $H_a: p_1 - p_2 > 0$ ),

$$\beta(p_1, p_2) = P \left[ \hat{p}_1 - \hat{p}_2 < z_{\alpha} \sqrt{\hat{p}\hat{q} \left( \frac{1}{m} + \frac{1}{n} \right)} \right]$$

$$= P \left[ \frac{(\hat{p}_1 - \hat{p}_2 - (p_1 - p_2))}{\sigma} < \frac{z_{\alpha} \sqrt{\hat{p}\hat{q} \left( \frac{1}{m} + \frac{1}{n} \right)} - (p_1 - p_2)}{\sigma} \right]$$

Quando  $m$  e  $n$  são grandes,

$$\hat{p} = (m\hat{p}_1 + n\hat{p}_2)/(m + n) \approx (mp_1 + np_2)/(m + n) = \bar{p}$$

e  $\hat{q} \approx \bar{q}$ , que produz a expressão (aproximada) anterior para  $\beta(p_1, p_2)$ . ■

De forma alternativa, para  $p_1, p_2$  especificado com  $p_1 - p_2 = d$ , os tamanhos das amostras necessários para chegar a  $\beta(p_1, p_2) = \beta$  podem ser determinados. Por exemplo: para o teste de cauda superior, igualamos  $-z_{\beta}$  ao argumento de  $\Phi(\cdot)$  (isto é, o que está dentro dos parênteses) no quadro precedente. Se  $m = n$ , existe uma expressão simples para o valor comum.

Para o caso  $m = n$ , o teste nível  $\alpha$  possui probabilidade  $\beta$  de erro tipo II nos valores alternativos  $p_1, p_2$  com  $p_1 - p_2 = d$  quando

$$n = \frac{[z_{\alpha} \sqrt{(p_1 + p_2)(q_1 + q_2)/2} + z_{\beta} \sqrt{p_1 q_1 + p_2 q_2}]^2}{d^2} \quad (9.7)$$

para um teste de cauda superior ou inferior, com  $\alpha/2$  substituindo  $\alpha$  para um teste bicaudal.

### Exemplo 9.12

Uma das aplicações verdadeiramente impressionantes de estatística ocorreu em relação ao projeto do experimento de Salk, em 1954, da vacina contra a poliomielite e à análise dos dados resultantes. Parte do experimento centralizou-se na eficácia da vacina em combater a pólio. Em virtude de se ter pensado que sem um grupo de controle formado por crianças não haveria uma boa base para avaliação da vacina, decidiu-se administrá-la a um grupo, e uma injeção de placebo (visualmente indistinguível da vacina, mas conhecida por não ter efeito) a um grupo de controle. Por razões éticas e também porque se pensou que o conhecimento da administração da vacina pudesse ter um efeito no tratamento e no diagnóstico, o experimento foi realizado de forma

**duplo-cego**, isto é, nem os indivíduos que recebem injeções nem os que as administram realmente sabiam quem estava recebendo vacina e quem estava recebendo o placebo (as amostras foram numericamente codificadas) – lembre-se de que, naquele momento, não estava muito claro se a vacina seria benéfica.

Sejam  $p_1$  e  $p_2$  as probabilidades de uma criança contrair pólio para as condições de controle e tratamento, respectivamente. O objetivo era testar  $H_0: p_1 - p_2 = 0$  versus  $H_a: p_1 - p_2 > 0$  (a alternativa afirma que é menos provável que uma criança vacinada contraia pólio que uma criança não-vacinada). Supondo que o valor real de  $p_1$  é 0,0003 (uma taxa de incidência de 30 por 100.000), a vacina seria uma melhora significativa se a taxa de incidência fosse reduzida à metade – isto é,  $p_2 = 0,00015$ . Usando um teste nível  $\alpha = 0,05$ , seria razoável, então, solicitar tamanhos das amostras para os quais  $\beta = 0,1$  quando  $p_1 = 0,0003$  e  $p_2 = 0,00015$ . Assumindo tamanhos das amostras iguais, o  $n$  necessário é obtido de (9.7) como

$$n = \frac{[1,645\sqrt{(0,5)(0,0045)(1,99955)} + 1,28\sqrt{(0,00015)(0,99985)} + (0,0003)(0,9997)]^2}{(0,0003 - 0,00015)^2}$$

$$= [(0,0349 + 0,0271)/0,00015]^2 \approx 171.000$$

Os dados reais para esse experimento são dados a seguir. Foram usados tamanhos das amostras de aproximadamente 200.000. O leitor pode verificar com facilidade que  $z = 6,43$ , valor altamente significativo. A vacina foi considerada um imenso sucesso!

Placebo:  $m = 201.229$ ,  $x =$  número de casos de pólio  $= 110$

Vacina:  $n = 200.745$ ,  $y = 33$  ■

### Intervalo de Confiança com Amostras Grandes para $p_1 - p_2$

Assim como no caso de médias, muitos problemas de duas amostras envolvem o objetivo de fazer comparações pelo teste de hipóteses, mas, às vezes, uma estimativa de intervalo para  $p_1 - p_2$  é apropriada. Tanto  $\hat{p}_1 = X/m$  como  $\hat{p}_2 = Y/n$  possuem distribuições normais aproximadas quando  $m$  e  $n$  são grandes. Se identificamos  $\theta$  com  $p_1 - p_2$ , então,  $\hat{\theta} = \hat{p}_1 - \hat{p}_2$  satisfaz as condições necessárias para obter um IC de amostra grande. Em particular, o desvio padrão estimado de  $\hat{\theta}$  é  $\sqrt{(\hat{p}_1\hat{q}_1/m) + (\hat{p}_2\hat{q}_2/n)}$ . O intervalo de  $100(1 - \alpha)\%$   $\hat{\theta} \pm z_{\alpha/2} \cdot \hat{\sigma}_{\hat{\theta}}$  então se torna

$$\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1\hat{q}_1}{m} + \frac{\hat{p}_2\hat{q}_2}{n}}$$

Observe que o desvio padrão estimado de  $\hat{p}_1 - \hat{p}_2$  (a expressão de raiz quadrada) aqui é diferente do que foi para o teste de hipóteses quando  $\Delta_0 = 0$ .

Uma pesquisa recente mostrou que o nível de confiança real para o IC tradicional dado anteriormente pode, às vezes, afastar-se muito do nível nominal (o nível que você pensa que está conseguindo ao usar um valor crítico  $z$  específico – por exemplo, 95% quando  $z_{\alpha/2} = 1,96$ ). A melhora sugerida é acrescentar um sucesso e uma falha a cada uma das duas amostras e então substituir os  $\hat{p}$ s e  $\hat{q}$ s na fórmula precedente por  $\tilde{p}$ 's e  $\tilde{q}$ 's onde  $\tilde{p}_1 = (x + 1)/(m + 2)$  etc. Esse intervalo também pode ser usado quando os tamanhos das amostras são bastante pequenos.

#### Exemplo 9.13

Os autores do artigo "Adjuvant Radiotherapy and Chemotherapy in Node-Positive Premenopausal Women with Breast Cancer" (*New Engl. J. of Med.*, 1997, p. 956-962) relataram os resultados de um experimento projetado para comparar o tratamento de pacientes com câncer utilizando somente quimioterapia com o tratamento que usa a combinação de quimioterapia e radiação. Dos 154 indivíduos que receberam o tratamento somente com quimioterapia, 76 sobreviveram pelo menos 15 anos, enquanto 98 dos 164 pacientes que receberam o tratamento híbrido sobreviveram pelo menos esse mesmo período. Com  $p_1$  representando a proporção de todas as mulheres que, quando tratadas apenas com quimioterapia, sobrevivem no mínimo 15 anos, e  $p_2$  representando a proporção análoga para o tratamento híbrido,  $\hat{p}_1 = 76/154 = 0,494$  e  $98/164 = 0,598$ . Um intervalo de confiança para a diferença entre as proporções com base na fórmula tradicional com um nível de confiança de aproximadamente 99% é



$$0,494 - 0,598 \pm (2,58) \sqrt{\frac{(0,494)(0,506)}{154} + \frac{(0,598)(0,402)}{164}} = -0,104 \pm 0,143 = (-0,247, 0,039)$$

No nível de confiança de 99%, é plausível que  $-0,247 < p_1 - p_2 < 0,039$ . Esse intervalo é razoavelmente extenso, um reflexo do fato de que os tamanhos de amostra não são terrivelmente grandes para esse tipo de intervalo. Observe que 0 é um dos valores plausíveis de  $p_1 - p_2$ , sugerindo que nenhum tratamento pode ser considerado superior ao outro. Usando  $\tilde{p}_1 = 77/156 = 0,494$ ,  $\tilde{q}_1 = 79/156 = 0,506$ ,  $\tilde{p}_2 = 0,596$ ,  $\tilde{q}_2 = 0,404$  com base nos tamanhos das amostras de 156 e 166, respectivamente, o intervalo "melhorado" aqui é idêntico ao intervalo anterior. ■

## Inferências com Amostras Pequenas

Algumas vezes, uma inferência relacionada a  $p_1 - p_2$  pode ter de se basear em amostras para as quais pelo menos um tamanho de amostra é pequeno. Os métodos apropriados para tais situações não são tão diretos quanto aqueles para amostras grandes, e há mais controvérsia entre os estatísticos com relação aos procedimentos recomendados. Um teste usado com frequência, chamado teste de Fisher-Irwin, baseia-se na distribuição hipergeométrica. Para mais informações, consulte seu amigo estatístico.

---

### Exercícios | Seção 9.4 (47-56)

47. Alguém que muda de marca induzido por motivos financeiros é menos provável de permanecer fiel do que alguém que muda sem essa indução? Sejam  $p_1$  e  $p_2$  as proporções reais de pessoas que mudam para uma determinada marca com e sem indução, respectivamente, que fazem subsequentemente uma segunda compra. Teste  $H_0: p_1 - p_2 = 0$  versus  $H_a: p_1 - p_2 < 0$  usando  $\alpha = 0,01$  e os dados a seguir:

$$m = 200 \quad \text{números de sucessos} = 30$$

$$n = 600 \quad \text{números de sucessos} = 180$$

(Dados semelhantes são fornecidos em "Impact of Deals and Deal Retraction on Brand Switching," *J. Marketing*, 1980, p. 62-70.)

48. Uma amostra de 300 adultos residentes em área urbana de um estado específico revelou que 63 favoreceram o aumento do limite de velocidade nas estradas de 55 para 65 mph, enquanto uma amostra de 180 habitantes rurais mostrou que 75 favoreceram o mesmo aumento. Esses dados indicam que a opinião sobre o aumento do limite de velocidade é diferente para os dois grupos de habitantes?
- a. Teste  $H_0: p_1 = p_2$  versus  $H_a: p_1 \neq p_2$  usando  $\alpha = 0,05$ , onde  $p_1$  refere-se à população urbana.
- b. Se as proporções reais que favorecem o aumento são realmente  $p_1 = 0,20$  (urbano) e  $p_2 = 0,40$  (rural), qual é a probabilidade de  $H_0$  ser rejeitada usando-se um teste nível 0,05 com  $m = 300$ ,  $n = 180$ ?
49. Acredita-se que a capa frontal e a natureza da primeira pergunta de pesquisas por correspondência influenciem o índice de respostas. O artigo "The Impact of Cover Design and First Questions on Response Rates for a Mail Survey of Skydivers" (*Leisure Sciences*, 1991, p. 67-76)

testou essa teoria, fazendo experimentos com diferentes modelos de capas. Uma capa era lisa, a outra tinha a figura de um pára-quedista. Os pesquisadores consideraram que o índice de retorno seria menor para a capa lisa.

Capa	Quantidade Enviada	Quantidade Desenvolvida
Lisa	207	104
Pára-quedista	213	109

Esses dados confirmam a hipótese dos pesquisadores? Teste as hipóteses relevantes, usando  $\alpha = 0,10$ , calculando primeiro um valor  $P$ .

50. Os professores consideram seu trabalho gratificante e satisfatório? O artigo "Work-Related Attitudes" (*Psychological Reports*, 1991, p. 443-450) relata os resultados de uma pesquisa de 395 professores de escola primária e 266 de escola secundária. Dos professores de escola primária, 224 disseram que estavam muito satisfeitos com seus empregos, e dos professores de escola secundária, 126 manifestaram a mesma opinião. Calcule a diferença entre a proporção de todos os professores de escola primária que estão satisfeitos e todos os de escola secundária que também estão satisfeitos, calculando um IC.
51. Uma amostra aleatória de 5726 números de telefone de uma região, adquiridos em março de 2002, mostrou que 1105 não estavam incluídos na lista e 1 ano depois, uma amostra de 5384 mostrou 980 números não-listados.
- a. Faça o teste no nível 0,10 para ver se há uma diferença nas proporções reais de números não-listados entre os dois anos.

- b. Se  $p_1 = 0,20$  e  $p_2 = 0,18$ , que tamanhos das amostras ( $m = n$ ) seriam necessários para detectar tal diferença com probabilidade de 0,90?
52. Está sendo dada atenção crescente à radiação por ionização como método de preservação de produtos de horticultura. O artigo "The Influence of Gamma-Irradiation on the Storage Life of Red Variety Garlic" (*J. of Food Processing and Preservation*, 1983, p. 179-183) relata que 153 dos 180 bulbos de alho irradiados foram comercializados (sem crescimento, deterioração ou amolecimento externo) 240 dias depois do tratamento, enquanto somente 119 dos 180 bulbos não-tratados foram vendidos depois desse período. Esses dados sugerem que a radiação por ionização é benéfica no que diz respeito à comercialização?
53. Em investigações médicas, a razão  $\theta = p_1/p_2$  geralmente é de maior interesse que a diferença  $p_1 - p_2$  (por exemplo: quantas vezes é mais provável que os indivíduos se recuperem ao receberem o tratamento 1 do que ao receberem o tratamento 2?). Seja  $\hat{\theta} = \hat{p}_1/\hat{p}_2$ . Quando  $m$  e  $n$  são grandes, a estatística  $\ln(\hat{\theta})$  possui distribuição aproximadamente normal com valor médio aproximado  $\ln(\theta)$  e desvio padrão aproximado  $[(m - x)/(mx) + (n - y)/(ny)]^{1/2}$ .
- a. Use esses fatos para obter uma fórmula de IC de 95% de amostra grande para estimar  $\ln(\theta)$ , e então um IC para o próprio  $\theta$ .
- b. Retorne aos dados de ataque cardíaco do Exemplo 1.3 e calcule um intervalo de valores plausíveis para  $\theta$  no nível de confiança de 95%. O que esse intervalo sugere sobre a eficácia do tratamento com aspirina?
54. Às vezes, os experimentos que utilizam respostas de sucesso ou fracasso são conduzidos de forma antes/depois ou emparelhada. Suponha que, antes do principal discurso político de um candidato,  $n$  indivíduos são selecionados e questionados se favorecem ( $S$ ) ou não ( $F$ ) o candidato. Então, depois do discurso, é feita a mesma pergunta às mesmas  $n$  pessoas. As respostas podem ser inseridas em uma tabela como segue:

		Depois	
		S	F
Antes	S	$X_1$	$X_2$
	F	$X_3$	$X_4$

onde  $X_1 + X_2 + X_3 + X_4 = n$ . Sejam  $p_1, p_2, p_3$  e  $p_4$  as probabilidades das quatro células, de modo que  $p_1 = P(S$  antes e  $S$  depois), e assim por diante. Queremos testar a hipótese de que a proporção real de defensores ( $S$ ) depois do discurso não aumentou contra a alternativa de que aumentou.

- a. Escreva as duas hipóteses de interesse em termos de  $p_1, p_2, p_3$  e  $p_4$ .
- b. Construa um estimador para a diferença antes/depois nas probabilidades de sucesso.
- c. Quando  $n$  é grande, é possível mostrar que a variável  $(X_1 - X_2)/n$  possui distribuição aproximadamente normal com variância dada por  $[p_1 + p_2 - (p_1 - p_2)^2]/n$ . Use isso para construir uma estatística de teste com uma distribuição aproximadamente normal padronizada quando  $H_0$  é verdadeira (o resultado é chamado teste de McNemar).
- d. Se  $x_1 = 350, x_2 = 150, x_3 = 200$  e  $x_4 = 300$ , a que conclusão você pode chegar?
55. Dois tipos diferentes de liga, A e B, foram usados para fabricar espécimes experimentais de uma junta de baixa tensão a ser usada em uma determinada aplicação de engenharia. A tensão de ruptura (ksi) de cada espécime foi determinada, e os resultados são resumidos na distribuição de frequência a seguir.

	A	B
26 – < 30	6	4
30 – < 34	12	9
34 – < 38	15	19
38 – < 42	7	10
	$m = 40$	$n = 42$

Calcule um IC de 95% para a diferença entre as proporções reais de todos os espécimes de ligas A e B que possuem uma tensão de ruptura de pelo menos 34 ksi.

56. Usando a fórmula tradicional, um IC de 95% para  $p_1 - p_2$  será construído com base em tamanhos das amostras iguais das duas populações. Para que valor de  $n$  ( $= m$ ) o intervalo resultante terá largura de no máximo 0,1, independentemente dos resultados da amostragem?

## 9.5 Inferências Sobre Variâncias de Duas Populações

Os métodos para comparar as variâncias de duas populações (ou desvios padrão) são ocasionalmente necessários, embora tais problemas surjam com muito menos frequência do que os que compreendem médias ou

proporções. Para o caso em que as populações sob investigação são normais, os procedimentos são feitos com base em uma nova família de distribuições de probabilidades.

## A Distribuição F

A distribuição de probabilidade  $F$  possui dois parâmetros, representados por  $\nu_1$  e  $\nu_2$ . O parâmetro  $\nu_1$  é chamado *número de graus de liberdade do numerador*, e  $\nu_2$  é o número de graus de liberdade do denominador; aqui  $\nu_1$  e  $\nu_2$  são números inteiros positivos. Uma variável aleatória que possui uma distribuição  $F$  não pode admitir um valor negativo. Uma vez que a função densidade é complicada e não será usada explicitamente, omitimos a fórmula. Existe uma ligação importante entre uma variável  $F$  e as variáveis qui-quadradas. Se  $X_1$  e  $X_2$  são variáveis independentes qui-quadradas com  $\nu_1$  e  $\nu_2$  gl, respectivamente, então, a va

$$F = \frac{X_1/\nu_1}{X_2/\nu_2} \quad (9.8)$$

formada à razão das duas variáveis qui-quadradas divididas por seus respectivos graus de liberdade apresenta distribuição  $F$ .

A Figura 9.8 ilustra o gráfico de uma função densidade  $F$  típica. Análogo à notação  $t_{\alpha, \nu}$  e  $\chi^2_{\alpha, \nu}$ , usamos  $F_{\alpha, \nu_1, \nu_2}$  para o ponto no eixo que inclui  $\alpha$  da área sob a curva de densidade  $F$  com  $\nu_1$  e  $\nu_2$  gl na cauda superior. A curva de densidade não é simétrica, então parece que os valores críticos de cauda superior e inferior devem ser tabulados. Entretanto isso não é necessário devido à propriedade a seguir.

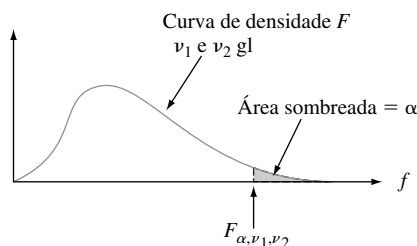


Figura 9.8 Uma curva de densidade  $F$  e um valor crítico

$$F_{1-\alpha, \nu_1, \nu_2} = 1/F_{\alpha, \nu_2, \nu_1} \quad (9.9)$$

A Tabela A.9 do Apêndice fornece  $F_{\alpha, \nu_1, \nu_2}$  para  $\alpha = 0,10, 0,05, 0,01$ , e vários valores de  $\nu_1$  (em colunas diferentes da tabela) e  $\nu_2$  (em grupos diferentes de linhas da tabela). Por exemplo:  $F_{0,05,6,10} = 3,22$  e  $F_{0,05,10,6} = 4,06$ . Para obter  $F_{0,95,6,10}$ , número que inclui 0,95 da área à sua direita (e, assim, 0,05 à esquerda) sob a curva  $F$  com  $\nu_1 = 6$  e  $\nu_2 = 10$ , usamos (9.9):  $F_{0,95,6,10} = 1/F_{0,05,10,6} = 1/4,06 = 0,246$ .

## Métodos Inferenciais

Um procedimento de teste de hipóteses relacionado à razão  $\sigma_1^2/\sigma_2^2$  bem como um IC para essa razão baseiam-se no seguinte resultado:

### TEOREMA

Seja  $X_1, \dots, X_m$  uma amostra aleatória de uma distribuição normal com variância  $\sigma_1^2$ , seja  $Y_1, \dots, Y_n$  outra amostra aleatória (independente dos  $X_i$ s) de uma distribuição normal com variância  $\sigma_2^2$ , e sejam  $S_1^2$  e  $S_2^2$  as variâncias das duas amostras.

$$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \quad (9.10)$$

possui distribuição  $F$  com  $\nu_1 = m - 1$  e  $\nu_2 = n - 1$ .

Esse teorema resulta da combinação de (9.8) com o fato de que as variáveis  $(m - 1)S_1^2/\sigma_1^2$  e  $(n - 1)S_2^2/\sigma_2^2$  têm distribuição qui-quadrada com  $m - 1$  e  $n - 1$  gl, respectivamente (consulte a Seção 7.4). Em virtude de  $F$  utilizar uma razão e não uma diferença, a estatística de teste é a razão de variâncias amostrais; a alegação de que  $\sigma_1^2 = \sigma_2^2$  é então rejeitada, se a razão variar muito de 1.

**O Teste  $F$  para Igualdade de Variâncias**

Hipótese nula:  $H_0: \sigma_1^2 = \sigma_2^2$

Valor da estatística de teste:  $f = s_1^2/s_2^2$

**Hipótese Alternativa**

$H_a: \sigma_1^2 > \sigma_2^2$

$H_a: \sigma_1^2 < \sigma_2^2$

$H_a: \sigma_1^2 \neq \sigma_2^2$

**Região de Rejeição para Teste Nível  $\alpha$**

$f \geq F_{\alpha, m-1, n-1}$

$f \leq F_{1-\alpha, m-1, n-1}$

tanto  $f \geq F_{\alpha/2, m-1, n-1}$  ou  $f \leq F_{1-\alpha/2, m-1, n-1}$

Uma vez que os valores críticos são tabelados somente para  $\alpha = 0,10, 0,05, 0,01$  e  $0,001$ , o teste bicaudal pode ser realizado somente nos níveis  $0,20, 0,10, 0,02$  e  $0,002$ . Tabulações mais extensas dos valores críticos  $F$  estão disponíveis em outras referências.

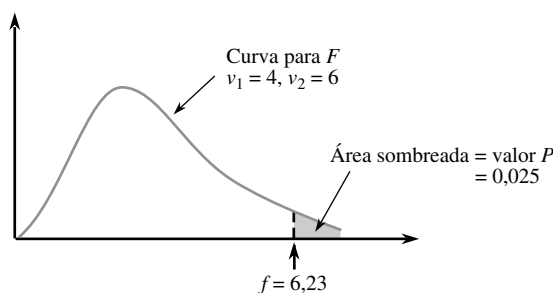
**Exemplo 9.14**

Com base nos dados relatados em um artigo de 1979 no *Journal of Gerontology* ("Serum Ferritin in an Elderly Population", p. 521-524), os autores concluíram que a distribuição de ferritina nos idosos teve uma variância menor que nos adultos jovens. (A ferritina do soro é usada no diagnóstico de deficiência de ferro.) Para uma amostra de 28 homens idosos, o desvio padrão da amostra de ferritina do soro (mg/L) foi  $s_1 = 52,6$ ; para 26 rapazes, o desvio padrão da amostra foi  $s_2 = 84,2$ . Esses dados sustentam a conclusão aplicada aos homens?

Sejam  $\sigma_1^2$  e  $\sigma_2^2$  a variância das distribuições de ferritina do soro dos idosos e dos rapazes, respectivamente. Desejamos testar  $H_0: \sigma_1^2 = \sigma_2^2$  versus  $H_a: \sigma_1^2 < \sigma_2^2$ . No nível  $0,01$ ,  $H_0$  será rejeitada se  $f \leq F_{0,99,27,25}$ . Para obter o valor crítico, precisamos de  $F_{0,01,25,27}$ . A partir da Tabela A.9 do Apêndice,  $F_{0,01,25,27} = 2,54$ , assim,  $F_{0,99,27,25} = 1/2,54 = 0,394$ . O valor calculado de  $F$  é  $(52,6)^2/(84,2)^2 = 0,390$ . Uma vez que  $0,390 \leq 0,394$ ,  $H_0$  é rejeitada no nível  $0,01$  em favor de  $H_a$ , assim, a variabilidade parece ser maior nos rapazes que nos idosos. ■

**Valores  $p$  para Testes  $F$**

Lembre-se de que o valor  $P$  para um teste  $t$  de cauda superior é a área sob a curva  $t$  relevante (com gl apropriado) à direita do  $t$  calculado. Da mesma forma, o valor  $P$  para um teste  $F$  de cauda superior é a área sob a curva  $F$  com gl do numerador e denominador apropriado à direita do  $f$  calculado. A Figura 9.9 ilustra isso para um teste com base em  $\nu_1 = 4$  e  $\nu_2 = 6$ .



**Figura 9.9** Valor  $P$  para teste  $F$  de cauda superior

Infelizmente, a tabulação das áreas de cauda superior da curva  $F$  é muito mais confusa que a das curvas  $t$ , pois os dois gls estão envolvidos. Para cada combinação de  $\nu_1$  e  $\nu_2$ , nossa tabela  $F$  fornece somente os quatro valores críticos que incluem as áreas 0,10, 0,05, 0,01 e 0,001. A Figura 9.10 mostra o que pode ser dito sobre o valor  $P$ , dependendo de onde  $f$  cair com relação aos quatro valores críticos.

Por exemplo: para um teste com  $\nu_1 = 4$  e  $\nu_2 = 6$ ,

$$f = 5,70 \Rightarrow 0,01 < \text{valor } P < 0,05$$

$$f = 2,16 \Rightarrow \text{Valor } P > 0,10$$

$$f = 25,03 \Rightarrow \text{valor } P < 0,001$$

Somente se  $f$  for igual a um valor tabulado, obteremos um valor  $P$  exato (por exemplo: se  $f = 4,53$ , então o valor  $P = 0,05$ ). Uma vez que sabemos que  $0,01 < \text{valor } P < 0,05$ ,  $H_0$  seria rejeitada em um nível de significância de 0,05 mas não em um nível de 0,01. Quando o valor  $P < 0,001$ ,  $H_0$  deveria ser rejeitada em qualquer nível de significância razoável.

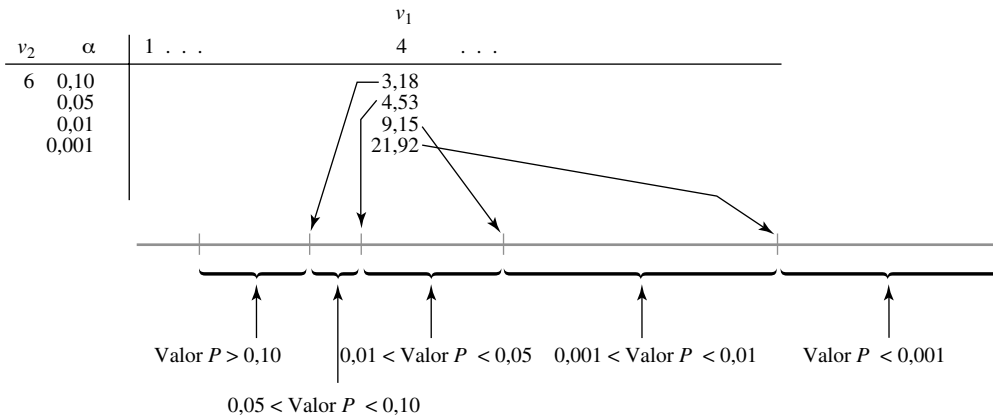


Figure 9.10 Obtenção de informações do valor  $P$  da tabela  $F$  para teste  $F$  de cauda superior

Os testes  $F$  discutidos nos capítulos a seguir serão todos de cauda superior. Se, no entanto, um teste  $F$  de cauda inferior for apropriado, então (9.9) deve ser usado para obter valores críticos de cauda inferior, de modo que possa ser estabelecido um limite (ou limites) no valor  $P$ . No caso de teste bicaudal, o limite ou os limites de um teste monocaudal devem ser multiplicados por 2. Por exemplo: se  $f = 5,82$  quando  $\nu_1 = 4$  e  $\nu_2 = 6$ , então, uma vez que 5,82 está entre os valores críticos 0,05 e 0,01,  $2(0,01) < \text{Valor } P < 2(0,05)$ , fornecendo  $0,02 < \text{Valor } P < 0,10$ .  $H_0$  seria então rejeitada, se  $\alpha = 0,10$ , mas não, se  $\alpha = 0,01$ . Nesse caso, não podemos dizer, a partir de nossa tabela, qual conclusão é apropriada quando  $\alpha = 0,05$  (uma vez que não sabemos se o valor  $P$  é menor ou maior que esse). Entretanto, o *software* estatístico mostra que a área à direita de 5,82 sob essa curva  $F$  é 0,029, assim o valor  $P$  é 0,058 e a hipótese nula não deve, portanto, ser rejeitada no nível 0,05 (0,058 é o menor  $\alpha$  para o qual  $H_0$  pode ser rejeitada e nosso  $\alpha$  escolhido é menor que esse).

### Intervalo de Confiança para $\sigma_1/\sigma_2$

O IC para  $\sigma_1^2/\sigma_2^2$  baseia-se na substituição de  $F$  na expressão de probabilidade

$$P(F_{1-\alpha/2, \nu_1, \nu_2} < F < F_{\alpha/2, \nu_1, \nu_2}) = 1 - \alpha$$

pela variável  $F$  (9.10) e na manipulação das desigualdades para isolar  $\sigma_1^2/\sigma_2^2$ . Um intervalo para  $\sigma_1/\sigma_2$  resulta da raiz quadrada de cada limite. Os detalhes são deixados como exercício.

**Exercícios** | Seção 9.5 (57–64)

57. Obtenha ou calcule as seguintes quantidades:  
**a.**  $F_{0,05,5,8}$     **b.**  $F_{0,05,8,5}$     **c.**  $F_{0,95,5,8}$     **d.**  $F_{0,95,8,5}$   
**e.** O 99° percentil da distribuição  $F$  com  $\nu_1 = 10$ ,  $\nu_2 = 12$   
**f.** O 1° percentil da distribuição  $F$  com  $\nu_1 = 10$ ,  $\nu_2 = 12$   
**g.**  $P(F \leq 6,16)$  para  $\nu_1 = 6$ ,  $\nu_2 = 4$   
**h.**  $P(0,177 \leq F \leq 4,74)$  para  $\nu_1 = 10$ ,  $\nu_2 = 5$
58. Forneça o máximo de informações que puder sobre o valor  $P$  do teste  $F$  em cada uma das situações a seguir:  
**a.**  $\nu_1 = 5$ ,  $\nu_2 = 10$ , teste de cauda superior,  $f = 4,75$   
**b.**  $\nu_1 = 5$ ,  $\nu_2 = 10$ , teste de cauda superior,  $f = 2,00$   
**c.**  $\nu_1 = 5$ ,  $\nu_2 = 10$ , teste bicaudal,  $f = 5,64$   
**d.**  $\nu_1 = 5$ ,  $\nu_2 = 10$ , teste de cauda inferior,  $f = 0,200$   
**e.**  $\nu_1 = 35$ ,  $\nu_2 = 20$ , teste de cauda superior,  $f = 3,24$
59. Retorne aos dados sobre o ângulo de inclinação máximo dado no Exercício 28 deste capítulo. Faça um teste no nível de significância 0,10 para ver se os desvios padrão da população dos grupos de duas idades são diferentes (os gráficos de probabilidade normal confirmam a suposição de normalidade necessária).
60. Consulte o Exemplo 9.7. Os dados sugerem que o desvio padrão da distribuição da tensão dos espécimes fundidos é menor que o dos espécimes não-fundidos? Faça um teste no nível de significância 0,01, obtendo o máximo de informações que puder sobre o valor  $P$ .
61. Toxafeno é um inseticida que foi identificado como um poluente no ecossistema dos Grandes Lagos. Para investigar o efeito da exposição ao toxafeno nos animais, foi dado toxafeno a grupos de ratos em sua dieta. O artigo "Reproduction Study of Toxaphene in the Rat" (*J. of Environ. Sci. Health*, 1988, p. 101-126) relata os ganhos de peso (em gramas) dos ratos que receberam uma dose baixa (4 ppm) e dos ratos de controle cuja dieta não incluiu o inseticida. O desvio padrão da amostra de controle de 23 ratos fêmeas foi de 32 g, e de 20 ratos

fêmeas com dose baixa, de 54 g. Esses dados sugerem que há mais variabilidade nos ganhos de peso de dose baixa que nos ganhos de peso de controle? Assumindo normalidade, faça um teste de hipóteses no nível de significância de 0,05.

62. Em um estudo da deficiência de cobre em gado, os valores de cobre ( $\mu\text{g Cu}/100 \text{ mL}$  de sangue) foram determinados para o pasto do gado em uma área conhecida por ter anomalias bem definidas de molibdênio (valores de metal em excesso do limite normal da variação regional) e para o pasto do gado em uma área normal ("An Investigation into Copper Deficiency in Cattle in the Southern Pennines", *J. Agricultural Soc. Cambridge*, 1972, p. 157-163), resultando em  $s_1 = 21,5$  ( $m = 48$ ) para a condição anômala e  $s_2 = 19,45$  ( $n = 45$ ) para a condição normal. Faça o teste para a igualdade versus desigualdade de variâncias da população no nível de significância de 0,10, usando a abordagem do valor  $P$ .
63. O artigo "Enhancement of Compressive Properties of Failed Concrete Cylinders with Polymer Impregnation" (*J. Testing and Evaluation*, 1977, p. 333-337) relata os dados a seguir sobre o módulo compressivo impregnado ( $\text{psi} \times 10^6$ ) quando dois polímeros diferentes foram usados para reparar rachaduras no concreto com defeito.

<i>Epóxi</i>	1,75	2,12	2,05	1,97
<i>Pré-polímero MMA</i>	1,77	1,59	1,70	1,69

Obtenha um IC de 90% para a razão de variâncias, usando primeiro o método sugerido no texto para obter uma fórmula geral de intervalo de confiança.

64. Reconsidere os dados do Exemplo 9.6 e calcule um limite de confiança superior de 95% para a razão do desvio padrão da distribuição de porosidade de triacetato ao da distribuição de porosidade do algodão.

Exercícios Suplementares (65–93)

65. Os dados resumidos a seguir sobre a resistência à compressão (lb) de caixas de  $12 \times 10 \times 8$  in. foram exibidos no artigo "Compression of Single-Wall Corrugated Shipping Containers Using Fixed and Floating Test Platens" (*J. Testing and Evaluation*, 1992, p. 318-320). Os autores afirmaram que "a diferença entre a resistência à compressão que utiliza o método de prensa com plataforma móvel e fixo foi considerada pequena, se comparada à variação normal da resistência à compressão entre caixas idênticas". Você concorda?

Método	Tamanho Amostral	Média Amostral	DP Amostral
Fixo	10	807	27
Móvel	10	757	41

66. Os autores do artigo "Dynamics of Canopy Structure and Light Interception in *Pinus elliotti*, North Florida" (*Ecological Monographs*, 1991, p. 33-51) planejaram um experimento para determinar o efeito de

fertilizantes em uma medida de área plantada. Diversos gráficos estavam disponíveis para o estudo e metade foi selecionada aleatoriamente para ser fertilizada. Para garantir que os quadros que receberiam o fertilizante e os quadros de controle fossem semelhantes, antes de começar o experimento, a densidade de árvores (o número de árvores por hectare) foi registrada para oito quadros a serem fertilizados e para oito quadros de controle, resultando nos dados fornecidos. A saída do MINITAB é exposta a seguir.

<i>Quadro de fertilizantes</i>	1024	1216	1312	1280
	1216	1312	992	1120

<i>Quadro de controle</i>	1104	1072	1088	1328
	1376	1280	1120	1200

Duas amostras T para fertilizantes X controle

	N	Média	Desv Pad	Média SE
fertilizante	8	1184	126	44
controle	8	1196	118	42

IC de 95% para  $\mu$  fertilizado -  $\mu$  de controle: (-144, 120)

- Construa um *boxplot* comparativo e comente sobre quaisquer características interessantes.
  - Você concluiria que há uma diferença significativa na densidade média de árvores dos quadros fertilizados e de controle? Use  $\alpha = 0,05$ .
  - Interprete o intervalo de confiança dado.
67. O índice de respostas a questionários é afetado pela inclusão de algum tipo de incentivo para responder junto ao questionário? Em um experimento, 110 questionários sem nenhum incentivo resultaram em 75 devolvidos, enquanto 98 questionários que incluíam uma chance de ganhar na loteria produziram 66 respostas ("Charities, No; Lotteries, No; Cash, Yes", *Public Opinion Quarterly*, 1996, p. 542-562). Esses dados sugerem que a inclusão de um estímulo aumenta a probabilidade de uma resposta? Determine e teste as hipóteses relevantes no nível de significância 0,10, usando o método do valor  $P$ .
68. Os dados a seguir foram obtidos em um estudo para avaliar o potencial de liquefação em uma estação de energia nuclear proposta ("Cyclic Strengths Compared for Two Sampling Techniques", *J. Geotechnical Division, Am. Soc. Civil Engrs. Proceedings*, 1981, p. 563-576). Antes do teste de resistência cíclico, foram obtidas amostras de solo, usando-se um método de tubo lançador e um método de bloco, resultando nos seguintes valores observados de densidade seca (lb/pés<sup>3</sup>):

<i>Amostragem do lançador</i>	101,1	111,1	107,6	98,1
	99,5	98,7	103,3	108,9
	109,1	104,1	110,0	98,4
	105,1	104,5	105,7	103,3
	100,3	102,6	101,7	105,4
	99,6	103,3	102,1	104,3

<i>Amostragem do bloco</i>	107,1	105,0	98,0	97,9
	103,3	104,6	100,1	98,2
	97,9	103,2	96,9	

Calcule e interprete um IC de 95% para a diferença entre as densidades secas médias reais para os dois métodos de amostragem.

69. O artigo "Quantitative MRI and Electrophysiology of Preoperative Carpal Tunnel Syndrome in a Female Population" (*Ergonomics*, 1997, p. 642-649) relatou que (-473,3, 1691,9) era um intervalo de confiança de 95% para amostra grande, para a diferença entre o volume muscular médio real do ténar (mm<sup>3</sup>) das pessoas que sofrem da síndrome do túnel carpal e o volume médio real das que não sofrem. Calcule um intervalo de confiança de 90% para essa diferença.
70. Os dados resumidos a seguir sobre a resistência a dobras (lb-in/in) das juntas foram extraídos do artigo "Bending Strength of Corner Joints Constructed with Injection Molded Splines" (*Forest Products J.*, Abril, 1997, p. 89-92).

Tipo	Tamanho Amostral	Média Amostral	DP Amostral
Sem revestimento lateral	10	80,95	9,59
Com revestimento lateral	10	63,23	5,96

- Calcule um limite de confiança inferior a 95% para a resistência média real das juntas com revestimento lateral.
  - Calcule um limite de previsão inferior de 95% para a resistência de uma única junta com revestimento lateral.
  - Calcule um intervalo que, com confiança de 95%, inclua os valores de resistência para pelo menos 95% da população de todas as juntas com revestimentos laterais.
  - Calcule um intervalo de confiança de 95% para a diferença entre as resistências médias reais dos dois tipos de juntas.
71. Foi realizado um experimento para comparar várias propriedades de fio torcido de algodão/poliéster acabado com *softener* somente, e fio acabado com *softener* mais 5% de resina DP ("Properties of a Fabric Made with Tandem Spun Yarns," *Textile Res. J.*, 1996, p. 607-611). Uma característica particularmente importante do tecido é sua durabilidade, isto é, sua habilidade de resistir ao desgaste. Para uma amostra de 40 espécimes somente com *softener*, a média amostral da resistência à abrasão de *stoll-flex* (ciclos) na direção do enchimento do fio era 3975,0, com um desvio padrão amostral de 245,1. Outra amostra de 40 espécimes com *softener* e resina forneceu uma média amostral e um desvio padrão amostral de 2795,0 e 293,7, respectivamente. Calcule um intervalo de confiança com nível de confiança de 99% para a diferença entre as médias reais das resistências à abrasão dos dois tipos de tecidos. Seu intervalo fornece evidên-

cia convincente de que as resistências médias reais são diferentes para os dois tipos de tecidos? Por quê?

72. O descarrilamento de um trem de carga devido à falha catastrófica de um mancal da armadura do motor de tração estimulou um estudo relatado no artigo "Locomotive Traction Motor Armature Bearing Life Study" (*Lubrication Engr.*, ago. 1997, p. 12-19). Uma amostra de 17 motores de tração de alta milhagem foi selecionada e a quantidade de penetração do cone (mm/10) foi determinada para o mancal de pinhão e para o mancal de armadura do comutador, resultando nos dados a seguir:

<i>Motor</i>	1	2	3	4	5	6
<i>Comutador</i>	211	273	305	258	270	209
<i>Pinhão</i>	226	278	259	244	273	236
<i>Motor</i>	7	8	9	10	11	12
<i>Comutador</i>	223	288	296	233	262	291
<i>Pinhão</i>	290	287	315	242	288	242
<i>Motor</i>	13	14	15	16	17	
<i>Comutador</i>	278	275	210	272	264	
<i>Pinhão</i>	278	208	281	274	268	

Calcule uma estimativa da diferença de médias populacionais entre a penetração do mancal de armadura do comutador e a penetração do mancal de pinhão. Faça a análise de forma que transmita informações sobre confiabilidade e precisão da estimativa. (Nota: um gráfico de probabilidade normal valida a suposição de normalidade necessária.) Você diria que a diferença entre as médias populacionais foi estimada com precisão? Parece que a penetração média populacional difere para os dois tipos de mancais? Explique.

73. *Headability* é a habilidade de uma peça cilíndrica de material ser moldada na cabeça de um pino, parafuso ou outro item fabricado a frio sem rachar. O artigo "New Methods for Assessing Cold Heading Quality" (*Wire J. Intl.*, out. 1996, p. 66-72) descreveu o resultado de um teste de impacto de *headability* aplicado a 30 espécimes de aço de alumínio neutralizado e a 30 espécimes de aço de silício neutralizado. A média amostral do número do grau de *headability* dos espécimes de aço foi 6,43, e a média amostral dos espécimes de alumínio foi 7,09. Suponha que os desvios padrão amostrais foram 1,08 e 1,19, respectivamente. Você concorda com os autores do artigo de que a diferença nos graus de *headability* é significativa no nível de 5% (assumindo que as duas distribuições de *headability* sejam normais)?
74. O artigo "Two Parameters Limiting the Sensitivity of Laboratory Tests of Condoms as Viral Barriers" (*J. of Testing and Eval.*, 1996, p. 279-286) relatou que, nos preservativos de marca A, entre 16 rasgos produzidos pela perfuração de uma agulha, o comprimento do rasgo médio da amostra foi 74,0  $\mu\text{m}$ , enquanto, para os 14 rasgos da marca B, o comprimento médio amostral foi 61,0  $\mu\text{m}$  (determinado pelo uso de microscópio de luz e micrógrafo de varredura eletrônica). Suponha que

os desvios padrão amostrais sejam 14,8 e 12,5, respectivamente (consistentes com os intervalos amostrais dados no artigo). Os autores comentaram que o preservativo de marca B, mais espesso, apresentou um comprimento médio do rasgo menor que o preservativo de marca A, mais fino. Essa diferença é, de fato, estatisticamente significativa? Determine as hipóteses apropriadas e teste em  $\alpha = 0,05$ .

75. As informações sobre a postura e a força da mão geradas pelos dedos durante a manipulação diária de vários objetos são necessárias para projetar dispositivos protéticos de alta tecnologia para mão. O artigo "Grip Posture and Forces During Holding Cylindrical Objects with Circular Grips" (*Ergonomics*, 1996, p. 1163-1176) relatou que, para uma amostra de 11 mulheres, a força média amostral à contração de quatro dedos (N) foi 98,1, e o desvio padrão amostral, 14,2. Para uma amostra de 15 homens, a média e o desvio padrão amostrais foram 129,2 e 39,1, respectivamente.
- Um teste realizado para ver se as resistências médias reais dos dois sexos eram diferentes resultou em  $t = 2,51$  e valor  $P = 0,019$ . O procedimento de teste apropriado descrito neste capítulo produz esse valor de  $t$  e o valor  $P$  determinado?
  - Existe evidência substancial de que a resistência média real dos homens excede a das mulheres por mais de 25 N? Expresse e teste as hipóteses relevantes.

76. O artigo "Pine Needles as Sensors of Atmospheric Pollution" (*Environ. Monitoring*, 1982, p. 273-286) relatou o uso da análise da atividade do nêutron para determinar a concentração de poluentes nas agulhas de pinheiro. De acordo com os autores do artigo, "Essas observações indicaram fortemente que, para esses elementos determinados pelos procedimentos analíticos, a distribuição da concentração é lognormal. Portanto, nos testes de significância, os logaritmos das concentrações serão usados". Os dados fornecidos referem-se à concentração de bromina nas agulhas tiradas de um local próximo a uma usina de vapor a óleo queimado e de um local relativamente limpo. Os valores resumidos são médias e desvios padrão das observações logtransformadas.

Local	Tamanho Amostral	Log da Concentração Média	DP do Log da Concentração
Usina de vapor	8	18,0	4,9
Limpo	9	11,0	4,6

Seja  $\mu_1^*$  a média real do log da concentração no primeiro local, e defina  $\mu_2^*$  analogamente ao segundo local.

- Use o teste  $t$  combinado (assumindo normalidade e desvios padrão iguais) para decidir, no nível de significância 0,05, se as duas médias da distribuição da concentração são iguais.



b. Se  $\sigma_1^*$  e  $\sigma_2^*$ , os desvios padrão das duas distribuições do log da concentração, não são iguais,  $\mu_1$  e  $\mu_2$ , as médias das distribuições da concentração, seriam as mesmas se  $\mu_1^* = \mu_2^*$ ? Explique o seu raciocínio.

77. Exposição a longo prazo de trabalhadores têxteis ao pó de algodão liberado durante o processamento pode resultar em problemas graves de saúde. Assim, os pesquisadores de fibra têxtil investigaram métodos que resultarão em riscos reduzidos enquanto preservam as propriedades importantes do tecido. Os dados a seguir sobre a tensão de coesão da fição ( $kN \cdot m/kg$ ) dos espécimes produzidos em cinco rotações múltiplas diferentes originam-se do artigo "Heat Treatment of Cotton: Effect on Endotoxin Content, Fiber and Yarn Properties and Processability" (*Textile Research J.*, 1996, p. 727-738).

Rotações múltiplas	1,054	1,141	1,245	1,370	1,481
Tensão de controle	0,45	0,60	0,61	0,73	0,69
Tensão aquecida	0,51	0,59	0,63	0,73	0,74

Os autores do artigo referido afirmaram que a resistência dos espécimes aquecidos pareceu, em média, ser levemente mais elevada que dos espécimes de controle. A diferença é estatisticamente significativa? Determine e teste as hipóteses relevantes, usando  $\alpha = 0,05$  para calcular o valor  $P$ .

78. Os dados resumidos a seguir sobre a razão da resistência à área transversal dos extensores de joelho são tirados do artigo "Knee Extensor and Knee Flexor Strength: Cross-Sectional Area Ratios in Young and Elderly Men" (*J. of Gerontology*, 1992, p. M204-M210).

Grupo	Tamanho Amostral	Média Amostral	Erro Amostral
Jovens	13	7,47	0,22
Idosos	12	6,71	0,28

Esses dados sugerem que a razão média real dos jovens excede a dos idosos? Faça um teste de hipóteses apropriadas, usando  $\alpha = 0,05$ . Certifique-se de determinar quaisquer hipóteses necessárias para sua análise.

79. Os dados a seguir sobre o tempo de resposta foram mostrados no artigo "The Extinguishment of Fires Using Low-Flow Water Hose Streams – Part II" (*Fire Technology*, 1991, p. 291-320).

Boa visibilidade	0,43	1,17	0,37	0,47	0,68	0,58	0,50	2,75
Visibilidade ruim	1,47	,80	1,58	1,53	4,33	4,23	3,25	3,22

Os autores analisaram os dados com o teste  $t$  combinado. O uso desse teste parece justificado? (*Sugestão*: verifique a normalidade. As pontuações normais de  $n = 8$  são  $-1,53, -0,89, -0,49, -0,15, 0,15, 0,49, 0,89, e 1,53$ .)

80. O artigo "The Relationship Between Distress and Delight in Males' and Females' Reactions to Frightening Films" (*Human Communication Research*, 1991, p. 625-637) relata que os investigadores mediram as reações emocionais de 50 homens e 60 mulheres, depois de assistirem a um trecho de um filme de terror. O artigo incluiu a seguinte declaração: "As mulheres eram muito mais propensas a se mostrarem aflitas do que os homens. Embora os homens tenham expressado níveis mais altos de prazer do que as mulheres, a diferença não foi estatisticamente significativa." Os dados resumidos a seguir também estavam contidos no artigo.

Sexo	Índice de Afição		Índice de Prazer	
	Média	DP	Média	DP
Homens	31,2	10,0	12,02	3,65
Mulheres	40,4	9,1	9,09	5,55
	Valor $P < 0,001$		Não-significativo (Valor $P > 0,05$ )	

Faça um breve comentário sobre as hipóteses relevantes e conclusões.

81. Em um experimento para comparar as resistências do suporte de pinos inseridos em dois tipos diferentes de suportes, uma amostra de 14 observações sobre o limite de tensão dos suportes de carvalho vermelho resultou em uma média amostral e um desvio padrão amostrais de 8,48 MPa e 0,79 MPa, respectivamente, enquanto uma amostra de 12 observações, quando os suportes de pinheiro de Douglas foram usados, forneceu uma média de 9,36 e um desvio padrão de 1,52 ("Bearing Strength of White Oak Pegs in Red Oak and Douglas Fir Timbers". *J. of Testing and Evaluation*, 1998, p. 109-114). Considere o teste para ver se os limites de tensão médios reais são idênticos ou não para os dois tipos de suportes. Compare os gls e os valores  $P$  dos testes  $t$  combinados e não-combinados.

82. Como a entrada de energia pode ser comparada ao seu consumo? Um aspecto dessa questão foi considerado no artigo "Measurement of Total Energy Expenditure by the Doubly Labelled Water Method in Professional Soccer Players" (*J. of Sports Sciences*, 2002, p. 391-397), que continha os dados seguintes (MJ/dia).

Jogador	1	2	3	4	5	6	7
Consumo	14,4	12,1	14,3	14,2	15,2	15,5	17,8
Entrada	14,6	9,2	11,8	11,6	12,7	15,0	16,3

Faça o teste para ver se há uma diferença significativa entre a entrada e o consumo. A conclusão depende do uso de um nível de significância de 0,05, 0,01 ou 0,001?

83. Um investigador deseja obter um IC para a diferença entre a tensão de ruptura média real dos cabos fabricados pela empresa I e pela empresa II. Suponha que a tensão de ruptura seja normalmente distribuída para ambos os tipos de cabo com  $\sigma_1 = 30$  psi e  $\sigma_2 = 20$  psi.

- a. Se os custos determinarem que o tamanho da amostra para o cabo tipo I deve ser três vezes o tamanho da amostra para o cabo tipo II, quantas observações são necessárias se o IC de 99% não for maior que 20 psi?
- b. Suponha que será feito um total de 400 observações. Quantas observações devem ser feitas sobre as amostras do cabo tipo I se a amplitude do intervalo resultante for mínima?

84. Um experimento para determinar os efeitos da temperatura na sobrevivência de ovos de inseto foi descrito no artigo "Development Rates and a Temperature-Dependent Model of Pales Weevil" (*Environ. Entomology*, 1987, p. 956-962). A 11 °C, 73 de 91 ovos sobreviveram para o próximo estágio do desenvolvimento. A 30°C, 102 de 110 ovos sobreviveram. Os resultados desse experimento sugerem que a taxa de sobrevivência (proporção que sobrevive) é diferente para as duas temperaturas? Calcule o valor *P* e use-o para testar as hipóteses apropriadas.

85. A capacidade de ligação da insulina (pmol/mg proteína) foi medida para uma amostra de ratos diabéticos tratados com uma dose baixa de insulina, e para outra amostra de ratos diabéticos tratados com uma dose alta, produzindo os dados a seguir:

Dose baixa  $m = 8, \bar{x} = 1,98, s_1 = 0,51$

Dose Alta  $n = 12, \bar{y} = 1,30, s_2 = 0,35$

(*J. Clinical Investigation*, 1978, p. 552–560.)

- a. Os dados indicam que há qualquer diferença na capacidade de ligação da insulina média real devida ao nível de dosagem? Use um teste com  $\alpha = 0,001$ .
  - b. O que se pode dizer sobre o valor *P*?
86. No artigo que serve de referência no Exercício 85, havia realmente quatro tratamentos experimentais: não-diabético de controle, diabético não-tratado, diabético de dose baixa e diabético de dose alta. Represente o tamanho da amostra do *i*-ésimo tratamento por  $n_i$  e a variância amostral por  $S_i^2$  ( $i = 1, 2, 3, 4$ ). Assumindo que a variância real de cada tratamento é  $\sigma^2$ , construa um estimador combinado de  $\sigma^2$  que seja não-tendencioso e verifique o uso das regras do valor esperado que é, de fato, não-tendencioso. Qual é sua estimativa para os dados reais a seguir? (*Sugestão*: modifique o estimador combinado na pág. 328.)

	Tratamento			
	1	2	3	4
Tamanho da amostra	16	18	8	12
DP amostral	0,64	0,81	0,51	0,35

87. Suponha que um teste nível 0,05 de  $H_0: \mu_1 - \mu_2 = 0$  versus  $H_a: \mu_1 - \mu_2 > 0$  será realizado, assumindo  $\sigma_1 = \sigma_2 = 10$  e normalidade de ambas as distribuições, usando tamanhos das amostras iguais ( $m = n$ ). Avalie a probabilidade de erro tipo II quando  $\mu_1 - \mu_2 = 1$  e  $n =$

25, 100, 2500, e 10.000. Você consegue imaginar problemas reais em que a diferença  $\mu_1 - \mu_2 = 1$  tenha pouca significância prática? Tamanhos das amostras de  $n = 10.000$  seriam desejáveis em tais problemas?

88. Os dados a seguir referem-se à contagem de bactérias transportadas pelo ar (número de colônias/pés<sup>3</sup>) para  $m = 8$  quartos de hospital atapetados e para  $n = 8$  quartos sem tapete ("Microbial Air Sampling in a Carpeted Hospital", *J. Environmental Health*, 1968, p. 405). Parece haver diferença na contagem média real de bactérias entre os quartos atapetados e os não-atapetados?

Atapetado	11,8	8,2	7,1	13,0	10,8	10,1	14,6	14,0
Não-atapetado	12,1	8,3	3,8	7,2	12,0	11,1	10,1	13,7

Suponha que você soube depois que todos os quartos atapetados estavam em um hospital de veteranos de guerra, enquanto todos os quartos não-atapetados estavam em um hospital infantil. Você seria capaz de avaliar o efeito da existência ou não do tapete? Comente.

89. Pesquisadores enviaram 5000 currículos em resposta a anúncios de emprego que apareceram no Boston Globe e no Chicago Tribune. Os currículos eram idênticos, exceto que 2500 deles tinham o primeiro nome "que soava ser de branco", como Brett e Emily, enquanto os outros 2500 tinham nomes "que soavam ser de negros", como Tamika e Rasheed. Dos currículos do primeiro tipo houve 250 respostas, e dos currículos do segundo tipo, somente 167 (esses números são muito consistentes com as informações que apareceram em um relatório de 15 de janeiro de 2003 feito pela Associated Press). Esses dados sugerem fortemente que um currículo com um nome "negro" seja menos provável de resultar em resposta que um currículo com um nome "branco"?

90. O teste de McNemar, desenvolvido no Exercício 54, também pode ser usado quando os indivíduos são pareados (unidos) para produzirem  $n$  pares e então a um membro de cada par é dado o tratamento 1, e a outro, o tratamento 2. Então,  $X_1$  é o número de pares em que ambos os tratamentos foram bem-sucedidos e, de maneira semelhante, para  $X_2, X_3,$  e  $X_4$ . A estatística de teste para testar a eficácia idêntica dos dois tratamentos é dada por  $(X_2 - X_3)/\sqrt{(X_2 + X_3)}$ , que possui uma distribuição aproximadamente normal padronizada quando  $H_0$  é verdadeira. Use isso para testar se a droga ergotamina é eficaz no tratamento de enxaquecas.

		Ergotamina	
		S	F
Placebo	S	44	34
	F	46	30

Os dados são fictícios, mas a conclusão está de acordo com a do artigo "Controlled Clinical Trial of Ergotamine Tartrate" (*British Med. J.*, 1970, p. 325-327).

91. O artigo "Evaluating Variability in Filling Operations" (*Food Tech.*, 1984, p. 51-55) descreve duas operações diferentes de enchimento usadas em uma instalação de empacotamento de carne moída. Ambas as operações foram definidas para encher pacotes com 1400 g de carne moída. Em uma amostra aleatória de tamanho 30, tirada de cada operação de enchimento, as médias e desvios padrão resultantes foram 1402,24 g e 10,97 g para a operação 1, e 1419,63 g e 9,96 g para a operação 2.
- Usando um nível de significância de 0,05, há evidência suficiente de que o peso médio real dos pacotes é diferente para as duas operações?
  - Os dados da operação 1 sugerem que o peso médio real dos pacotes produzidos pela operação 1 é superior a 1400 g? Use um nível de significância de 0,05.
92. Seja  $X_1, \dots, X_m$ , uma amostra aleatória de uma distribuição de Poisson com parâmetro  $\lambda_1$  e por  $Y_1, \dots, Y_n$  uma amostra aleatória de outra distribuição de Poisson com parâmetro  $\lambda_2$ . Desejamos testar  $H_0: \lambda_1 - \lambda_2 = 0$  contra uma das três alternativas-padrão. Uma vez que  $\mu = \lambda$  para uma distribuição de Poisson, quando  $m$  e  $n$  são grandes, pode ser usado o teste  $z$  de amostra grande

da Seção 9.1. Entretanto, o fato de  $V(\bar{X}) = \lambda/n$  sugere que um denominador diferente deve ser usado na padronização de  $\bar{X} - \bar{Y}$ . Desenvolva um procedimento de teste de amostra grande apropriado para este problema e aplique-o aos dados a seguir para testar se as densidades de mudas de uma determinada espécie são iguais em duas regiões diferentes (onde cada observação é o número de mudas encontradas em um quadrante de amostragem quadrada localizado aleatoriamente com área de 1 m<sup>2</sup>; assim, para a região 1, existem 40 quadrantes em que uma muda foi observada etc.):

	Frequência								
	0	1	2	3	4	5	6	7	
Região 1	28	40	28	17	8	2	1	1	$m = 125$
Região 2	14	25	30	18	49	2	1	1	$n = 140$

93. Consultando o Exercício 92, desenvolva uma fórmula de intervalo de confiança de amostra grande para  $\lambda_1 - \lambda_2$ . Calcule o intervalo para os dados + lá fornecidos usando nível de confiança de 95%.

---

## Bibliografia

Consulte a bibliografia no final do Capítulo 7.

# A Análise de Variância

## Introdução

Ao estudar métodos para a análise de dados quantitativos, primeiramente nos concentramos nos problemas que incluem uma única amostra de números e, em seguida, efetuamos uma análise comparativa de duas dessas diferentes amostras. Nos problemas com uma única amostra, os dados consistiam em observações sobre respostas de indivíduos ou objetos experimentais selecionados aleatoriamente de uma única população. Nos problemas com duas amostras, ambas foram extraídas de duas populações diferentes e os parâmetros de interesse foram as médias populacionais. Ou então, dois tratamentos diferentes foram aplicados a unidades experimentais (indivíduos ou objetos) selecionadas de uma única população. Nesse último caso, os parâmetros de interesse são denominados médias do tratamento real.

A **análise de variância**, ou mais resumidamente **ANOVA**, refere-se de maneira ampla a um conjunto de situações experimentais e procedimentos estatísticos para a análise de respostas quantitativas de unidades experimentais. O problema ANOVA mais simples é conhecido por vários nomes, como **fator único**, **classificação única** ou **ANOVA unidirecional**, e envolve a análise de dados obtidos de amostras de mais de duas populações numéricas (distribuições) ou dados de experimentos em que foram empregados mais de dois tratamentos. A característica que diferencia os tratamentos ou populações entre si é chamada **fator** em estudo, e os tratamentos ou populações diferentes são chamados **níveis** do fator. Exemplos de situações desse tipo incluem:

1. Um experimento para estudar os efeitos de cinco diferentes marcas de gasolina na eficiência operacional (milhas/galão) de motores de automóvel.
2. Um experimento para estudar os efeitos da presença de quatro diferentes soluções de açúcar (glicose, sacarose, frutose e uma combinação das três) no crescimento de bactérias.
3. Um experimento para investigar se a concentração de madeira de lei em polpa; (%) tem efeito sobre a resistência à tração dos sacos fabricados da polpa.

4. Um experimento para determinar se a densidade da cor de amostras de tecido depende da quantidade de tintura usada.

Em (1), o fator de interesse é a marca da gasolina e há cinco níveis diferentes. Em (2), o fator é o açúcar, com quatro níveis (ou cinco, se for usada uma solução de controle sem açúcar). Em ambos, (1) e (2), o fator é de natureza qualitativa e os níveis correspondem a possíveis categorias do fator. Em (3) e (4), os fatores são concentração de madeira de lei e quantidade de tintura, respectivamente; como ambos os fatores são de natureza quantitativa, os níveis encontram diferentes aplicações. Quando o fator de interesse é quantitativo, técnicas estatísticas de análise de regressão (discutidas nos capítulos 12 e 13) também podem ser usadas para analisar os dados.

Este capítulo concentra-se na ANOVA de fator único. A Seção 10.1 apresenta o teste  $F$  para verificar a hipótese nula de que as médias populacionais ou de tratamento são idênticas. A Seção 10.2 leva em conta uma análise adicional dos dados quando  $H_0$  for rejeitado. A Seção 10.3 abrange alguns outros aspectos de ANOVA de fator único. O Capítulo 11 apresenta experimentos ANOVA com mais de um fator único.

## 10.1 | ANOVA de fator único

A ANOVA de fator único concentra-se na comparação de mais de duas médias populacionais ou de tratamento. Seja

$I$  = o número de populações ou tratamentos que estão sendo comparados

$\mu_1$  = a média populacional 1 ou a resposta média real quando se aplica o tratamento 1

⋮

$\mu_I$  = a média populacional  $I$  ou a resposta média real quando se aplica o tratamento  $I$

Então as hipóteses de interesse são

$$H_0: \mu_1 = \mu_2 = \cdots = \mu_I$$

versus

$H_a$ : pelo menos dois  $\mu_i$ s são diferentes

Se  $I = 4$ ,  $H_0$  será verdadeiro somente se os quatro  $\mu_i$ s forem idênticos.  $H_a$  seria verdadeiro, por exemplo, se  $\mu_1 = \mu_2 \neq \mu_3 = \mu_4$ , se  $\mu_1 = \mu_3 = \mu_4 \neq \mu_2$ , ou se os quatro  $\mu_i$ s fossem diferentes entre si.

Para testar essas hipóteses, devemos ter em mãos uma amostra aleatória de cada população ou do tratamento.

### Exemplo 10.1

O artigo “Compression of Single-Wall Corrugated Shipping Containers Using Fixed and Floating Test Platens” (*J. Testing and Evaluation*, 1992, p. 318-320) relata um experimento em que se comparou a resistência à compressão (lb) de vários tipos diferentes de caixas. A Tabela 10.1 apresenta os resultados de um experimento ANOVA de fator único, envolvendo  $I = 4$  tipos de caixas (as médias amostrais e os desvios padrão estão coerentes com os valores especificados no artigo).

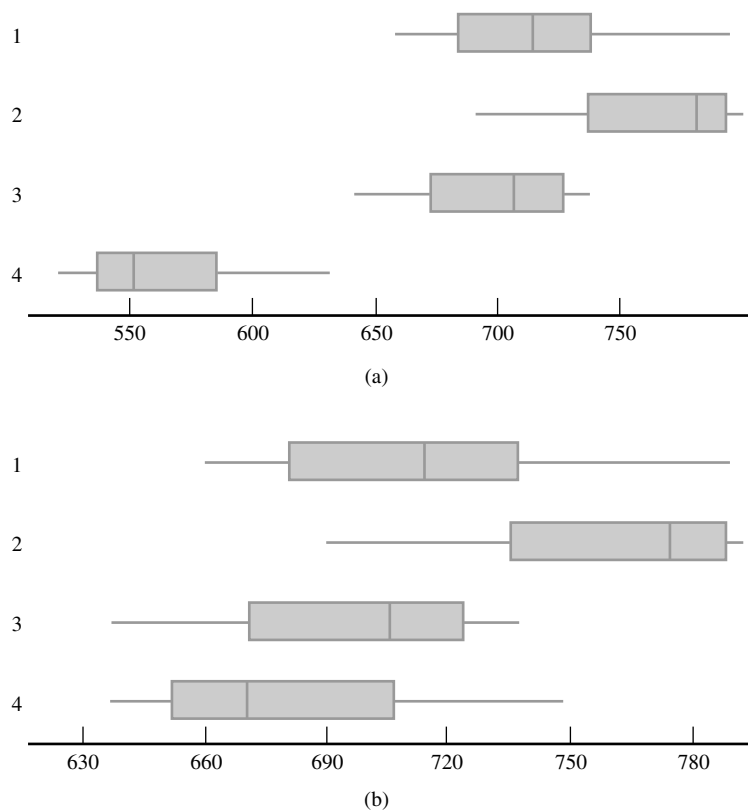
**Tabela 10.1** Dados e quantidades resumidas do Exemplo 10.1

Tipo de Caixa	Resistência de Compressão (lb)	Médias Amostrais	DP Amostral
1	655,5 788,3 734,3 721,4 679,1 699,4	713,00	46,55
2	789,2 772,5 786,9 686,1 732,1 774,8	756,93	40,34
3	737,1 639,0 696,3 671,7 717,2 727,1	698,07	37,20
4	535,1 628,7 542,4 559,0 586,9 520,0	562,02	39,87
	Média geral =	682,50	

Onde  $\mu_i$  denota a resistência de compressão média real das caixas do tipo  $i$  ( $i = 1, 2, 3, 4$ ), a hipótese nula é  $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ . A Figura 10.1(a) mostra um *boxplot* comparativo das quatro amostras. Nos três primeiros tipos de caixa há uma quantidade substancial de superposição entre as observações, mas as resistências de compressão do quarto tipo parecem, consideravelmente, menores que as dos outros tipos. O que sugere que  $H_0$  não é verdadeiro. O *boxplot* comparativo na Figura 10.1(b) tem por base adicionar 120 a cada observação na quarta amostra (o que leva a média a 682,02 e mantém o mesmo desvio padrão), mantendo as outras observações inalteradas. Não está mais claro se  $H_0$  é verdadeiro ou falso. Em situações desse tipo, precisamos de um procedimento formal de teste. ■

### Notação e Suposições

Nos problemas de duas amostras usamos, as letras  $X$  e  $Y$  para diferenciar as observações entre as amostras. Como isso pode ser confuso, no caso de três ou mais amostras, costuma-se usar uma única letra com dois índices.

**Figura 10.1** *Boxplots* do Exemplo 10.1: (a) dados originais; (b) dados alterados

O primeiro índice identifica o número amostral, que corresponde à população ou ao tratamento que faz parte da amostra, e o segundo indica a posição da observação dentro daquela amostra. Sejam

$X_{i,j}$  = a variável aleatória (va) que representa a  $j$ -ésima medida obtida da  $i$ -ésima população ou a medida obtida na  $j$ -ésima unidade experimental que recebe o  $i$ -ésimo tratamento.

$x_{i,j}$  = o valor observado de  $X_{i,j}$  no momento em que se realiza o experimento.

Os dados observados normalmente são exibidos em uma tabela retangular, conforme a Tabela 10.1. Nessa tabela, as amostras das diferentes populações são mostradas em diferentes linhas, e  $x_{i,j}$  é o  $j$ -ésimo número na  $i$ -ésima linha. Por exemplo:  $x_{2,3} = 786,9$  (a terceira observação da segunda população) e  $x_{4,1} = 535,1$ . Quando não houver nenhuma ambigüidade, escreveremos  $x_{ij}$ , em vez de  $x_{i,j}$  (por exemplo: se houvesse 15 observações em cada um dos 12 tratamentos,  $x_{112}$  poderia significar  $x_{1,12}$  ou  $x_{11,2}$ ). Supõe-se que os  $X_{ij}$ s em qualquer amostra em particular sejam independentes – uma amostra aleatória da distribuição da  $i$ -ésima população ou tratamento – e que diferentes amostras sejam independentes entre si.

Em alguns experimentos, amostras diferentes contêm números diferentes de observações. Entretanto os conceitos e métodos ANOVA de fator único são mais facilmente desenvolvidos no caso de amostras de tamanhos iguais. Limitando-nos por enquanto a esse caso, consideremos  $J$  o número de observações em cada amostra ( $J = 6$  no Exemplo 10.1). O conjunto de dados consiste em  $IJ$  observações. As médias amostrais individuais serão representadas por  $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_I$ . Isto é,

$$\bar{X}_i = \frac{\sum_{j=1}^J X_{ij}}{J} \quad i = 1, 2, \dots, I$$

O ponto colocado no lugar do segundo índice significa que adicionamos todos os valores desse índice, mantendo fixo o valor do outro índice, e a barra horizontal indica uma divisão por  $J$  para obter a média. Da mesma forma, a média de todas as  $IJ$  observações, denominada **média geral**, é

$$\bar{X}_{..} = \frac{\sum_{i=1}^I \sum_{j=1}^J X_{ij}}{IJ}$$

Em relação aos dados de resistência na Tabela 10.1,  $\bar{x}_1 = 713,00$ ,  $\bar{x}_2 = 756,93$ ,  $\bar{x}_3 = 698,07$ ,  $\bar{x}_4 = 562,02$ , e  $\bar{x}_{..} = 682,50$ . Além disso, sejam  $S_1^2, S_2^2, \dots, S_I^2$  as variâncias amostrais:

$$S_i^2 = \frac{\sum_{j=1}^J (X_{ij} - \bar{X}_i)^2}{J - 1} \quad i = 1, 2, \dots, I$$

Pelo Exemplo 10.1,  $s_1 = 46,55$ ,  $s_1^2 = 2166,90$ , e assim por diante.

#### SUPOSIÇÕES

As  $I$  distribuições populacionais ou de tratamento são todas normais e têm a mesma variância  $\sigma^2$ . Isto é, cada  $X_{ij}$  é normalmente distribuído com

$$E(X_{ij}) = \mu_i \quad V(X_{ij}) = \sigma^2$$

Os  $I$  desvios padrão da amostra geralmente serão diferentes, mesmo quando os  $\sigma$ s correspondentes forem idênticos. No Exemplo 10.1, o maior entre  $s_1, s_2, s_3$  e  $s_4$  é em torno de 1,25 vez o menor. Uma regra prática simples é que, se o maior  $s$  não for maior do que duas vezes o menor, é razoável supor  $\sigma^2$  iguais.

Em capítulos anteriores, um gráfico de probabilidade normal foi sugerido para verificação da normalidade. Os tamanhos da amostra individuais na ANOVA normalmente são muito pequenos para que  $I$  em gráficos sepa-

rados sejam informativos. Pode-se criar um único gráfico subtraindo  $\bar{x}_1$  de cada observação na primeira amostra,  $\bar{x}_2$  de cada observação na segunda, e assim por diante, e em seguida dispor no gráfico esses  $IJ$  desvios em comparação com os percentis  $z$ . A Figura 10.2 apresenta esse gráfico usando os dados do Exemplo 10.1. A linearidade do modelo apóia em grande medida a suposição de normalidade.

Se a suposição de normalidade ou a de variâncias iguais forem consideradas implausíveis, deve-se empregar um método de análise diferente do usual teste  $F$ . É aconselhável procurar a opinião de especialistas em situações como essa (uma das possibilidades, transformar os dados, é proposta na Seção 10.3).

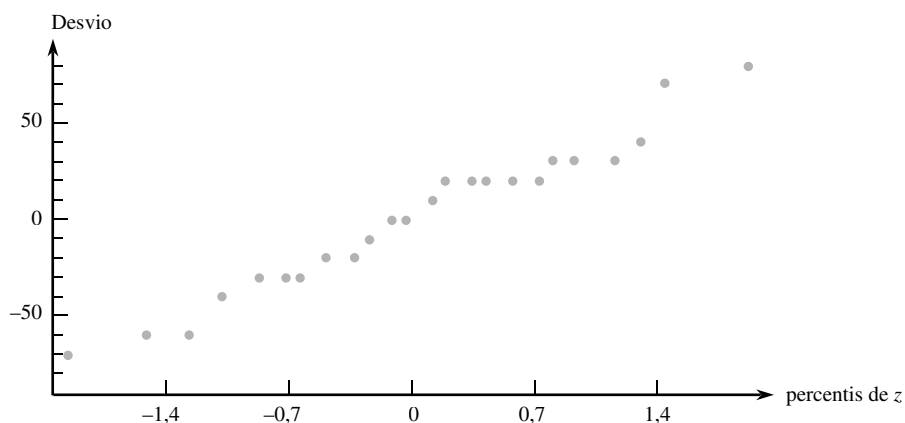


Figura 10.2 Gráfico de probabilidade normal baseado nos dados do Exemplo 10.1

## Estatística de Teste

Se  $H_0$  for verdadeiro, as  $J$  observações em cada amostra são provenientes de uma distribuição de população normal com o mesmo valor da média  $\mu$ . Nesse caso, é provável que as médias amostrais  $\bar{x}_1, \dots, \bar{x}_I$  estejam razoavelmente próximas. O procedimento de teste tem por base comparar a medida das diferenças entre os  $\bar{x}_i$ s (variação “entre amostras”) com a medida de variação calculada dentro de cada uma das amostras.

### DEFINIÇÃO

O **quadrado médio dos tratamentos** é dado por

$$\begin{aligned} \text{QMTr} &= \frac{J}{I-1} [(\bar{X}_1 - \bar{X}_{..})^2 + (\bar{X}_2 - \bar{X}_{..})^2 + \dots + (\bar{X}_I - \bar{X}_{..})^2] \\ &= \frac{J}{I-1} \sum_i (\bar{X}_i - \bar{X}_{..})^2 \end{aligned}$$

e o **quadrado médio do erro** é

$$\text{QME} = \frac{S_1^2 + S_2^2 + \dots + S_I^2}{I}$$

A estatística de teste ANOVA de fator único é  $F = \text{QMTr}/\text{QME}$ .

A terminologia “quadrado médio” será explicada resumidamente. Observe que são usados  $\bar{X}$ s e  $S^2$ s maiúsculos, de modo que o QMTr e o QME são definidos como estatísticas. Manteremos a tradição e também usaremos QMTr e QME (em vez de qmtr e qme) para representar os valores calculados dessas estatísticas. Cada  $S_i^2$  avalia a variação em cada amostra particular; portanto, QME é a medida de variação dentro das amostras.



Que tipo de valor de  $F$  apresenta evidência a favor ou contra  $H_0$ ? Primeiro, observe que, se  $H_0$  for verdadeiro (todos os  $\mu_i$ s forem iguais), os valores das médias amostrais individuais devem estar próximos um do outro e, portanto, da média geral, o que produz um valor de QMTr relativamente pequeno. Entretanto, se os  $\mu_i$ s forem muito diferentes, alguns  $\bar{x}_i$ s provavelmente serão bem diferentes de  $\bar{x}_..$ . Logo, o valor de QMTr é influenciado pela condição de  $H_0$  (verdadeiro ou falso). O mesmo não ocorre com o QME, porque os  $s_i^2$ s dependem apenas do valor de  $\sigma^2$  e não do lugar onde as várias distribuições estão centradas. O quadro a seguir apresenta uma importante propriedade de  $E(\text{QMTr})$  e  $E(\text{QME})$ , os valores esperados dessas duas estatísticas.

**PROPOSIÇÃO**

Quando  $H_0$  for verdadeiro,

$$E(\text{QMTr}) = E(\text{QME}) = \sigma^2$$

ao passo que, quando  $H_0$  for falso,

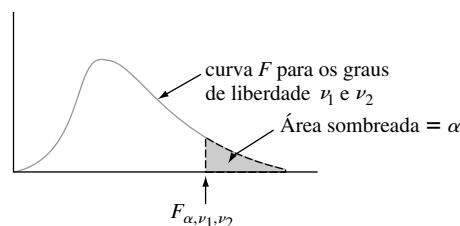
$$E(\text{QMTr}) > E(\text{QME}) = \sigma^2$$

Ou seja, ambas as estatísticas são não tendenciosas para estimar a variância comum da população  $\sigma^2$  quando  $H_0$  for verdadeiro, mas QMTr tende a superestimar  $\sigma^2$  quando  $H_0$  for falso.

A não-tendenciosidade de QME é uma consequência de  $E(S_i^2) = \sigma^2$ , quer  $H_0$  seja verdadeiro ou falso. Quando  $H_0$  for verdadeiro, todo  $\bar{X}_i$  tem o mesmo valor médio  $\mu$  e a mesma variância  $\sigma^2/J$ , de modo que  $\sum(\bar{X}_i - \bar{X}_..)^2/(I - 1)$ , a “variância amostral” dos  $\bar{X}_i$ s estima  $\sigma^2/J$  de maneira não-tendenciosa; multiplicando-se esse valor por  $J$ , obtém-se QMTr como um estimador não-tendencioso do próprio  $\sigma^2$ . Os  $\bar{X}_i$ s tendem a dispersar mais quando  $H_0$  é falso, do que quando é verdadeiro, o que, nesse caso, tende a aumentar o valor de QMTr. Logo, um valor de  $F$  muito maior que 1, que corresponda a QMTr bem maior que QME, põe  $H_0$  consideravelmente em dúvida. A forma da região de rejeição é  $f \geq c$ . O corte  $c$  deve ser escolhido para obter  $P(F \geq c \text{ quando } H_0 \text{ for verdadeiro}) = \alpha$ , que é o nível de significância desejado. Portanto precisamos da distribuição de  $F$  quando  $H_0$  for verdadeiro.

## Distribuições $F$ e Teste $F$

No Capítulo 9, apresentamos uma família de distribuições de probabilidade denominada distribuições  $F$ . A distribuição  $F$  surge junto de uma razão em que existe um número de graus de liberdade (gl) associado ao numerador e outro número de graus de liberdade associado ao denominador. Considere  $\nu_1$  e  $\nu_2$  o número de graus de liberdade do numerador e do denominador de uma variável com distribuição  $F$ . Tanto  $\nu_1$  quanto  $\nu_2$  são inteiros positivos. A Figura 10.3 apresenta uma curva de densidade  $F$  e o valor crítico da cauda superior  $F_{\alpha, \nu_1, \nu_2}$  correspondente. A Tabela A.9 do Apêndice fornece esses valores críticos para  $\alpha = 0,10, 0,05, 0,01, \text{ e } 0,001$ . Valores de  $\nu_1$  são identificados em colunas diferentes da tabela e as linhas são associadas aos vários valores de  $\nu_2$ . Por exemplo: o valor crítico de  $F$  que inclui a área da cauda superior 0,05 abaixo da curva  $F$ , com  $\nu_1 = 4$  e  $\nu_2 = 6$ , é  $F_{0,05,4,6} = 4,53$ , ao passo que  $F_{0,05,6,4} = 6,16$ . O principal resultado teórico é que a estatística de teste  $F$  tem distribuição  $F$  quando  $H_0$  for verdadeiro.



**Figura 10.3** Curva  $F$  e valor crítico  $F_{\alpha, \nu_1, \nu_2}$

## TEOREMA

Seja  $F = \text{QMTr}/\text{QME}$  a estatística de teste em um problema ANOVA de fator único que contém  $I$  populações ou tratamentos com uma amostra aleatória de  $J$  observações de cada população ou tratamento. Quando  $H_0$  for verdadeiro e as suposições básicas dessa seção forem satisfeitas,  $F$  tem distribuição  $F$  com  $\nu_1 = I - 1$  e  $\nu_2 = I(J - 1)$ . Quando  $f$  representa o valor calculado de  $F$ , a região de rejeição  $f \geq F_{\alpha, I-1, I(J-1)}$  especifica um teste com nível de significância  $\alpha$ . Consulte a Seção 9.5 para saber como obter informações sobre o valor  $P$  para testes  $F$ .

O raciocínio lógico para  $\nu_1 = I - 1$  é que, embora QMTr tenha por base os  $I$  desvios  $\bar{X}_1. - \bar{X}_{..}, \dots, \bar{X}_I. - \bar{X}_{..}$ ,  $\sum(\bar{X}_i. - \bar{X}_{..}) = 0$ , então somente  $I - 1$  desses desvios são determinados livremente. Pelo fato de cada amostra contribuir com  $J - 1$  gl para o QME e essas amostras serem independentes,  $\nu_2 = (J - 1) + \dots + (J - 1) = I(J - 1)$ .

**Exemplo 10.2** (Continuação do exemplo 10.1)

Os valores de  $I$  e  $J$  para os dados de resistência são 4 e 6, respectivamente; portanto, o numerador gl =  $I - 1 = 3$  e o denominador gl =  $I(J - 1) = 20$ . No nível de significância 0,05,  $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$  será rejeitado, favorecendo a conclusão de que pelo menos dois  $\mu_i$ s são diferentes se  $f \geq F_{0,05,3,20} = 3,10$ . A média geral é  $\bar{x}_{..} = \sum \sum x_{ij} / (IJ) = 682,50$ .

$$\begin{aligned} \text{QMTr} &= \frac{6}{4 - 1} [(713,00 - 682,50)^2 + (756,93 - 682,50)^2 \\ &\quad + (698,07 - 682,50)^2 + (562,02 - 682,50)^2] = 42.455,86 \end{aligned}$$

$$\text{QME} = \frac{1}{4} [(46,55)^2 + (40,34)^2 + (37,20)^2 + (39,87)^2] = 1691,92$$

$$f = \text{QMTr}/\text{QME} = 42.455,86/1691,92 = 25,09$$

Visto que  $25,09 \geq 3,10$ ,  $H_0$  é enfaticamente rejeitada ao nível de significância 0,05. A resistência de compressão média real parece depender do tipo de caixa. De fato, o valor  $P = \text{área sob a curva } F \text{ à direita de } 25,09 = 0,000$ , de modo que  $H_0$  seria rejeitada em qualquer nível de significância aceitável. ■

**Fórmulas de Cálculo para ANOVA**

Os cálculos que conduzem a  $f$  podem ser efetuados eficazmente por meio de fórmulas semelhantes à fórmula de cálculo usada para a variância de amostra única  $s^2$ . Seja  $x_i.$  a soma (não a média, visto que não existe nenhuma barra) dos  $x_{ij}$ s para  $i$  fixo (soma dos números na  $i$ -ésima linha da tabela) e  $x_{..}$  a soma de todos os  $x_{ij}$ s (o total geral).

## DEFINIÇÃO

A soma dos quadrados totais (SQT), a soma dos quadrados dos tratamentos (SQTr) e a soma dos quadrados dos erros (SQE) são dadas por

$$\text{SQT} = \sum_{i=1}^I \sum_{j=1}^J (x_{ij} - \bar{x}_{..})^2 = \sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 - \frac{1}{IJ} x_{..}^2$$

$$\text{SQTr} = \sum_{i=1}^I \sum_{j=1}^J (\bar{x}_i. - \bar{x}_{..})^2 = \frac{1}{J} \sum_{i=1}^I x_i.^2 - \frac{1}{IJ} x_{..}^2$$

$$\text{SQE} = \sum_{i=1}^I \sum_{j=1}^J (x_{ij} - \bar{x}_i.)^2 \quad \text{onde } x_i. = \sum_{j=1}^J x_{ij} \quad x_{..} = \sum_{i=1}^I \sum_{j=1}^J x_{ij}$$

A soma de quadrados SQTr aparece no numerador de  $F$  e SQE aparece no denominador de  $F$ ; o motivo de se definir SQT em breve ficará óbvio.

As expressões na extrema direita de SQT e SQTr são fórmulas de cálculo para essas somas de quadrados. Tanto SQT como SQTr contêm  $x^2/(IJ)$  (o quadrado total geral dividido por  $IJ$ ), que é normalmente denominado **fator de correção da média** (FC). Após o cálculo do fator de correção, obtém-se a SQT elevando-se ao quadrado cada um dos números na tabela de dados, adicionando-se esses quadrados e subtraindo-se o fator de correção. Obtém-se a SQTr elevando-se ao quadrado o total de cada linha, somando-os, dividindo-se por  $J$  e subtraindo-se o fator de correção.

A fórmula de cálculo para a SQE é obtida de uma relação simples entre as três somas de quadrados.

#### Identidade Fundamental

$$SQT = SQTr + SQE \quad (10.1)$$

Logo, se nenhuma das duas somas de quadrados forem calculadas, a terceira pode ser obtida por meio de (10.1); SQT e SQTr são as mais fáceis de calcular; portanto  $SQE = SQT - SQTr$ . Obtém-se a prova elevando-se ao quadrado ambos os lados da relação

$$x_{ij} - \bar{x}_{..} = (x_{ij} - \bar{x}_{i.}) + (\bar{x}_{i.} - \bar{x}_{..}) \quad (10.2)$$

e somando para todos os  $i$  e  $j$ . Dessa manipulação resultam SQT, à esquerda, e SQTr e SQE como os dois termos extremos à direita. O termo do produto cruzado é facilmente percebido visto como zero.

A interpretação da identidade fundamental é essencial para se compreender a ANOVA. A SQT é a medida da variação total nos dados – a soma de todos os desvios ao quadrado em torno da média geral. Segundo a identidade, essa variação total pode ser dividida em duas partes. A SQE mede a variação que existiria (dentro das linhas), mesmo se  $H_0$  fosse verdadeiro e, portanto, é a parte da variação total *inexplicada* em ambos os casos, tanto quando  $H_0$  for verdadeiro como quando for falso. A SQTr é a quantidade de variação (entre linhas) que *pode ser explicada* por possíveis diferenças nos  $\mu_s$ . Se a variação explicada for grande em relação à variação inexplicada,  $H_0$  será rejeitada em favor de  $H_a$ .

Após SQTr e SQE terem sido calculadas, cada uma é dividida por seu gl correspondente para obter a média dos quadrados. Então,  $F$  é a razão de ambas as médias dos quadrados.

$$QMT_r = \frac{SQTr}{I - 1} \quad QME = \frac{SQE}{I(J - 1)} \quad F = \frac{QMT_r}{QME} \quad (10.3)$$

Os cálculos em geral são resumidos em um formato tabular, denominado **tabela ANOVA**, como mostrado na Tabela 10.2. As tabelas produzidas por *softwares* de estatística incluem uma coluna de valor  $P$  à direita de  $f$ .

**Tabela 10.2** Tabela ANOVA

Fonte de Variação	gl	Soma dos Quadrados	Média dos Quadrados	$f$
Tratamentos	$I - 1$	SQTr	$QMT_r = SQTr/(I - 1)$	$QMT_r/QME$
Erros	$I(J - 1)$	SQE	$QME = SQE/[I(J - 1)]$	
Total	$IJ - 1$	SQT		

## Exemplo 10.3

Os dados a seguir provêm de um experimento de comparação do grau de resíduos em tecidos copolimerizados com três diferentes misturas de ácido metacrílico (dados semelhantes foram apresentados no artigo "Chemical Factors Affecting Soiling and Soil Release from Cotton DP Fabric", *American Dyestuff Reporter*, 1983, p. 25-30).

						$x_i$	$\bar{x}_i$
Mistura 1	0,56	1,12	0,90	1,07	0,94	4,59	0,918
Mistura 2	0,72	0,69	0,87	0,78	0,91	3,97	0,794
Mistura 3	0,62	1,08	1,07	0,99	0,93	4,69	0,938
						$x_{..} = 13,25$	

Seja  $\mu_i$  o grau de resíduos médio real quando se usa a mistura  $i$  ( $i = 1, 2, 3$ ). A hipótese nula  $H_0: \mu_1 = \mu_2 = \mu_3$  expressa que o grau de resíduos médio real é idêntico para as três misturas. Efetuaremos um teste usando o nível de significância 0,01 para verificar se  $H_0$  deve ser rejeitada em favor da afirmação de que o grau de resíduos médio real não é o mesmo para as três misturas. Visto que  $I - 1 = 2$  e  $I(J - 1) = 12$ , o valor crítico de  $F$  para a região de rejeição é  $F_{0,01,2,12} = 6,93$ . Elevando ao quadrado cada uma das 15 observações e somando-as, obtém-se  $\sum \sum x_{ij}^2 = (0,56)^2 + (1,12)^2 + \dots + (0,93)^2 = 12,1351$ . Os valores das três somas de quadrados são

$$SQT = 12,1351 - (13,25)^2/15 = 12,1351 - 11,7042 = 0,4309$$

$$SQTr = \frac{1}{5} [(4,59)^2 + (3,97)^2 + (4,69)^2] - 11,7042$$

$$= 11,7650 - 11,7042 = 0,0608$$

$$SQE = 0,4309 - 0,0608 = 0,3701$$

Os cálculos restantes estão resumidos na tabela ANOVA a seguir. Pelo fato de  $f = 0,99$  não ser maior do que  $F_{0,01,2,12} = 6,93$ ,  $H_0$  não é rejeitada no nível de significância 0,01. As misturas parecem indistinguíveis em relação ao grau de resíduos ( $F_{0,10,2,12} = 2,81 \Rightarrow \text{valor } P > 0,10$ ).

Fonte de Variação	gl	Soma dos Quadrados	Média dos Quadrados	$f$
Tratamentos	2	0,0608	0,0304	0,99
Erros	12	0,3701	0,0308	
Total	14	0,4309		

Quando o teste  $F$  faz com que  $H_0$  seja rejeitada, o investigador em geral ficará interessado em efetuar uma análise adicional para determinar quais  $\mu_i$ s diferem entre si. Os procedimentos para isso são chamados procedimentos de comparação múltipla. Vários deles são descritos nas duas seções a seguir.

## Exercícios | Seção 10.1 (1–10)

- Em um experimento para comparar a resistência à tração de  $I = 5$  em tipos diferentes de fio de cobre foram usadas  $J = 4$  amostras de cada tipo. As estimativas entre as amostras e dentro das amostras de  $\sigma^2$  foram calculadas como  $QMTr = 2673,3$  e  $QME = 1094,2$ , respectivamente.
  - Use o teste  $F$  no nível 0,05 para testar  $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$  versus  $H_a$ : pelo menos dois  $\mu_i$  são desiguais.
  - O que se pode dizer sobre o valor  $P$  do teste?
- Suponha que as observações de resistência de compressão sobre o quarto tipo de caixa no Exemplo 10.1 tenham

sido 655,1, 748,7, 662,4, 679,0. 706,9 e 640,0 (obtidas adicionando-se 120 a cada  $x_{4j}$  anterior). Supondo que não haja nenhuma mudança nas observações restantes, efetue um teste  $F$  usando  $\alpha = 0,05$ .

3. A quantidade de luz produzida (lúmen) foi determinada para cada uma das  $I = 3$  diferentes marcas de lâmpadas incandescentes brancas de 60 watts, com  $J = 8$  lâmpadas de cada marca testada. As somas de quadrados foram calculadas como  $SQE = 4773,3$  e  $SQTr = 591,2$ . Expresse as hipóteses de interesse (incluindo definições dos parâmetros) e use o teste  $F$  ANOVA ( $\alpha = 0,05$ ) para determinar se há alguma diferença na quantidade média real de luz produzida entre as três marcas para esse tipo de lâmpada, obtendo o maior número possível de informações sobre o valor  $P$ .
4. Em um estudo de avaliação dos efeitos da infecção por malária por mosquitos hospedeiros ("Plasmodium Cynomolgi: Effects of Malaria Infection on Laboratory Flight Performance of Anopheles Stephensi Mosquitos", *II Experimental Parasitology*, 1977, p. 397-404), os mosquitos alimentaram-se de macacos resos infectados e não-infectados. Subseqüentemente, foi medida a distância que eles voavam durante um período de 24 horas, usando-se um medidor de vôo. Os mosquitos foram divididos em quatro grupos de oito: resos infectados e presença de esporocistos (IRS), resos infectados e presença de oocistos (IRD), resos infectados sem desenvolvimento da infecção (IRN) e resos não-infectados (C). Os valores de dados de resumo são  $\bar{x}_{1.} = 4,39$  (IRS),  $\bar{x}_{2.} = 4,52$  (IRD),  $\bar{x}_{3.} = 5,49$  (IRN),  $\bar{x}_{4.} = 6,36$  (C),  $\bar{x}_{..} = 5,19$  e  $\sum \sum x_{ij}^2 = 911,91$ . Use o teste  $F$  ANOVA, no nível 0,05, para determinar se há alguma diferença entre o tempo de vôo médio real nos quatro tratamentos.
5. Considere os dados resumidos a seguir sobre o módulo de elasticidade ( $\times 10^6$  libras por polegada quadrada ou psi) para madeiras de três diferentes categorias (bem próximos dos valores usados no artigo "Bending Strength and Stiffness of Second-Growth Douglas-Fir Dimension Lumber" (*Forest Products J.*, 1991, p. 35-43), com a exceção de que o tamanho das amostras no artigo era maior):

Categoria	$J$	$\bar{x}_{i.}$	$s_i$
1	10	1,63	0,27
2	10	1,56	0,24
3	10	1,42	0,26

Use esses dados e o nível de significância de 0,01 para testar a hipótese nula de que não há nenhuma diferença na média do módulo de elasticidade nas três categorias.

6. O artigo "Origin of Precambrian Iron Formations" (*Econ. Geology*, 1964, p. 1025-1057) relata os seguintes dados sobre o total de Fe de quatro tipos de formação de ferro ( $I =$  carbonato,  $2 =$  silicato,  $3 =$  magnetita,  $4 =$  hematita).

1:	20,5	28,1	27,8	27,0	28,0
	25,2	25,3	27,1	20,5	31,3
2:	26,3	24,0	26,2	20,2	23,7
	34,0	17,1	26,8	23,7	24,9
3:	29,5	34,0	27,5	29,4	27,9
	26,2	29,9	29,5	30,0	35,6
4:	36,5	44,2	34,1	30,3	31,4
	33,1	34,1	32,9	36,3	25,5

Implemente uma análise do teste  $F$  de variância no nível de significância 0,01 e apresente o resumo dos resultados em uma tabela ANOVA.

7. Em um experimento para investigar o desempenho de quatro marcas diferentes de velas de ignição que seriam usadas em motos de 125 cc de dois tempos, cinco velas de cada marca foram testadas, observando-se o número de milhas percorridas (com velocidade constante) até o momento em que apresentassem falha. Uma tabela de dados ANOVA parcial é apresentada aqui. Preencha os campos vazios, expresse as hipóteses relevantes e faça um teste, obtendo o máximo possível de informações sobre o valor  $P$ .

Fonte	gl	Soma dos Quadrados	Média dos Quadrados	$f$
Marca				
Erro			14.713,69	
Total				
310.500,76				

8. Um estudo sobre as propriedades das vigas de conexão por chapa metálica para serem usadas como suporte de telhados ("Modeling Joints Made with Light-Gauge Metal Connector Plates", *Forest Products J.*, 1979, p. 39-44) apresentaram as seguintes observações sobre o índice de rigidez axial (kips/in.) para chapas de 4, 6, 8, 10 e 12 pol. de comprimento:

4:	309,2	409,5	311,0	326,5	316,8
	349,8	309,7			
6:	402,1	347,2	361,0	404,5	331,0
	348,9	381,7			
8:	392,4	366,2	351,0	357,1	409,9
	367,3	382,0			
10:	346,7	452,9	461,4	433,1	410,6
	384,2	362,6			
12:	407,4	441,8	419,9	410,7	473,4
	441,2	465,8			

A variação no comprimento da chapa tem algum efeito sobre a rigidez axial média real? Expresse e teste as hipóteses relevantes, usando análise de variância com  $\alpha = 0,01$ . Apresente seus resultados em uma tabela ANOVA. (*Sugestão*:  $\sum \sum x_{ij}^2 = 5.241,420,79$ .)

9. Foram analisadas seis amostras de cada um dos quatro tipos de grão de cereal cultivados em uma região para determinar o teor de tiamina, o que gerou os seguintes dados ( $\mu\text{g/g}$ ):

Trigo	5,2	4,5	6,0	6,1	6,7	5,8
Cevada	6,5	8,0	6,1	7,5	5,9	5,6
Milho	5,8	4,7	6,4	4,9	6,0	5,2
Aveia	8,3	6,1	7,8	7,0	5,5	7,2

Será que esses dados sugerem que pelo menos dois tipos de grão são diferentes em relação ao teor médio real de tiamina? Use um nível de teste  $\alpha = 0,05$  com base no método do valor  $P$ .

10. Em uma ANOVA de fator único com  $I$  tratamentos e  $J$  observações por tratamento, considere  $\mu = (1/I)\sum\mu_i$ .

- a. Expresse  $E(\bar{X}_{..})$  em função de  $\mu$ . [Sugestão:  $\bar{X}_{..} = (1/I)\sum\bar{X}_{i.}$ ]
- b. Calcule  $E(\bar{X}_{i.}^2)$ . [Sugestão: para qualquer  $Y$ ,  $E(Y^2) = V(Y) + [E(Y)]^2$ .]
- c. Calcule  $E(\bar{X}_{..}^2)$ .
- d. Calcule  $E(\text{SQTr})$  e, em seguida, mostre que

$$E(\text{QMTr}) = \sigma^2 + \frac{J}{I-1}\sum(\mu_i - \mu)^2$$

- e. Usando o resultado da parte (d), quanto vale  $E(\text{QMTr})$  quando  $H_0$  for verdadeiro? Quando  $H_0$  for falso, quanto vale  $E(\text{QMTr})$  comparado com  $\sigma^2$ ?

## 10.2 | Comparações Múltiplas em ANOVA

Quando o valor calculado da estatística  $F$  na ANOVA de fator único não é significativo, a análise é finalizada porque não se identificou nenhuma diferença entre os  $\mu_i$ s. Porém, quando  $H_0$  for rejeitada, o investigador normalmente terá interesse em saber quais dos  $\mu_i$ s são diferentes entre si. Um método para efetuar uma análise adicional é o **procedimento de comparações múltiplas**.

Vários desses procedimentos mais freqüentemente usados baseiam-se na seguinte idéia central: primeiro, calcule um intervalo de confiança para cada diferença  $\mu_i - \mu_j$  com  $i < j$ . Logo, se  $I = 4$ , os seis ICs necessários seriam para  $\mu_1 - \mu_2$  (mas não igualmente para  $\mu_2 - \mu_1$ ),  $\mu_1 - \mu_3$ ,  $\mu_1 - \mu_4$ ,  $\mu_2 - \mu_3$ ,  $\mu_2 - \mu_4$ , e  $\mu_3 - \mu_4$ . Então, se o intervalo de  $\mu_1 - \mu_2$  não incluir 0, conclua que  $\mu_1$  e  $\mu_2$  diferem significativamente entre si; se o intervalo incluir 0, ambos os  $\mu$ s não são considerados significativamente diferentes. Se seguirmos a mesma linha de raciocínio para cada um dos demais intervalos, podemos avaliar, para cada par de  $\mu$ , se são ou não significativamente diferentes entre si.

Os procedimentos que se baseiam nessa idéia diferem no método usado para calcular os vários ICs. Apresentamos aqui um método popular que controla o nível de confiança *simultâneo* de todos os  $I(I-1)/2$  intervalos calculados.

### Procedimento de Tukey (o Método T)

O procedimento de Tukey usa outra distribuição de probabilidade denominada **Distribuição de amplitude studentizada**. Essa distribuição depende de dois parâmetros: os gl do numerador  $m$  e os gl do denominador  $\nu$ . Seja  $Q_{\alpha,m,\nu}$  o valor crítico  $\alpha$  da cauda superior da distribuição de amplitude studentizada com  $m$  gl no numerador e  $\nu$  gl no denominador gl (análogo a  $F_{\alpha,\nu_1,\nu_2}$ ). Os valores de  $Q_{\alpha,m,\nu}$  são apresentados na Tabela A.10 do Apêndice.

PROPOSIÇÃO

Com probabilidade  $1 - \alpha$ ,

$$\begin{aligned} \bar{X}_{i.} - \bar{X}_{j.} - Q_{\alpha,I,(J-1)} \sqrt{\text{QME}/J} &\leq \mu_i - \mu_j \\ &\leq \bar{X}_{i.} - \bar{X}_{j.} + Q_{\alpha,I,(J-1)} \sqrt{\text{QME}/J} \end{aligned} \tag{10.4}$$

para todo  $i$  e  $j$  ( $i = 1, \dots, I$  e  $j = 1, \dots, I$ ) em que  $i < j$ .

Observe que o número de gl do numerador do valor crítico apropriado  $Q_\alpha$  é  $I$ , o número de médias de população ou tratamento que estão sendo comparadas, e não  $I - 1$ , como no teste  $F$ . Quando  $\bar{x}_i$ ,  $\bar{x}_j$ , e QME calculados são substituídos em (10.4), o resultado é um conjunto de intervalos de confiança com um nível de confiança *simultâneo*  $100(1 - \alpha)\%$  para todas as diferenças do padrão  $\mu_i - \mu_j$  com  $i < j$ . Todo intervalo que não inclui 0 leva à conclusão que os valores correspondentes de  $\mu_i$  e  $\mu_j$  diferem significativamente entre si.

Visto que não estamos de fato interessados nos limites inferior e superior dos vários intervalos, mas apenas em saber qual inclui e qual não inclui 0, grande parte dos cálculos aritméticos de (10.4) pode ser evitada. O quadro a seguir detalha e descreve como essas diferenças podem ser identificadas.

#### O método $T$ para identificar $\mu$ s significativamente diferentes

Selecione  $\alpha$ , obtenha  $Q_{\alpha, I, I(J-1)}$  da Tabela A.10 do Apêndice e calcule  $w = Q_{\alpha, I, I(J-1)} \cdot \sqrt{\text{QME}/J}$ . Em seguida, relacione as médias amostrais em ordem crescente e sublinhe os pares que diferem por menos de  $w$ . Qualquer par de médias amostrais não-sublinhado pela mesma linha corresponde a um par de médias de população ou tratamento considerados significativamente diferentes.

Suponha, por exemplo, que  $I = 5$  e que

$$\bar{x}_2. < \bar{x}_5. < \bar{x}_4. < \bar{x}_1. < \bar{x}_3.$$

Então

1. Considere primeiro a menor média  $\bar{x}_2.$ . Se  $\bar{x}_5. - \bar{x}_2. \geq w$ , vá para o Passo 2. Contudo, se  $\bar{x}_5. - \bar{x}_2. < w$ , relacione essas duas primeiras duas médias com um segmento de linha. Em seguida, se possível, estenda esse segmento de linha ainda mais para a direita, para o maior  $\bar{x}_i.$  que diferir de  $\bar{x}_2.$  de um valor menor que  $w$  (de forma que a linha possa unir duas, três ou mais médias).
2. Agora, considere  $\bar{x}_5.$  e estenda novamente um segmento de linha para o maior  $\bar{x}_i.$  à direita que difira de  $\bar{x}_5.$  de um valor menor que  $w$  (talvez não seja possível traçar essa linha; alternativamente, é possível sublinhar apenas duas, três ou até as quatro médias restantes).
3. Continue considerando  $\bar{x}_4.$  e procedendo da mesma forma e finalmente considerando  $\bar{x}_1.$

Em resumo, na lista ordenada, a partir de cada média é estendido um segmento de linha o máximo possível para a direita desde que a diferença entre as médias seja menor que  $w$ . É fácil notar que determinado intervalo da fórmula (10.4) conterá 0 se, e somente se, o par de médias amostrais correspondente estiver sublinhado pelo mesmo segmento de linha.

#### Exemplo 10.4

Um experimento foi implementado para comparar cinco marcas diferentes de filtro de óleo de automóvel em relação à capacidade de capturar material estranho. Seja  $\mu_i$  a quantidade média real de material capturado pelos filtros da marca  $i$  ( $i = 1, \dots, 5$ ) sob condições controladas. Utilizou-se uma amostra de nove filtros de cada marca, resultando nos seguintes valores de média amostral:  $\bar{x}_1. = 14,5$ ,  $\bar{x}_2. = 13,8$ ,  $\bar{x}_3. = 13,3$ ,  $\bar{x}_4. = 14,3$ , e  $\bar{x}_5. = 13,1$ . A Tabela 10.3 é o resumo da tabela ANOVA da primeira parte da análise.

**Tabela 10.3** Tabela ANOVA do Exemplo 10.4

Fonte de Variação	gl	Soma dos Quadrados	Média dos Quadrados	$f$
Tratamentos (marcas)	4	13,32	3,33	37,84
Erros	40	3,53	0,088	
Total	44	16,85		

Visto que  $F_{0,05,4,40} = 2,61$ ,  $H_0$  é rejeitada (decisivamente) ao nível 0,05. Usaremos agora o procedimento de Tukey para procurar diferenças significativas entre os  $\mu_i$ s. Na Tabela A.10 do Apêndice,  $Q_{0,05,5,40} = 4,04$  (o segundo índice de  $Q$  é  $I$  e não  $I - 1$  como em  $F$ ), de modo que  $w = 4,04\sqrt{0,088/9} = 0,4$ . Depois de organizar as cinco médias amostrais em ordem crescente, as duas menores podem ser relacionadas por um segmento de linha porque diferem de um valor menor que 0,4. Contudo esse segmento não pode ser mais estendido para a direita, visto que  $13,8 - 13,1 = 0,7 \geq 0,4$ . Ao mover uma média para a direita, o par  $\bar{x}_3$  e  $\bar{x}_2$  não pode ser sublinhado porque essas médias diferem de um valor superior a 0,4. Ao mover-se novamente para a direita, a média seguinte 13,8 não pode ser relacionada a nenhum valor adicional à direita. E, finalmente, as duas últimas médias podem ser sublinhadas com o mesmo segmento de linha.

$$\begin{array}{ccccc} \bar{x}_5. & \bar{x}_3. & \bar{x}_2. & \bar{x}_4. & \bar{x}_1. \\ \hline 13,1 & 13,3 & 13,8 & 14,3 & 14,5 \end{array}$$

Logo, as marcas 1 e 4 não são significativamente diferentes entre si, mas são muito maiores do que as três outras marcas em relação ao teor médio real. A marca 2 é substancialmente melhor que a 3 e a 5, mas pior que a 1 e a 4, e as marcas 3 e 5 não apresentam diferenças significativas.

Se  $\bar{x}_2 = 14,15$ , em vez de 13,8, usando o mesmo  $w$  calculado, a disposição das médias sublinhadas seria

$$\begin{array}{ccccc} \bar{x}_5. & \bar{x}_3. & \bar{x}_2. & \bar{x}_4. & \bar{x}_1. \\ \hline 13,1 & 13,3 & 14,15 & 14,3 & 14,5 \end{array} \quad \blacksquare$$

### Exemplo 10.5

Um biólogo desejava estudar os efeitos do etanol durante o sono. Foi selecionada uma amostra de 20 ratos, classificados por idade e outras características, e cada um recebeu uma injeção oral contendo determinada concentração de etanol de acordo com o respectivo peso. Registrou-se então o tempo de sono associado ao movimento rápido dos olhos (*Rapid Eye Moviment* – REM) de cada rato durante um período de 24 horas, obtendo-se os seguintes resultados:

Tratamento (concentração de etanol)		$x_i$	$\bar{x}_i$
0 (controle)	88,6    73,2    91,4    68,0    75,2	396,4	79,28
1 g/kg	63,0    53,9    69,2    50,1    71,5	307,7	61,54
2 g/kg	44,9    59,5    40,2    56,3    38,7	239,6	47,92
4 g/kg	31,0    39,6    45,3    25,2    22,7	163,8	32,76

$$x_{..} = 1107,5 \quad \bar{x}_{..} = 55,375$$

Será que os dados indicam que o valor real do tempo médio de sono REM depende da concentração de etanol? (Esse exemplo se baseia em um experimento relatado em “Relationship of Ethanol Blood Level to REM and Non-REM Sleep Time and Distribution in the Rat”, *Life Sciences*, 1978, p. 839-846.)

Os  $\bar{x}_i$ s diferem substancialmente um do outro, mas há ainda muita variabilidade dentro de cada amostra, de forma que, para responder precisamente à pergunta, devemos proceder a ANOVA. Sendo  $\sum \sum x_{ij}^2 = 68.697,6$  e o fator de correção  $x_{..}^2/(IJ) = (1107,5)^2/20 = 61.327,8$ , as fórmulas de cálculos geram o resultado

$$SQT = 68.697,6 - 61.327,8 = 7369,8$$

$$\begin{aligned} SQT_r &= \frac{1}{5}[(396,40)^2 + (307,70)^2 + (239,60)^2 + (163,80)^2] - 61.327,8 \\ &= 67.210,2 - 61.327,8 = 5882,4 \end{aligned}$$

e

$$SQE = 7369,8 - 5882,4 = 1487,4$$



A Tabela 10.4 é uma tabela ANOVA SAS. A última coluna fornece o valor  $P$ , que é 0,0001. Usando um nível de significância de 0,05, rejeitamos a hipótese nula  $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ , visto que o valor  $P = 0,0001 < 0,05 = \alpha$ . O real valor do tempo médio de sono REM parece depender do nível de concentração.

**Tabela 10.4** Tabela ANOVA SAS

Variável Dependente: TEMPO		Procedimento de Análise de Variância			
Fonte	GL	Soma de Quadrados	Média de Quadrados	Valor F	Pr . F
Modelo	3	5882,35750	1960,78583	21,09	0,0001
Erro Corrigido	16	1487,40000	92,96250		
Total	19	7369,75750			

Há  $I = 4$  tratamentos e 16 gl para os erros, de modo que  $Q_{0,05,4,16} = 4,05$  e  $w = 4,05\sqrt{93,0/5} = 17,47$ . Se ordenarmos e sublinharmos as médias, obteremos

$\bar{x}_4$	$\bar{x}_3$	$\bar{x}_2$	$\bar{x}_1$
32,76	47,92	61,54	79,28

Deve-se interpretar os sublinhados com cuidado, visto que aparentemente concluímos que os tratamentos 2 e 3 não diferem, 3 e 4 não diferem, embora 2 e 4 difiram. A forma sugerida para expressar isso é dizer que, embora existam evidências que nos levem a concluir que os tratamentos 2 e 4 diferem um do outro, nem um nem outro se mostrou significativamente diferente de 3. O tratamento 1 apresentou um real valor do médio de sono REM significativamente maior que os outros tratamentos.

A Figura 10.4 mostra o resultado SAS da aplicação do procedimento de Tukey.

```

Alfa 5 0,05 gl 5 16 QME 5 92.9625
Valor Crítico da Amplitude Studentizada 5 4,046
Diferença Mínima de Significância 5 17,446

As médias com a mesma letra não são significativamente diferentes.

Agrupamento de Tukey      Média      N      TRATAMENTO
A                          79,280     5      0 (controle)
B                          61,540     5      1 gm/kg
C                          47,920     5      2 gm/kg
C                          32,760     5      4 gm/kg

```

**Figura 10.4** Método de Tukey usando o SAS

## Interpretação de $\alpha$ no método de Tukey

Afirmamos anteriormente que o nível de confiança *simultâneo* é controlado pelo método de Tukey. Que significado tem a palavra "simultâneo" no nosso contexto? Considere o cálculo do IC de 95% para a média populacional  $\mu$  com base em uma amostra dessa população e, em seguida, o IC de 95% para uma proporção populacional  $p$  com base em outra amostra selecionada independentemente da primeira. Antes de obter os dados, a probabilidade de o primeiro intervalo incluir  $\mu$  é de 0,95, e essa é também a probabilidade de o segundo intervalo incluir  $p$ . Pelo fato de ambas as amostras serem selecionadas de maneira independente uma da outra, a probabilidade de *ambos* os intervalos incluírem os valores dos respectivos parâmetros é de  $(0,95)(0,95) = (0,95)^2 \approx 0,90$ . Portanto, o nível de confiança *simultâneo* ou *conjunto* dos dois intervalos é aproximadamente 90% — se pares de intervalos amostrais independentes forem calculados várias vezes, a longo prazo, em aproximadamente 90% das vezes, o primeiro intervalo incluirá  $\mu$  e o segundo incluirá  $p$ . Da mesma maneira, se três ICs forem calculados com base em amostras independentes, o nível de confiança simultâneo será  $100(0,95)^3\% \approx 86\%$ . Obviamente,

à medida que o número de intervalos aumenta, diminuirá o nível de confiança segundo o qual todos os intervalos incluam seus respectivos parâmetros.

Suponha agora que queiramos manter o nível de confiança simultâneo em 95%. Portanto, em duas amostras independentes, o nível de confiança individual para cada uma teria de ser  $100\sqrt{0,95}\% \approx 97,5\%$ . Quanto maior o número de intervalos, maior terá de ser o nível de confiança individual para que se mantenha o nível simultâneo de 95%.

Uma questão delicada em relação aos intervalos de Tukey é que eles não se baseiam em amostras independentes — o QME aparece em todos eles e vários compartilham os mesmos  $\bar{x}_i$ s. (por exemplo: no caso de  $I = 4$ , três intervalos diferentes usam  $\bar{x}_1$ ). Isso significa que não há nenhum argumento de probabilidade preciso que nos permita determinar o nível de confiança simultâneo com base nos níveis de confiança individuais. Entretanto é possível demonstrar que, se  $Q_{0,05}$  for usado, o nível de confiança simultâneo é controlado em 95%, ao passo que usando  $Q_{0,01}$  obtém-se um nível simultâneo de 99%. Para obter um nível simultâneo de 95%, o nível individual de cada intervalo deve ser consideravelmente maior que 95%. Dizendo a mesma coisa de maneira ligeiramente diferente, para se obter 5% de taxa de erro relativa a todo o *experimento* (*experimentwise*) ou à *família*, a taxa de erro individual ou por comparação deve ser consideravelmente menor que 0,05. No MINITAB, o usuário é solicitado a especificar a taxa de erro relativa à família (por exemplo, 5%) e, em seguida, inclui no resultado a taxa de erro individual (consulte o Exercício 16).

## Intervalos de Confiança de Outras Funções Paramétricas

Em algumas situações, um IC é desejável para uma função dos  $\mu_i$ s mais complicada do que a diferença  $\mu_i - \mu_j$ . Considere  $\theta = \sum c_i \mu_i$ , onde os  $c_i$ s são constantes. Uma dessas funções é  $\frac{1}{2}(\mu_1 + \mu_2) - \frac{1}{3}(\mu_3 + \mu_4 + \mu_5)$ , que no contexto do Exemplo 10.4 mede a diferença entre o grupo que consiste nas primeiras duas marcas e o grupo das três últimas marcas. Pelo fato de os  $X_{ij}$ s serem normalmente distribuídos com  $E(X_{ij}) = \mu_i$  e  $V(X_{ij}) = \sigma^2$ ,  $\hat{\theta} = \sum c_i \bar{X}_i$  é normalmente distribuído, não-tendencioso para  $\theta$  e

$$V(\hat{\theta}) = V(\sum c_i \bar{X}_i) = \sum c_i^2 V(\bar{X}_i) = \frac{\sigma^2}{J} \sum c_i^2$$

Calcular  $\sigma^2$  por meio do QME e criar  $\hat{\sigma}_\theta$  resulta em uma variável  $t$   $(\hat{\theta} - \theta)/\hat{\sigma}_\theta$  que pode ser manipulada para obter o seguinte intervalo de confiança  $100(1 - \alpha)\%$  para  $\sum c_i \mu_i$ :

$$\sum c_i \bar{x}_i \pm t_{\alpha/2, I(J-1)} \sqrt{\frac{\text{QME} \sum c_i^2}{J}} \quad (10.5)$$

**Exemplo 10.6** (Continuação do exemplo 10.4)

A função paramétrica para comparar as primeiras duas marcas (de fábrica) de filtro de óleo com as três últimas marcas (nacionais) é  $\theta = \frac{1}{2}(\mu_1 + \mu_2) - \frac{1}{3}(\mu_3 + \mu_4 + \mu_5)$ , da qual

$$\sum c_i^2 = \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(-\frac{1}{3}\right)^2 + \left(-\frac{1}{3}\right)^2 + \left(-\frac{1}{3}\right)^2 = \frac{5}{6}$$

Com  $\hat{\theta} = \frac{1}{2}(\bar{x}_1 + \bar{x}_2) - \frac{1}{3}(\bar{x}_3 + \bar{x}_4 + \bar{x}_5) = 0,583$  e  $\text{QME} = 0,088$ , um intervalo de 95% é

$$0,583 \pm 2,021 \sqrt{5(0,088)/[(6)(9)]} = 0,583 \pm 0,182 = (0,401, 0,765) \quad \blacksquare$$

Às vezes um experimento é realizado para comparar cada um dos vários "novos" tratamentos com um tratamento de controle. Nessas circunstâncias, é apropriado empregar uma técnica de comparações múltiplas denominada método de Dunnett.

**Exercícios** | Seção 10.2 (11–21)

11. Um experimento de comparação das taxas de dispersão de cinco marcas diferentes de tinta látex amarela para ambientes internos disponíveis em uma determinada área usou quatro galões ( $J = 4$ ) de cada tinta. As taxas de dispersão da média amostral (pés<sup>2</sup>/gal.) das cinco marcas foi  $\bar{x}_1 = 462,0$ ,  $\bar{x}_2 = 512,8$ ,  $\bar{x}_3 = 437,5$ ,  $\bar{x}_4 = 469,3$ , e  $\bar{x}_5 = 532,1$ . O valor calculado de  $F$  foi considerado significativo no nível  $\alpha = 0,05$ . Com  $QME = 272,8$ , use o procedimento de Tukey para investigar diferenças significativas nas taxas médias reais de dispersão entre as marcas.
12. No Exercício 11, considere  $\bar{x}_3 = 427,5$ . Que taxas médias reais de dispersão diferem significativamente entre si? Não deixe de usar o sublinhado para ilustrar suas conclusões e escreva um parágrafo resumindo os resultados obtidos.
13. Repita o Exercício 12, supondo  $\bar{x}_2 = 502,8$  e  $\bar{x}_3 = 427,5$ .
14. Use o procedimento de Tukey nos dados do Exercício 4 para identificar diferenças no tempo de vôo médio real entre os quatro tipos de mosquito.
15. Use o procedimento de Tukey nos dados do Exercício 6 para identificar diferenças no tempo de vôo médio real entre os quatro tipos de mosquito.
16. Analise novamente os dados de rigidez axial fornecidos no Exercício 8. Seguem os resultados ANOVA obtidos no MINITAB:

Análise de Variância de Rigidez					
Fonte	GL	SS	MS	F	P
Comprimento	4	43993	10998	10,48	0,000
Erro	30	31475	1049		
Total	34	75468			

Nível	N	Média	Desvio padrão
4	7	333,21	36,59
6	7	368,06	28,57
8	7	375,13	20,83
10	7	407,36	44,51
12	7	437,17	26,00

Desvio padrão Combinado 532,39

Comparações par a par de Tukey

Taxa de erro relativa à família 50,0500

Taxa de erro individual 50,00693

Valor crítico 54,10

Intervalos de (média em nível de coluna) - (média em nível de linha)

4	6	8	10
6	] 85,0		
	15,4		

8	] 92,1	] 57,3		
	8,3	43,1		
10	] 124,3	] 89,5	] 82,4	
	] 23,9	10,9	18,0	
12	] 154,2	] 119,3	] 112,2	] 80,0
	] 53,8	] 18,9	] 11,8	20,4

- a. É razoável afirmar que as variâncias das cinco distribuições do índice de rigidez axial são idênticas? Explique.
  - b. Use o resultado (sem consultar nossa tabela  $F$ ) para testar as hipóteses relevantes.
  - c. Use os intervalos de Tukey fornecidos nos resultados para determinar as médias que diferem e crie o padrão sublinhado correspondente.
17. Refira-se ao Exercício 5. Calcule o IC de 95%  $t$  para  $\theta = \frac{1}{2}(\mu_1 + \mu_2) - \mu_3$ .
  18. Considere os dados a seguir sobre o crescimento de plantas após a aplicação de diferentes tipos de hormônio de crescimento.

	1	13	17	7	14
2	21	13	20	17	
Hormônio	3	18	15	20	17
4	7	11	18	10	
5	6	11	15	8	

- a. Execute um teste  $F$  ao nível  $\alpha = 0,05$ .
  - b. O que ocorre quando se aplica o procedimento de Tukey?
19. Considere um experimento ANOVA de fator único no qual  $I = 3$ ,  $J = 5$ ,  $\bar{x}_1 = 10$ ,  $\bar{x}_2 = 12$ , e  $\bar{x}_3 = 20$ . Encontre um valor de SQE para o qual  $f > F_{0,05,2,12}$ , de modo que  $H_0: \mu_1 = \mu_2 = \mu_3$  seja rejeitada, embora, quando se aplique o procedimento de Tukey, nenhum dos  $\mu_i$  possa ser considerado significativamente diferente entre si.
  20. Refira-se ao Exercício 19 e considere  $\bar{x}_1 = 10$ ,  $\bar{x}_2 = 15$ , e  $\bar{x}_3 = 20$ . Será que agora você conseguiria encontrar um valor de SQE que gere uma contradição desse tipo entre o teste  $F$  e o procedimento de Tukey?
  21. O artigo "The Effect of Enzyme Inducing Agents on the Survival Times of Rats Exposed to Lethal Levels of Nitrogen Dioxide" (*Toxicology and Applied Pharmacology*, 1978, p. 169-174) relata os seguintes dados sobre o tempo de sobrevivência de ratos que receberam dióxido de nitrogênio (70 ppm) por meio de diferentes dietas. Havia  $J = 14$  ratos em cada grupo.

Dieta	$\bar{x}_i$ (min)	$s_i$
1. Controle	166	32
2. 3-Metilcolantrene	303	53
3. Alilisopropilacetamida	266	54
4. Fenobarbital	212	35
5. Clopromazine	202	34
6. Ácido <i>p</i> -aminobenzóico	184	31

- Teste a hipótese nula, que postula que o tempo de sobrevivência médio real não depende da dieta, em relação à alternativa de que existe alguma dependência na dieta, usando  $\alpha = 0,01$ .
- Suponha que  $100(1 - \alpha)\%$  ICs para  $k$  funções paramétricas diferentes sejam calculados com base no mesmo conjunto de dados ANOVA. Então, verifica-se facilmente que o nível de confiança simultâneo é de pelo menos  $100(1 - k\alpha)\%$ . Calcule os ICs com um nível de confiança simultâneo de pelo menos 98% para  $\mu_1 - \frac{1}{5}(\mu_2 + \mu_3 + \mu_4 + \mu_5 + \mu_6)$  e  $\frac{1}{4}(\mu_2 + \mu_3 + \mu_4 + \mu_5) - \mu_6$ .

## 10.3 | Mais sobre ANOVA de fator único

Nesta seção, consideramos brevemente outras questões relacionadas com a ANOVA de fator único. Dentre elas, inclui-se uma descrição alternativa dos parâmetros do modelo,  $\beta$  para o teste  $F$ , a relação do teste com procedimentos anteriormente analisados, transformação de dados, um modelo de efeitos aleatórios e fórmulas para o caso de tamanhos de amostras desiguais.

### Descrição alternativa do modelo ANOVA

As suposições ANOVA de fator único podem ser descritas, sucintamente, por meio da “equação do modelo”

$$X_{ij} = \mu_i + \epsilon_{ij}$$

em que  $\epsilon_{ij}$  representa um desvio aleatório em relação à população ou à média do tratamento real  $\mu_i$ . Os  $\epsilon_{ij}$  são considerados variáveis aleatórias independentes, normalmente distribuídas (o que significa que os  $X_{ij}$  também são), com  $E(\epsilon_{ij}) = 0$  [de modo que  $E(X_{ij}) = \mu_i$  e  $V(\epsilon_{ij}) = \sigma^2$  [da qual  $V(X_{ij}) = \sigma^2$  para todo  $i$  e  $j$ ]]. Uma descrição alternativa da ANOVA de fator único permite uma percepção mais ampla e sugere generalizações apropriadas a modelos com mais de um fator. Defina um parâmetro  $\mu$  por

$$\mu = \frac{1}{I} \sum_{i=1}^I \mu_i$$

e os parâmetros  $\alpha_1, \dots, \alpha_I$  por

$$\alpha_i = \mu_i - \mu \quad (i = 1, \dots, I)$$

Então, a média do tratamento  $\mu_i$  pode ser escrita como  $\mu + \alpha_i$ , onde  $\mu$  representa a resposta média real global no experimento e  $\alpha_i$  é o efeito medido como um desvio de  $\mu$  e provocado pelo  $i$ -ésimo tratamento. Embora inicialmente tivéssemos  $I$  parâmetros, no momento temos  $I + 1$  ( $\mu, \alpha_1, \dots, \alpha_I$ ). Contudo, pelo fato de  $\sum \alpha_i = 0$  (o desvio médio da resposta média global é zero), somente  $I$  desses novos parâmetros são determinados de modo independente, de modo que existem tantos parâmetros independentes quanto antes. Em função de  $\mu$  e  $\alpha_i$ , o modelo passa a ser

$$X_{ij} = \mu + \alpha_i + \epsilon_{ij} \quad (i = 1, \dots, I, \quad j = 1, \dots, J)$$

No Capítulo 11, desenvolveremos modelos análogos de ANOVA multifatores. A afirmação de que os  $\mu_i$  são idênticos equivale à igualdade dos  $\alpha_i$ , e, pelo fato de  $\sum \alpha_i = 0$ , a hipótese nula passa a ser

$$H_0: \alpha_1 = \alpha_2 = \dots = \alpha_I = 0$$

Na Seção 10.1, afirmou-se que o QMTr é um estimador não-tendencioso de  $\sigma^2$  quando  $H_0$  for verdadeiro, mas que em outras circunstâncias tende a superestimar  $\sigma^2$ . Mais precisamente,

$$E(\text{QMTr}) = \sigma^2 + \frac{J}{I-1} \sum \alpha_i^2$$

Quando  $H_0$  for verdadeiro,  $\sum \alpha_i^2 = 0$  de modo que  $E(\text{QMTr}) = \sigma^2$  (a não-tendenciosidade de QME se mantém, independentemente de  $H_0$  ser verdadeiro ou falso). Se  $\sum \alpha_i^2$  for usado para avaliar até que ponto  $H_0$  é falso, um valor maior de  $\sum \alpha_i^2$  aumentará ainda mais a probabilidade de QMTr superestimar  $\sigma^2$ . No capítulo seguinte, serão usadas fórmulas para quadrados médios esperados de modelos de multifatores, a fim de propor uma maneira de formular quocientes  $F$  para testar várias hipóteses.

**Demonstração da Fórmula para  $E(\text{QMTr})$**  Para qualquer va  $Y$ ,  $E(Y^2) = V(Y) + [E(Y)]^2$ , de modo que

$$\begin{aligned} E(\text{SQTr}) &= E\left(\frac{1}{J} \sum_i X_i^2 - \frac{1}{IJ} X_{..}^2\right) = \frac{1}{J} \sum_i E(X_i^2) - \frac{1}{IJ} E(X_{..}^2) \\ &= \frac{1}{J} \sum_i \{V(X_i) + [E(X_i)]^2\} - \frac{1}{IJ} \{V(X_{..}) + [E(X_{..})]^2\} \\ &= \frac{1}{J} \sum_i \{J\sigma^2 + [J(\mu + \alpha_i)]^2\} - \frac{1}{IJ} [IJ\sigma^2 + (IJ\mu)^2] \\ &= I\sigma^2 + IJ\mu^2 + 2\mu J \sum_i \alpha_i + J \sum_i \alpha_i^2 - \sigma^2 - IJ\mu^2 \\ &= (I-1)\sigma^2 + J \sum_i \alpha_i^2 \quad (\text{visto que } \sum \alpha_i = 0) \end{aligned}$$

Portanto, o resultado segue da relação  $\text{QMTr} = \text{SQTr}/(I-1)$ . ■

## $\beta$ para o Teste $F$

Considere um conjunto de valores de parâmetro  $\alpha_1, \alpha_2, \dots, \alpha_I$  para o qual  $H_0$  não seja verdadeiro. A probabilidade de um erro tipo II,  $\beta$ , é a probabilidade de  $H_0$  não ser rejeitada quando esse conjunto for o conjunto de valores reais. Alguém pode pensar que  $\beta$  teria de ser determinado separadamente para cada configuração diferente de  $\alpha_i$ . Felizmente, visto que  $\beta$  para o teste  $F$  depende dos  $\alpha_i$ s e de  $\sigma^2$ , somente por meio de  $\sum \alpha_i^2/\sigma^2$ , ele pode ser avaliado simultaneamente para várias alternativas diferentes. Por exemplo:  $\sum \alpha_i^2 = 4$  para cada um dos seguintes conjuntos de  $\alpha_i$  para os quais  $H_0$  é falso, de modo que  $\beta$  é idêntico para as três alternativas:

1.  $\alpha_1 = -1, \alpha_2 = -1, \alpha_3 = 1, \alpha_4 = 1$
2.  $\alpha_1 = -\sqrt{2}, \alpha_2 = \sqrt{2}, \alpha_3 = 0, \alpha_4 = 0$
3.  $\alpha_1 = -\sqrt{3}, \alpha_2 = \sqrt{1/3}, \alpha_3 = \sqrt{1/3}, \alpha_4 = \sqrt{1/3}$

A quantidade  $J\sum \alpha_i^2/\sigma^2$  é denominada **parâmetro de não-centralidade** para a ANOVA unidirecional (porque, quando  $H_0$  for falso, a estatística de teste tem distribuição  $F$  não-central com isso como um de seus parâmetros), e  $\beta$  é uma função decrescente do valor desse parâmetro. Logo, para valores fixos de  $\sigma^2$  e  $J$ , é mais provável que a hipótese nula seja rejeitada para alternativas distantes de  $H_0$  ( $\sum \alpha_i^2$  grande) do que para alternativas próximas de  $H_0$ . Para um valor fixo de  $\sum \alpha_i^2$ ,  $\beta$  diminui à medida que o tamanho da amostra  $J$  em cada tratamento aumenta, e aumenta à medida que a variância  $\sigma^2$  aumenta (visto que a variabilidade subjacente torna ainda mais difícil detectar qualquer desvio de  $H_0$ ).

Pelo fato de o cálculo manual de  $\beta$  e a determinação do tamanho da amostra para o teste  $F$  serem muito difíceis (como no caso dos testes  $t$ ), os estatísticos criaram conjuntos de curvas das quais  $\beta$  pode ser obtido.

Conjuntos de curvas para  $\nu_1 = 3$  e  $\nu_1 = 4$  gl no numerador são mostrados na Figura 10.5 e na Figura 10.6<sup>1</sup>, respectivamente. Depois que os valores de  $\sigma^2$  e dos  $\alpha$ s, para os quais  $\beta$  é desejado, são especificados, eles são usados para calcular o valor de  $\phi$ , onde  $\phi^2 = (J/I)\sum\alpha_i^2/\sigma^2$ . Em seguida, entramos no conjunto apropriado de curvas com o valor de  $\phi$  no eixo horizontal, movemo-nos para a curva associada com o gl do erro  $\nu_2$  e movendo novamente para o valor do poder do teste no eixo vertical. Finalmente,  $\beta = 1 - \text{poder do teste}$ .

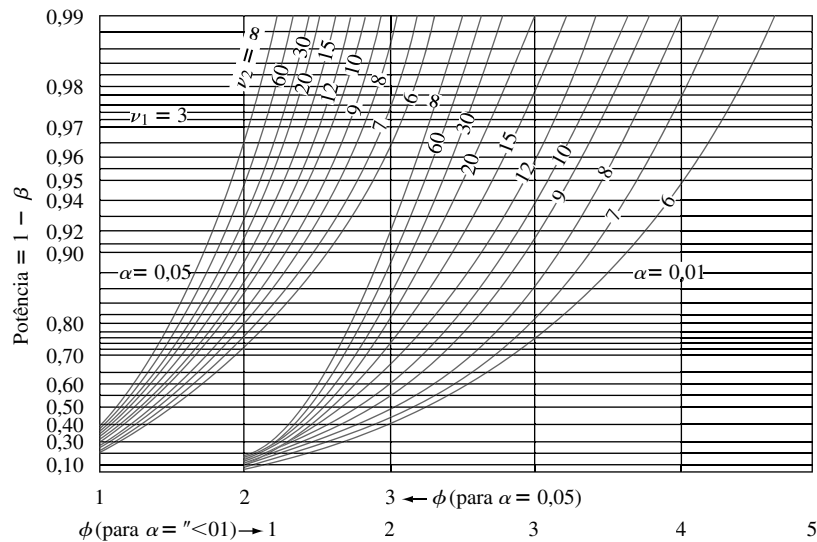


Figura 10.5 Curvas de poder do teste para o teste F ANOVA ( $\nu_1 = 3$ )

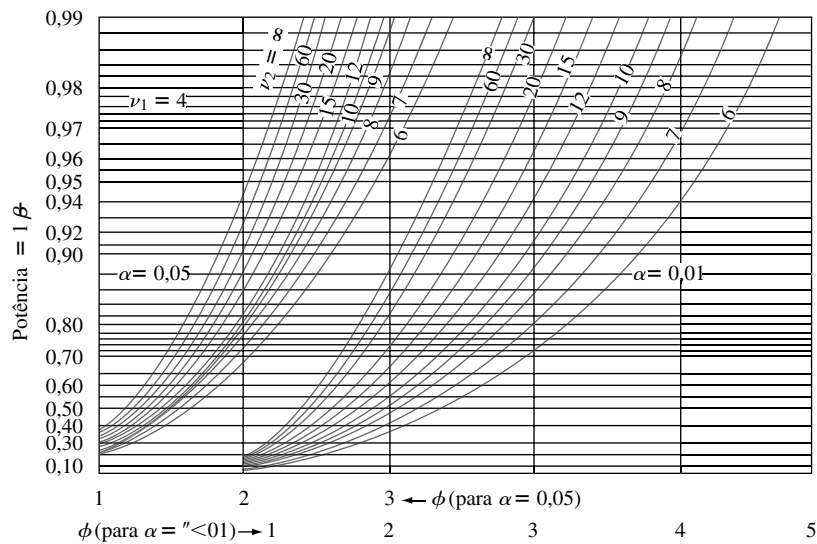


Figura 10.6 Curvas de poder do teste para o teste F ANOVA ( $\nu_1 = 4$ )

Exemplo 10.7

Os efeitos de quatro diferentes tratamentos térmicos sobre o ponto de ruptura ( $t/pol.^2$ ) de barras de aço serão investigados. Oito barras no total serão fundidas por meio de “tratamentos diferentes”. Suponha que o desvio padrão real do ponto de ruptura para qualquer um dos quatro tratamentos seja  $\sigma = 1$ . Qual a probabilidade de  $H_0$  não ser rejeitada no nível 0,05, se três dos tratamentos tiverem o mesmo ponto de ruptura esperado e o outro

<sup>1</sup> Extraído de E. S. Pearson e H. O. Hartley, "Charts of the Power Function for Analysis of Variance Tests, Derived from the Non-central F Distribution" *Biometrika*, vol. 38, 1951, p.112, com a permissão da Biometrika Trustees.

tratamento tiver um ponto de ruptura esperado  $1/t$  maior que o valor comum dos outros três (isto é, o quarto é, em média, 1 desvio padrão superior aos dos três primeiros tratamentos)?

Suponha que  $\mu_1 = \mu_2 = \mu_3$  e  $\mu_4 = \mu_1 + 1$ ,  $\mu = (\sum \mu_i)/4 = \mu_1 + \frac{1}{4}$ . Logo,  $\alpha_1 = \mu_1 - \mu = -\frac{1}{4}$ ,  $\alpha_2 = -\frac{1}{4}$ ,  $\alpha_3 = -\frac{1}{4}$ ,  $\alpha_4 = \frac{3}{4}$ , de modo que

$$\phi^2 = \frac{8}{4} \left[ \left( \frac{1}{4} \right)^2 + \left( -\frac{1}{4} \right)^2 + \left( -\frac{1}{4} \right)^2 + \left( \frac{3}{4} \right)^2 \right] = \frac{3}{2}$$

e  $\phi = 1,22$ . Os graus de liberdade são  $\nu_1 = I - 1 = 3$  e  $\nu_2 = I(J - 1) = 28$ ; portanto, interpolando-se visualmente entre  $\nu_2 = 20$  e  $\nu_2 = 30$ , obtém-se um poder do teste  $\approx 0,47$  e  $\beta \approx 0,53$ . Esse  $\beta$  é muito grande, de modo que devemos aumentar o valor de  $J$ . Quantas barras de cada tipo seriam necessárias para obter  $\beta \approx 0,05$  para a alternativa que está sendo avaliada? Experimentando diferentes valores de  $J$ , podemos observar que  $J = 24$  satisfaz a necessidade, o que não é verdadeiro para qualquer  $J$  menor que esse valor. ■

Como alternativa ao uso de curvas de poder do teste, o pacote de *software* estatístico SAS dispõe de uma função que calcula a área acumulada sob curva  $F$  não-central (entra com  $F_{\alpha}$ ,  $gl$  no numerador,  $gl$  no denominador e  $\phi^2$ ), e a área é  $\beta$ . Na versão 13 do MINITAB, isso é bem diferente. O usuário é solicitado a especificar a diferença máxima entre  $\mu_i$ , em vez de as médias individuais. Por exemplo: queremos calcular o poder do teste quando  $I = 4$ ,  $\mu_1 = 100$ ,  $\mu_2 = 101$ ,  $\mu_3 = 102$ , e  $\mu_4 = 106$ . Então, a diferença máxima é  $106 - 100 = 6$ . Entretanto, o poder do teste depende não apenas dessa diferença máxima, como também dos valores de todos os  $\mu_i$ s. Nessa circunstância, o MINITAB calcula o menor valor possível do poder do teste na condição  $\mu_1 = 100$  e  $\mu_4 = 106$ , o que ocorre quando os dois outros  $\mu$ s estão no intervalo entre 100 e 106. Se esse poder do teste for 0,85, podemos dizer que o poder do teste é pelo menos 0,85 e  $\beta$  é no máximo 0,15, quando os dois  $\mu$ s mais extremos estão separados por 6 (o tamanho comum da amostra,  $\alpha$  e  $\sigma$  também devem ser especificados). Além disso, o *software* determinará o tamanho comum da amostra necessário, se forem especificados uma diferença máxima e um poder do teste mínimo.

## Relação do Teste $F$ com Teste $t$

Quando o número de tratamentos ou de populações for  $I = 2$ , todos os resultados e fórmulas associados com o teste  $F$  ainda farão sentido, de modo que a ANOVA pode ser usada para testar  $H_0: \mu_1 = \mu_2$  versus  $H_a: \mu_1 \neq \mu_2$ . Nesse caso, pode ser usado um teste  $t$  bicaudal de duas amostras. Na Seção 9.3, mencionamos o teste  $t$  combinado, que exige variâncias iguais, como alternativa ao procedimento  $t$  de duas amostras. É possível demonstrar que os testes  $F$  ANOVA de fator único e  $t$  bicaudal combinado são equivalentes; para qualquer conjunto de dados, os valores  $P$  em ambos os testes serão idênticos, de modo que tanto um como o outro chegarão à mesma conclusão.

O teste  $t$  de duas amostras é mais flexível que o teste  $F$  quando  $I = 2$ , por dois motivos: primeiro, ele é válido sem a suposição de que  $\sigma_1 = \sigma_2$ ; segundo, ele pode ser usado para testar  $H_a: \mu_1 > \mu_2$  (teste  $t$  de cauda superior) ou  $H_a: \mu_1 < \mu_2$  e igualmente  $H_a: \mu_1 \neq \mu_2$ . No caso de  $I \geq 3$ , infelizmente não há nenhum procedimento de teste geral conhecido e com propriedades satisfatórias que não assuma variâncias iguais.

## ANOVA de Fator Único quando os Tamanhos das Amostras são Desiguais

Quando os tamanhos das amostras de cada população ou tratamento não são iguais, sejam  $J_1, J_2, \dots, J_I$  os tamanhos das amostras  $I$  e  $n = \sum_i J_i$  o número total de observações. O quadro a seguir apresenta fórmulas ANOVA e o procedimento de teste.

$$SQT = \sum_{i=1}^I \sum_{j=1}^{J_i} (X_{ij} - \bar{X}_{i.})^2 = \sum_{i=1}^I \sum_{j=1}^{J_i} X_{ij}^2 - \frac{1}{n} X_{..}^2 \quad \text{gl} = n - 1$$

$$SQTr = \sum_{i=1}^I \sum_{j=1}^{J_i} (\bar{X}_{i.} - \bar{X}_{..})^2 = \sum_{i=1}^I \frac{1}{J_i} X_{i.}^2 - \frac{1}{n} X_{..}^2 \quad \text{gl} = I - 1$$

$$SQE = \sum_{i=1}^I \sum_{j=1}^{J_i} (X_{ij} - \bar{X}_{i.})^2 = SQT - SQTr \quad \text{gl} = \sum (J_i - 1) = n - I$$

Valor do teste estatístico:

$$f = \frac{QMTr}{QME} \quad \text{onde } QMTr = \frac{SQTr}{I - 1} \quad QME = \frac{SQE}{n - I}$$

Região de rejeição:  $f \geq F_{\alpha, I-1, n-I}$

### Exemplo 10.8

O artigo "On the Development of a New Approach for the Determination of Yield Strength in Mg-based Alloys" (*Light Metal Age*, p. 51-53, out. 1998) apresentou os seguintes dados sobre módulo de elasticidade (GPa) obtido por um novo método ultra-sônico para espécimes de uma determinada liga produzida por meio de três diferentes processos de fundição:

									$J_i$	$x_i$	$\bar{x}_i$
<i>Moldagem permanente</i>	45,5	45,3	45,4	44,4	44,6	43,9	44,6	44,0	8	357,7	44,71
<i>Fundição</i>	44,2	43,9	44,7	44,2	44,0	43,8	44,6	43,1	8	352,5	44,06
<i>Moldagem de gesso</i>	46,0	45,9	44,8	46,2	45,1	45,5			6	273,5	45,58
									22	983,7	

Sejam  $\mu_1$ ,  $\mu_2$ , e  $\mu_3$  os módulos elásticos médios reais para os três diferentes processos sob as circunstâncias determinadas. As hipóteses relevantes são  $H_0: \mu_1 = \mu_2 = \mu_3$  versus  $H_a$ : pelo menos dois dos  $\mu_i$ s são diferentes. A estatística de teste é, obviamente,  $F = QMTr/QME$ , com base em  $I - 1 = 2$  gl no numerador e  $n - I = 22 - 3 = 19$  gl no denominador. As quantidades correspondentes incluem

$$\sum \sum x_{ij}^2 = 43.998,73 \quad CF = \frac{983,7^2}{22} = 43.984,80$$

$$SQT = 43.998,73 - 43.984,80 = 13,93$$

$$SQTr = \frac{357,7^2}{8} + \frac{352,5^2}{8} + \frac{273,5^2}{6} - 43.984,80 = 7,93$$

$$SQE = 13,93 - 7,93 = 6,00$$

Os cálculos restantes são apresentados na tabela ANOVA a seguir. Visto que  $F_{0,001,2,19} = 10,16 < 12,56 = f$ , o valor  $P$  é menor que 0,001. Logo, a hipótese nula deve ser rejeitada em qualquer nível de significância razoável; há evidências convincentes de que o módulo elástico médio real depende de alguma forma do processo de fundição usado.



Fonte de Variação	gl	Soma dos Quadrados	Média dos Quadrados	$f$
Tratamentos	2	7,93	3,965	12,56
Erro	19	6,00	0,3158	
Total	21	13,93		

## Comparações Múltiplas quando os Tamanhos Amostrais são Desiguais

Há maior controvérsia entre os estatísticos com relação ao procedimento de comparação múltipla a ser utilizado quando os tamanhos das amostras são desiguais do que quando são iguais. O procedimento que apresentamos aqui é recomendado no excelente livro *Beyond ANOVA: Basics of Applied Statistics* (consulte a Bibliografia do capítulo) para ser usado quando os  $I$  tamanhos das amostras  $J_1, J_2, \dots, J_I$  são razoavelmente próximos entre si ("desequilíbrio moderado"). Isso modifica o método de Tukey porque usa as médias dos pares  $1/J_i$  no lugar de  $1/J$ .

Seja

$$w_{ij} = Q_{\alpha, I, n-I} \cdot \sqrt{\frac{\text{QME}}{2} \left( \frac{1}{J_i} + \frac{1}{J_j} \right)}$$

Logo, a probabilidade é de *aproximadamente*  $1 - \alpha$  de que

$$\bar{X}_i - \bar{X}_j - w_{ij} \leq \mu_i - \mu_j \leq \bar{X}_i - \bar{X}_j + w_{ij}$$

para todos  $i$  e  $j$  ( $i = 1, \dots, I$  e  $j = 1, \dots, I$ ) com  $i \neq j$ .

O nível de confiança simultâneo  $100(1 - \alpha)\%$  é apenas aproximado, e não exato, como no caso de tamanhos das amostras desiguais. O método do sublinhado ainda pode ser usado, mas agora o fator  $w_{ij}$  usado para determinar se  $\bar{x}_i$  e  $\bar{x}_j$  podem ser conectados depende de  $J_i$  e  $J_j$ .

**Exemplo 10.9** (Continuação do Exemplo 10.8)

Os tamanhos das amostras dos dados do módulo elástico foram  $J_1 = 8, J_2 = 8, J_3 = 6$ , e  $I = 3, n - I = 19$ ,  $\text{QME} = 0,316$ . Um nível de confiança simultâneo de aproximadamente 95% exige  $Q_{0,05,3,19} = 3,59$ , de onde

$$w_{12} = 3,59 \sqrt{\frac{0,316}{2} \left( \frac{1}{8} + \frac{1}{8} \right)} = 0,713, \quad w_{13} = 0,771 \quad w_{23} = 0,771$$

Visto que  $\bar{x}_1 - \bar{x}_2 = 44,71 - 44,06 = 0,65 < w_{12}$ ,  $\mu_1$  e  $\mu_2$  são considerados não significativamente diferentes. O esquema sublinhado a seguir mostra que  $\mu_1$  e  $\mu_3$  parecem significativamente diferentes, como o são  $\mu_2$  e  $\mu_3$ .

2. Fundição	1. Permanente	3. Gesso
44,06	44,71	45,58

## Transformação de Dados

O uso de métodos ANOVA pode ser invalidado por diferenças substanciais nas variâncias  $\sigma_1^2, \dots, \sigma_I^2$  (as quais, até agora, foram consideradas iguais com valor  $\sigma^2$  comum). Às vezes ocorre de  $V(X_{ij}) = \sigma_i^2 = g(\mu_i)$ , função conhecida de  $\mu_i$  (de modo que, quando  $H_0$  for falso, as variâncias não são iguais). Por exemplo: se  $X_{ij}$  tiver distribuição de Poisson com parâmetro  $\lambda_i$  (aproximadamente normal, se  $\lambda_i \geq 10$ ), então  $\mu_i = \lambda_i$  e  $\sigma_i^2 = \lambda_i$ , de modo

que  $g(\mu_i) = \mu_i$  é a função conhecida. Nessas circunstâncias, em geral é possível transformar os  $X_{ij}$ s em  $h(X_{ij})$  de maneira que tenham variâncias iguais próximas (ao mesmo tempo mantendo as variáveis transformadas aproximadamente normais) e em seguida, o teste  $F$  pode ser usado nas observações transformadas. A principal idéia de escolher uma transformação  $h(\cdot)$  é que geralmente  $V[h(X_{ij})] \approx V(X_{ij}) \cdot [h'(\mu_i)]^2 = g(\mu_i) \cdot [h'(\mu_i)]^2$ . Gostariamos de encontrar a função  $h(\cdot)$  para a qual  $g(\mu_i) \cdot [h'(\mu_i)]^2 = c$  (uma constante) para todo  $i$ .

**PROPOSIÇÃO**

Se  $V(X_{ij}) = g(\mu_i)$ , função conhecida de  $\mu_i$ , então a transformação  $h(X_{ij})$  que "estabiliza a variância", de modo que  $V[h(X_{ij})]$  é aproximadamente a mesma para todo  $i$  dado por  $h(x) \propto \int [g(x)]^{-1/2} dx$ .

No caso de Poisson,  $g(x) = x$ , de modo que  $h(x)$  deve ser proporcional a  $\int x^{-1/2} dx = 2x^{1/2}$ . Logo, os dados de Poisson devem ser transformados em  $h(x_{ij}) = \sqrt{x_{ij}}$  antes da análise.

### Modelo de Efeitos Aleatórios

Os problemas de fator único considerados até aqui foram concebidos como exemplos de um modelo ANOVA de **efeitos fixos**. Com isso queremos dizer que os níveis escolhidos do fator em estudo são os únicos considerados relevantes pelo investigador. O modelo de efeitos fixos de fator único é

$$X_{ij} = \mu + \alpha_i + \epsilon_{ij} \quad \sum \alpha_i = 0 \quad (10.6)$$

onde os  $\epsilon_{ij}$  são aleatórios e tanto  $\mu$  quanto os  $\alpha_i$  são parâmetros fixos cujos valores são desconhecidos.

Em alguns problemas de fator único, os níveis específicos estudados pelo investigador são escolhidos de acordo com o propósito ou por amostragem de uma ampla população de níveis. Por exemplo: para estudar os efeitos sobre o tempo de execução da tarefa de usar diferentes operadores em uma determinada máquina, pode-se escolher uma amostra de cinco operadores de um amplo grupo deles. De modo semelhante, o efeito do pH do solo sobre a produção de milho pode ser estudado usando solos com quatro valores específicos de pH escolhidos dentre os vários níveis de pH possíveis. Quando os níveis usados são selecionados aleatoriamente de uma ampla população de níveis possíveis, diz-se que o fator é aleatório em vez de fixo. Nesse caso, o modelo de efeitos fixos (10.6) não é mais apropriado. Pode-se obter um modelo de **efeitos aleatórios** análogo substituindo-se os  $\alpha_i$  fixos em (10.6) por variáveis aleatórias. A descrição do modelo resultante é

$$X_{ij} = \mu + A_i + \epsilon_{ij} \quad \text{com } E(A_i) = E(\epsilon_{ij}) = 0$$

$$V(\epsilon_{ij}) = \sigma^2 \quad V(A_i) = \sigma_A^2 \quad (10.7)$$

todos os  $A_i$ s e  $\epsilon_{ij}$ s são normalmente distribuídos e independentes entre si.

A condição  $E(A_i) = 0$  em (10.7) é semelhante à condição  $\sum \alpha_i = 0$  em (10.6); ela estabelece que o efeito esperado ou médio do  $i$ -ésimo nível medido como um desvio de  $\mu$  é zero.

No modelo de efeitos aleatórios (10.7), a hipótese de não haver nenhum efeito em virtude de níveis diferentes é  $H_0: \sigma_A^2 = 0$ , expressando que diferentes níveis do fator em nada contribuem para a variabilidade da resposta. *Embora as hipóteses nos modelos de efeitos fixos e aleatórios de fator único sejam diferentes, elas são testadas exatamente da mesma forma*, calculando  $F = \text{QMTr}/\text{QME}$  e rejeitando  $H_0$ , se  $f \geq F_{\alpha, I-1, n-I}$ . Isso pode ser demonstrado intuitivamente, observando-se que  $E(\text{QME}) = \sigma^2$  (da mesma maneira que para efeitos fixos), ao passo que

$$E(\text{QMTr}) = \sigma^2 + \frac{1}{I-1} \left( n - \frac{\sum J_i^2}{n} \right) \sigma_A^2 \quad (10.8)$$

onde  $J_1, J_2, \dots, J_l$  são tamanhos das amostras e  $n = \sum J_i$ . O fator entre parênteses no lado direito de (10.8) é não-negativo, de maneira que novamente  $E(QMTr) = \sigma^2$ , se  $H_0$  for verdadeiro, e  $E(QMTr) > \sigma^2$ , se  $H_0$  for falso.

**Exemplo 10.10**

O estudo sobre forças e pressões não-destrutivas em materiais oferece informações importantes para projetos de engenharia eficazes. O artigo "Zero-Force Travel-Time Parameters for Ultrasonic Head-Waves in Railroad Rail" (*Materials Evaluation*, 1985 p. 854-858) relata um estudo sobre o tempo de percurso de certo tipo de onda resultante da pressão longitudinal dos trilhos usados nas estradas de ferro. Três medidas foram feitas em cada um dos trilhos selecionados aleatoriamente de uma população do referido grupo. Os investigadores usaram ANOVA de efeitos aleatórios para determinar se algumas variações no tempo de percurso podiam ser atribuídas a uma "variabilidade entre trilhos". Os dados são fornecidos na tabela a seguir (todos os valores, em nanossegundos, resultaram da subtração de 36,1  $\mu s$  da observação original) com a tabela ANOVA derivada. O valor da razão de  $F$  é altamente significativo, de modo que  $H_0: \sigma_A^2 = 0$  é rejeitada em favor da conclusão de que as diferenças entre trilhos são uma fonte de variabilidade no tempo de percurso.

				$x_i$
1:	55	53	54	162
2:	26	37	32	95
3:	78	91	85	254
4:	92	100	96	288
5:	49	51	50	150
6:	80	85	83	<u>248</u>
				$x_{..} = 1197$

Fonte de Variação	gl	Soma dos Quadrados	Média dos Quadrados	$f$
Tratamentos	5	9310,5	1862,1	115,2
Erro	12	194,0	16,17	
Total	17	9504,5		



**Exercícios** | Seção 10.3 (22–34)

22. Os dados a seguir referem-se à produção de tomates (kg/lote) para quatro níveis diferentes de salinidade; aqui, o nível de salinidade refere-se à condutividade elétrica (CE), onde os níveis escolhidos foram CE = 1,6; 3,8; 6,0 e 10,2 nmhos/cm:

1,6	59,5	53,3	56,8	63,1	58,7
3,8	55,2	59,1	52,8	54,5	
6,0	51,7	48,8	53,9	49,0	
10,2	44,6	48,5	41,0	47,3	46,1

Use o teste  $F$  no nível  $\alpha = 0,05$  para testar quaisquer diferenças na produção média real provocadas por níveis de salinidade diferentes.

23. Aplique o método de Tukey modificado para os dados do Exercício 22, a fim de identificar diferenças significativas entre os  $\mu_i$ .

24. A tabela ANOVA parcial apresentada a seguir é reproduzida do artigo "Perception of Spatial Incongruity" (*J. Nervous and Mental Disease*, 1961 p. 222), no qual a capacidade de cada um dos três diferentes grupos para identificar uma incongruência de percepção foi analisada e comparada. Todos os indivíduos no experimento foram hospitalizados para receber tratamento psiquiátrico. Havia 21 indivíduos no grupo depressivo, 32 no grupo funcional classificado como "outros" e 21 indivíduos no grupo com lesões cerebrais. Complete a tabela ANOVA e realize o teste  $F$  no nível  $\alpha = 0,01$ .

Fonte	gl	Soma dos Quadrados	Média dos Quadrados	$f$
Grupos			76,09	
Erro				
Total		1.123,14		

25. Os lipídios fornecem a maior parte da energia diária de bebês e crianças. Há um interesse cada vez maior pela qualidade do suprimento diário de lipídios durante a infância por ser o principal determinante do crescimento, do desenvolvimento neural, da visão e da saúde a longo prazo. O artigo "Essential Fat Requirements of Preterm Infants" (*Amer. J. of Clinical Nutrition*, 2000 p. 245S-250S) relatou os dados a seguir sobre a porcentagem total de gordura poliinsaturada para bebês classificados aleatoriamente em quatro diferentes dietas alimentares: leite do peito, fórmula à base de Óleo de Milho (OM), a base de Óleo de Soja (OS), ou à base de Óleo de Soja e de Peixe (OSP):

Dieta	Tamanho da Amostra	Média Amostral	DP Amostral
Leite do peito	8	43,0	1,5
OM	13	42,4	1,3
OS	17	43,1	1,2
OSP	14	43,5	1,2

- a. Que suposições devem ser feitas sobre o total das quatro distribuições de gordura poliinsaturada antes de realizar uma ANOVA de fator único para determinar se há alguma diferença no teor de gordura médio real?
- b. Faça o teste sugerido na parte (a). O que se pode dizer do valor  $P$ ?

26. Amostras de seis diferentes marcas de imitação da margarina *diet* foram analisadas para determinar o nível de ácidos graxos poliinsaturados fisiologicamente ativos (AGPFA, em porcentagem), gerando os seguintes resultados:

<i>Imperial</i>	14,1	13,6	14,4	14,3	
<i>Parkay</i>	12,8	12,5	13,4	13,0	12,3
<i>Blue Bonnet</i>	13,5	13,4	14,1	14,3	
<i>Chiffon</i>	13,2	12,7	12,6	13,9	
<i>Mazola</i>	16,8	17,2	16,4	17,3	18,0
<i>Fleischmann's</i>	18,1	17,2	18,7	18,4	

(Os números anteriores são fictícios, mas as médias amostrais estão de acordo com os dados relatados na edição de janeiro de 1975 do periódico *Consumer Reports*.)

- a. Use ANOVA para testar as diferenças entre as porcentagens médias reais de AGPFA das diferentes marcas.
- b. Calcule o IC para todas as diferenças  $(\mu_i - \mu_j)$ .

c. As margarinas Mazola e Fleischmann's são à base de milho, ao passo que as outras são à base de soja. Calcule o IC para

$$(\mu_1 + \mu_2 + \mu_3 + \mu_4) - (\mu_5 + \mu_6)$$

[Sugestão: modifique a expressão de  $V(\hat{\theta})$  que levou à (10.5) na seção anterior.]

27. Embora, depois da água, o chá seja a bebida mais amplamente consumida do mundo, pouco se sabe sobre seu valor nutricional. Folacina é a única vitamina B presente em qualquer quantidade significativa de chá. Avanços recentes nos métodos de teste possibilitaram a determinação precisa do teor de folacina. Considere os dados a seguir do teor de folacina em amostras aleatoriamente selecionadas de quatro marcas predominantes de chá verde.

Marca	Observações							
1	7,9	6,2	6,6	8,6	8,9	10,1	9,6	
2	5,7	7,5	9,8	6,1	8,4			
3	6,8	7,5	5,0	7,4	5,3	6,1		
4	6,4	7,1	7,9	4,5	5,0	4,0		

(Esses dados baseiam-se no artigo "Folacin Content of Tea", *J. Amer. Dietetic Assoc.*, 1983 p. 627-632.) Será que eles sugerem que o teor médio real de folacina é o mesmo em todas as marcas?

- a. Faça um teste usando  $\alpha = 0,05$  pelo método do valor  $P$ .
- b. Avalie a plausibilidade de qualquer suposição necessária para a sua análise na parte (a).
- c. Faça uma análise de comparações múltiplas para identificar diferenças significativas entre as marcas.

28. Para uma ANOVA de fator único com tamanhos das amostras  $J_i$  ( $i = 1, 2, \dots, I$ ), demonstre que  $SQTr = \sum J_i(\bar{X}_i - \bar{X}_{..})^2 = \sum_i J_i \bar{X}_i^2 - n\bar{X}_{..}^2$ , onde  $n = \sum J_i$ .

29. Quando os tamanhos das amostras são iguais ( $J_i = J$ ), os parâmetros  $\alpha_1, \alpha_2, \dots, \alpha_I$  da parametrização alternativa são limitados por  $\sum \alpha_i = 0$ . Para tamanhos das amostras desiguais, a restrição mais natural é  $\sum J_i \alpha_i = 0$ . Use isso para mostrar que

$$E(QMTr) = \sigma^2 + \frac{1}{I-1} \sum J_i \alpha_i^2$$

Qual é o  $E(QMTr)$  quando  $H_0$  for verdadeiro? [Esse valor esperado é correto, se  $\sum J_i \alpha_i = 0$  for substituída pela restrição  $\sum \alpha_i = 0$  (ou por qualquer outra restrição linear exclusiva nos  $\alpha_i$ s usados para reduzir o modelo a  $I$  parâmetros independentes), mas  $\sum J_i \alpha_i = 0$  simplifica os cálculos e produz estimativas naturais dos parâmetros do modelo (particularmente,  $\hat{\alpha}_i = \bar{x}_i - \bar{x}_{..}$ ).]

30. Reconsidere o Exemplo 10.7 sobre a investigação dos efeitos de diferentes tratamentos térmicos sobre o ponto de ruptura de barras de aço.
- Se  $J = 8$  e  $\sigma = 1$ , qual é o  $\beta$  para um nível de teste  $F$  de 0,05 quando  $\mu_1 = \mu_2$ ,  $\mu_3 = \mu_1 - 1$ , e  $\mu_4 = \mu_1 + 1$ ?
  - Para a alternativa da parte (a), que valor de  $J$  é necessário para obter  $\beta = 0,05$ ?
  - Se houver  $I = 5$  tratamentos térmicos,  $J = 10$  e  $\sigma = 1$ , qual será o  $\beta$  para um nível de teste  $F$  0,05 quando quatro dos  $\mu_i$ s são iguais e a diferença do quinto em relação aos outros quatro é 1?
31. Quando os tamanhos das amostras não são iguais, o parâmetro de não-centralidade é  $\sum J_i \alpha_i^2 / \sigma^2$  e  $\phi^2 = (1/I) \sum J_i \alpha_i^2 / \sigma^2$ . Em relação ao Exercício 22, qual é o poder do teste quando  $\mu_2 = \mu_3$ ,  $\mu_1 = \mu_2 - \sigma$ , e  $\mu_4 = \mu_2 + \sigma$ ?
32. Em um experimento para comparar a qualidade de quatro diferentes marcas de fitas de gravação magnética, foram selecionados cinco carretéis de 2400 pés de cada marca (A-D), determinando-se em seguida o número de defeitos.

A:	10	5	12	14	8
B:	14	12	17	9	8
C:	13	18	10	15	18
D:	17	16	12	22	14

Acredita-se que o número de defeitos tenha, aproximadamente, uma distribuição Poisson para cada marca. Analise os dados ao nível de significância 0,01 para verificar se o número esperado de defeitos por carretel é o mesmo para cada marca.

33. Suponha que  $X_{ij}$  seja uma variável binomial com parâmetros  $n$  e  $p_i$  (portanto, aproximadamente normal quando  $np_i \geq 5$  e  $nq_i \geq 5$ ). Então, visto que  $\mu_i = np_i$ ,  $V(X_{ij}) = \sigma_i^2 = np_i(1 - p_i) = \mu_i(1 - \mu_i/n)$ , como os  $X_{ij}$ s devem ser transformados para estabilizar a variância? [Sugestão:  $g(\mu_i) = \mu_i(1 - \mu_i/n)$ .]
34. Simplifique  $E(QMTr)$  para o modelo de efeitos aleatórios quando  $J_1 = J_2 = \dots = J_I = J$ .

Exercícios Suplementares (35–46)

35. Um experimento foi feito para comparar as taxas de fluxo de quatro diferentes tipos de bocal.
- Os tamanhos das amostras foram 5, 6, 7 e 6, respectivamente e os cálculos geraram  $f = 3,68$ . Expresse e teste as hipóteses de interesse usando  $\alpha = 0,01$ .
  - A análise dos dados usando um *software* estatístico forneceu um valor  $P = 0,029$ . No nível 0,01, o que você concluiria e por quê?
36. O artigo "Computer-Assisted Instruction Augmented with Planned Teacher/Student Contacts" (*J. Exp. Educ.*, inverno, 1980-1981, p. 120-126) comparou cinco diferentes métodos de ensino de estatística descritiva. Os cinco métodos foram: Aula e Discussão tradicionais (A/D), instrução programada por meio de Livros de referência (L), texto programado com aulas (L/A), instrução por Computador (C) e instrução por Computador com Aula (C/A). Quarenta e cinco alunos foram escolhidos aleatoriamente, 9 para cada método. Após a conclusão do curso, os alunos foram submetidos a um exame de 1 hora de duração. Além disso, um teste de memória de 10 minutos foi administrado seis semanas depois. Dados resumidos estão na tabela a seguir:

Método	Exame		Teste de Memória	
	$\bar{x}_i$	$s_i$	$\bar{x}_i$	$s_i$
A/D	29,3	4,99	30,20	3,82
L	28,0	5,33	28,80	5,26
L/A	30,2	3,33	26,20	4,66
C	32,4	2,94	31,10	4,91
C/A	34,2	2,74	30,20	3,53

A média geral do exame foi 30,82 e a média geral do teste de memória foi 29,30.

- Os dados sugerem que há uma diferença entre os cinco métodos de ensino em relação à pontuação média real do exame? Use  $\alpha = 0,05$ .
  - Usando um nível de significância de 0,05, teste a hipótese nula de que não há nenhuma diferença entre as pontuações médias reais do teste de memória para os cinco diferentes métodos de ensino.
37. Inúmeros fatores contribuem para o funcionamento silencioso de um motor elétrico ("Increasing Market Share Through Improved Product and Process Design: An Experimental Approach", *Quality Engineering*, 1991, p. 361-369). É particularmente desejável reduzir ao máximo o ruído e a vibração do motor. Para estudar o efeito que a marca do mancal tem sobre a vibração do motor, foram examinadas cinco marcas diferentes de mancal de motor, instalando-se cada tipo de mancal em diferentes amostras aleatórias de seis motores. A quantidade de vibração do motor (medida em microns) foi registrada durante o funcionamento de cada um dos 30 motores. Apresentam-se a seguir os dados para esse estudo. Expresse e teste as hipóteses de interesse no nível de significância de 0,05 e, em seguida, efetue uma análise de comparações múltiplas, se apropriado.

	Média						
Marca 1	13,1	15,0	14,0	14,4	14,0	11,6	13,68
Marca 2	16,3	15,7	17,2	14,9	14,4	17,2	15,95
Marca 3	13,7	13,9	12,4	13,8	14,9	13,3	13,67
Marca 4	15,7	13,7	14,4	16,0	13,9	14,7	14,73
Marca 5	13,5	13,4	13,2	12,7	13,4	12,3	13,08

38. Um artigo publicado no periódico científico britânico *Nature* ("Sucrose Induction of Hepatic Hyperplasia in the Rat", 25 de agosto de 1972, p. 461) relata um experimento no qual cinco grupos compostos por seis ratos cada foram submetidos a uma dieta com diferentes carboidratos. Ao final do experimento, determinou-se o teor de DNA (mg/g do fígado) no fígado de cada rato, gerando os seguintes resultados:

Carboidrato	$\bar{x}_i$
Amido	2,58
Sacarose	2,63
Frutose	2,13
Glicose	2,41
Maltose	2,49

Supondo ainda que  $\sum \sum x_{ij}^2 = 183,4$ , os dados indicam que o teor de DNA médio real é influenciado pelo tipo de carboidrato na dieta? Faça um tabela ANOVA e use um nível de significância de 0,05.

39. Referindo-se ao Exercício 38, crie um IC  $t$  para

$$\theta = \mu_1 - (\mu_2 + \mu_3 + \mu_4 + \mu_5)/4$$

que meça a diferença entre o teor médio de DNA da dieta de amido e a média conjunta das outras quatro dietas. O intervalo resultante inclui zero?

40. Refira-se ao Exercício 38. Qual o  $\beta$  do teste quando o teor médio real de DNA é idêntico para três das dietas e fica 1 desvio padrão ( $\sigma$ ) abaixo desse valor comum para as outras duas dietas?

41. Quatro laboratórios (1-4) são selecionados aleatoriamente de uma ampla população e cada um é solicitado a determinar três porcentagens de álcool metílico em amostras de um composto extraído de um único lote. Com base nos dados a seguir, as diferenças entre os laboratórios são uma fonte de variação na porcentagem de álcool metílico? Expresse e teste as hipóteses de interesse, usando um nível de significância de 0,05.

1:	85,06	85,25	84,87
2:	84,99	84,28	84,88
3:	84,48	84,72	85,10
4:	84,10	84,55	84,05

42. A frequência crítica de oscilação (fco) é a frequência mais alta (em ciclos/s) na qual uma pessoa pode detectar a oscilação em uma fonte de luz oscilante. Em frequências acima de fco, a fonte de luz parece ser contínua, embora na verdade esteja oscilante. Uma investigação realizada para verificar se a fco média real depende da cor da íris gerou os seguintes dados (com base no artigo "The Effects of Iris Color on Critical Flicker Frequency", *J. General Psych.*, 1973, p. 91-95):

	Cor da Íris		
	1. Marrom	2. Verde	3. Azul
	26,8	26,4	25,7
	27,9	24,2	27,2
	23,7	28,0	29,9
	25,0	26,9	28,5
	26,3	29,1	29,4
	24,8		28,3
	25,7		
	24,5		
$J_i$	8	5	6
$x_i$	204,7	134,6	169,0
$\bar{x}_i$	25,59	26,92	28,17

$n = 19$      $x_{..} = 508,3$

a. Expresse e teste as hipóteses de interesse no nível de significância 0,05 usando a tabela  $F$  para obter um limite superior e/ou inferior no valor  $P$ . (Sugestão:  $\sum \sum x_{ij}^2 = 13.659,67$  e  $OC = 13.598,36$ .)

b. Investigue as diferenças entre as cores da íris em relação à fco média.

43. Sejam  $c_1, c_2, \dots, c_I$  os números que satisfazem  $\sum c_i = 0$ . Então,  $\sum c_i \mu_i = c_1 \mu_1 + \dots + c_I \mu_I$  é chamado de *contraste* nos  $\mu_i$ s. Observe que, com  $c_1 = 1, c_2 = -1, c_3 = \dots = c_I = 0, \sum c_i \mu_i = \mu_1 - \mu_2$ , o que significa que toda diferença par a par entre  $\mu_i$  é um contraste (da mesma forma que, por exemplo,  $\mu_1 - 0,5\mu_2 - 0,5\mu_3$ ). Um método atribuído a Scheffé fornece ICs simultâneos com nível de confiança simultâneo  $100(1 - \alpha)\%$  para *todos* os contrastes possíveis (um número infinito deles!). O intervalo para  $\sum c_i \mu_i$  é

$$\sum c_i \bar{x}_i \pm (\sum c_i^2 / J_i)^{1/2} \cdot [(I - 1) \cdot QME \cdot F_{\alpha, I-1, n-I}]^{1/2}$$

Usando os dados sobre frequência crítica de oscilação do Exercício 42, calcule os intervalos de Scheffé para os contrastes  $\mu_1 - \mu_2, \mu_1 - \mu_3, \mu_2 - \mu_3$ , e  $0,5\mu_1 + 0,5\mu_2 - \mu_3$  (o último contraste compara o azul com a média de marrom e verde). Que contrastes parecem diferir significativamente de 0 e por quê?

44. Quatro tipos de argamassa – de cimento comum (ACC), composta de polímero (ACP), de resina (AR) e composta de cimento e polímero (ACP) – foram submetidas a um teste de compressão para medir a resistência (MPa). Três observações de resistência para cada tipo de argamassa são apresentadas no artigo "Polymer Mortar Composite Matrices for Maintenance-Free Highly Durable Ferrocement" (*J. Ferrocement*, 1984 p. 337-345) e são reproduzidas aqui. Elabore uma tabela ANOVA. Usando um nível de significância de 0,05, determine se os dados sugerem que a resistência média real não é a mesma para os quatro tipos de argamassa. Se concluir que todas as resistências médias reais não são iguais, use o método de Tukey para identificar as diferenças significativas.

ACC	32,15	35,53	34,20
ACP	126,32	126,80	134,79
AR	117,91	115,02	114,58
ACP	29,09	30,87	29,80

45. Suponha que os  $x_{ij}$ s sejam "codificados" por  $y_{ij} = cx_{ij} + d$ . Até que ponto o valor da estatística  $F$

calculada compara-se ao valor calculado pelos  $x_{ij}$ s? Justifique sua afirmação.

46. No Exemplo 10.10, subtraia  $\bar{x}_i$  de cada observação na  $i$ -ésima amostra ( $i = 1, \dots, 6$ ) para obter um conjunto de 18 resíduos. Em seguida, elabore um gráfico de probabilidade normal e comente a plausibilidade da suposição de normalidade.

---

## Bibliografia

MILLER, R. *Beyond ANOVA: The Basics of Applied Statistics*. Nova York: Wiley, 1986. Uma excelente fonte de informação sobre verificação de suposições e métodos alternativos de análise.

MONTGOMERY, D. *Design and Analysis of Experiments*. Nova York: Wiley, 2001. 5 ed. Uma apresentação extremamente atualizada de modelos e da metodologia ANOVA.

NETER, J., WASSERMAN, W., KUTNER, M. *Applied Linear Statistical Models*. Homewood, IL, Irwin, 1996. 4 ed. A segunda metade desse livro contém uma pesquisa

muito bem apresentada sobre a ANOVA; seu nível é comparável ao do presente texto, mas a discussão é mais abrangente, o que torna a obra uma excelente fonte de consulta.

OTT, R. L. e LONGNECKER, M. *An Introduction to Statistical Methods and Data Analysis*. Belmont, CA: Duxbury Press, 2001. 5 ed. Inclui vários capítulos sobre a metodologia ANOVA, que pode ser lido por estudantes que desejam uma exposição substancialmente não-matemática; contém um ótimo capítulo sobre vários métodos de comparação múltipla.

# ANOVA

## Multifatores

### Introdução

No capítulo anterior, usamos a análise de variância (ANOVA) para testar a igualdade tanto de  $I$  médias populacionais diferentes como as respostas médias reais correspondentes aos  $I$  níveis diferentes de um fator único (alternativamente chamados de  $I$  tratamentos diferentes). Em várias situações experimentais, existem dois ou mais fatores concomitantemente importantes. Este capítulo dá continuidade aos métodos do Capítulo 10 para investigar essas situações.

Nas duas primeiras seções, concentramo-nos no caso de dois fatores de interesse. Usaremos  $I$  para representar o número de níveis do primeiro fator ( $A$ ) e  $J$  para representar o número de níveis do segundo fator ( $B$ ). Portanto, há  $IJ$  combinações possíveis que consistem de um nível do fator  $A$  e um do fator  $B$ ; cada uma dessas combinações é chamada de tratamento, de modo que há  $IJ$  tratamentos diferentes. O número de observações feitas no tratamento  $(i, j)$  será representado por  $K_{ij}$ . Na seção 11.1, apresentamos o modelo e a análise quando  $K_{ij} = 1$ . Um caso de particular importância desse tipo é o modelo em blocos aleatorizados, no qual o fator único  $A$  é o mais importante, mas outro fator, "blocos", é criado para controlar variabilidades externas em unidades ou sujeitos experimentais. Na Seção 11.2, concentramo-nos no caso  $K_{ij} = K > 1$  e apresentamos brevemente as dificuldades associadas a  $K_{ij}$ s desiguais.

A Seção 11.3 avalia experimentos envolvendo mais de dois fatores, incluindo o modelo quadrado latino, que controla os efeitos de dois fatores externos considerados influentes sobre a variável de resposta. Quando o número de fatores é grande, um experimento constituído de pelo menos uma observação para cada tratamento é caro e trabalhoso. Um caso especial importante que discutimos na Seção 11.4 é aquele em que existem  $p$  fatores com dois níveis cada. Logo, existem  $2^p$  tratamentos diferentes. Avaliamos tanto o caso em que se fazem observações sobre todos esses tratamentos (projeto completo) como o caso em que se fazem observações apenas para um subconjunto de tratamentos selecionado (projeto incompleto).



## 11.1 ANOVA de Dois Fatores com $K_{ij} = 1$

Quando o fator  $A$  constitui-se de  $I$  níveis e o fator  $B$  de  $J$  níveis, existem  $IJ$  diferentes combinações (pares) de níveis de dois fatores, cada uma denominada tratamento. Sendo  $K_{ij} =$  o número de observações do tratamento constituído pelo fator  $A$  no nível  $i$  e pelo fator  $B$  no nível  $j$ , nesta seção concentramo-nos no caso de  $K_{ij} = 1$ , de modo que os dados consistem em  $IJ$  observações. Primeiro, discutiremos o modelo de efeitos fixos, no qual os únicos níveis de interesse para os dois fatores são aqueles de fato representados no experimento. O caso em que um ou ambos os fatores são aleatórios é discutido brevemente no fim da seção.

### Exemplo 11.1

Será que é tão fácil remover manchas de caneta apagável de tecidos quanto a palavra *apagável* pretende significar? Considere os dados a seguir de um experimento para comparar três diferentes marcas de caneta e quatro diferentes tratamentos de lavagem em relação à capacidade de remover manchas em um determinado tipo de tecido (baseado em “An Assessment of the Effects of Treatment, Time, and Heat on the Removal of Erasable Pen Marks from Cotton and Cotton/Polyester Blend Fabrics”, *J. of Testing and Evaluation*, 1991, p. 394-397). A variável da resposta é um indicador quantitativo da mudança global da cor da amostra; quanto menor o valor, mais manchas foram removidas.

		Tratamento de Lavagem				Total
		1	2	3	4	
Marca da Caneta	1	0,97	0,48	0,48	0,46	2,39
	2	0,77	0,14	0,22	0,25	1,38
	3	0,67	0,39	0,57	0,19	1,82
Total		2,41	1,01	1,27	0,90	5,59

Há alguma diferença na quantidade média real de mudança de cor devido às diferentes marcas de caneta ou aos diferentes tratamentos de lavagem? ■

Como na ANOVA de fator único, os subscritos duplos são usados para identificar variáveis aleatórias e valores observados. Sejam

$X_{ij}$  = variável aleatória (*va*) que representa a medida quando o fator  $A$  é mantido no nível  $i$  e o fator  $B$  no nível  $j$

$x_{ij}$  = valor observado de  $X_{ij}$

Os  $x_{ij}$ s normalmente são apresentados em uma tabela bidirecional na qual a  $i$ -ésima linha contém os valores observados quando o fator  $A$  é mantido no nível  $i$  e a  $j$ -ésima coluna contém os valores observados quando o fator  $B$  é mantido no nível  $j$ . No experimento da caneta apagável do Exemplo 11.1, o número de níveis do fator  $A$  é  $I = 3$ , o número de níveis do fator  $B$  é  $J = 4$ ,  $x_{13} = 0,48$ ,  $x_{22} = 0,14$ , e assim por diante.

Embora na ANOVA de fator único estivéssemos interessados apenas nas médias da linha e na média geral, aqui também estamos interessados nas médias da coluna. Sejam

$$\begin{aligned} \bar{X}_{i.} &= \text{m\u00e9dia das medidas obtidas quando o} &= \frac{\sum_{j=1}^J X_{ij}}{J} \\ &\quad \text{fator } A \text{ \u00e9 mantido no n\u00edvel } i \\ \bar{X}_{.j} &= \text{m\u00e9dia das medidas obtidas quando o} &= \frac{\sum_{i=1}^I X_{ij}}{I} \\ &\quad \text{fator } B \text{ \u00e9 mantido no n\u00edvel } j \\ \bar{X}_{..} &= \text{m\u00e9dia geral} &= \frac{\sum_{i=1}^I \sum_{j=1}^J X_{ij}}{IJ} \end{aligned}$$

com valores observados  $\bar{x}_{i.}$ ,  $\bar{x}_{.j}$  e  $\bar{x}_{..}$ . Os totais, e não as médias, são representados por meio da omissão da barra horizontal (portanto,  $x_{.j} = \sum_i x_{ij}$  etc.). Para examinarmos, intuitivamente, se há algum efeito decorrente dos níveis do fator  $A$ , devemos comparar os  $\bar{x}_{i.}$ s observados entre si; as informações sobre os diferentes níveis do fator  $B$  devem ser produzidas pelos  $\bar{x}_{.j}$ s.

## O Modelo

Procedendo analogamente à ANOVA de fator único, nossa primeira inclinação ao especificar um modelo é considerar  $\mu_{ij}$  = resposta média real quando o fator  $A$  está no nível  $i$  e o fator  $B$  no nível  $j$ , gerando  $IJ$  parâmetros médios. Então, seja

$$X_{ij} = \mu_{ij} + \epsilon_{ij}$$

onde  $\epsilon_{ij}$  é a quantidade aleatória pela qual os valores observados diferem de seus valores esperados; os  $\epsilon_{ij}$ s são considerados normais e independentes com variância comum  $\sigma^2$ . Infelizmente, não existe procedimento de teste válido para essa escolha de parâmetros. A razão é que, para a hipótese de interesse alternativa, os  $\mu_{ij}$ s são livres para assumir quaisquer valores, enquanto  $\sigma^2$  pode ter qualquer valor maior do que zero; há portanto,  $IJ + 1$  parâmetros que variam livremente. No entanto, existem apenas  $IJ$  observações; desse modo, depois de usar cada  $x_{ij}$  como estimativa de  $\mu_{ij}$ , não há como estimar  $\sigma^2$ .

Para resolver esse problema de um modelo que tenha mais parâmetros do que valores observados, devemos especificar um modelo que, embora realista, envolva relativamente poucos parâmetros.

Suponha que existam  $I$  parâmetros  $\alpha_1, \alpha_2, \dots, \alpha_I$  e  $J$  parâmetros  $\beta_1, \beta_2, \dots, \beta_J$ , de modo que

$$X_{ij} = \alpha_i + \beta_j + \epsilon_{ij} \quad (i = 1, \dots, I, \quad j = 1, \dots, J) \quad (11.1)$$

portanto,

$$\mu_{ij} = \alpha_i + \beta_j \quad (11.2)$$

Incluindo  $\sigma^2$ , existem agora  $I + J + 1$  parâmetros no modelo, de modo que, se  $I \geq 3$  e  $J \geq 3$ , haverá menos parâmetros que observações (de fato, em breve modificaremos (11.2) para que mesmo  $I = 2$  e/ou  $J = 2$  se adaptem).

O modelo especificado em (11.1) e (11.2) denomina-se **modelo aditivo** porque toda resposta média  $\mu_{ij}$  é a soma de um efeito decorrente do fator  $A$  no nível  $i$  ( $\alpha_i$ ) e um efeito decorrente do fator  $B$  no nível  $j$  ( $\beta_j$ ). A diferença entre as respostas médias para o fator  $A$  no nível  $i$  e no nível  $i'$  quando  $B$  é mantido no nível  $j$  é  $\mu_{ij} - \mu_{i'j}$ . Quando o modelo é aditivo,

$$\mu_{ij} - \mu_{i'j} = (\alpha_i + \beta_j) - (\alpha_{i'} + \beta_j) = \alpha_i - \alpha_{i'}$$

que é independente do nível  $j$  do segundo fator. Um resultado semelhante mantém-se para  $\mu_{ij} - \mu_{ij'}$ . Logo, aditividade significa que a diferença nas respostas médias nos dois níveis de um dos fatores é a mesma para todos os níveis do outro fator. A Figura 11.1(a) mostra um conjunto de respostas médias que satisfazem a condição de aditividade, e a Figura 11.1(b) mostra uma disposição não-aditiva de respostas médias.

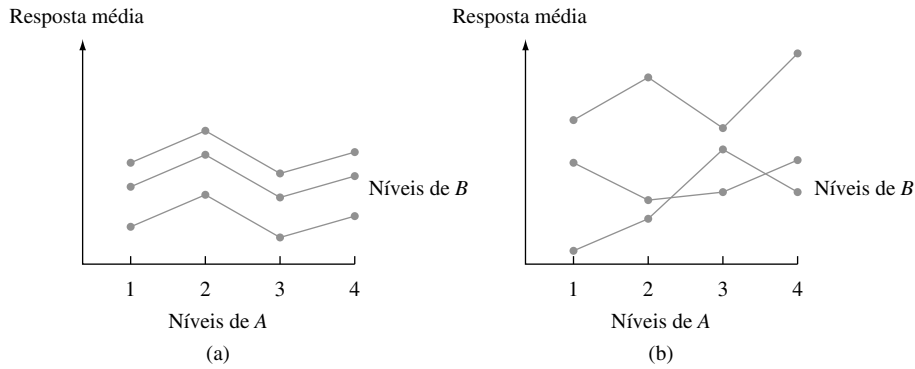


Figura 11.1 Respostas médias de dois tipos de modelo: (a) aditivo; (b) não-aditivo

Exemplo 11.2 (Continuação do Exemplo 11.1)

Quando plotamos os  $x_{ij}$ s observados de modo análogo ao da Figura 11.1, obtemos o resultado mostrado na Figura 11.2. Embora haja alguns “cruzamentos” nos  $x_{ij}$ s observados, a configuração é razoavelmente representativa do que se espera da aditividade com apenas uma observação por tratamento.

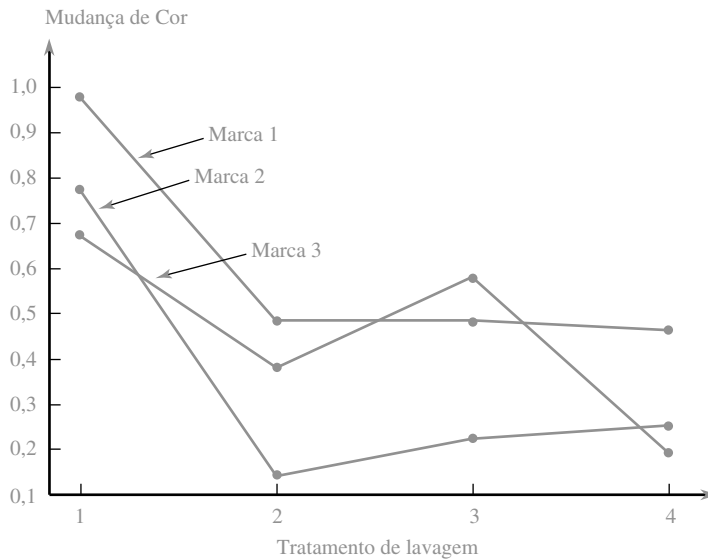


Figura 11.2 Gráfico dos dados do Exemplo 11.1

A expressão (11.2) não é bem a descrição final do modelo porque os  $\alpha_i$ s e  $\beta_j$ s não são determinados unicamente. A seguir encontram-se duas diferentes configurações dos  $\alpha_i$ s e dos  $\beta_j$ s que geram os mesmos  $\mu_{ij}$ s aditivos.

	$\beta_1 = 1$	$\beta_2 = 4$
$\alpha_1 = 1$	$\mu_{11} = 2$	$\mu_{12} = 5$
$\alpha_2 = 2$	$\mu_{21} = 3$	$\mu_{22} = 6$

	$\beta_1 = 2$	$\beta_2 = 5$
$\alpha_1 = 0$	$\mu_{11} = 2$	$\mu_{12} = 5$
$\alpha_2 = 1$	$\mu_{21} = 3$	$\mu_{22} = 6$

Subtraindo qualquer constante  $c$  de todos os  $\alpha_i$ s e adicionando  $c$  a todos os  $\beta_j$ s, obtêm-se outras configurações correspondentes ao mesmo modelo aditivo. Essa não-unicidade é eliminada usando-se o modelo a seguir.

$$X_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij} \quad (11.3)$$

onde  $\sum_{i=1}^I \alpha_i = 0$ ,  $\sum_{j=1}^J \beta_j = 0$ , e os  $\epsilon_{ij}$ s são considerados independentes, normalmente distribuídos com média 0 e variância comum  $\sigma^2$ .

Isso é análogo à escolha alternativa de parâmetros na ANOVA de fator único discutida na Seção 10.3. Não é difícil observar que (11.3) é um modelo aditivo no qual os parâmetros são unicamente determinados (por exemplo: para os  $\mu_{ij}$ s mencionados anteriormente,  $\mu = 4$ ,  $\alpha_1 = -0,5$ ,  $\alpha_2 = 0,5$ ,  $\beta_1 = -1,5$ , e  $\beta_2 = 1,5$ ). Observe que há apenas  $I - 1$   $\alpha_i$ s e  $J - 1$   $\beta_j$ s determinados de maneira independente, de modo que (incluindo  $\mu$ ) (11.3) especifica  $I + J - 1$  parâmetros médios.

A interpretação dos parâmetros de (11.3) é direta:  $\mu$  é a média geral real (resposta média obtida pelo cálculo da média de todos os níveis de ambos os fatores),  $\alpha_i$  é o efeito do fator  $A$  no nível  $i$  (medido como desvio de  $\mu$ ) e  $\beta_j$  é o efeito do fator  $B$  no nível  $j$ . Estimadores não-tendenciosos (e de máxima verossimilhança) para esses parâmetros são

$$\hat{\mu} = \bar{X}_{..} \quad \hat{\alpha}_i = \bar{X}_{i.} - \bar{X}_{..} \quad \hat{\beta}_j = \bar{X}_{.j} - \bar{X}_{..}$$

Há duas diferentes hipóteses de interesse em um experimento de dois fatores com  $K_{ij} = 1$ . A primeira, representada por  $H_{0A}$  expressa que os diferentes níveis do fator  $A$  não exercem nenhum efeito sobre a resposta média real. A segunda, representada por  $H_{0B}$  expressa que não há nenhum efeito do fator  $B$ .

$$\begin{aligned} H_{0A}: \alpha_1 = \alpha_2 = \cdots = \alpha_I = 0 \\ \text{versus } H_{aA}: \text{ pelo menos um } \alpha_i \neq 0 \\ H_{0B}: \beta_1 = \beta_2 = \cdots = \beta_J = 0 \\ \text{versus } H_{aB}: \text{ pelo menos um } \beta_j \neq 0 \end{aligned} \quad (11.4)$$

(Nenhum efeito do fator  $A$  significa que todos os  $\alpha_i$ s são iguais, de maneira que todos devem ser 0, visto que somam 0. O mesmo é verdadeiro para os  $\beta_j$ s.)

## Os Procedimentos de Teste

A descrição e a análise, a partir de agora, seguem estreitamente a da ANOVA de fator único. As somas relevantes dos quadrados e suas formas de cálculo são dadas por

$$\begin{aligned} \text{SQT} &= \sum_{i=1}^I \sum_{j=1}^J (X_{ij} - \bar{X}_{..})^2 = \sum_{i=1}^I \sum_{j=1}^J X_{ij}^2 - \frac{1}{IJ} X_{..}^2 & \text{df} &= IJ - 1 \\ \text{SQA} &= \sum_{i=1}^I \sum_{j=1}^J (\bar{X}_{i.} - \bar{X}_{..})^2 = \frac{1}{J} \sum_{i=1}^I X_{i.}^2 - \frac{1}{IJ} X_{..}^2 & \text{df} &= I - 1 \\ \text{SQB} &= \sum_{i=1}^I \sum_{j=1}^J (\bar{X}_{.j} - \bar{X}_{..})^2 = \frac{1}{I} \sum_{j=1}^J X_{.j}^2 - \frac{1}{IJ} X_{..}^2 & \text{df} &= J - 1 \\ \text{SQE} &= \sum_{i=1}^I \sum_{j=1}^J (X_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X}_{..})^2 & \text{df} &= (I - 1)(J - 1) \end{aligned} \quad (11.5)$$

e a identidade fundamental

$$\text{SQT} = \text{SQA} + \text{SQB} + \text{SQE} \quad (11.6)$$

permite que SQE seja determinado por subtração.

A expressão de SQE resulta da substituição de  $\mu$ ,  $\alpha_i$  e  $\beta_j$  em  $\sum[X_{ij} - (\mu + \alpha_i + \beta_j)]^2$  por seus estimadores. O gl do erro é  $IJ - \text{número de parâmetros médios estimados} = IJ - [1 + (I - 1) + (J - 1)] = (I - 1)(J - 1)$ . Como na ANOVA de fator único, a variação total é dividida em uma parte (SQE) que não é explicada, independentemente de  $H_{0A}$  e  $H_{0B}$  serem verdadeiras ou falsas, e em duas partes que podem ser explicadas por uma possível falsidade das duas hipóteses nulas.

A teoria estatística estabelece que, se formarmos  $F$  razões (ou quocientes) como na ANOVA de fator único, quando  $H_{0A}$  ( $H_{0B}$ ) é verdadeira, a razão  $F$  correspondente terá uma distribuição  $F$  com  $I - 1$  ( $J - 1$ ) gl no numerador e  $(I - 1)(J - 1)$  gl no denominador.

Hipóteses	Valor do Teste Estatístico	Região de Rejeição
$H_{0A}$ versus $H_{aA}$	$f_A = \frac{QMA}{QME}$	$f_A \geq F_{\alpha, I-1, (I-1)(J-1)}$
$H_{0B}$ versus $H_{aB}$	$f_B = \frac{QMB}{QME}$	$f_B \geq F_{\alpha, J-1, (I-1)(J-1)}$

**Exemplo 11.3** (Continuação do Exemplo 11.2)

Os  $x_{i.}$ s (totais de linha) e os  $x_{.j}$ s (totais de coluna) dos dados de mudança de cor são exibidos nas margens direita e inferior da tabela de dados apresentada previamente. Além disso,  $\sum\sum x_{ij}^2 = 3,2987$  e o fator de correção é  $x_{..}^2/(IJ) = (5,59)^2/12 = 2,6040$ . As somas dos quadrados são, portanto,

$$SQT = 3,2987 - 2,6040 = 0,6947$$

$$SQA = \frac{1}{4}[(2,39)^2 + (1,38)^2 + (1,82)^2] - 2,6040 = 0,1282$$

$$SQB = \frac{1}{3}[(2,41)^2 + (1,01)^2 + (1,27)^2 + (0,90)^2] - 2,6040 = 0,4797$$

$$SQE = 0,6947 - (0,1282 + 0,4797) = 0,0868$$

A tabela ANOVA a seguir (Tabela 11.1) apresenta um resumo dos cálculos.

**Tabela 11.1** Tabela ANOVA do Exemplo 11.3

Fonte de Variação	gl	Soma dos Quadrados	Quadrado Médio	$f$
Fator A (marca)	$I - 1 = 2$	SQA = 0,1282	QMA = 0,0641	$f_A = 4,43$
Fator B (tratamento de lavagem)	$J - 1 = 3$	SQB = 0,4797	QMB = 0,1599	$f_B = 11,05$
Erro	$(I - 1)(J - 1) = 6$	SQE = 0,0868	QME = 0,01447	
Total	$IJ - 1 = 11$	SQT = 0,6947		

O valor crítico para teste de  $H_{0A}$  a um nível de significância de 0,05 é de  $F_{0,05,2,6} = 5,14$ . Visto que  $4,43 < 5,14$ ,  $H_{0A}$  não pode ser rejeitada ao nível de significância 0,05. A mudança de cor média real não parece depender da marca da caneta. Pelo fato de  $F_{0,05,3,6} = 4,76$  e  $11,05 \geq 4,76$ ,  $H_{0B}$  é rejeitada ao nível de significância 0,05, em favor da afirmação de que a mudança de cor varia de acordo com o tratamento de lavagem. Um *software* estatístico fornece os valores  $P$  0,066 e 0,007 para esses dois testes. ■

**Quadrados Médios Esperados**

A plausibilidade do uso dos testes  $F$  que acabamos de descrever é demonstrada calculando-se os quadrados médios esperados. Após alguma álgebra entediante,

$$E(QME) = \sigma^2 \quad (\text{quando o modelo é aditivo})$$

$$E(\text{QMA}) = \sigma^2 + \frac{J}{I-1} \sum_{i=1}^I \alpha_i^2$$

$$E(\text{QMB}) = \sigma^2 + \frac{I}{J-1} \sum_{j=1}^J \beta_j^2$$

Quando  $H_{0A}$  é verdadeira, QMA é um estimador não-tendencioso de  $\sigma^2$ , de modo que  $F$  é uma razão de dois estimadores não-tendenciosos de  $\sigma^2$ . Quando  $H_{0A}$  é falsa, QMA tende a superestimar  $\sigma^2$ ; portanto,  $H_{0A}$  deve ser rejeitada quando a razão  $F$  for muito grande. Raciocínio semelhante se aplica a QMB e  $H_{0B}$ .

## Comparações Múltiplas

Quando  $H_{0A}$  ou  $H_{0B}$  forem rejeitadas, pode-se usar o procedimento de Tukey para identificar diferenças significativas entre os níveis do fator em estudo. As etapas da análise são idênticas às da ANOVA de fator único:

1. Para comparar níveis do fator  $A$ , obtenha  $Q_{\alpha, I, (I-1)(J-1)}$ .  
Para comparar níveis do fator  $B$ , obtenha  $Q_{\alpha, J, (I-1)(J-1)}$ .
2. Calcule

$$w = Q \cdot (\text{desvio padrão estimado das médias amostrais que estão sendo comparadas})$$

$$= \begin{cases} Q_{\alpha, I, (I-1)(J-1)} \cdot \sqrt{\text{QME}/J} & \text{para comparações do fator } A \\ Q_{\alpha, J, (I-1)(J-1)} \cdot \sqrt{\text{QME}/I} & \text{para comparações do fator } B \end{cases}$$

(porque, por exemplo, o desvio padrão de  $\bar{X}_i$  é  $\sigma/\sqrt{J}$ ).

3. Organize as médias amostrais em ordem crescente, sublinhe os pares cuja diferença é menor que  $w$  e identifique pares não-sublinhados pela mesma linha, como se correspondessem a níveis significativamente diferentes do fator especificado.

### Exemplo 11.4 (Continuação do Exemplo 11.3)

Identificação de diferenças significativas entre os quatro tratamentos de lavagem exige que  $Q_{0.05, 4, 6} = 4,90$  e  $w = 4,90 \sqrt{(0,01447)/3} = 0,340$ . As quatro médias amostrais do fator  $B$  (médias das colunas) agora são relacionadas em ordem crescente e qualquer par cuja diferença é menor que 0,340 é sublinhado por um segmento de linha:

$$\begin{array}{cccc} \bar{x}_4 & \bar{x}_2 & \bar{x}_3 & \bar{x}_1 \\ \hline 0,300 & 0,337 & 0,423 & 0,803 \end{array}$$

O tratamento de lavagem 1 parece diferir substancialmente dos outros três tratamentos, mas não se identifica nenhuma outra diferença significativa. Em particular, não está claro entre os tratamentos 2, 3 e 4 qual é melhor para remover manchas. ■

## Experimentos em Blocos Aleatorizados

Ao usar a ANOVA de fator único para testar a existência de efeitos decorrentes de  $I$  tratamentos diferentes em estudo, uma vez que  $IJ$  sujeitos ou unidades experimentais sejam escolhidos, os tratamentos devem ser alocados de maneira completamente aleatória. Isto é,  $J$  sujeitos devem ser escolhidos aleatoriamente para o primeiro tratamento; em seguida, para o segundo tratamento, outra amostra de  $J$  deve ser escolhida de maneira aleatória entre os  $IJ - J$  sujeitos restantes, e assim por diante.

Contudo, não é raro os sujeitos ou as unidades experimentais exibirem heterogeneidade em relação a outras variáveis que podem influenciar as respostas observadas. Quando isso ocorrer, a presença ou a ausência de um valor  $F$  significativo pode ocorrer devido a essa variação externa, e não à presença ou ausência de efeitos do

fator. Esse foi o motivo que nos levou a introduzir um experimento pareado (ou com dados emparelhados) no Capítulo 9. A analogia com um experimento pareado, quando  $I > 2$ , denomina-se experimento em **blocos aleatorizados**. Produz-se um fator externo, "blocos", dividindo-se as  $IJ$  unidades em  $J$  grupos com  $I$  unidades em cada grupo. Esse agrupamento ou formação de blocos se dá de tal maneira que, dentro de cada bloco, as  $I$  unidades sejam homogêneas em relação a outros fatores considerados influentes nas respostas. Portanto, dentro de cada bloco homogêneo, os  $I$  tratamentos são atribuídos aleatoriamente aos  $I$  sujeitos ou unidades no bloco.

### Exemplo 11.5

Uma organização de teste de produtos de consumo queria comparar o consumo de energia anual de cinco marcas diferentes de desumidificador. Pelo fato de o consumo de energia depender do nível de umidade prevalente, decidiu-se monitorar cada marca em quatro níveis diferentes de umidade, do moderado ao alto (portanto, dispondo em blocos o nível de umidade). Dentro de cada nível, as marcas foram atribuídas aleatoriamente aos cinco locais selecionados. O consumo de energia resultante (kWh anual) é exibido na Tabela 11.2.

**Tabela 11.2** Dados do Consumo de Energia para o Exemplo 11.5

Tratamentos (marcas)	Blocos (nível de umidade)				$x_{.i}$	$\bar{x}_{.i}$
	1	2	3	4		
1	685	792	838	875	3190	797,50
2	722	806	893	953	3374	843,50
3	733	802	880	941	3356	839,00
4	811	888	952	1005	3656	914,00
5	828	920	978	1023	3749	937,25
$x_{.j}$	3779	4208	4541	4797	17325	

Visto que  $\sum \sum x_{ij}^2 = 15.178.901,00$  e  $x_{..}^2/(IJ) = 15.007.781,25$ ,

$$SQT = 15.178.901,00 - 15.007.781,25 = 171.119,75$$

$$SQA = \frac{1}{4} [60.244.049] - 15.007.781,25 = 53.231,00$$

$$SQB = \frac{1}{5} [75.619.995] - 15.007.781,25 = 116.217,75$$

e

$$SQE = 171.119,75 - 53.231,00 - 116.217,75 = 1671,00$$

Os cálculos ANOVA estão resumidos na Tabela 11.3.

**Tabela 11.3** Tabela ANOVA do Exemplo 11.5

Fonte de Variação	gl	Soma dos Quadrados	Quadrado Médio	$f$
Tratamentos (marcas)	4	53.231,00	13.307,75	$f_A = 95,57$
Blocos	3	116.217,75	38.739,25	$f_B = 278,20$
Erro	12	1671,00	139,25	
Total	19	171.119,75		

Visto que  $F_{0,05,4,12} = 3,26$  e  $f_A = 95,57 \geq 3,26$ ,  $H_0$  é rejeitada em favor de  $H_a$  e concluímos que o consumo de energia depende da marca do umidificador. Para identificar marcas significativamente diferentes, usamos o procedimento de Tukey.  $Q_{0,05,5,12} = 4,51$  e  $w = 4,51\sqrt{139,25/4} = 26,6$ .

$\bar{x}_{1.}$	$\bar{x}_{3.}$	$\bar{x}_{2.}$	$\bar{x}_{4.}$	$\bar{x}_{5.}$
797,50	839,00	843,50	914,00	937,25

Os sublinhados indicam que as marcas podem ser divididas em três grupos em relação ao consumo de energia.

Pelo fato de o fator do bloco ter importância secundária,  $F_{0,05,3,12}$  não é necessário, embora o valor calculado de  $F_B$  seja, sem dúvida, altamente significativo. A Figura 11.3 mostra o resultado SQA para esses dados. Observe que, na primeira parte da tabela ANOVA, as somas dos quadrados (SQs) dos tratamentos (marcas) e blocos (níveis de umidade) estão agrupadas em um único “modelo” SQ.

Procedimento de Análise da Variância

Variável Dependente: USO DA ENERGIA

FONTE	Gl	Soma dos Quadrados	Quadrado Médio	Valor F	Pr > F
Modelo	7	169.448,750	24.206,964	173,84	0,0001
Erro	12	1.671,000	139,250		
Total Corrigido	19	171.119,750			

	R-Quadrado	C.V.	QME-raiz	Média USO ENERGIA
	0,990235	1,362242	11,8004	866,25000

Fonte	GL	Anova SQ	Quadrado Médio	Valor F	PR > F
MARCA	4	53.231,000	13.307,750	95,57	0,0001
UMIDADE	3	116.217,750	38.739,250	278,20	0,0001

Alfa = 0,05 gl = 12 QME = 139,25  
 Valor Crítico da Amplitude Studentizada = 4,508  
 Diferença Significativa Mínima = 26,597

As médias com a mesma letra não são significativamente diferentes.

Agrupamento Tukey	Média	N	MARCA
A	937,250	4	5
A			
A	914,000	4	4
B	843,500	4	2
B			
B	839,000	4	3
C	797,500	4	1

**Figura 11.3** Resultado SQA dos dados de consumo de energia

Em muitas situações experimentais em que se aplicam tratamentos a sujeitos, um único sujeito pode receber todos os  $I$  tratamentos. Portanto, a disposição em blocos é geralmente aplicada nos próprios sujeitos para controlar a variabilidade entre eles; logo, diz-se que cada sujeito age de acordo com o próprio controle. Os cientistas sociais às vezes chamam esses experimentos de modelos de medidas repetidas. As “unidades” dentro de um bloco são, portanto, as diferentes “instâncias” de aplicação do tratamento. De forma similar, os blocos em geral são considerados como períodos de tempo, locais e observadores diferentes.

### Exemplo 11.6

Os dados da Tabela 11.4 foram reproduzidos do artigo “Compounding of Discriminative Stimuli from the Same and Different Sensory Modalities” (*J. Experimental Analysis Behavior*, 1971, p. 337-342). A reação do rato foi mantida segundo uma programação de reforço em intervalos fixos, acompanhado de um tom ou de duas luzes distintas. As luzes eram moderadas (L1) ou de baixa intensidade (L2). As observações são especificadas como o número médio de reações de cada sujeito durante demonstrações de estímulo únicas ou compostas, por um período de quatro dias.

**Tabela 11.4** Dados e quantidades resumidas do Exemplo 11.6

Estímulo	Sujeito				$x_i$	$\bar{x}_i$
	1	2	3	4		
L1	8,0	17,3	52,0	22,0	99,3	24,8
L2	6,9	19,3	63,7	21,6	111,5	27,9
Tom (T)	9,3	18,8	60,0	28,3	116,4	29,1
L1 + L2	9,2	24,9	82,4	44,9	161,4	40,3
L1 + T	12,0	31,7	83,8	37,4	164,9	41,2
L2 + T	9,4	33,6	96,6	40,6	180,2	45,1
$x_j$	54,8	145,6	438,5	194,8	833,7	



Com  $\sum\sum x_{ij}^2 = 44.614,21$ ,  $SQT = 15.653,56$ ,  $SQA = 1428,28$ ,  $SQB = 13.444,63$  e  $SQE = 780,65$ . Os cálculos adicionais estão resumidos na Tabela 11.5.

**Tabela 11.5** Tabela ANOVA do Exemplo 11.6

Fonte de Variação	gl	Soma dos Quadrados	Quadrado Médio	$f$
Estímulos (A)	5	1428,28	285,66	$f_A = 5,49$
Sujeitos (B)	3	13.444,63	4.481,54	$f_B = 86,12$
Erro	15	780,65	52,04	
Total	23	15.653,56		

Visto que  $F_{0,05,5,15} = 2,90$  e  $5,49 \geq 2,90$ , concluímos que existem diferenças nas respostas médias reais associadas aos diferentes estímulos. Para o procedimento de Tukey,  $w = 4,59\sqrt{52,04/4} = 16,56$ .

$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	$\bar{x}_4$	$\bar{x}_5$	$\bar{x}_6$
24,8	27,9	29,1	40,3	41,2	45,1

Desse modo, L1 e L2 são significativamente diferentes de L2 + T e não há nenhuma outra diferença significativa entre os estímulos. ■

Na maioria dos experimentos em blocos aleatorizados em que os sujeitos servem como blocos, os sujeitos que de fato participam do experimento são selecionados de uma grande população. Desse modo, os sujeitos contribuem para efeitos aleatórios e não para efeitos fixos. Isso não afeta o procedimento de comparação de tratamentos quando  $K_{ij} = 1$  (uma observação por "célula" como nesta seção), mas o procedimento será alterado se  $K_{ij} = K > 1$ . Avaliaremos em breve modelos de dois fatores nos quais os efeitos são aleatórios.

**Mais sobre experimentos em bloco** Quando  $I = 2$ , tanto o teste  $F$  como o teste  $t$  de diferenças pareadas podem ser usados para analisar os dados. A conclusão resultante não dependerá do procedimento empregado, visto que  $T^2 = F$  e  $t_{\alpha/2, \nu}^2 = F_{\alpha, 1, \nu}$ .

Assim como o pareamento, o experimento em blocos necessita de um ganho potencial e de uma perda potencial de precisão. Se a heterogeneidade for muito grande nas unidades experimentais, o valor do parâmetro de variância  $\sigma^2$  no modelo unidirecional será grande. O efeito da disposição em bloco é filtrar a variação representada por  $\sigma^2$  no modelo bidirecional apropriado para um experimento em bloco aleatorizado. Se todo o resto permanecer igual, um valor menor de  $\sigma^2$  resulta em um teste mais propenso a detectar desvios de  $H_0$  (isto é, em um teste com um poder maior).

Contudo, outras questões não permanecem iguais aqui, visto que o teste  $F$  de fator único baseia-se em  $I(J - 1)$  graus de liberdade (gl) do erro, enquanto o teste  $F$  de dois fatores baseia-se em  $(I - 1)(J - 1)$  gl do erro. Uma quantidade menor de graus de liberdade para erros diminui o poder do teste, fundamentalmente porque o estimador do denominador de  $\sigma^2$  não é tão preciso. Essa perda em graus de liberdade pode ser particularmente séria se o investigador só tiver acesso a um número pequeno de observações. Entretanto, se há indício de que a disposição em blocos reduzirá de maneira significativa a variabilidade, essa perda em graus de liberdade provavelmente valha a pena.

## Modelos de Efeitos Aleatórios

Em vários experimentos, os níveis reais de um fator usado no experimento, em vez de serem apenas aqueles que interessam ao investigador, foram selecionados de uma população bem maior de possíveis níveis do fator. Em uma circunstância de dois fatores, quando isso ocorre para ambos os fatores, é apropriado empregar um **modelo de efeitos aleatórios**. Na situação em que os níveis de um fator são apenas aqueles que interessam e os níveis do outro fator são selecionados de uma população de níveis, tem-se um **modelo de efeitos mistos**. O modelo de efeitos aleatórios de dois fatores, quando  $K_{ij} = 1$ , é

$$X_{ij} = \mu + A_i + B_j + \epsilon_{ij} \quad (i = 1, \dots, I, \quad j = 1, \dots, J)$$

onde os  $A_i$ s,  $B_j$ s, e  $\epsilon_{ij}$ s são todas as variáveis independentes e normalmente distribuídas com média 0 e variâncias  $\sigma_A^2$ ,  $\sigma_B^2$  e  $\sigma^2$ , respectivamente. As hipóteses de interesse são, portanto,  $H_{0A}: \sigma_A^2 = 0$  (o nível do fator A não contribui para a variação na resposta) versus  $H_{aA}: \sigma_A^2 > 0$  e  $H_{0B}: \sigma_B^2 = 0$  versus  $H_{aB}: \sigma_B^2 > 0$ . Embora  $E(QME) = \sigma^2$  como antes, os quadrados médios esperados dos fatores A e B agora são

$$E(QMA) = \sigma^2 + J\sigma_A^2 \quad E(QMB) = \sigma^2 + I\sigma_B^2$$

Portanto, quando  $H_{0A}$  ( $H_{0B}$ ) é verdadeira,  $F_A$  ( $F_B$ ) continua sendo uma razão (quociente) de dois estimadores não-tendenciosos de  $\sigma^2$ . É possível mostrar que um teste de nível  $\alpha$  para  $H_{0A}$  versus  $H_{aA}$  ainda rejeitará  $H_{0A}$  se  $f_A \geq F_{\alpha, I-1, (I-1)(J-1)}$ , e, da mesma forma, se o mesmo procedimento anterior for usado para optar por  $H_{0B}$  ou  $H_{aB}$ . No caso em que o fator A é fixo e o fator B é aleatório, o modelo de efeitos mistos é

$$X_{ij} = \mu + \alpha_i + B_j + \epsilon_{ij} \quad (i = 1, \dots, I, \quad j = 1, \dots, J)$$

onde  $\sum \alpha_i = 0$  e os  $B_j$ s e  $\epsilon_{ij}$ s são normalmente distribuídos com média 0 e variâncias  $\sigma_B^2$  e  $\sigma^2$ , respectivamente. Agora, as duas hipóteses nulas são

$$H_{0A}: \alpha_1 = \dots = \alpha_I = 0 \quad \text{e} \quad H_{0B}: \sigma_B^2 = 0$$

com os quadrados médios esperados

$$E(QME) = \sigma^2 \quad E(QMA) = \sigma^2 + \frac{J}{I-1} \sum \alpha_i^2 \quad E(QMB) = \sigma^2 + I\sigma_B^2$$

Os procedimentos de teste para  $H_{0A}$  versus  $H_{aA}$  e  $H_{0B}$  versus  $H_{aB}$  são exatamente iguais aos anteriores. Por exemplo: na análise dos dados sobre mudança de cor do Exemplo 11.1, se os quatro tratamentos de lavagem fossem selecionados aleatoriamente, então, pelo fato de  $f_B = 11,05$  e  $F_{0,05,3,6} = 4,76$ ,  $H_{0B}: \sigma_B^2 = 0$  seria rejeitada em favor de  $H_{aB}: \sigma_B^2 > 0$ . Uma estimativa do “componente da variância”  $\sigma_B^2$  é portanto especificada por  $(QMB - QME)/I = 0,0485$ .

Em resumo, quando  $K_{ij} = 1$ , embora as hipóteses e os quadrados médios esperados sejam diferentes do caso em que ambos os efeitos são fixos, os procedimentos de teste são idênticos.

## Exercícios | Seção 11.1 (1–15)

- Determinou-se o número de milhas de desgaste útil da banda de rodagem (em milhares) de pneus de cinco diferentes modelos de carros subcompactos (fator A, com  $I = 5$ ) juntamente ao de pneus radiais de quatro diferentes marcas (fator B, com  $J = 4$ ), o que resultou em  $IJ = 20$  observações. Os valores  $SQA = 30,6$ ,  $SQB = 44,1$  e  $SQE = 59,2$  foram então calculados. Suponha que um modelo aditivo seja apropriado.
  - Teste  $H_0: \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 0$  (não há nenhuma diferença no tempo de vida médio real do pneu decorrente dos modelos de carro) versus  $H_a$ : pelo menos um  $\alpha_i \neq 0$  usando um teste de nível 0,05.
  - $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$  (não há nenhuma diferença no tempo de vida médio real do pneu decorrente das marcas de pneu) versus  $H_a$ : pelo menos um  $\beta_j \neq 0$  usando um teste de nível 0,05.
- Quatro diferentes revestimentos estão sendo avaliados para protegerem tubulações de metal contra corrosão.

As tubulações serão enterradas em três diferentes tipos de solo. Para investigar se a quantidade de corrosão depende do revestimento ou do tipo de solo, foram selecionadas 12 peças de tubulação. Cada uma é revestida com um dos quatro revestimentos e enterrada em um dos três tipos de solo durante um período de tempo fixo, após o qual se determina a quantidade de corrosão (profundidade das covas maiores, em 0,0001 pol.). Os dados são apresentados na tabela a seguir.

		Tipo de Solo (B)		
		1	2	3
Revestimento (A)	1	64	49	50
	2	53	51	48
	3	47	45	50
	4	51	43	52

- a. Assumindo a validade do modelo aditivo, efetue a análise ANOVA por meio de uma tabela ANOVA para verificar se a quantidade de corrosão depende do tipo de revestimento usado ou do tipo de solo. Use  $\alpha = 0,05$ .
- b. Calcule  $\hat{\mu}$ ,  $\hat{\alpha}_1$ ,  $\hat{\alpha}_2$ ,  $\hat{\alpha}_3$ ,  $\hat{\alpha}_4$ ,  $\hat{\beta}_1$ ,  $\hat{\beta}_2$  e  $\hat{\beta}_3$ .

3. O artigo "Adiabatic Humidification of Air with Water in a Packed Tower" (*Chem. Eng. Prog.*, 1952, p. 362-370) apresenta dados sobre o coeficiente de condutividade térmica da película gasosa [Btu/(hr.pés<sup>2</sup>.°F)] como função da taxa de gás (fator A) e da taxa de líquido (fator B).

		B			
		1(190)	2(250)	3(300)	4(400)
A	1(200)	200	226	240	261
	2(400)	278	312	330	381
	3(700)	369	416	462	517
	4(1100)	500	575	645	733

- a. Depois de elaborar uma tabela ANOVA, teste, no nível 0,01, tanto a hipótese de ausência de efeito da taxa de gás contra a alternativa apropriada, como a hipótese de ausência de efeito da taxa de líquido contra a alternativa apropriada.
  - b. Use o procedimento de Tukey para investigar diferenças no coeficiente de condutividade térmica esperado decorrentes de diferentes taxas de gás.
  - c. Repita a parte (b) para as taxas de líquido.
4. Em um experimento para verificar se a quantidade de demãos de tinta látex azul-clara para ambientes internos depende da marca da tinta ou da marca do rolo, 1 galão de cada uma das quatro marcas de tinta foi aplicado, usando-se cada uma das três marcas de rolo, gerando os dados a seguir (número de pés quadrados cobertos).

		Marca do Rolo		
		1	2	3
Marca da Tinta	1	454	446	451
	2	446	444	447
	3	439	442	444
	4	444	437	443

- a. Elabore a tabela ANOVA. [*Sugestão*: os cálculos podem ser feitos mais rapidamente subtraindo-se 400 (ou qualquer outro número conveniente) de cada observação. Isso não afeta os resultados finais.]
  - b. Expresse e teste as hipóteses apropriadas para determinar se a marca da tinta tem algum efeito sobre a demão. Use  $\alpha = 0,05$ .
  - c. Repita a parte (b) para a marca do rolo.
  - d. Use o método de Tukey para identificar diferenças significativas entre as marcas. Há alguma marca aparentemente preferível às outras?
5. Em um experimento para avaliar o efeito do ângulo de tração sobre a força necessária para separar conectores

elétricos, quatro diferentes ângulos (fator A) foram usados e cada uma das amostras de cinco conectores (fator B) foi puxada uma vez em cada ângulo ("A Mixed Model Factorial Experiment in Testing Electrical Connectors", *Industrial Quality Control*, 1960, p. 12-16). Os dados são exibidos na tabela a seguir.

		B				
		1	2	3	4	5
A	0°	45,3	42,2	39,6	36,8	45,8
	2°	44,1	44,1	38,4	38,0	47,2
	4°	42,7	42,7	42,6	42,2	48,9
	6°	43,5	45,8	47,9	37,9	56,4

Os dados sugerem que a força de separação média real é influenciada pelo ângulo de tração? Expresse e teste as hipóteses apropriadas, ao nível 0,01, elaborando primeiro uma tabela ANOVA (SQT = 396,13, SQA = 58,16 e SQB = 246,97).

- 6. Uma determinada cidade emprega três avaliadores, os quais são responsáveis pela determinação do valor das propriedades residenciais na cidade. Para verificar se esses avaliadores diferem sistematicamente em relação às suas avaliações, foram selecionadas cinco casas e solicitou-se a cada avaliador que determinasse o valor de mercado de cada casa. Sendo o fator A os avaliadores ( $I = 3$ ) e o fator B as casas ( $J = 5$ ), suponha que SQA = 11,7, SQB = 113,5 e SQE = 25,6.
  - a. Teste  $H_0: \alpha_1 = \alpha_2 = \alpha_3 = 0$  ao nível 0,05. ( $H_0$  expressa que não há diferenças sistemáticas entre os avaliadores.)
  - b. Explique por que foi usado um experimento em bloco aleatorizado, com apenas cinco casas, em vez de uma forma de experimento ANOVA envolvendo um total de 15 diferentes casas e um avaliador para cinco diferentes casas (um grupo diferente de 5 para cada avaliador).
- 7. O artigo "Rate of Stuttering Adaptation Under Two Electro-Shock Conditions" (*Behavior Research Therapy*, 1967, p. 49-54) apresenta pontuações de adaptação para três diferentes tratamentos: (1) nenhum choque, (2) choque após cada palavra gaguejada e (3) choque em cada instante de gagueira. Esses tratamentos foram usados em cada um de 18 gagos.
  - a. As estatísticas resumidas incluem  $x_{1.} = 905$ ,  $x_{2.} = 913$ ,  $x_{3.} = 936$ ,  $x_{..} = 2754$ ,  $\sum_j x_{.j}^2 = 430.295$  e  $\sum_i \sum_j x_{ij}^2 = 143.930$ . Elabore a tabela ANOVA e efetue um teste ao nível 0,05 para verificar se a pontuação de adaptação média real depende de um tratamento específico.
  - b. Com base na razão F dos sujeitos (fator B), você acredita que dispor os sujeitos em bloco foi eficaz nesse experimento? Explique.
- 8. A tabela a seguir fornece a concentração plasmática de epinefrina em 10 sujeitos experimentais durante

anestésias de (1) isoflurano, (2) halotano e (3) ciclopropano (“Sympathoadrenal and Hemodynamic Effects of Isoflurane, Halothane and Cyclopropane in Dogs”, *Anesthesiology*, 1974, p. 465-470).

a. A opção de efeito anestésico influencia na concentração média real? Teste  $H_0: \alpha_1 = \alpha_2 = \alpha_3 = 0$  ao nível 0,05 após a elaboração da tabela ANOVA.

		Sujeito (B)				
		1	2	3	4	5
Anestésico (A)	1	0,28	0,51	1,00	0,39	0,29
	2	0,30	0,39	0,63	0,38	0,21
	3	1,07	1,35	0,69	0,28	1,24
		6	7	8	9	10
Anestésico (A)	1	0,36	0,32	0,69	0,17	0,33
	2	0,88	0,39	0,51	0,32	0,42
	3	1,53	0,49	0,56	1,02	0,30

$\sum \sum x_{ij}^2 = 13,7980$

b. Use o procedimento de Tukey para investigar diferenças significativas entre os anestésicos.

9. O artigo “The Effects of a Pneumatic Stool and a One-Legged Stool on Lower Limb Joint Load and Muscular Activity During Sitting and Rising” (*Ergonomics*, 1993, p. 519-535) apresenta os dados a seguir sobre o esforço necessário para um sujeito levantar-se de quatro diferentes tipos de assento (escala de Borg). Investigue a variância usando  $\alpha = 0,05$  e, em seguida, realize uma análise de comparações múltiplas, se apropriado.

		Sujeito									$\bar{x}_i$
		1	2	3	4	5	6	7	8	9	
Tipo de Assento	1	12	10	7	7	8	9	8	7	9	8,56
	2	15	14	14	11	11	11	12	11	13	12,44
	3	12	13	13	10	8	11	12	8	10	10,78
	4	10	12	9	9	7	10	11	7	8	9,22

10. A resistência do concreto usado em construções comerciais tende a variar entre um lote e outro. Conseqüentemente, pequenos cilindros de teste de concreto amostrados de um lote são "tratados" durante períodos de no máximo 28 dias, mais ou menos, em ambientes de temperatura e umidade controladas antes de avaliar a resistência. O concreto é em seguida “comprado e vendido de acordo com os cilindros de teste de resistência” (Método de Teste Padrão ASTM C 31 para Fabricação e Tratamento de Amostras de Teste de Concreto em Campo). Os dados a seguir resultam de um experimento de comparação de três diferentes métodos de tratamento em relação à resistência à compressão (MPa). Analise esses dados.

Lote	Método A	Método B	Método C
1	30,7	33,7	30,5
2	29,1	30,6	32,6
3	30,0	32,2	30,5
4	31,9	34,6	33,5
5	30,5	33,0	32,4
6	26,9	29,3	27,8
7	28,2	28,4	30,7
8	32,4	32,4	33,6
9	26,6	29,5	29,2
10	28,6	29,4	33,2

11. Os "resíduos" de uma ANOVA de dois fatores, com  $K_{ij} = 1$  são as quantidades  $x_{ij} - (\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j)$ . Um gráfico de probabilidade normal desses resíduos pode ser usado para verificar a plausibilidade da hipótese de normalidade. Elabore esse gráfico usando os dados do Exemplo 11.1 e comente-o.

12. Suponha que no experimento descrito no Exercício 6 as cinco casas tivessem sido de fato selecionadas aleatoriamente entre as casas de um determinado tamanho ou tempo de vida, de modo que o fator B seja aleatório, em vez de fixo. Teste  $H_0: \sigma_B^2 = 0$  versus  $H_a: \sigma_B^2 > 0$  usando um teste de nível 0,01.

13. a. Mostre que uma constante  $d$  pode ser adicionada a (ou subtraída de) de cada  $x_{ij}$  sem influenciar nenhuma das somas de quadrados ANOVA.

b. Suponha que cada  $x_{ij}$  seja multiplicado por uma constante  $c$  diferente de zero. Como isso influencia as somas de quadrados ANOVA? Como isso influencia os valores da estatística  $F$ ,  $F_A$  e  $F_B$ ? Que efeito tem “codificar” os dados por  $y_{ij} = cx_{ij} + d$  sobre as conclusões resultantes dos procedimentos ANOVA?

14. Use a constatação de que  $E(X_{ij}) = \mu + \alpha_i + \beta_j$  com  $\sum \alpha_i = \sum \beta_j = 0$  para mostrar que  $E(\bar{X}_i - \bar{X}_{..}) = \alpha_i$ , de modo que  $\hat{\alpha}_i = \bar{X}_i - \bar{X}_{..}$  seja um estimador não-tendencioso de  $\alpha_i$ .

15. As curvas de potência das Figuras 10.5 e 10.6 podem ser usadas para obter  $\beta = P$  (erro do tipo II) no teste  $F$  na ANOVA de dois fatores. Para valores fixos de  $\alpha_1, \alpha_2, \dots, \alpha_p$ , a quantidade  $\phi^2 = (J/I) \sum \alpha_i^2 / \sigma^2$  é calculada. Então, o número correspondente a  $\nu_1 = I - 1$  é inserido no eixo horizontal, no valor  $\phi$ , a potência é lida no eixo vertical a partir da curva denominada  $\nu_2 = (I - 1)(J - 1)$  e  $\beta = 1 - \text{potência}$ .

a. Em relação ao experimento sobre corrosão descrito no Exercício 2, encontre  $\beta$  quando  $\alpha_1 = 4, \alpha_2 = 0, \alpha_3 = \alpha_4 = -2$ , e  $\sigma = 4$ . Repita para  $\alpha_1 = 6, \alpha_2 = 0, \alpha_3 = \alpha_4 = -3$  e  $\sigma = 4$ .

b. Por simetria, qual é o  $\beta$  para o teste de  $H_{0B}$  versus  $H_{aB}$  no Exemplo 11.1, quando  $\beta_1 = 0,3, \beta_2 = \beta_3 = \beta_4 = -0,1$ , e  $\sigma = 0,3$ ?

## 11.2 ANOVA de Dois Fatores com $K_{ij} > 1$

Na Seção 11.1 analisamos os dados provenientes de um experimento de dois fatores em que havia uma observação para cada uma das  $IJ$  combinações de níveis de dois fatores. Para obter procedimentos de teste válidos, considerou-se que os  $\mu_{ij}$  têm uma estrutura aditiva com  $\mu_{ij} = \mu + \alpha_i + \beta_j$ ,  $\sum \alpha_i = \sum \beta_j = 0$ . Aditividade significa que a diferença nas respostas médias reais para quaisquer dois níveis dos fatores é a mesma para cada nível do outro fator. Por exemplo,  $\mu_{ij} - \mu_{i'j} = (\mu + \alpha_i + \beta_j) - (\mu + \alpha_{i'} + \beta_j) = \alpha_i - \alpha_{i'}$ , independentemente do nível de  $j$  do segundo fator. Isso é apresentado na Figura 11.1(a), na qual as linhas que ligam as respostas médias reais são paralelas.

A Figura 11.1(b) detalha um conjunto de respostas médias reais que não têm estrutura aditiva. As linhas que ligam esses  $\mu_{ij}$ s não são paralelas, o que significa que a diferença nas respostas médias reais para diferentes níveis de um fator depende do nível do outro fator. Quando a aditividade não se mantém, dizemos que existe interação entre os diferentes níveis dos fatores. A suposição de aditividade nos permitiu obter, na Seção 11.1, um estimador da variância de erro aleatório  $\sigma^2$  (QME), cuja não-tendenciosidade não dependia da veracidade da hipótese nula de interesse. Sendo  $K_{ij} > 1$  para pelo menos um par  $(i, j)$ , um estimador válido de  $\sigma^2$  pode ser obtido sem assumir a aditividade. Ao especificar o modelo apropriado e derivar os procedimentos de teste, vamos nos concentrar no caso  $K_{ij} = K > 1$ , de maneira que o número de observações por “célula” (para cada combinação de níveis) seja constante.

### Parâmetros para o Modelo de Efeitos Fixos com Interação

Em vez de usar os próprios  $\mu_{ij}$ s como parâmetros do modelo, é comum usar um conjunto equivalente que revela mais nitidamente o papel da interação. Seja

$$\mu = \frac{1}{IJ} \sum_i \sum_j \mu_{ij} \quad \mu_{i.} = \frac{1}{J} \sum_j \mu_{ij} \quad \mu_{.j} = \frac{1}{I} \sum_i \mu_{ij} \quad (11.7)$$

Logo,  $\mu$  é a resposta esperada da média de todos os níveis de ambos os fatores (a média geral real),  $\mu_{i.}$  é a resposta esperada da média dos níveis do segundo fator quando o primeiro fator  $A$  é mantido no nível  $i$  e, no caso de  $\mu_{.j}$ , pode-se afirmar algo semelhante. Agora, defina

$$\begin{aligned} \alpha_i &= \mu_{i.} - \mu = \text{o efeito do fator } A \text{ no nível } i \\ \beta_j &= \mu_{.j} - \mu = \text{o efeito do fator } B \text{ no nível } j \end{aligned} \quad (11.8)$$

$$\gamma_{ij} = \mu_{ij} - (\mu + \alpha_i + \beta_j)$$

de onde

$$\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} \quad (11.9)$$

O modelo é aditivo se, e somente se, todos os  $\gamma_{ij} = 0$ . Os  $\gamma_{ij}$ s são chamados de parâmetros de interação. Os  $\alpha_i$ s são chamados de **efeitos principais do fator A** e os  $\beta_j$ s são os **efeitos principais do fator B**. Embora haja  $I$   $\alpha_i$ s,  $J$   $\beta_j$ s, e  $IJ$   $\gamma_{ij}$ s, além de  $\mu$ , as condições  $\sum \alpha_i = 0$ ,  $\sum \beta_j = 0$ ,  $\sum_j \gamma_{ij} = 0$  para qualquer  $i$ , e  $\sum_i \gamma_{ij} = 0$  para qualquer  $j$  [todos em virtude de (11.7) e (11.8)] implicam que somente  $IJ$  desses novos parâmetros são determinados de modo independente:  $\mu$ ,  $I - 1$  dos  $\alpha_i$ s,  $J - 1$  dos  $\beta_j$ s, e  $(I - 1)(J - 1)$  dos  $\gamma_{ij}$ s.

Agora existem três conjuntos de hipóteses a serem considerados:

$$\begin{array}{ll}
 H_{0AB}: \gamma_{ij} = 0 \text{ para todos os } i, j & \text{versus } H_{aAB}: \text{ pelo menos um } \gamma_{ij} \neq 0 \\
 H_{0A}: \alpha_1 = \dots = \alpha_I = 0 & \text{versus } H_{aA}: \text{ pelo menos um } \alpha_i \neq 0 \\
 H_{0B}: \beta_1 = \dots = \beta_J = 0 & \text{versus } H_{aB}: \text{ pelo menos um } \beta_j \neq 0
 \end{array}$$

A hipótese de não-interação  $H_{0AB}$  normalmente é testada primeiro. Se  $H_{0AB}$  não for rejeitada, as outras duas podem ser testadas para verificar se os efeitos principais são significativos. Se  $H_{0AB}$  for rejeitada e  $H_{0A}$  for em seguida testada e não rejeitada, o modelo resultante  $\mu_{ij} = \mu + \beta_j + \gamma_{ij}$  não serve para uma interpretação direta e clara. Nesse caso, é melhor elaborar uma figura semelhante à Figura 11.1(b) para tentar visualizar uma forma pela qual os fatores interagem.

## Notação, Modelo e Análise

No momento usamos subscritos triplos tanto para variáveis aleatórias como para valores observados, onde  $X_{ijk}$  e  $x_{ijk}$  referem-se à  $k$ -ésima observação (replicação) quando o fator  $A$  está no nível  $i$  e o fator  $B$  está no nível  $j$ . O modelo é, portanto,

$$X_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk} \quad (11.10)$$

$$i = 1, \dots, I, \quad j = 1, \dots, J, \quad k = 1, \dots, K$$

onde os  $\epsilon_{ij}$ s são independentes e normalmente distribuídos, cada um com média 0 e variância  $\sigma^2$ .

Novamente, um ponto no lugar do subscrito significa que somamos todos os seus valores, ao passo que uma barra horizontal representa o cálculo da média. Portanto,  $X_{i\cdot}$  é o total de todas as  $K$  observações feitas para o fator  $A$  no nível  $i$  e para o fator  $B$  no nível  $j$  [todas as observações na  $(i, j)$ -ésima célula] e  $\bar{X}_{i\cdot}$  é a média dessas  $K$  observações.

### Exemplo 11.7

Três diferentes variedades de tomate (Harvester, Pusa Early Dwarf e Ife N<sup>o</sup> 1) e quatro diferentes densidades de muda (10, 20, 30 e 40 mil plantas por hectare) estão sendo analisadas para serem plantadas em uma determinada região. Para verificar se tanto a variedade como a densidade da muda influenciam na produção, cada combinação de variedade e densidade de muda é usada em três diferentes gráficos, gerando os dados sobre produção apresentados na Tabela 11.6 (baseada no artigo “Effects of Plant Density on Tomato Yields in Western Nigeria”, *Experimental Agriculture*, 1976, p. 43-47).

**Tabela 11.6** Dados sobre produção do Exemplo 11.7

Variedade	Densidade de Plantio												$x_{i\cdot}$	$\bar{x}_{i\cdot}$
	10,000			20,000			30,000			40,000				
<b>H</b>	10,5	9,2	7,9	12,8	11,2	13,3	12,1	12,6	14,0	10,8	9,1	12,5	136,0	11,33
<b>Ife</b>	8,1	8,6	10,1	12,7	13,7	11,5	14,4	15,4	13,7	11,3	12,5	14,5	146,5	12,21
<b>P</b>	16,1	15,3	17,5	16,6	19,2	18,5	20,8	18,0	21,0	18,4	18,9	17,2	217,5	18,13
$x_{\cdot j}$	103,3			129,5			142,0			125,2			500,00	
$\bar{x}_{\cdot j}$	11,48			14,39			15,78			13,91				13,89

Aqui,  $I = 3$ ,  $J = 4$  e  $K = 3$ , para um total de  $IJK = 36$  observações. ■

Para testar as hipóteses de interesse, novamente definimos as somas de quadrados e apresentamos fórmulas computacionais:

$$\begin{aligned} \text{SQT} &= \sum_i \sum_j \sum_k (X_{ijk} - \bar{X}_{...})^2 = \sum_i \sum_j \sum_k X_{ijk}^2 - \frac{1}{IJK} X^2_{...} & \text{gl} &= IJK - 1 \\ \text{SQE} &= \sum_i \sum_j \sum_k (X_{ijk} - \bar{X}_{ij.})^2 \\ &= \sum_i \sum_j \sum_k X_{ijk}^2 - \frac{1}{K} \sum_i \sum_j X_{ij.}^2 & \text{gl} &= IJ(K - 1) \\ \text{SQA} &= \sum_i \sum_j \sum_k (\bar{X}_{i..} - \bar{X}_{...})^2 = \frac{1}{JK} \sum_i X_{i..}^2 - \frac{1}{IJK} X^2_{...} & \text{gl} &= I - 1 \\ \text{SQB} &= \sum_i \sum_j \sum_k (\bar{X}_{.j.} - \bar{X}_{...})^2 = \frac{1}{JK} \sum_j X_{.j.}^2 - \frac{1}{IJK} X^2_{...} & \text{gl} &= J - 1 \\ \text{SQAB} &= \sum_i \sum_j \sum_k (\bar{X}_{ij.} - \bar{X}_{i..} - \bar{X}_{.j.} + \bar{X}_{...})^2 & \text{gl} &= (I - 1)(J - 1) \end{aligned}$$

A identidade fundamental

$$\text{SQT} = \text{SQA} + \text{SQB} + \text{SQAB} + \text{SQE}$$

significa que a **soma dos quadrados da interação** SQAB pode ser obtida por subtração.

A variação total é, portanto, dividida em quatro partes: uma inexplicada (SQE – que estaria presente independentemente da veracidade das três hipóteses nulas) e três que podem ser explicadas pela veracidade ou inverossimilhança das três  $H_0$ s. Cada um dos quatro quadrados médios é definido por  $QM = SQ/\text{gl}$ . Os quadrados médios esperados sugerem que cada conjunto de hipóteses deve ser testado por meio de uma razão apropriada de quadrados médios com QME no denominador:

$$\begin{aligned} E(\text{SQE}) &= \sigma^2 \\ E(\text{SQA}) &= \sigma^2 + \frac{JK}{I - 1} \sum_{i=1}^I \alpha_i^2 \\ E(\text{SQB}) &= \sigma^2 + \frac{IK}{J - 1} \sum_{j=1}^J \beta_j^2 \\ E(\text{SQAB}) &= \sigma^2 + \frac{K}{(I - 1)(J - 1)} \sum_{i=1}^I \sum_{j=1}^J \gamma_{ij}^2 \end{aligned}$$

É possível mostrar que cada uma das três razões de quadrado médio tem uma distribuição  $F$  quando a  $H_0$  associada é verdadeira, o que produz os procedimentos de teste de nível  $\alpha$ .

Hipóteses	Valor do Teste Estatístico	Região de Rejeição
$H_{0A}$ versus $H_{aA}$	$f_A = \frac{\text{SQA}}{\text{SQE}}$	$f_A \geq F_{\alpha, I-1, IJ(K-1)}$
$H_{0B}$ versus $H_{aB}$	$f_B = \frac{\text{SQB}}{\text{SQE}}$	$f_B \geq F_{\alpha, J-1, IJ(K-1)}$
$H_{0AB}$ versus $H_{aAB}$	$f_{AB} = \frac{\text{SQAB}}{\text{SQE}}$	$f_{AB} \geq F_{\alpha, (I-1)(J-1), IJ(K-1)}$

Como antes, os resultados da análise estão resumidos em uma tabela ANOVA.

**Exemplo 11.8** (Continuação do Exemplo 11.7)

Com base nos dados especificados,  $x_{...}^2 = (500)^2 = 250.000$

$$\sum_i \sum_j \sum_k x_{ijk}^2 = (10,5)^2 + (9,2)^2 + \dots + (18,9)^2 + (17,2)^2 = 7404,80$$

$$\sum_i x_{i..}^2 = (136,0)^2 + (146,5)^2 + (217,5)^2 = 87.264,50$$

e

$$\sum_j x_{.j.}^2 = 63.280,18$$

Os totais de célula ( $x_{ij.}$ s) são

	10.000	20.000	30.000	40.000
H	27,6	37,3	38,7	32,4
Ife	26,8	37,9	43,5	38,3
P	48,9	54,3	59,8	54,5

de onde  $\sum_i \sum_j x_{ij.}^2 = (27,6)^2 + \dots + (54,5)^2 = 22.100,28$ . Portanto

$$SQT = 7404,80 - \frac{1}{36}(250.000) = 7404,80 - 6944,44 = 460,36$$

$$SQA = \frac{1}{12}(87.264,50) - 6944,44 = 327,60$$

$$SQB = \frac{1}{9}(63.280,18) - 6944,44 = 86,69$$

$$SQE = 7404,80 - \frac{1}{3}(22.100,28) = 38,04$$

e

$$SQAB = 460,36 - 327,60 - 86,69 - 38,04 = 8,03$$

A Tabela 11.7 resume os cálculos

**Tabela 11.7** Tabela ANOVA do Exemplo 11.8

Fonte de Variação	gl	Soma dos Quadrados	Quadrado Médio	$f$
Variedades	2	327,60	163,8	$f_A = 103,02$
Densidade	3	86,69	28,9	$f_B = 18,18$
Interação	6	8,03	1,34	$f_{AB} = 0,84$
Erro	24	38,04	1,59	
Total	35	460,36		

Visto que  $F_{0,01,6,24} = 3,67$  e  $f_{AB} = 0,84$  não é  $\geq 3,67$ ,  $H_{0AB}$  não pode ser rejeitada no nível 0,01, de modo que concluímos que os efeitos da interação não são significativos. Agora, a presença ou a ausência dos efeitos principais pode ser investigada. Visto que  $F_{0,01,2,24} = 5,61$  e  $f_A = 103,02 \geq 5,61$ ,  $H_{0A}$  é rejeitada no nível 0,01, confirmando a conclusão de que as diferentes variedades influenciam nas produções médias reais. De modo semelhante,  $f_B = 18,18 \geq 4,72 = F_{0,01,3,24}$ , o que nos leva a concluir que a produção média real também depende da densidade da muda. ■



## Comparações Múltiplas

Quando a hipótese de não- interação  $H_{0AB}$  não for rejeitada e pelo menos uma entre as duas hipóteses nulas do efeito principal for rejeitada, o método de Tukey pode ser usado para identificar diferenças significativas nos níveis. Para identificar diferenças entre os  $\alpha_i$ s quando  $H_{0A}$  for rejeitada,

1. obtenha  $Q_{\alpha, I, IJ(K-1)}$ , onde o segundo subscrito  $I$  identifica o número de níveis que estão sendo comparados e o terceiro subscrito refere-se ao número de graus de liberdade dos erros;
2. calcule  $w = Q\sqrt{QME/(JK)}$ , onde  $JK$  é o número de observações resultante do cálculo da média para obter cada um dos  $\bar{x}_{i..}$ s comparados na Etapa 3;
3. ordene os  $\bar{x}_{i..}$ s do menor para o maior e, como antes, sublinhe todos os pares cuja diferença é menor que  $w$ . Os pares não-sublinhados correspondem a níveis significativamente diferentes do fator  $A$ .

Para identificar níveis diferentes do fator  $B$  quando  $H_{0B}$  for rejeitada, substitua o segundo subscrito em  $Q$  por  $J$ , substitua  $JK$  por  $IK$  em  $w$  e substitua  $\bar{x}_{i..}$  por  $\bar{x}_{.j.}$ .

**Exemplo 11.9** (Continuação do Exemplo 11.8)

Para o fator  $A$  (variedades),  $I = 3$ , de modo que  $\alpha = 0,01$  e  $IJ(K-1) = 24$ ,  $Q_{0,01,3,24} = 4,55$ . Logo,  $w = 4,55\sqrt{1,59/12} = 1,66$ , de modo que, ao ordenarmos e sublinharmos, obteremos

$\bar{x}_{1..}$	$\bar{x}_{2..}$	$\bar{x}_{3..}$
<u>11,33</u>	<u>12,21</u>	18,13

As variedades Harvester e Ife não parecem diferir significativamente entre si em relação ao efeito sobre a produção média real, mas ambas diferem da variedade Pusa.

Para o fator  $B$  (densidade),  $J = 4$ , de modo que  $Q_{0,01,4,24} = 4,91$  e  $w = 4,91\sqrt{1,59/9} = 2,06$ .

$\bar{x}_{.1.}$	$\bar{x}_{.4.}$	$\bar{x}_{.2.}$	$\bar{x}_{.3.}$
11,48	<u>13,91</u>	<u>14,39</u>	15,78

Portanto, com a taxa de erro 0,01 relativa ao experimento, que é bem moderada, somente a densidade menor parece diferir significativamente das outras. Mesmo sendo  $\alpha = 0,05$  (de modo que  $w = 1,64$ ), as densidades 2 e 3 não podem ser consideradas significativamente diferentes entre si em relação a seu efeito sobre a produção. ■

## Modelos com Efeitos Mistos e Aleatórios

Em alguns problemas, os níveis de ambos os fatores podem ter sido escolhidos de uma ampla população de possíveis níveis, de maneira que os efeitos do fator são aleatórios em vez de fixos. Como na Seção 11.1, se ambos os fatores produzem efeitos aleatórios, o modelo é denominado modelo de efeitos aleatórios, ao passo que, se um fator é fixo e o outro aleatório, o modelo resultante é o de efeitos mistos. Consideraremos agora a análise de um modelo de efeitos mistos na qual o fator  $A$  (linhas) é o fator fixo e o fator  $B$  (colunas) é o fator aleatório. O caso em que ambos são aleatórios é abordado no Exercício 26.

O modelo de efeitos mistos, nesta situação, é

$$X_{ijk} = \mu + \alpha_i + B_j + G_{ij} + \epsilon_{ijk}$$

$$i = 1, \dots, I, \quad j = 1, \dots, J, \quad k = 1, \dots, K$$

Aqui,  $\mu$  e os  $\alpha_i$ s são constantes com  $\sum \alpha_i = 0$  e os  $B_j$ s,  $G_{ij}$ s, e  $\epsilon_{ijk}$ s são variáveis aleatórias independentes e normalmente distribuídas com um valor esperado 0 e variâncias  $\sigma_B^2$ ,  $\sigma_G^2$ , e  $\sigma^2$ , respectivamente.<sup>1</sup>

<sup>1</sup> Isso é chamado de modelo "irrestrito". Um modelo alternativo "restrito" exige que  $\sum G_{ij} = 0$  para cada  $j$  (de modo que os  $G_{ij}$ s não sejam mais independentes). Os quadrados médios esperados e as razões  $F$  apropriadas para testar determinadas hipóteses dependem da opção do modelo. A opção-padrão do MINITAB apresenta dados para o modelo irrestrito.

$$\begin{array}{ll}
 H_{0A}: \alpha_1 = \alpha_2 = \dots = \alpha_I = 0 & \text{versus } H_{aA}: \text{pelo menos um } \alpha_i \neq 0 \\
 H_{0B}: \sigma_B^2 = 0 & \text{versus } H_{aB}: \sigma_B^2 > 0 \\
 H_{0G}: \sigma_G^2 = 0 & \text{versus } H_{aG}: \sigma_G^2 > 0
 \end{array}$$

É comum testar  $H_{0A}$  e  $H_{0B}$  somente se as hipóteses de não-interação  $H_{0G}$  não puderem ser rejeitadas.

As somas de quadrados relevantes e os quadrados médios necessários para os procedimentos de teste são definidos e calculados exatamente como no caso dos efeitos fixos. Os quadrados médios esperados para o modelo misto são

$$E(\text{QME}) = \sigma^2$$

$$E(\text{QMA}) = \sigma^2 + K\sigma_G^2 + \frac{JK}{I-1} \sum \alpha_i^2$$

$$E(\text{QMB}) = \sigma^2 + K\sigma_G^2 + IK\sigma_B^2$$

e

$$E(\text{QMAB}) = \sigma^2 + K\sigma_G^2$$

Logo, para testar a hipótese de não-interação, a razão  $f_{AB} = \text{QMA}/\text{QME}$  é novamente apropriada, em que  $H_{0G}$  é rejeitada se  $f_{AB} \geq F_{\alpha, (I-1)(J-1), IJ(K-1)}$ . Entretanto, para testar  $H_{0A}$  versus  $H_{aA}$ , os quadrados médios esperados sugerem que, embora o numerador da razão  $F$  deva ser ainda QMA, o denominador deve ser QMAB em vez de QME. QMAB é também o denominador da razão  $F$  para testar  $H_{0B}$ .

Para testar  $H_{0A}$  versus  $H_{aA}$  (fator  $A$  fixo, fator  $B$  aleatório), o valor da estatística de teste é  $f_A = \text{QMA}/\text{QMAB}$  e a região de rejeição é  $f_A \geq F_{\alpha, I-1, (I-1)(J-1)}$ . O teste de  $H_{0B}$  versus  $H_{aB}$  utiliza  $f_B = \text{QMB}/\text{QMAB}$ , e a região de rejeição é  $f_B \geq F_{\alpha, J-1, (I-1)(J-1)}$ .

### Exemplo 11.10

Um engenheiro de processos identificou duas causas possíveis da vibração presente nos motores elétricos: o material usado para a carcaça do motor (fator  $A$ ) e a fonte de energia dos mancais usados no motor (fator  $B$ ). Os dados a seguir sobre a quantidade de vibração (mícrons) foram obtidos em um experimento em que os motores com carcaças de aço, alumínio e plástico foram fabricados com mancais fornecidos por cinco fontes selecionadas aleatoriamente.

		Fonte de Energia									
		1		2		3		4		5	
Material	Aço	13,1	13,2	16,3	15,8	13,7	14,3	15,7	15,8	13,5	12,5
	Alumínio	15,0	14,8	15,7	16,4	13,9	14,3	13,7	14,2	13,4	13,8
	Plástico	14,0	14,3	17,2	16,7	12,4	12,3	14,4	13,9	13,2	13,1

Apenas o material das três carcaças usadas no experimento está sendo avaliado para ser usado na fabricação, de modo que o fator  $A$  é fixo. Contudo, as cinco fontes de energia foram selecionadas aleatoriamente de uma população bem maior, de maneira que o fator  $B$  é aleatório. As hipóteses nulas de interesse são

$$H_{0A}: \alpha_1 = \alpha_2 = \alpha_3 = 0 \quad H_{0B}: \sigma_B^2 = 0 \quad H_{0AB}: \sigma_G^2 = 0$$

Os resultados do MINITAB são apresentados na Figura 11.4. A coluna de valor  $P$  na tabela ANOVA indica que as últimas duas hipóteses nulas devem ser rejeitadas no nível de significância 0,05. Os diferentes materiais usados na carcaça por si mesmos não parecem influenciar na vibração, mas a interação entre o material e fornecedor é uma fonte significativa de variação na vibração.

Fator	Tipo	Níveis	Valores				
casmater	fixo	3	1	2	3		
fonte	aleatório	5	1	2	3	4	5

	GL	SS	MS	F	P
Fonte casmater	2	0,7047	0,3523	0,24	0,790
fonte	4	36,6747	9,1687	6,32	0,013
fonte*casmater	8	11,6053	1,4507	13,03	0,000
Erro	15	1,6700	0,1113		
Total	29	50,6547			

Fonte	Componente de variância	Termo do erro	Quadrado Médio Esperado para Cada Termo (usando um modelo irrestrito)
1 casmater		3	(4) + 2(3) + Q[1]
2 fonte	1,2863	3	(4) + 2(3) + 6(2)
3 fonte*casmater	0,6697	4	(4) + 2(3)
4 Erro	0,1113		(4)

**Figura 11.4** Resultados gerados pela opção ANOVA mista do MINITAB para os dados do Exemplo 11.10

Quando pelo menos dois dos  $K_{ij}$ s são desiguais, os cálculos ANOVA são bem mais complexos que no caso em que  $K_{ij} = K$ , e não existe nenhuma fórmula ideal para a estatística de teste apropriada. Uma das referências do capítulo pode ser consultada para obter mais informações.

**Exercícios** | Seção 11.2 (16–26)

16. Em um experimento para avaliar os efeitos do tempo de tratamento (fator  $A$ ) e o tipo de mistura (fator  $B$ ) na resistência à compressão de cubos de cimento endurecido, três diferentes tempos de tratamento foram usados em conjunto com quatro diferentes misturas, obtendo-se três observações para cada uma das 12 combinações de tempo de tratamento/mistura. As somas de quadrados resultantes foram calculadas para ser  $SQA = 30.763,0$ ,  $SQB = 34.185,6$ ,  $SQE = 97.436,8$  e  $SQT = 205.966,6$ .
  - a. Elabore uma tabela ANOVA.
  - b. Teste no nível 0,05 a hipótese nula  $H_{0AB}$ : todos os  $\gamma_{ij} = 0$  (sem interação de fatores) em relação à  $H_{aAB}$ : pelo menos um  $\gamma_{ij} \neq 0$ .
  - c. Teste no nível 0,05 a hipótese nula  $H_{0A}$ :  $\alpha_1 = \alpha_2 = \alpha_3 = 0$  (não existem efeitos principais do fator  $A$ ) em relação a  $H_{aA}$ : pelo menos um  $\alpha_i \neq 0$ .
  - d. Teste  $H_{0B}$ :  $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$  versus  $H_{aB}$ : pelo menos um  $\beta_j \neq 0$  usando um nível de teste 0,05.
  - e. Os valores dos  $\bar{x}_{i..}$ s foram  $\bar{x}_{1..} = 4010,88$ ,  $\bar{x}_{2..} = 4029,10$ , e  $\bar{x}_{3..} = 3960,02$ . Use o procedimento de Tukey para investigar diferenças significativas entre os três tempos de tratamento.
17. O artigo "Towards Improving the Properties of Plaster Moulds and Castings" (*J. Engr. Manuf.*, 1991, p. 265-269) descreve várias ANOVAs realizadas para investigar como a quantidade acrescida de fibra de carbono e

de areia influencia em várias características do processo de moldagem. Aqui fornecemos dados sobre a dureza da fundição e a resistência do molde molhado.

Acréscimo de Areia (%)	Acréscimo de Fibra de Carbono (%)	Dureza de Fundição	Resistência do Molde Molhado
0	0	61,0	34,0
0	0	63,0	16,0
15	0	67,0	36,0
15	0	69,0	19,0
30	0	65,0	28,0
30	0	74,0	17,0
0	0,25	69,0	49,0
0	0,25	69,0	48,0
15	0,25	69,0	43,0
15	0,25	74,0	29,0
30	0,25	74,0	31,0
30	0,25	72,0	24,0
0	0,50	67,0	55,0
0	0,50	69,0	60,0
15	0,50	69,0	45,0
15	0,50	74,0	43,0
30	0,50	74,0	22,0
30	0,50	74,0	48,0

- a. Uma ANOVA da resistência do molde molhado obtém  $SQ_{\text{Areia}} = 705$ ,  $SQ_{\text{Fibra}} = 1278$ ,  $SQE = 843$  e  $SQT = 3105$ . Execute um teste para verificar se existe algum efeito usando  $\alpha = 0,05$ .
  - b. Execute uma ANOVA das observações da dureza da fundição usando  $\alpha = 0,05$ .
  - c. Plote um gráfico com a dureza média amostral versus percentagens de areia para diferentes níveis de fibra de carbono. O gráfico está consistente com suas análises na parte (b)?
18. Os dados a seguir provêm de um experimento realizado para investigar se o produto de um determinado processo químico dependia da formulação de um insumo específico ou da velocidade do misturador.

		Velocidade		
		60	70	80
Formulação	1	189,7	185,1	189,0
		188,6	179,4	193,0
		190,1	177,3	191,1
	2	165,1	161,7	163,3
		165,9	159,8	166,6
		167,6	161,6	170,3

Um pacote de software estatístico forneceu  $SQ(\text{Formulação}) = 2253,44$ ,  $SQ(\text{Velocidade}) = 230,81$ ,  $SQ(\text{Formulação} \times \text{Velocidade}) = 18,58$  e  $SQE = 71,87$ .

- a. De alguma forma parece haver interação entre os fatores?
  - b. O produto parece depender tanto da formulação como da velocidade?
  - c. Calcule as estimativas dos principais efeitos.
  - d. Os valores ajustados são  $\hat{x}_{ijk} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_{ij}$  e os resíduos são  $x_{ijk} - \hat{x}_{ijk}$ . Verifique que os resíduos são 0,23, -0,87, 0,63, 4,50, -1,20, -3,30, -2,03, 1,97, 0,07, -1,10, -0,30, 1,40, 0,67, -1,23, 0,57, -3,43, -0,13 e 3,57.
  - e. Elabore um gráfico de probabilidade normal com os resíduos fornecidos na parte (d). Os  $\epsilon_{ijk}$ s parecem ter distribuição normal?
19. A tabela de dados a seguir fornece observações sobre a acidez total de amostras de carvão de três diferentes tipos, em que as determinações foram feitas usando-se três diferentes concentrações de etanólico NaOH ("Chemistry of Brown Coals", *Australian J. Applied Science*, 1958, p. 375-379).

		Tipo de Carvão		
		Morwell	Yallourn	Maddingley
Concentração de NaOH	0,404N	8,27, 8,17	8,66, 8,61	8,14, 7,96
	0,626N	8,03, 8,21	8,42, 8,58	8,02, 7,89
	0,786N	8,60, 8,20	8,61, 8,76	8,13, 8,07

Além disso,  $\sum_i \sum_j \sum_k x_{ijk}^2 = 1240,1525$  e  $\sum_i \sum_j x_{ij.}^2 = 2479,9991$ .

- a. Supondo que ambos os efeitos sejam fixos, elabore uma tabela ANOVA, verifique se existe interação e, em seguida, verifique se existem efeitos principais em cada um dos fatores (para todos, use o nível 0,01).
  - b. Use o procedimento de Tukey para identificar diferenças significativas entre os tipos de carvão.
20. A corrente (em  $\mu A$ ) necessária para produzir um determinado nível de brilho em um tubo de televisão foi medida para dois tipos diferentes de vidro e três diferentes tipos de fósforo, gerando os dados a seguir ("Fundamentals of Analysis of Variance", *Industrial Quality Control*, 1956, p. 5-8).

		Tipo de Fósforo		
		1	2	3
Vidro1		280, 290, 285	300, 310, 295	270, 285, 290
Vidro2		230, 235, 240	260, 240, 235	220, 225, 230

Supondo que ambos os fatores sejam fixos, teste  $H_{0AB}$  versus  $H_{aAB}$  no nível 0,01. Em seguida, se  $H_{0AB}$  não puder ser rejeitada, teste os dois conjuntos de hipóteses de efeito principal.

- 21. Em um experimento para investigar o efeito do "fator cimento" (número de sacos de cimento por jarda cúbica) sobre a resistência à flexão do concreto resultante ("Studies of Flexural Strength of Concrete. Part 3: Effects of Variation in Testing Procedure", *Proceedings ASTM*, 1957, p. 1127-1139),  $I = 3$  diferentes valores de fator foram usados,  $J = 5$  diferentes lotes de cimento foram selecionados e  $K = 2$  vigas foram modeladas de cada combinação entre fator cimento/lote. Os valores resumidos incluem  $\sum \sum \sum x_{ijk}^2 = 12.280.103$ ,  $\sum \sum x_{ij.}^2 = 24.529.699$ ,  $\sum x_{i..}^2 = 122.380.901$ ,  $\sum x_{.j.}^2 = 73.427.483$  e  $x_{...} = 19.143$ .
  - a. Elabore a tabela ANOVA.
  - b. Supondo um modelo misto de fator cimento (A) fixo e lotes (B) aleatórios, teste os três pares de hipóteses de interesse, no nível 0,05.
22. Um estudo foi executado para comparar a vida útil de quatro marcas de caneta esferográfica especial. Por se ter acreditado que a superfície de escrita podia influenciar na vida útil, três diferentes superfícies foram selecionadas aleatoriamente. Foi usada uma máquina de escrever para garantir que as condições fossem homogêneas em outras circunstâncias (por exemplo, pressão constante e ângulo fixo). A tabela a seguir mostra os dois tempos de vida útil (min.) obtidos para cada combinação entre marca-superfície. Além disso,  $\sum \sum \sum x_{ijk}^2 = 11.499.492$  e  $\sum \sum x_{ij.}^2 = 22.982.552$ .

		Superfície de Escrita			$x_{i..}$
		1	2	3	
Marca da Caneta	1	709, 659	713, 726	660, 645	4112
	2	668, 685	722, 740	692, 720	4227
	3	659, 685	666, 684	678, 750	4122
	4	698, 650	704, 666	686, 733	4137
	$x_{.j.}$	5413	5621	5564	16,598

Execute uma ANOVA apropriada e expresse suas conclusões.

23. Os dados a seguir foram obtidos em um experimento realizado para investigar se a resistência à compressão de cilindros de concreto depende do tipo de material de revestimento usado ou da variabilidade em lotes diferentes (“The Effect of Type of Capping Material on the Compressive Strength of Concrete Cylinders”, *Proceedings ASTM*, 1958, p. 1166–1186). Cada um dos números é um total da célula ( $x_{ij}$ ) tendo por base  $K = 3$  observações.

		Lote				
		1	2	3	4	5
Material de Revestimento	1	1847	1942	1935	1891	1795
	2	1779	1850	1795	1785	1626
	3	1806	1892	1889	1891	1756

Além disso,  $\sum\sum x_{ijk}^2 = 16.815.853$  e  $\sum\sum x_{ij}^2 = 50.443.409$ . Obtenha a tabela ANOVA e, em seguida, teste as hipóteses  $H_{0G}$  versus  $H_{aG}$ ,  $H_{0A}$  versus  $H_{aA}$  e  $H_{0B}$  versus  $H_{aB}$ , no nível 0,01, supondo que o revestimento seja um efeito fixo e os lotes, um efeito aleatório.

24. a. Mostre que  $E(\bar{X}_{i..} - \bar{X}_{...}) = \alpha_i$ , de modo que  $\bar{X}_{i..} - \bar{X}_{...}$  é um estimador não-tendencioso de  $\alpha_i$  (no modelo de efeitos fixos).  
 b. Sendo  $\hat{\gamma}_{ij} = \bar{X}_{ij.} - \bar{X}_{i..} - \bar{X}_{.j.} + \bar{X}_{...}$ , mostre que  $\hat{\gamma}_{ij}$  é um estimador não-tendencioso para  $\gamma_{ij}$  (no modelo de efeitos fixos).
25. Mostre como um IC  $t$   $100(1 - \alpha)\%$  para  $\alpha_i - \alpha_j$ , pode ser obtido. Em seguida, calcule um intervalo de 95% para  $\alpha_2 - \alpha_3$  usando os dados do Exercício 19. [Sugestão: com  $\theta = \alpha_2 - \alpha_3$ , o resultado do Exercício 24a indica como obter  $\hat{\theta}$ . Em seguida, calcule  $V(\hat{\theta})$  e  $\sigma_{\hat{\theta}}$  e obtenha uma estimativa de  $\sigma_{\hat{\theta}}$  usando  $\sqrt{QME}$  para estimar  $\sigma$  (que identifica o número apropriado de gl).]
26. Quando ambos os fatores são aleatórios em um experimento ANOVA bidirecional, com  $K$  replicações por combinação de níveis do fator, os quadrados médios esperados são  $E(QME) = \sigma^2$ ,  $E(QMA) = \sigma^2 + K\sigma_G^2 + JK\sigma_A^2$ ,  $E(QMB) = \sigma^2 + K\sigma_G^2 + IK\sigma_B^2$  e  $E(QMAB) = \sigma^2 + K\sigma_G^2$ .
- a. Que razão  $F$  é apropriada para testar  $H_{0G}$ :  $\sigma_G^2 = 0$  versus  $H_{aG}$ :  $\sigma_G^2 > 0$ ?  
 b. Responda à parte (a) para testar  $H_{0A}$ :  $\sigma_A^2 = 0$  versus  $H_{aA}$ :  $\sigma_A^2 > 0$  e  $H_{0B}$ :  $\sigma_B^2 = 0$  versus  $H_{aB}$ :  $\sigma_B^2 > 0$ .

### 11.3 ANOVA de Três Fatores

Para indicar a natureza dos modelos e análises quando os experimentos ANOVA contêm mais de dois fatores, vamos nos concentrar aqui no caso de três fatores fixos –  $A$ ,  $B$  e  $C$ . O número de níveis dos três fatores será representado por  $I$ ,  $J$  e  $K$ , respectivamente, e  $L_{ijk}$  = o número de observações feitas com o fator  $A$  no nível  $i$ , com o fator  $B$  no nível  $j$  e com o fator  $C$  no nível  $k$ . Como na ANOVA de dois fatores, a análise fica bem complicada quando os  $L_{ijk}$ s não são todos iguais, então especificamos ainda mais  $L_{ijk} = L$ . Portanto,  $X_{ijkl}$  e  $x_{ijkl}$  representam o valor observado, antes e depois que o experimento é realizado, da  $l$ -ésima replicação ( $l = 1, 2, \dots, L$ ) quando os três fatores são fixos nos níveis  $i, j$  e  $k$ .

Para compreender os parâmetros que aparecerão no modelo ANOVA de três fatores, primeiro lembre-se de que na ANOVA de dois fatores com replicações,  $E(X_{ijk}) = \mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}$ , onde as restrições  $\sum_i \alpha_i = \sum_j \beta_j = 0$ ,  $\sum_i \gamma_{ij} = 0$  para todo  $j$  e  $\sum_j \gamma_{ij} = 0$  para todo  $i$  foram necessárias para obter um conjunto único de parâmetros. Se usarmos pontos subscritos nos  $\mu_{ij}$ s para representar o cálculo da média (em vez da adição), então

$$\mu_{i.} - \mu_{..} = \frac{1}{J} \sum_j \mu_{ij} - \frac{1}{IJ} \sum_i \sum_j \mu_{ij} = \alpha_i$$

será o efeito do fator  $A$  no nível  $i$  calculado pela média de níveis do fator  $B$ , ao passo que

$$\mu_{ij} - \mu_{.j} = \mu_{ij} - \frac{1}{I} \sum_i \mu_{ij} = \alpha_i + \gamma_{ij}$$

será o efeito do fator  $A$  no nível  $i$  específico do fator  $B$  no nível  $j$ . Se o efeito de  $A$  no nível  $i$  depender do nível de  $B$ , haverá interação entre os fatores e os  $\gamma_{ij}$ s não serão todos zero. Em particular,

$$\mu_{ij} - \mu_{.j} - \mu_{i.} + \mu_{..} = \gamma_{ij} \tag{11.11}$$

## O Modelo de Efeitos Fixos de Três Fatores

O modelo para a ANOVA de três fatores com  $L_{ijk} = L$  é

$$X_{ijkl} = \mu_{ijk} + \epsilon_{ijkl} \quad \begin{array}{l} i = 1, \dots, I, \quad j = 1, \dots, J \\ k = 1, \dots, K, \quad l = 1, \dots, L \end{array} \quad (11.12)$$

onde os  $\epsilon_{ijkl}$ s têm distribuição normal com média 0 e variância  $\sigma^2$ , e

$$\mu_{ijk} = \mu + \alpha_i + \beta_j + \delta_k + \gamma_{ij}^{AB} + \gamma_{ik}^{AC} + \gamma_{jk}^{BC} + \gamma_{ijk} \quad (11.13)$$

As restrições necessárias para obter parâmetros definidos exclusivamente são as de que as somas de qualquer subscrito de qualquer parâmetro no lado direito de (11.13) sejam iguais a 0.

Os parâmetros  $\gamma_{ij}^{AB}$ ,  $\gamma_{ik}^{AC}$  e  $\gamma_{jk}^{BC}$  são chamados de interação de dois fatores e  $\gamma_{ijk}$  é chamado de interação de três fatores; os  $\alpha_i$ s,  $\beta_j$ s e  $\delta_k$ s são os parâmetros dos efeitos principais. Para qualquer nível  $k$  fixo do terceiro fator, analogamente a (11.11),

$$\mu_{ijk} - \mu_{i..} - \mu_{.jk} + \mu_{..k} = \gamma_{ij}^{AB} + \gamma_{ijk}$$

é a interação do  $i$ -ésimo nível de  $A$  com o  $j$ -ésimo nível de  $B$  específico ao  $k$ -ésimo nível de  $C$ , ao passo que

$$\mu_{ij.} - \mu_{i..} - \mu_{.j.} + \mu_{...} = \gamma_{ij}^{AB}$$

é a interação entre  $A$  no nível  $i$  e  $B$  no nível  $j$  calculados pela média de níveis de  $C$ . Se a interação de  $A$  no nível  $i$  e de  $B$  no nível  $j$  não depender de  $k$ , todos os  $\gamma_{ijk}$ s serão iguais a 0. Portanto,  $\gamma_{ijk}$ s não-zero denotam a não-aditividade dos dois fatores  $\gamma_{ij}^{AB}$ s nos vários níveis do terceiro fator  $C$ . Se o experimento incluísse mais de três fatores, haveria termos correspondentes de interação de ordem superior com interpretações análogas. Observe que, no argumento anterior, se tivéssemos resolvido fixar o nível de  $A$  ou de  $B$  (em vez de o nível de  $C$ , como foi feito) e examinar os  $\gamma_{ijk}$ s, as respectivas interpretações seriam as mesmas — se qualquer das interações dos dois fatores depender do nível do terceiro fator, então há  $\gamma_{ijk}$ s não-zero.

## Análise de um Experimento de Três Fatores

Quando  $L > 1$ , há uma soma de quadrados para cada efeito principal, para cada interação de dois fatores e para a interação de três fatores. Para escrevê-las de forma que indique como as somas de quadrados são definidas quando há mais de três fatores, observe que nenhum dos parâmetros e modelos em (11.13) pode ser estimado de maneira não-tendenciosa, calculando-se a média de  $X_{ijkl}$  ao longo de subscritos apropriados e obtendo diferenças. Dessa forma,

$$\begin{aligned} \hat{\mu} &= \bar{X}_{....} & \hat{\alpha}_i &= \bar{X}_{i...} - \bar{X}_{....} & \hat{\gamma}_{ij}^{AB} &= \bar{X}_{ij..} - \bar{X}_{i...} - \bar{X}_{.j.} + \bar{X}_{....} \\ \hat{\gamma}_{ijk} &= \bar{X}_{ijk.} - \bar{X}_{ij..} - \bar{X}_{i.k.} - \bar{X}_{.jk.} + \bar{X}_{i..} + \bar{X}_{.j.} + \bar{X}_{..k} - \bar{X}_{....} \end{aligned}$$

e os outros efeitos principais e estimadores de interação são obtidos por simetria. As somas de quadrados são

$$\begin{aligned} \text{SQT} &= \sum_i \sum_j \sum_k \sum_l (X_{ijkl} - \bar{X}_{....})^2 & \text{gl} &= IJKL - 1 \\ \text{SQA} &= \sum_i \sum_j \sum_k \sum_l \hat{\alpha}_i^2 = JKL \sum_i (\bar{X}_{i...} - \bar{X}_{....})^2 & \text{gl} &= I - 1 \\ \text{SQAB} &= \sum_i \sum_j \sum_k \sum_l (\hat{\gamma}_{ij}^{AB})^2 = KL \sum_i \sum_j (\bar{X}_{ij..} - \bar{X}_{i...} - \bar{X}_{.j.} + \bar{X}_{....})^2 & \text{gl} &= (I - 1)(J - 1) \\ \text{SQABC} &= \sum_i \sum_j \sum_k \sum_l \hat{\gamma}_{ijk}^2 = L \sum_i \sum_j \sum_k \hat{\gamma}_{ijk}^2 & \text{gl} &= (I - 1)(J - 1)(K - 1) \\ \text{SQE} &= \sum_i \sum_j \sum_k \sum_l (X_{ijkl} - \bar{X}_{ijk.})^2 & \text{gl} &= IJK(L - 1) \end{aligned}$$

e os outros efeitos principais e as somas de quadrados da interação de dois fatores são obtidos por simetria. SQT é a soma das oito SQs restantes.

Mesmo as fórmulas computacionais para essas SQs são extremamente entediante. Por esse motivo, nós as evitamos em favor dos dados produzidos por um pacote de *software* estatístico. A versão atual do MINITAB, por exemplo, adaptará um modelo de três fatores com efeitos fixos, mistos ou aleatórios.

Cada soma de quadrados (exceto a SQT) quando dividida por seu gl resulta em um quadrado médio, com

$$E(QME) = \sigma^2$$

$$E(QMA) = \sigma^2 + \frac{JKL}{I-1} \sum_i \alpha_i^2$$

$$E(QMAB) = \sigma^2 + \frac{KL}{(I-1)(J-1)} \sum_i \sum_j (\gamma_{ij}^{AB})^2$$

$$E(QMABC) = \sigma^2 + \frac{L}{(I-1)(J-1)(K-1)} \sum_i \sum_j \sum_k (\gamma_{ijk})^2$$

e expressões semelhantes para os outros quadrados médios esperados. As hipóteses do efeito principal e da interação são testadas, dispondo-se as razões *F* com o QME em cada denominador.

Hipótese Nula	Valor do Teste Estatístico	Região de Rejeição
$H_{0A}$ : todos os $\alpha_i = 0$	$f_A = \frac{QMA}{QME}$	$f_A \geq F_{\alpha, I-1, JK(L-1)}$
$H_{0AB}$ : todos os $\gamma_{ij}^{AB} = 0$	$f_{AB} = \frac{QMAB}{QME}$	$f_{AB} \geq F_{\alpha, (I-1)(J-1), JK(L-1)}$
$H_{0ABC}$ : todos os $\gamma_{ijk} = 0$	$f_{ABC} = \frac{QMABC}{QME}$	$f_{ABC} \geq F_{\alpha, (I-1)(J-1)(K-1), JK(L-1)}$

Em geral, as hipóteses do efeito principal são testadas apenas se todas as interações não forem consideradas significativas.

Essa análise pressupõe que  $L_{ijk} = L > 1$ . Se  $L = 1$ , portanto, como no caso de dois fatores, as interações da mais alta ordem devem ser consideradas iguais a 0 para se obter um QME que calcule  $\sigma^2$ . Estabelecendo  $L = 1$  e desconsiderando a soma do quarto subscrito sobre *l*, as fórmulas precedentes para as somas de quadrados ainda permanecem válidas, e  $SQE = \sum_i \sum_j \sum_k \hat{\gamma}_{ijk}^2$  com  $\bar{X}_{ijk.} = X_{ijk}$  na expressão para  $\hat{\gamma}_{ijk}$ .

**Exemplo 11.11**

As observações a seguir (temperatura corporal – 100°F) foram relatadas em um experimento realizado para investigar a tolerância do gado ao calor (“The Significance of the Coat in Heat Tolerance of Cattle”, *Australian J. Agriculture Research*, 1959, p. 744-748). As avaliações foram feitas em quatro períodos diferentes (fator *A*, com  $I = 4$ ) em duas diferentes linhagens de gado (fator *B*, com  $J = 2$ ), tendo quatro tipos distintos de pelagem (fator *C*, com  $K = 4$ );  $L = 3$  observações foram feitas para cada uma das  $4 \times 2 \times 4 = 32$  combinações de níveis dos três fatores.

	$B_1$				$B_2$			
	$C_1$	$C_2$	$C_3$	$C_4$	$C_1$	$C_2$	$C_3$	$C_4$
$A_1$	3,6	3,4	2,9	2,5	4,2	4,4	3,6	3,0
	3,8	3,7	2,8	2,4	4,0	3,9	3,7	2,8
	3,9	3,9	2,7	2,2	3,9	4,2	3,4	2,9
$A_2$	3,8	3,8	2,9	2,4	4,4	4,2	3,8	2,0
	3,6	3,9	2,9	2,2	4,4	4,3	3,7	2,9
	4,0	3,9	2,8	2,2	4,6	4,7	3,4	2,8
$A_3$	3,7	3,8	2,9	2,1	4,2	4,0	4,0	2,0
	3,9	4,0	2,7	2,0	4,4	4,6	3,8	2,4
	4,2	3,9	2,8	1,8	4,5	4,5	3,3	2,0
$A_4$	3,6	3,6	2,6	2,0	4,0	4,0	3,8	2,0
	3,5	3,7	2,9	2,0	4,1	4,4	3,7	2,2
	3,8	3,9	2,9	1,9	4,2	4,2	3,5	2,3

A tabela de totais de célula ( $x_{ijk}$ 's) de todas as combinações dos três fatores é

$x_{ijk}$	$B_1$				$B_2$			
	$C_1$	$C_2$	$C_3$	$C_4$	$C_1$	$C_2$	$C_3$	$C_4$
$A_1$	11,3	11,0	8,4	7,1	12,1	12,5	10,7	8,7
$A_2$	11,4	11,6	8,6	6,8	13,4	13,2	10,9	7,7
$A_3$	11,8	11,7	8,4	5,9	13,1	13,1	11,1	6,4
$A_4$	10,9	11,2	8,4	5,9	12,3	12,6	11,0	6,5

A Figura 11.5 exibe os gráficos das médias das células correspondentes  $\bar{x}_{ijk} = x_{ijk}/3$ . Retornaremos a esses gráficos depois de examinarmos os testes de várias hipóteses. O ponto de partida para esses testes é a tabela ANOVA apresentada na Tabela 11.8.

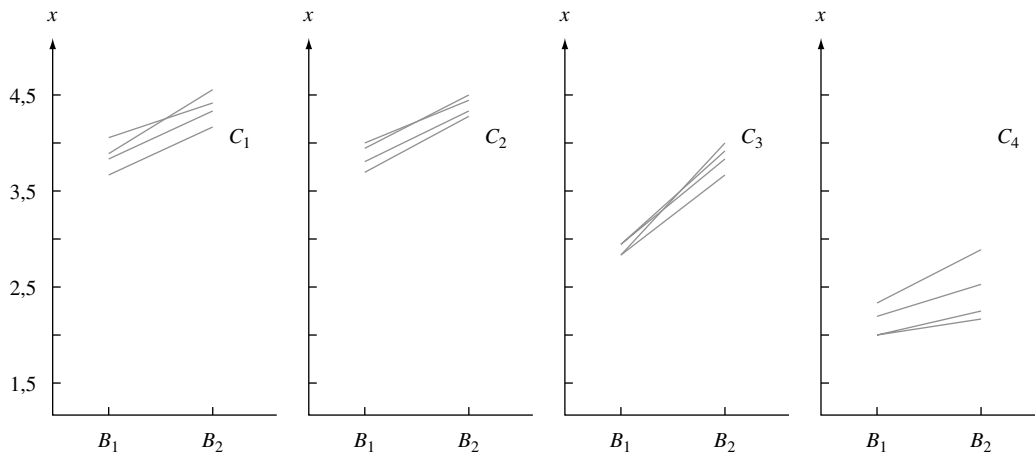


Figura 11.5 Gráficos de  $x_{ijk}$ , para o Exemplo 11.11

Tabela 11.8 Tabela ANOVA do Exemplo 11.11

Fonte	gl	Soma dos Quadrados	Quadrado Médio	$f$
A	$I - 1 = 3$	0,49	0,163	4,13
B	$J - 1 = 1$	6,45	6,45	163,29
C	$K - 1 = 3$	48,93	16,31	412,91
AB	$(I - 1)(J - 1) = 3$	0,02	0,0067	0,170
AC	$(I - 1)(K - 1) = 9$	1,61	0,179	4,53
BC	$(J - 1)(K - 1) = 3$	0,88	0,293	7,42
ABC	$(I - 1)(J - 1)(K - 1) = 9$	0,25	0,0278	0,704
Erro	$IJK(L - 1) = 64$	2,53	0,0395	
Total	$IJKL - 1 = 95$	61,16		

Visto que  $F_{0,01,9,64} \approx 2,70$  e  $f_{ABC} = \text{QM}_{ABC}/\text{QME} = 0,704$  não ultrapassa 2,70, concluímos que as interações de três fatores não são significativas. Contudo, embora as interações  $AB$  também não sejam significativas, tanto as interações  $AC$  e  $BC$  como os efeitos principais parecem necessários no modelo. Quando não houver nenhuma interação  $ABC$  ou  $AB$ , um gráfico dos  $\bar{x}_{ijk}$ 's ( $= \hat{\mu}_{ijk}$ ) separadamente de cada nível de  $C$  não deve revelar nenhuma interação substancial (se apenas as interações  $ABC$  forem iguais a zero, os gráficos serão mais difíceis de interpretar; consulte o artigo "Two-Dimensional Plots for Interpreting Interactions in the Three-Factor Analysis of Variance Model", *Amer. Statistician*, maio de 1979, p. 63-69). ■

O procedimento de Tukey pode ser usado na ANOVA de três (ou mais) fatores. O segundo subscrito em  $Q$  é o número de médias amostrais que estão sendo comparadas e o terceiro são os graus de liberdade do erro.



Também podem ser analisados modelos com efeitos aleatórios e mistos. As somas de quadrados e os graus de liberdade são idênticos no caso dos efeitos fixos, mas os quadrados médios esperados são, obviamente, diferentes para os efeitos principais aleatórios ou as interações. Uma boa referência é o livro de Douglas Montgomery relacionado na bibliografia deste capítulo.

### Modelos em Quadrado Latino

Quando vários fatores precisam ser investigados simultaneamente, o experimento em que há pelo menos uma observação para cada combinação possível de níveis é chamado de **layout completo**. Se os fatores forem  $A, B$  e  $C$  com  $I, J$  e  $K$  níveis, respectivamente, um *layout* completo exige pelo menos  $IJK$  observações. Em geral, um experimento dessa proporção é impraticável por limitações de custo, tempo ou espaço ou é praticamente impossível. Por exemplo: se a variável de resposta for a venda de um determinado produto e os fatores forem diferentes formas de exposição, diferentes lojas e diferentes períodos de tempo, apenas uma forma de exposição pode ser usada de modo realista em uma determinada loja durante um período de tempo específico.

O experimento de três fatores em que são feitas menos de  $IJK$  observações é denominado **layout incompleto**. Há alguns *layouts* incompletos em que o padrão de combinação de fatores é tal que a análise é direta. Um modelo de três fatores desse tipo é o **quadrado latino**. Esse modelo é apropriado quando  $I = J = K$  (por exemplo: quatro formas de exposição, quatro lojas e quatro períodos de tempo) e todas as interações de dois e três fatores são consideradas inexistentes. Se os níveis do fator  $A$  forem identificados com as linhas de uma tabela bidimensional e os níveis do fator  $B$  com as colunas dessa tabela, a característica definidora de um modelo em quadrado latino é que *todo nível do fator C aparecerá exatamente uma vez em cada linha e exatamente uma vez em cada coluna*. Na Figura 11.6 são apresentados exemplos de quadrados latinos  $3 \times 3, 4 \times 4$  e  $5 \times 5$ . Há diferentes quadrados latinos  $3 \times 3$  e o número de diferentes quadrados latinos  $N \times N$  aumenta rapidamente com  $N$  (por exemplo: toda permutação de linha de um determinado quadrado latino gera um quadrado latino, e isso ocorre de maneira semelhante para permutações de coluna). É recomendável que o quadrado de fato usado em um determinado experimento seja escolhido aleatoriamente do conjunto de quadrados possíveis da dimensão desejada; para obter mais detalhes, consulte uma das referências deste capítulo.

**Modelo e Análise para os Quadrados Latinos** A letra  $N$  será usada para representar o valor comum de  $I, J$  e  $K$ . Portanto, um *layout* completo com uma observação por combinação exigiria  $N^3$  observações, ao passo que um quadrado latino exige apenas  $N^2$ . Assim que um quadrado específico for escolhido, o valor de  $k$  (o nível do fator  $C$ ) é determinado completamente pelos valores de  $i$  e  $j$ .

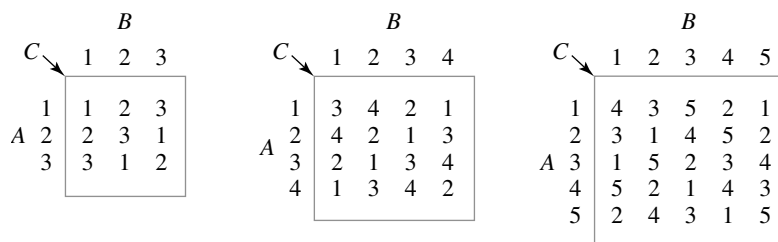


Figura 11.6 Exemplos de quadrados latinos

Para enfatizar isso, usamos  $x_{ij(k)}$  para representar o valor observado quando os três fatores estão nos níveis  $i, j$  e  $k$ , respectivamente, onde  $k$  assume um único valor para cada par  $i, j$ . Portanto, o modelo é

$$X_{ij(k)} = \mu + \alpha_i + \beta_j + \delta_k + \epsilon_{ij(k)} \quad i, j, k = 1, \dots, N$$

onde  $\sum \alpha_i = \sum \beta_j = \sum \delta_k = 0$  e os  $\epsilon_{ij(k)}$ s são independentes e normalmente distribuídos com média 0 e variância  $\sigma^2$ .

Empregamos a notação a seguir para totais e médias:

$$X_{i..} = \sum_j X_{ij(k)} \quad X_{.j.} = \sum_i X_{ij(k)} \quad X_{..k} = \sum_{ij} X_{ij(k)} \quad X_{...} = \sum_i \sum_j X_{ij(k)}$$

$$\bar{X}_{i..} = \frac{X_{i..}}{N} \quad \bar{X}_{.j.} = \frac{X_{.j.}}{N} \quad \bar{X}_{..k} = \frac{X_{..k}}{N} \quad \bar{X}_{...} = \frac{X_{...}}{N^2}$$

Observe que, embora  $X_{i..}$  anteriormente sugerisse uma soma dupla, agora corresponde a uma única soma sobre todos os  $j$  (e os valores associados de  $k$ ). As somas dos quadrados usados na análise são a seguir apresentadas.

$$\begin{aligned} \text{SQT} &= \sum_i \sum_j (X_{ij(k)} - \bar{X}_{...})^2 = \sum_i \sum_j X_{ij(k)}^2 - \frac{X_{...}^2}{N^2} & \text{gl} &= N^2 - 1 \\ \text{SQA} &= \sum_i \sum_j (\bar{X}_{i..} - \bar{X}_{...})^2 = \frac{1}{N} \sum_i X_{i..}^2 - \frac{X_{...}^2}{N^2} & \text{gl} &= N - 1 \\ \text{SQB} &= \sum_i \sum_j (\bar{X}_{.j.} - \bar{X}_{...})^2 = \frac{1}{N} \sum_j X_{.j.}^2 - \frac{X_{...}^2}{N^2} & \text{gl} &= N - 1 \\ \text{SQC} &= \sum_i \sum_j (\bar{X}_{..k} - \bar{X}_{...})^2 = \frac{1}{N} \sum_k X_{..k}^2 - \frac{X_{...}^2}{N^2} & \text{gl} &= N - 1 \\ \text{SQE} &= \sum_i \sum_j [X_{ij(k)} - (\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\delta}_k)]^2 \\ &= \sum_i \sum_j (X_{ij(k)} - \bar{X}_{i..} - \bar{X}_{.j.} - \bar{X}_{..k} + 2\bar{X}_{...})^2 & \text{gl} &= (N - 1)(N - 2) \\ \text{SQT} &= \text{SQA} + \text{SQB} + \text{SQC} + \text{SQE} \end{aligned}$$

Cada quadrado médio é, obviamente, a razão SQ/gl. Para testar  $H_{0C}$ :  $\delta_1 = \delta_2 = \dots = \delta_N = 0$ , o valor da estatística de teste é  $f_C = \text{QMC}/\text{QME}$ , em que  $H_{0C}$  é rejeitada se  $f_C \geq F_{\alpha, N-1, (N-1)(N-2)}$ . As outras duas hipóteses do efeito principal são também rejeitadas se a razão  $F$  correspondente ultrapassar  $F_{\alpha, N-1, (N-1)(N-2)}$ .

Se nenhuma das hipóteses nulas for rejeitada, diferenças significativas podem ser identificadas por meio do procedimento de Tukey. Após o cálculo de  $w = Q_{\alpha, N, (N-1)(N-2)} \cdot \sqrt{\text{QME}/N}$ , os pares de médias amostrais (os  $\bar{x}_{i..}$ s,  $\bar{x}_{.j.}$ s, ou  $\bar{x}_{..k}$ s) que diferirem mais do que  $w$  correspondem a diferenças significativas entre os efeitos associados do fator (os  $\alpha_i$ s,  $\beta_j$ s ou  $\delta_k$ s).

A hipótese  $H_{0C}$  geralmente é a de maior interesse. O modelo em quadrado latino é usado para controlar variações externas nos fatores  $A$  e  $B$ , como no modelo em bloco aleatorizado para o caso de um único fator externo. Logo, no exemplo de venda do produto mencionado anteriormente, a variação decorrente das lojas e dos períodos de tempo é controlada por um modelo em quadrado latino, que permite que um investigador tente identificar efeitos decorrentes de formas diferentes de exposição do produto.

### Exemplo 11.12

Em um experimento para investigar o efeito da umidade relativa sobre a resistência à abrasão do corte de couro em um molde retangular (“The Abrasion of Leather”, *J. Inter. Soc. Leather Trades' Chemists*, 1946, p. 287), foi usado um quadrado latino 6 X 6 para controlar uma possível variabilidade decorrente da posição da linha ou coluna no molde. Os seis níveis de umidade relativa investigados foram 1 = 25%, 2 = 37%, 3 = 50%, 4 = 62%, 5 = 75% e 6 = 87%, gerando os seguintes resultados:

		B (colunas)						
		1	2	3	4	5	6	$x_{i..}$
A (linhas)	1	<sup>3</sup> 7,38	<sup>4</sup> 5,39	<sup>6</sup> 5,03	<sup>2</sup> 5,50	<sup>5</sup> 5,01	<sup>1</sup> 6,79	35,10
	2	<sup>2</sup> 7,15	<sup>1</sup> 8,16	<sup>5</sup> 4,96	<sup>4</sup> 5,78	<sup>3</sup> 6,24	<sup>6</sup> 5,06	37,35
	3	<sup>4</sup> 6,75	<sup>6</sup> 5,64	<sup>3</sup> 6,34	<sup>5</sup> 5,31	<sup>1</sup> 7,81	<sup>2</sup> 8,05	39,90
	4	<sup>1</sup> 8,05	<sup>3</sup> 6,45	<sup>2</sup> 6,31	<sup>6</sup> 5,46	<sup>4</sup> 6,05	<sup>5</sup> 5,51	37,83
	5	<sup>6</sup> 5,65	<sup>5</sup> 5,44	<sup>1</sup> 7,27	<sup>3</sup> 6,54	<sup>2</sup> 7,03	<sup>4</sup> 5,96	37,89
	6	<sup>5</sup> 6,00	<sup>2</sup> 6,55	<sup>4</sup> 5,93	<sup>1</sup> 8,02	<sup>6</sup> 5,80	<sup>3</sup> 6,61	38,91
$x_{.j.}$		40,98	37,63	35,84	36,61	37,94	37,98	

Além disso,  $x_{..1} = 46,10$ ,  $x_{..2} = 40,59$ ,  $x_{..3} = 39,56$ ,  $x_{..4} = 35,86$ ,  $x_{..5} = 32,23$ ,  $x_{..6} = 32,64$ ,  $x_{...} = 226,98$ , e  $\sum_i \sum_j x_{ij(k)}^2 = 1462,89$ . Cálculos adicionais estão resumidos na Tabela 11.9.

**Tabela 11.9** Tabela ANOVA do Exemplo 11.12

Fonte de Variação	gl	Soma dos Quadrados	Quadrado Médio	$f$
A (linhas)	5	2,19	0,438	2,50
B (colunas)	5	2,57	0,514	2,94
C (tratamentos)	5	23,53	4,706	26,89
Erro	20	3,49	0,175	
Total	35	31,78		

Visto que  $F_{0,05,5,20} = 2,71$  e  $26,89 \geq 2,71$ ,  $H_{0C}$  é rejeitada em favor da hipótese de que, em média, a umidade relativa influencia na resistência à abrasão.

Para aplicar o procedimento de Tukey,  $w = Q_{0,05,6,20} \cdot \sqrt{QME/6} = 4,45\sqrt{0,175/6} = 0,76$ . Ordenando os  $\bar{x}_{..k}$ s e sublinhando, obtemos

75%	87%	62%	50%	37%	25%
<u>5,37</u>	5,44	<u>5,98</u>	<u>6,59</u>	6,77	7,68

Particularmente, a menor umidade relativa parece resultar em uma resistência à abrasão média real significativamente maior do que a de qualquer outra umidade relativa investigada. ■

## Exercícios | Seção 11.3 (27–37)

**27.** A produção de uma máquina de extrusão contínua de revestimento de tubos de aço com plástico foi investigada como função do perfil de temperatura do termostato ( $A$ , em três níveis), do tipo de plástico ( $B$ , em três níveis) e da velocidade do pino giratório que força o plástico por meio de uma matriz de fabricação de tubos ( $C$ , em três níveis). Havia duas replicações ( $L = 2$ ) em cada combinação de níveis dos fatores, resultando um total de 54 observações sobre a produção. As somas dos quadrados foram  $SQA = 14.144,44$ ,  $SQB = 5511,27$ ,  $SQC = 244.696,39$ ,  $SQAB = 1069,62$ ,  $SQAC = 62,67$ ,  $SQBC = 331,67$ ,  $SQE = 3127,50$  e  $SQT = 270.024,33$ .

- Elabore a tabela ANOVA.
- Use testes  $F$  apropriados para mostrar que nenhuma das razões  $F$  para duas ou três interações é significativa ao nível 0,05.
- Que efeitos principais parecem significativos?
- Com  $x_{..1} = 8242$ ,  $x_{..2} = 9732$ , e  $x_{..3} = 11.210$ , use o procedimento de Tukey para identificar diferenças significativas entre os níveis do fator  $C$ .

**28.** Para verificar se a força de empuxo na perfuração é influenciada pela velocidade da perfuração ( $A$ ), pela taxa de alimentação ( $B$ ) ou pelo material usado ( $C$ ), foi reali-

zando um experimento usando quatro velocidades, três taxas e dois tipos de material, com duas amostras ( $L = 2$ ) perfuradas em cada combinação de níveis dos três fatores. As somas dos quadrados foram calculadas como a seguir:  $SQA = 19.149,73$ ,  $SQB = 2.589.047,62$ ,  $SQC = 157.437,52$ ,  $SQAB = 53.238,21$ ,  $SQAC = 9033,73$ ,  $SQBC = 91.880,04$ ,  $SQE = 56.819,50$  e  $SQT = 2.983.164,81$ . Elabore uma tabela ANOVA e identifique interações significativas usando  $\alpha = 0,01$ . Por acaso algum fator isolado parece não exercer nenhum efeito sobre a força de empuxo? (Em outras palavras, algum fator parece não-significativo em todos os efeitos em que aparece?)

**29.** O artigo "An Analysis of Variance Applied to Screw Machines" (*Industrial Quality Control*, 1956, p. 8-9) descreve um experimento realizado para investigar como o comprimento das barras de aço era influenciado pelo horário do dia ( $A$ ), pelo tratamento térmico aplicado ( $B$ ) e pela máquina de fabricar parafusos utilizada ( $C$ ). Os três horários especificados foram 8h, 11h e 15h, e havia dois tratamentos e quatro máquinas (um experimento fatorial  $3 \times 2 \times 4$ ), cujos resultados são apresentados a seguir [codificado como 1000 (comprimento – 4,380), o que não afeta a análise].

		$B_1$			
		$C_1$	$C_2$	$C_3$	$C_4$
$A_1$	6, 9, 1, 3	7, 9, 5, 5	1, 2, 0, 4	6, 6, 7, 3	
$A_2$	6, 3, 1, -1	8, 7, 4, 8	3, 2, 1, 0	7, 9, 11, 6	
$A_3$	5, 4, 9, 6	10, 11, 6, 4	-1, 2, 6, 1	10, 5, 4, 8	

		$B_2$			
		$C_1$	$C_2$	$C_3$	$C_4$
$A_1$	4, 6, 0, 1	6, 5, 3, 4	-1, 0, 0, 1	4, 5, 5, 4	
$A_2$	3, 1, 1, -2	6, 4, 1, 3	2, 0, -1, 1	9, 4, 6, 3	
$A_3$	6, 0, 3, 7	8, 7, 10, 0	0, -2, 4, -4	4, 3, 7, 0	

As somas dos quadrados incluem SQAB = 1,646, SQAC = 71,021, SQBC = 1,542, SQE = 447,500 e SQT = 1037,833.

- a. Elabore a tabela ANOVA para esses dados.
- b. Verifique se algum efeito da interação é significativo ao nível 0,05.
- c. Verifique se algum dos efeitos principais é significativo ao nível 0,05 (isto é,  $H_{0A}$  versus  $H_{aA}$  etc.).
- d. Use o procedimento de Tukey para investigar as diferenças significativas entre as quatro máquinas.

30. As quantidades resumidas a seguir foram calculadas com base em um experimento envolvendo quatro níveis de nitrogênio (A), dois horários de plantio (B) e dois níveis de potássio (C) ("Use and Misuse of Multiple Comparison Procedures", *Agronomy J.*, 1977, p. 205-208). Somente uma observação (conteúdo N, em porcentagem, de grão de milho) foi feita para cada uma das 16 combinações de níveis.

SQA = 0,22625 SQB = 0,000025  
 SQC = 0,0036 SQAB = 0,04325  
 SQAC = 0,00065 SQBC = 0,000625  
 SQT = 0,2384.

- a. Elabore a tabela ANOVA.
- b. Suponha que não exista nenhum efeito de interação tripla, de modo que QMABC é uma estimativa válida de  $\sigma^2$ , e use o nível 0,05 para tentar encontrar interações e efeitos principais.
- c. As médias de nitrogênio são  $\bar{x}_{1..} = 1,1200$ ,  $\bar{x}_{2..} = 1,3025$ ,  $\bar{x}_{3..} = 1,3875$ , e  $\bar{x}_{4..} = 1,4300$ . Use o método de Tukey para examinar diferenças na porcentagem N entre os níveis de nitrogênio ( $Q_{0,05,4,3} = 6,82$ ).

31. O artigo "Kolbe-Schmitt Carbonation of 2-Naphthol" (*Industrial and Eng. Chemistry: Process and Design Development*, 1969, p. 165-173) apresentou os dados a

seguir sobre a porcentagem de produção de ácido BON como uma função do tempo de reação (1, 2 e 3 horas), da temperatura (30, 70 e 100°C) e da pressão (30, 70 e 100 psi). Supondo que não houvesse nenhuma interação de três fatores, de modo que SQE = SQABC desse uma estimativa de  $\sigma^2$ , o MINITAB forneceu a tabela ANOVA a seguir. Realize os testes apropriados.

		$B_1$		
		$C_1$	$C_2$	$C_3$
$A_1$	68,5	73,0	68,7	
$A_2$	74,5	75,0	74,6	
$A_3$	70,5	72,5	74,7	

		$B_2$		
		$C_1$	$C_2$	$C_3$
$A_1$	72,8	80,1	72,0	
$A_2$	72,0	81,5	76,0	
$A_3$	69,5	84,5	76,0	

		$B_3$		
		$C_1$	$C_2$	$C_3$
$A_1$	72,5	72,5	73,1	
$A_2$	75,5	70,0	76,0	
$A_3$	65,0	66,5	70,5	

Análise de Variância da Produção

Fonte	GL	SQ	QM	F	P
horário	2	42,112	21,056	8,76	0,010
temperatura	2	110,732	55,366	23,04	0,000
pressão	2	68,136	34,068	14,18	0,002
hor.*temp.	4	67,761	16,940	7,05	0,010
hor.*pressão	4	35,184	8,796	3,66	0,056
temp.*pressão	4	136,437	34,109	14,20	0,001
Erro	8	19,223	2,403		
Total	26	479,585			

32. Quando os fatores A e B são fixos, mas o fator C é aleatório e o modelo restrito é usado (consulte a nota de rodapé na página 402; aqui, há uma complicação técnica com o modelo irrestrito),

$$E(QME) = \sigma^2$$

$$E(QMA) = \sigma^2 + JL\sigma_{AC}^2 + \frac{JKL}{I-1} \sum \alpha_i^2$$

$$E(QMB) = \sigma^2 + IL\sigma_{BC}^2 + \frac{IKL}{J-1} \sum \beta_j^2$$

$$E(QMC) = \sigma^2 + IJL\sigma_C^2$$

$$E(QMAB) = \sigma^2 + L\sigma_{ABC}^2 + \frac{IKL}{(I-1)(J-1)} \sum_i \sum_j (\gamma_{ij}^{AB})^2$$

$$E(QMAC) = \sigma^2 + JL\sigma_{AC}^2$$

$$E(QMBC) = \sigma^2 + IL\sigma_{BC}^2$$

$$E(QMABC) = \sigma^2 + L\sigma_{ABC}^2$$

- a. Com base nesses quadrados médios esperados, que razões  $F$  você usaria para testar  $H_0: \sigma_{ABC}^2 = 0$ ;  $H_0: \sigma_C^2 = 0$ ;  $H_0: \gamma_{ij}^{AB} = 0$  para todos os  $i, j$ ; e  $H_0: \alpha_1 = \dots = \alpha_j = 0$ ?
- b. Em um experimento para avaliar os efeitos do período, do tipo de solo e dos dias de produção sobre a resistência à compressão de misturas de cimento/solo, dois períodos (A), quatro tipos de terra (B) e três dias (C, considerado aleatório) foram usados, com  $L = 2$  observações feitas para cada combinação de níveis de fator. As somas dos quadrados resultantes foram  $SQA = 14.318,24$ ,  $SQB = 9.656,40$ ,  $SQC = 2270,22$ ,  $SQAB = 3408,93$ ,  $SQAC = 1442,58$ ,  $SQBC = 3096,21$ ,  $SQABC = 2832,72$  e  $SQE = 8655,60$ . Obtenha a tabela ANOVA e realize os testes usando o nível 0,01.

33. Por causa de uma possível variabilidade no envelhecimento, decorrente das diferentes fundições e seus segmentos, foi usado um modelo em quadrado latino com  $N = 7$  para investigar o efeito do tratamento térmico sobre o envelhecimento. Sendo  $A =$  fundições,  $B =$  segmentos,  $C =$  tratamentos térmicos, as estatísticas resumidas incluem  $x_{...} = 3.815,8$ ,  $\sum x_{i..}^2 = 297.216,90$ ,  $\sum x_{.j.}^2 = 297.200,64$ ,  $\sum x_{..k}^2 = 297.155,01$ , e  $\sum \sum x_{ijk}^2 = 297.317,65$ . Obtenha uma tabela ANOVA e, ao nível 0,05, teste a hipótese de que o tratamento térmico não tem nenhum efeito sobre o envelhecimento.

34. O artigo “The Responsiveness of Food Sales to Shelf Space Requirements” (*J. Marketing Research*, 1964, p. 63-67) descreve o uso do modelo em quadrado latino para investigar o efeito do espaço na prateleira sobre a venda de alimentos. O experimento, realizado no período de seis semanas, usando seis diferentes lojas, resultou nos dados a seguir sobre a venda de creme de café em pó (o índice do espaço na prateleira está entre parênteses):

		Semana		
		1	2	3
Loja	1	27 (5)	14 (4)	18 (3)
	2	34 (6)	31 (5)	34 (4)
	3	39 (2)	67 (6)	31 (5)
	4	40 (3)	57 (1)	39 (2)
	5	15 (4)	15 (3)	11 (1)
	6	16 (1)	15 (2)	14 (6)

		Semana		
		4	5	6
Loja	1	35 (1)	28 (6)	22 (2)
	2	46 (3)	37 (2)	23 (1)
	3	49 (4)	38 (1)	48 (3)
	4	70 (6)	37 (4)	50 (5)
	5	9 (2)	18 (5)	17 (6)
	6	12 (5)	19 (3)	22 (4)

Elabore a tabela ANOVA e expresse e teste, ao nível 0,01, a hipótese de que o espaço na prateleira não influencia nas vendas, comparativamente à alternativa apropriada.

35. O artigo “Variation in Moisture and Ascorbic Acid Content from Leaf to Leaf and Plant to Plant in Turnip Greens” (*Southern Cooperative Services Bull.*, 1951, p. 13-17) emprega um modelo em quadrado latino em que o fator A é uma planta, o fator B é o tamanho da folha (da menor à maior), o fator C (entre parênteses) é o tempo de pesagem e a variável de resposta é a quantidade de umidade.

		Tamanho da Folha (B)		
		1	2	3
Planta (A)	1	6,67 (5)	7,15 (4)	8,29 (1)
	2	5,40 (2)	4,77 (5)	5,40 (4)
	3	7,32 (3)	8,53 (2)	8,50 (5)
	4	4,92 (1)	5,00 (3)	7,29 (2)
	5	4,88 (4)	6,16 (1)	7,83 (3)

		Tamanho da Folha (B)	
		4	5
Planta (A)	1	8,95 (3)	9,62 (2)
	2	7,54 (1)	6,93 (3)
	3	9,99 (4)	9,68 (1)
	4	7,85 (5)	7,08 (4)
	5	5,83 (2)	8,51 (5)

Quando os três fatores são aleatórios, os quadrados médios esperados são  $E(QMA) = \sigma^2 + N\sigma_A^2$ ,  $E(QMB) = \sigma^2 + N\sigma_B^2$ ,  $E(QMC) = \sigma^2 + N\sigma_C^2$  e  $E(QME) = \sigma^2$ . Isso significa que as razões  $F$  para testar  $H_{0A}: \sigma_A^2 = 0$ ,  $H_{0B}: \sigma_B^2 = 0$  e  $H_{0C}: \sigma_C^2 = 0$  são idênticas às razões para efeitos fixos. Obtenha a tabela ANOVA e teste para verificar se, ao nível 0,05, há alguma variação na quantidade de umidade decorrente dos fatores.

36. O artigo “An Assessment of the Effects of Treatment, Time, and Heat on the Removal of Erasable Pen Marks from Cotton and Cotton/Polyester Blend Fabrics” (*J. Testing and Eval.*, 1991, p. 394-397) descreve as somas de quadrados a seguir para a variável de resposta grau de remoção de manchas:  $SQA = 39,171$ ,  $SQB = 0,665$ ,  $SQC = 21,508$ ,  $SQAB = 1,432$ ,  $SQAC = 15,953$ ,  $SQBC = 1,382$ ,  $SQABC = 9,016$  e  $SQE = 115,820$ . Quatro diferentes tratamentos de lavagem, três diferentes tipos de caneta e seis tecidos distintos foram usados no experimento e havia três observações para cada combinação de tratamento-caneta-tecido. Faça uma análise de variância usando  $\alpha = 0,01$  para cada teste e expresse suas conclusões (admita efeitos fixos para os três fatores).

37. Um experimento ANOVA de quatro fatores foi realizado para investigar os efeitos do tecido (A), do tipo de exposição (B), do nível de exposição (C) e da direção das

fibras ( $D$ ) na intensidade de mudança da cor em tecidos expostos, medidos com um espectrocolorímetro. Duas observações foram feitas para cada um dos três tecidos, dos dois tipos de exposição, dos três níveis de exposição e dos dois níveis de direção, gerando  $QMA = 2207,329$ ,  $QMB = 47,255$ ,  $QMC = 491,783$ ,  $QMD = 0,044$ ,  $QMAB = 15,303$ ,  $QMAC = 275,446$ ,  $QMAD = 0,470$ ,

$QMBC = 2,141$ ,  $QMBD = 0,273$ ,  $QMCD = 0,247$ ,  $QMABC = 3,714$ ,  $QMABD = 4,072$ ,  $QMACD = 0,767$ ,  $QMBCD = 0,280$ ,  $QME = 0,977$  e  $QMT = 93,621$  ("Accelerated Weathering of Marine Fabrics", *J. Testing and Eval.*, 1992, p. 139-143). Supondo efeitos fixos para todos os fatores, realize uma análise de variância usando  $\alpha = 0,01$  para todos os testes e resuma suas conclusões.

## 11.4 | Experimentos Fatoriais $2^p$

Se um investigador desejar estudar simultaneamente o efeito dos diferentes fatores de  $p$  sobre a variável de resposta e os fatores tiverem  $I_1, I_2, \dots, I_p$  níveis, respectivamente, o experimento exigirá pelo menos  $I_1; I_2; \dots; I_p$  observações. Nessas circunstâncias, o investigador em geral pode realizar um "experimento peneira" com cada fator em apenas dois níveis para obter informações preliminares sobre os efeitos dos fatores. O experimento em que há  $p$  fatores, cada um em dois níveis, é chamado de **experimento fatorial  $2^p$** . No que se refere aos cálculos, a análise dos dados de um experimento como esse é mais simples do que a de experimentos fatoriais mais gerais. Além disso, o experimento  $2^p$  oferece uma forma simples para introduzir os importantes conceitos de confusão (*confounding*) e replicação fracionária.

### Experimentos $2^3$

Como na Seção 11.3,  $X_{ijkl}$  e  $x_{ijkl}$  referem-se à observação da  $l$ -ésima replicação com os fatores  $A, B$  e  $C$  nos níveis  $i, j$  e  $k$ , respectivamente. O modelo para essa situação é

$$X_{ijkl} = \mu + \alpha_i + \beta_j + \delta_k + \gamma_{ij}^{AB} + \gamma_{ik}^{AC} + \gamma_{jk}^{BC} + \gamma_{ijk} + \epsilon_{ijkl} \quad (11.14)$$

para  $i = 1, 2; j = 1, 2; k = 1, 2; l = 1, \dots, n$ . Os  $\epsilon_{ijkl}$ s são considerados independentes e normalmente distribuídos, com média 0 e variância  $\sigma^2$ . Por haver apenas dois níveis de cada fator, as condições de contorno nos parâmetros de (11.14) que especificam unicamente o modelo são expressas de maneira lógica:  $\alpha_1 + \alpha_2 = 0, \dots, \gamma_{11}^{AB} + \gamma_{21}^{AB} = 0, \gamma_{12}^{AB} + \gamma_{22}^{AB} = 0, \gamma_{11}^{AC} + \gamma_{21}^{AC} = 0, \gamma_{12}^{AC} + \gamma_{22}^{AC} = 0$ , e assim por diante. Essas condições implicam a existência de apenas um parâmetro funcionalmente independente de cada tipo (para cada efeito principal e interação). Por exemplo:  $\alpha_2 = -\alpha_1$ , ao passo que  $\gamma_{21}^{AB} = -\gamma_{11}^{AB}$ ,  $\gamma_{12}^{AB} = -\gamma_{22}^{AB}$  e  $\gamma_{22}^{AB} = \gamma_{11}^{AB}$ . Por isso, cada soma de quadrados na análise terá 1 gl.

Os parâmetros do modelo podem ser estimados calculando-se a média de vários subscritos dos  $X_{ijkl}$ s e, em seguida, formando combinações lineares apropriadas das médias. Por exemplo:

$$\hat{\alpha}_1 = \bar{X}_{1\dots} - \bar{X}_{\dots} = \frac{(X_{111\cdot} + X_{121\cdot} + X_{112\cdot} + X_{122\cdot} - X_{211\cdot} - X_{212\cdot} - X_{221\cdot} - X_{222\cdot})}{8n}$$

e

$$\hat{\gamma}_{11}^{AB} = \bar{X}_{11\dots} - \bar{X}_{1\dots} - \bar{X}_{\cdot 1\dots} + \bar{X}_{\dots} = \frac{(X_{111\cdot} - X_{121\cdot} - X_{211\cdot} + X_{221\cdot} + X_{112\cdot} - X_{122\cdot} - X_{212\cdot} + X_{222\cdot})}{8n}$$

Cada estimador é, exceto para o fator  $1/(8n)$ , uma função linear dos totais de célula ( $X_{ijk\cdot}$ s) em que cada coeficiente é +1 ou -1, com um número igual de cada; essas funções são chamadas de **contrastes** nos  $X_{ijk\cdot}$ s. Além disso, os estimadores satisfazem as mesmas condições de contorno em função dos próprios parâmetros. Por exemplo:

$$\begin{aligned} \hat{\alpha}_1 + \hat{\alpha}_2 &= \bar{X}_{1\dots} - \bar{X}_{\dots} + \bar{X}_{2\dots} - \bar{X}_{\dots} = \bar{X}_{1\dots} + \bar{X}_{2\dots} - 2\bar{X}_{\dots} \\ &= \frac{1}{4n}X_{1\dots} + \frac{1}{4n}X_{2\dots} - \frac{2}{8n}X_{\dots} = \frac{1}{4n}X_{\dots} - \frac{1}{4n}X_{\dots} = 0 \end{aligned}$$

**Exemplo 11.13**

Em um experimento para investigar as propriedades da resistência à compressão de misturas de cimento/solo, dois diferentes períodos de envelhecimento foram usados em conjunto com duas diferentes temperaturas de envelhecimento e dois solos diferentes. Duas replicações foram feitas para cada combinação de níveis dos três fatores, resultando nos dados a seguir:

Período	Temperatura	Solo	
		1	2
1	1	471.413	385.434
	2	485.552	530.593
2	1	712.637	770.705
	2	712.789	741.806

Os totais de célula calculados são  $x_{111} = 884$ ,  $x_{211} = 1349$ ,  $x_{121} = 1037$ ,  $x_{221} = 1501$ ,  $x_{112} = 819$ ,  $x_{212} = 1475$ ,  $x_{122} = 1123$ , e  $x_{222} = 1547$ , de modo que  $x_{\dots} = 9735$ . Então

$$\begin{aligned} \hat{\alpha}_1 &= (884 - 1349 + 1037 - 1501 + 819 - 1475 + 1123 - 1547)/16 \\ &= -125,5625 = -\hat{\alpha}_2 \\ \hat{\gamma}_{11}^{AB} &= (884 - 1349 - 1037 + 1501 + 819 - 1475 - 1123 + 1547)/16 \\ &= -14,5625 = -\hat{\gamma}_{12}^{AB} = -\hat{\gamma}_{21}^{AB} = \hat{\gamma}_{22}^{AB} \end{aligned}$$

As estimativas dos outros parâmetros podem ser calculadas da mesma maneira. ■

**Análise de um Experimento 2<sup>3</sup>** O motivo que nos leva a calcular estimativas dos parâmetros é que as somas dos quadrados dos vários efeitos são facilmente obtidas por meio dessas estimativas. Por exemplo:

$$SQA = \sum_i \sum_j \sum_k \sum_l \hat{\alpha}_i^2 = 4n \sum_{i=1}^2 \hat{\alpha}_i^2 = 4n[\hat{\alpha}_1^2 + (-\hat{\alpha}_1)^2] = 8n\hat{\alpha}_1^2$$

e

$$\begin{aligned} SQA_B &= \sum_i \sum_j \sum_k \sum_l (\hat{\gamma}_{ij}^{AB})^2 \\ &= 2n \sum_{i=1}^2 \sum_{j=1}^2 (\hat{\gamma}_{ij}^{AB})^2 = 2n[(\hat{\gamma}_{11}^{AB})^2 + (-\hat{\gamma}_{11}^{AB})^2 + (-\hat{\gamma}_{11}^{AB})^2 + (\hat{\gamma}_{11}^{AB})^2] \\ &= 8n(\hat{\gamma}_{11}^{AB})^2 \end{aligned}$$

Visto que cada estimativa é um contraste nos totais de célula multiplicados por  $1/(8n)$ , cada soma de quadrados tem o formato  $(\text{contraste})^2/(8n)$ . Logo, para calcular as várias somas de quadrados, precisamos conhecer os coeficientes (+1 ou -1) dos contrastes apropriados. Os sinais (+ ou -) em cada  $x_{ijk}$ , em cada contraste do efeito, são exibidos de maneira mais conveniente em uma tabela. Usaremos a notação (1) para a condição experimental  $i = 1, j = 1, k = 1$ ,  $ab$  para  $i = 2, j = 2, k = 1$ , e assim por diante. Se o nível 1 for considerado "baixo" e o nível 2 "alto", qualquer letra que aparecer indicará um nível alto do fator correspondente. Na Tabela 11.10, cada coluna fornece os sinais para um determinado contraste do efeito nos  $x_{ijk}$ s, correspondente às diferentes condições experimentais.

**Tabela 11.10** Sinais para calcular contrastes do efeito

Condição Experimental	Total da Célula	Efeito Fatorial						
		A	B	C	AB	AC	BC	ABC
(1)	$x_{111}$	-	-	-	+	+	+	-
a	$x_{211}$	+	-	-	-	-	+	+
b	$x_{121}$	-	+	-	-	+	-	+
ab	$x_{221}$	+	+	-	+	-	-	-
c	$x_{112}$	-	-	+	+	-	-	+
ac	$x_{212}$	+	-	+	-	+	-	-
bc	$x_{122}$	-	+	+	-	-	+	-
abc	$x_{222}$	+	+	+	+	+	+	+

Em cada uma das primeiras três colunas, o sinal será +, se o fator correspondente estiver no nível alto, e -, se estiver no nível baixo. Todo sinal na coluna AB é, portanto, "produto" dos sinais nas colunas A e B, sendo (+)(+) = (-)(-) = + e (+)(-) = (-)(+) = -, e algo semelhante pode-se afirmar para as colunas AC e BC. Finalmente, os sinais na coluna ABC são os produtos de AB com C (ou B com AC ou A com BC). Portanto, por exemplo:

$$\text{Contraste AC} = +x_{111} - x_{211} + x_{121} - x_{221} - x_{112} + x_{212} - x_{122} + x_{222}.$$

Uma vez que os sete contrastes dos efeitos são calculados,

$$\text{SQ(efeito)} = \frac{(\text{contraste do efeito})^2}{8n}$$

Mesmo com uma tabela de sinais, o cálculo de contrastes é tedioso. Uma técnica computacional eficaz, atribuída a Yates, é a seguinte: escreva em uma coluna os oito totais de célula na **ordem-padrão**, como apresentado na tabela de sinais, e forme três colunas adicionais. Em cada uma dessas três colunas, as quatro primeiras entradas são as somas das entradas 1 e 2, 3 e 4, 5 e 6, 7 e 8 das colunas anteriores. As quatro últimas entradas são as diferenças entre as entradas 2 e 1, 4 e 3, 6 e 5 e 8 e 7 das colunas anteriores. A última coluna contém  $x_{\dots}$  e os sete contrastes dos efeitos na ordem-padrão. Ao elevarmos cada contraste ao quadrado e dividirmos por  $8n$ , obteremos as sete somas de quadrados.

**Exemplo 11.14** (Continuação do Exemplo 11.13)

Visto que  $n = 2$ ,  $8n = 16$ . O método de Yates é demonstrado na Tabela 11.11.

**Tabela 11.11** Método de cálculo de Yates

Condição do tratamento	$x_{ijk}$	1	2	Contraste dos Efeitos	SS = (contraste) <sup>2</sup> /16
(1) = $x_{111}$	884	2233	4771	9735	
a = $x_{211}$	1349	2538	4964	2009	252.255,06
b = $x_{121}$	1037	2294	929	681	28.985,06
ab = $x_{221}$	1501	2670	1080	-233	3.393,06
c = $x_{112}$	819	465	305	193	2.328,06
ac = $x_{212}$	1475	464	376	151	1.425,06
bc = $x_{122}$	1123	656	-1	71	315,06
abc = $x_{222}$	1547	424	-232	-231	3.335,06
					292.036,42

Com base nos dados originais,  $\sum_i \sum_j \sum_k \sum_l x_{ijkl}^2 = 6.232.289$ , e

$$\frac{x_{\dots}^2}{16} = 5.923.139,06$$



de modo que

$$SQT = 6.232.289 - 5.923.139,06 = 309.149,94$$

$$SQE = SQT - [SQA + \dots + SQABC] = 309.149,94 - 292.036,42 \\ = 17.113,52$$

Os cálculos ANOVA estão resumidos na Tabela 11.12.

**Tabela 11.12** Tabela ANOVA do Exemplo 11.14

Fonte de Variação	gl	Soma dos Quadrados	Quadrado Médio	<i>f</i>
A	1	252.255,06	252.255,06	117,92
B	1	28.985,06	28.985,06	13,55
C	1	2.328,06	2.328,06	1,09
AB	1	3.393,06	3.393,06	1,59
AC	1	1.425,06	1.425,06	0,67
BC	1	315,06	315,06	0,15
ABC	1	3.335,06	3.335,06	1,56
Erro	8	17.113,52	2.139,19	
Total	15	309.149,94		

A Figura 11.7 mostra o resultado SQA para esse exemplo. Somente os valores *P* da idade (A) e da temperatura (B) são menores que 0,01, de modo que apenas esses efeitos são considerados significativos.

Procedimento de Análise de Variância					
Variável Dependente: RESISTÊNCIA					
Fonte	GL	Soma dos Quadrados	Quadrado Médio	Valor F	Pr > F
Modelo	7	292036,4375	41719,4911	19,50	0,0002
Erro	8	17113,5000	2139,1875		
Total Corrigido	15	309149,9375			
R Quadrado		C.V.	QME-RAIZ	Média	USO ENERGIA
0,944643		7,601660	46,25135	608,437500	
Fonte	GL	SQ Anova	Quadrado Médio	Valor F	Pr > F
IDADE	1	252255,0625	252255,0625	117,92	0,0001
TEMPERATURA	1	28985,0625	28985,0625	13,55	0,0062
IDADE* TEMP.	1	3393,0625	3393,0625	1,59	0,2434
SOLO	1	2328,0625	2328,0625	1,09	0,3273
IDADE* SOLO	1	1425,0625	1425,0625	0,67	0,4380
TEMP.* SOLO	1	315,0625	315,0625	0,15	0,7111
IDADE* TEMP.* SOLO	1	3335,0625	3335,0625	1,56	0,2471

**Figura 11.7** Resultado do SQA dos dados de resistência do Exemplo 11.14 ■

## Experimentos 2<sup>p</sup> para *p* > 3

Embora os cálculos sejam muito tediosos, quando  $p > 3$ , a análise é análoga à do caso de três fatores. Por exemplo: se houver quatro fatores *A*, *B*, *C* e *D*, haverá 16 diferentes condições experimentais. As primeiras oito na ordem-padrão são exatamente as já relacionadas para um experimento de três fatores. As oito restantes são obtidas colocando-se a letra *d* ao lado de cada condição no primeiro grupo. O método de Yates é portanto iniciado calculando-se os totais por meio das replicações, relacionando-os na ordem-padrão e procedendo como antes; com *p* fatores, a *p*-ésima coluna à direita dos totais do tratamento fornecerá os contrastes dos efeitos.

Para  $p > 3$ , em geral não haverá nenhuma replicação do experimento (portanto, somente uma única replicação completa é obtível). Uma possibilidade de testar as hipóteses é supor que não existem

determinados efeitos de ordem superior e, em seguida, adicionar as somas dos quadrados correspondentes para obter uma SQE. Entretanto essa suposição pode ser enganosa, quando não se tem um conhecimento prévio (consulte o livro de Montgomery relacionado na bibliografia deste capítulo). Um método alternativo é trabalhar diretamente com os contrastes dos efeitos. Cada contraste tem uma distribuição normal, com a mesma variância. Quando um determinado efeito não existe, o valor esperado do contraste correspondente é 0, mas isso não é verdade quando o efeito existe. O método de análise sugerido é elaborar um gráfico de probabilidade normal dos contrastes dos efeitos (ou, de maneira equivalente, as estimativas dos parâmetros dos efeitos, visto que a estimativa = contraste/ $2^p$  quando  $n = 1$ ). Os pontos correspondentes a efeitos ausentes provavelmente estejam próximos de uma linha reta, ao passo que os pontos associados com efeitos importantes normalmente ficarão longe dessa linha.

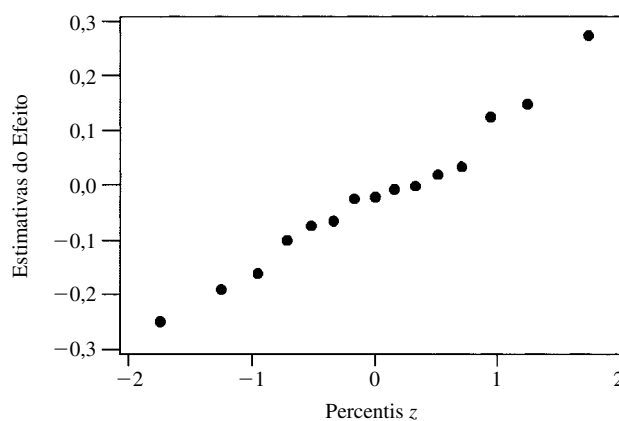
### Exemplo 11.15

Os dados a seguir provêm do artigo “Quick and Easy Analysis of Unreplicated Factorials” (*Technometrics*, 1989, p. 469-473). Os quatro fatores são  $A$  = resistência ao ácido,  $B$  = tempo,  $C$  = quantidade de ácido e  $D$  = temperatura, e a variável de resposta é a produção de isatina. As observações, na ordem-padrão, são 0,08, 0,04, 0,53, 0,43, 0,31, 0,09, 0,12, 0,36, 0,79, 0,68, 0,73, 0,08, 0,77, 0,38, 0,49 e 0,23. A Tabela 11.13 apresenta as estimativas dos efeitos exatamente como consta do artigo (que usa contraste/8 em vez de contraste/16).

**Tabela 11.13** Estimativas dos efeitos do Exemplo 11.15

Efeito	$A$	$B$	$AB$	$C$	$AC$	$BC$	$ABC$	$D$
estimativa	-0,191	-0,021	-0,001	-0,076	0,034	-0,066	0,149	0,274
Efeito	$AD$	$BD$	$ABD$	$CD$	$ACD$	$BCD$	$ABCD$	
estimativa	-0,161	-0,251	-0,101	-0,026	-0,066	0,124	0,019	

A Figura 11.8 é um gráfico de probabilidade normal das estimativas dos efeitos. Todos os pontos no gráfico estão próximos da mesma linha reta, sugerindo que não existe absolutamente nenhum efeito (em breve daremos um exemplo em que isso não ocorre).



**Figura 11.8** Gráfico de probabilidade normal das estimativas dos efeitos do Exemplo 11.15 ■

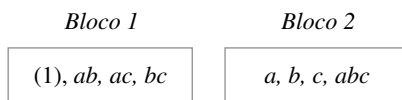
As avaliações visuais sobre o desvio da reta em um gráfico de probabilidade normal são muito subjetivas. O artigo citado no Exemplo 11.15 descreve uma técnica mais objetiva para identificar efeitos significativos em um experimento não-replicável.

## Confusão

Em geral, não é possível conduzir todas as condições experimentais  $2^p$  de um experimento fatorial  $2^p$  em um ambiente experimental homogêneo. Nessas circunstâncias, talvez seja possível separar as condições experimentais em  $2^r$  blocos homogêneos ( $r < p$ ), de modo que haja  $2^{p-r}$  condições experimentais em cada bloco. Os blocos podem, por exemplo, corresponder a diferentes laboratórios, a diferentes períodos de tempo ou a diferentes operadores ou grupos de trabalho. No caso mais simples,  $p = 3$  e  $r = 1$ , de maneira que haja dois blocos, em que cada um é formado por quatro das oito condições experimentais.

Como sempre, a disposição em bloco é eficaz para reduzir variações associadas com fontes externas. Contudo, quando as condições experimentais  $2^p$  são dispostas em dois blocos, o preço pago por essa disposição é que  $2^r - 1$  dos efeitos do fator não podem ser estimados. Isso ocorre porque  $(R^r - 1)$  efeitos do fator (efeitos principais e/ou interações) são misturados ou **confundidos** com os efeitos do bloco. Portanto, a designação de condições experimentais a blocos normalmente é feita de tal maneira que apenas as interações de nível superior são confundíveis, enquanto os efeitos principais e as interações de ordem inferior permanecem estimáveis e as hipóteses podem ser testadas.

Para verificar como é realizada a designação de blocos, considere primeiro um experimento  $2^3$  com dois blocos ( $r = 1$ ) e quatro tratamentos por bloco. Suponha que selecionemos  $ABC$  como o efeito de confusão com os blocos. Portanto, qualquer condição experimental que tenha um número ímpar de letras em comum com  $ABC$ , como  $b$  (uma única letra) ou  $abc$  (três letras), é colocada em um bloco, ao passo que qualquer condição que tenha um número par de letras em comum com  $ABC$  (onde 0 é par) é disposta no outro bloco. A Figura 11.9 mostra essa designação de tratamentos aos dois blocos.



**Figura 11.9** Confusão de  $ABC$  em um experimento  $2^3$

Na ausência de replicações, os dados de um experimento como esse, em geral, seriam analisados supondo-se que não havia nenhuma interação de dois fatores (aditividade) e usando-se  $SQE = SQAB + SSQC + SQBC$  com 3 gl para tentar encontrar efeitos principais. Alternativamente, um gráfico de probabilidade normal de contrastes do efeito ou de estimativas de parâmetro do efeito pode ser examinado. Contudo, é mais freqüente haver replicações quando estão sendo estudados apenas três fatores. Suponha que haja  $u$  replicações, que resultam em um total de  $2^r \cdot u$  blocos no experimento. Portanto, após subtrair de  $SQT$  todas as somas de quadrados associadas com os efeitos não confundidos com blocos (usa-se o método de Yates para efetuar os cálculos), a soma dos quadrados do bloco é calculada usando-se os totais  $2^r \cdot u$  do bloco e, em seguida, subtraída para obter  $SQE$  (de modo que haja  $2^r \cdot u - 1$  gl para os blocos).

### Exemplo 11.16

O artigo “Factorial Experiments in Pilot Plant Studies” (*Industrial and Eng. Chemistry*, 1951, p. 1300-1306) apresenta os resultados de um experimento realizado para avaliar os efeitos da temperatura do reator ( $A$ ), a emissão de gás ( $B$ ) e a concentração do componente ativo ( $C$ ) sobre a resistência da solução do produto (medidas baseadas em unidades arbitrárias) em uma unidade de recirculação. Foram usados dois blocos, confundindo-se o efeito  $ABC$  com os blocos, e ocorreram duas replicações, gerando os dados da Figura 11.10. Os totais de replicação para os quatro blocos  $\times$  são 288, 212, 88 e 220, com um total geral de 808, de modo que

$$SQBI = \frac{(288)^2 + (212)^2 + (88)^2 + (220)^2}{4} - \frac{(808)^2}{16} = 5204,00$$

As outras somas de quadrados são calculadas por meio do método de Yates, usando-se os oito totais da condição experimental, resultando na tabela ANOVA representada pela Tabela 11.14. Ao comparar com  $F_{0,05,1,6} = 5,99$ , concluímos que somente os efeitos principais de  $A$  e  $C$  diferem significativamente de zero.

Replicação 1				Replicação 2			
Bloco 1		Bloco 2		Bloco 1		Bloco 2	
(1)	99	a	18	(1)	46	a	18
ab	52	b	51	ab	-47	b	62
ac	42	c	108	ac	22	c	104
bc	95	abc	35	bc	67	abc	36

Figura 11.10 Dados do Exemplo 11.16

Tabela 11.14 Tabela ANOVA do Exemplo 11.16

Fonte de Variação	gl	Soma dos Quadrados	Quadrado Médio	f
A	1	12.996	12.996	39,82
B	1	702,25	702,25	2,15
C	1	2.756,25	2.756,25	8,45
AB	1	210,25	210,25	0,64
AC	1	30,25	30,25	0,093
BC	1	25	25	0,077
Blocos	3	5.204	1.734,67	5,32
Erro	6	1.958	326,33	
Total	15	23.882		

### Confusão usando mais de Dois Blocos

No caso de  $r = 2$  (quatro blocos), três efeitos são confundidos com blocos. O investigador primeiro escolhe dois efeitos definidos para que sejam confundidos. Por exemplo: em um experimento de cinco fatores ( $A, B, C, D$  e  $E$ ), as duas interações de três fatores  $BCD$  e  $CDE$  deviam ser escolhidas para confusão. O terceiro efeito de confusão é, portanto, a **interação generalizada** de ambas, obtida escrevendo-se lado a lado os dois efeitos escolhidos e, em seguida, eliminando-se qualquer letra em comum a ambos:  $(BCD)(CDE) = BE$ . Observe que, se  $ABC$  e  $CDE$  forem escolhidos para confusão, a respectiva interação generalizada será  $(ABC)(CDE) = ABDE$ , de modo que nenhum efeito principal ou interação de dois fatores seja confundido.

Depois que os dois efeitos definidos forem selecionados para confusão, o primeiro bloco conterá todas as condições de tratamento com um número par de letras em comum com ambos os efeitos definidos. O segundo bloco consistirá de todas as condições com um número par de letras em comum com o primeiro contraste definido e um número ímpar de letras em comum com o segundo contraste, e o terceiro e quarto blocos consistirão nos contrastes “ímpares/pares” e “ímpares/ímpares”. Em um experimento de cinco fatores com os efeitos definidos  $ABC$  e  $CDE$ , isso resulta na designação de blocos mostrada na Figura 11.11 (em que o número de letras em comum com cada contraste definido aparece ao lado de cada condição experimental).

Bloco 1		Bloco 2		Bloco 3		Bloco 4	
(1)	(0, 0)	d	(0, 1)	a	(1, 0)	c	(1, 1)
ab	(2, 0)	e	(0, 1)	b	(1, 0)	ad	(1, 1)
de	(0, 2)	ac	(2, 1)	cd	(1, 2)	ae	(1, 1)
acd	(2, 2)	bc	(2, 1)	ce	(1, 2)	bd	(1, 1)
ace	(2, 2)	abd	(2, 1)	ade	(1, 2)	be	(1, 1)
bcd	(2, 2)	abe	(2, 1)	bde	(1, 2)	abc	(3, 1)
bce	(2, 2)	acde	(2, 3)	abcd	(3, 2)	cde	(1, 3)
abde	(2, 2)	bcde	(2, 3)	abce	(3, 2)	abcde	(3, 3)

Figura 11.11 Quatro blocos em um experimento fatorial  $2^5$  com os efeitos definidos  $ABC$  e  $CDE$

O bloco contendo (1) é chamado de **bloco principal**. Assim que ele é formado, um segundo bloco pode ser conseguido, selecionando-se qualquer condição experimental que não esteja no bloco principal e obtendo-se a respectiva interação generalizada com qualquer condição no bloco principal. Os outros blocos são, portanto, formados da mesma maneira, primeiramente selecionando uma condição não-presente em um bloco já formado e em seguida identificando-se interações generalizadas com o bloco principal.

Em situações experimentais em que  $p > 3$ , em geral não há nenhuma replicação, de modo que as somas dos quadrados associados com interações de ordem superior em que não há confusão normalmente são combinadas para obter uma soma dos quadrados do erro que possa ser usada nos denominadores das várias estatísticas  $F$ . Todos os cálculos podem ser realizados uma vez mais por meio da técnica de Yates, sendo  $SQB^1$  a soma de quadrados associada com efeitos de confusão.

Quando  $r > 2$ , primeiro selecionam-se  $r$  efeitos determinantes para serem confundidos com os blocos, tomando-se o cuidado para que nenhum dos efeitos escolhidos seja a interação generalizada de quaisquer outros dois efeitos selecionados. Os  $2^r - r - 1$  efeitos adicionais confundidos com os blocos são portanto as interações generalizadas de todos os efeitos no conjunto definido (incluindo não apenas as interações generalizadas dos pares de efeitos, como também as dos conjuntos de três, quatro, e assim por diante). Consulte o livro de Montgomery para obter detalhes.

## Replicação fracionária

Quando o número  $p$  de fatores é grande, uma única replicação de um experimento  $2^p$  já é suficientemente dispendiosa e trabalhosa. Por exemplo: uma replicação de um experimento fatorial  $2^6$  contém uma observação para cada uma das 64 diferentes condições experimentais. Uma estratégia atraente em situações desse tipo é fazer observações para apenas uma fração das condições  $2^p$ . Desde que se tome cuidado na escolha das condições a serem observadas, ainda é possível obter mais informações sobre os efeitos do fator.

Suponha que decidamos incluir apenas  $2^{p-1}$  (metade) das condições  $2^p$  possíveis em nosso experimento; isso é comumente conhecido por **replicar pela metade**. O preço pago por essa economia é dobrado. Em primeiro lugar, as informações sobre um único efeito (determinadas pelas  $2^{p-1}$  condições selecionadas para a observação) de nada servem para o investigador no sentido de que não é possível obter nenhuma estimativa razoável do efeito. Segundo, os  $2^p - 2$  efeitos principais e interações remanescentes são pareados de modo que nenhum efeito em um determinado par seja confundido com o outro efeito no mesmo par. Por exemplo: um desses pares pode ser  $\{A, BCD\}$ , de maneira que estimativas distintas do efeito principal de  $A$  e da interação  $BCD$  não são possíveis. Portanto, vale a pena selecionar uma replicação pela metade para a qual os efeitos principais e as interações de ordem inferior são pareadas (confundidas) somente com interações de ordem superior e não entre si.

A primeira etapa para selecionar uma replicação pela metade é escolher um efeito determinante como o efeito não-estimável. Suponha que, em um experimento de cinco fatores,  $ABCDE$  seja escolhido como o efeito determinante. Agora, as  $2^5 = 32$  possíveis condições de tratamento são divididas em dois grupos com 16 condições em cada, um grupo formado por todas as condições com um número ímpar de letras em comum com  $ABCDE$ , outro contendo um número par de letras em comum com o contraste definido. Logo, ambos os grupos de 16 condições são usados como a replicação pela metade. O grupo “ímpar” é

$$a, b, c, d, e, abc, abd, abe, acd, ace, ade, bcd, bce, bde, cde, abcde$$

Todo efeito principal e toda interação diferente de  $ABCDE$  são, portanto, confundidos com (**considerados análogos a**) sua interação generalizada com  $ABCDE$ . Logo,  $(AB)(ABCDE) = CDE$ , de forma que a interação  $AB$  e a interação  $CDE$  são confundidas uma com a outra. Os **pares análogos** resultantes são

$$\begin{array}{ccccc} \{A, BCDE\} & \{B, ACDE\} & \{C, ABDE\} & \{D, ABCE\} & \{E, ABCD\} \\ \{AB, CDE\} & \{AC, BDE\} & \{AD, BCE\} & \{AE, BCD\} & \{BC, ADE\} \\ \{BD, ACE\} & \{BE, ACD\} & \{CD, ABE\} & \{CE, ABD\} & \{DE, ABC\} \end{array}$$

Observe especialmente que todo efeito principal é considerado análogo a uma interação de quatro fatores. Admitir que essas interações são insignificantes nos possibilita verificar se existem efeitos principais.

Para selecionar uma replicação pelo quarto de um experimento fatorial  $2^p$  ( $2^{p-2}$  das condições de tratamento  $2^p$  possíveis), dois efeitos determinantes devem ser selecionados. Esses dois efeitos e a respectiva interação generalizada passam a ser os efeitos não-estimáveis. Em vez de usar pares análogos como na replicação pela metade, todo efeito remanescente é então confundido com três outros efeitos, sendo cada um a respectiva interação generalizada com um dos três efeitos não-estimáveis.

### Exemplo 11.17

O artigo “More on Planning Experiments to Increase Research Efficiency” (*Industrial and Eng. Chemistry*, 1970, p. 60-65) descreve os resultados de uma replicação pelo quarto de um experimento  $2^p$  em que os cinco fatores empregados foram  $A$  = temperatura de condensação,  $B$  = quantidade de material  $B$ ,  $C$  = volume de solvente,  $D$  = tempo de condensação e  $E$  = quantidade de material  $E$ . A variável de resposta foi o resultado do processo químico. Os contrastes determinantes escolhidos foram  $ACE$  e  $BDE$ , sendo a interação generalizada  $(ACE)(BDE) = ABCD$ . Os 28 efeitos principais e interações remanescentes podem agora ser divididos em sete grupos de quatro efeitos cada, de modo que os efeitos dentro de um grupo não possam ser avaliados separadamente. Por exemplo: as interações generalizadas de  $A$  com efeitos não-estimáveis são  $(A)(ACE) = CE$ ,  $(A)(BDE) = ABDE$  e  $(A)(ABCD) = BCD$ , de maneira que um grupo análogo é  $\{A, CE, ABDE, BCD\}$ . O conjunto completo dos grupos análogos é:

$$\begin{array}{lll} \{A, CE, ABDE, BCD\} & \{B, ABCE, DE, ACD\} & \{C, AE, BCDE, ABD\} \\ \{D, ACDE, BE, ABC\} & \{E, AC, BD, ABCDE\} & \{AB, BCE, ADE, CD\} \\ \{AD, CDE, ABE, BC\} & & \blacksquare \end{array}$$

**Análise de uma Replicação Fracionária** Depois que os contrastes definidos forem selecionados para uma replicação pelo quarto, eles são usados como na discussão sobre confusão para dividir as  $2^p$  condições de tratamento em quatro grupos de  $2^{p-2}$  condições. Então, qualquer um dos quatro grupos é selecionado como o conjunto de condições para o qual serão coletados dados. Comentários semelhantes aplicam-se a replicações  $1/2^r$  de um experimento fatorial  $2^p$ .

Feitas as observações para as combinações de tratamento selecionadas, será elaborada uma tabela de sinais semelhante à Tabela 11.10. Essa tabela contém uma linha exclusiva para cada uma das combinações de tratamento, realmente observadas, e não todas as linhas  $2^p$ , e uma coluna exclusiva para cada grupo análogo (visto que cada efeito no grupo teria o mesmo conjunto de sinais para as condições de tratamento selecionadas para serem observadas). Os sinais em cada coluna indicam, como de costume, de que modo os contrastes das várias somas de quadrados são calculados. O método de Yates pode ainda ser usado, mas a regra para organizar as condições observadas na ordem-padrão deve ser modificada.

A parte difícil de uma análise de replicação fracionária normalmente está relacionada com a decisão sobre o que usar para a soma dos quadrados do erro. Visto que em geral não haverá nenhuma replicação (embora seja possível observar, por exemplo, duas replicações de uma replicação pelo quarto), algumas somas de quadrados do efeito devem ser combinadas para obter uma soma de quadrados do erro. Em uma replicação pela metade, em um experimento  $2^8$ , por exemplo, uma estrutura análoga pode ser escolhida de modo que os oito efeitos principais e as 28 interações de dois fatores sejam, cada um, confundidos apenas com interações de ordem superior e haja 27 grupos análogos adicionais envolvendo apenas interações de ordem superior. Supondo que não haja nenhum efeito de interação de ordem superior, as 27 somas de quadrados resultantes podem então ser adicionadas para gerar uma soma de quadrados do erro, o que nos possibilita realizar testes de 1 gl para verificar todos os efeitos principais e interações de dois fatores. Contudo, em várias situações, os testes para verificação de efeitos principais podem ser obtidos apenas se combinarmos algumas ou todas as somas de quadrados associadas com grupos análogos contendo interações de dois fatores, e as interações de dois fatores correspondentes não podem ser investigadas.

### Exemplo 11.18 (Continuação do Exemplo 11.17)

O conjunto de condições de tratamento escolhido e os valores resultantes da replicação pelo quarto do experimento  $2^5$  foram

$e$	$ab$	$ad$	$bc$	$cd$	$ace$	$bde$	$abcde$
23,2	15,5	16,9	16,2	23,8	23,4	16,8	18,1

A tabela de sinais resumida é exibida na Tabela 11.15.

Sendo SQA a soma de quadrados de efeitos no grupo análogo  $\{A, CE, ABDE, BCD\}$ ,

$$SQA = \frac{(-23,2 + 15,5 + 16,9 - 16,2 - 23,8 + 23,4 - 16,8 + 18,1)^2}{8} = 4,65$$

De forma similar,  $SQB = 53,56$ ,  $SQC = 10,35$ ,  $SQD = 0,91$ ,  $SQE' = 10,35$  (' distingue essa quantidade entre as somas dos quadrados do erro SQE),  $SQAB = 6,66$  e  $SQAD = 3,25$ , fornecendo  $SQT = 4,65 + 53,56 + \dots + 3,25 = 89,73$ . Para testar os efeitos principais, usamos  $SQE = SQAB + SQAD = 9,91$  com 2 gl. A tabela ANOVA encontra-se na Tabela 11.16.

**Tabela 11.15** Tabela de sinais do Exemplo 11.18

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>AB</i>	<i>AD</i>
<i>e</i>	-	-	-	-	+	+	+
<i>ab</i>	+	+	-	-	-	+	-
<i>ad</i>	+	-	-	+	-	-	+
<i>bc</i>	-	+	+	-	-	-	+
<i>cd</i>	-	-	+	+	-	+	-
<i>ace</i>	+	-	+	-	+	-	-
<i>bde</i>	-	+	-	+	+	-	-
<i>abcde</i>	+	+	+	+	+	+	+

**Tabela 11.16** Tabela ANOVA do Exemplo 11.18

Fonte	gl	Soma dos Quadrados	Quadrado Médio	<i>f</i>
<i>A</i>	1	4,65	4,65	0,94
<i>B</i>	1	53,56	53,56	10,80
<i>C</i>	1	10,35	10,35	2,09
<i>D</i>	1	0,91	0,91	0,18
<i>E</i>	1	10,35	10,35	2,09
Erro	2	9,91	4,96	
Total	7	89,73		

Visto que  $F_{0,05,1,2} = 18,51$ , nenhum dos cinco efeitos principais pode ser considerado significativo. Obviamente, com apenas 2 gl de erro, o teste não é tão poderoso (isto é, é bem provável que não se consiga detectar a presença de efeitos). O artigo *Industrial and Engineering Chemistry*, do qual os dados se originaram, na verdade tinha uma estimativa independente do erro-padrão dos efeitos do tratamento que se baseava em experiências anteriores, de maneira que ele usou uma análise ligeiramente diferente. Aqui, o objetivo foi apenas ilustrativo, visto que alguém normalmente podia querer muito mais do que 2 gl de erro. ■

Alternativamente aos testes *F* baseados em somas de quadrados combinadas para obter SQE, pode-se examinar um gráfico de probabilidade normal de contrastes do efeito.

### Exemplo 11.19

Foi realizado um experimento para investigar o grau de encolhimento do material plástico de revestimento usado nos cabos de velocímetro ("An Explanation and Critique of Taguchi's Contribution to Quality Engineering", *Quality and Reliability Engr: Intl.*, 1988, p. 123-131). Os engenheiros começaram com 15 fatores: diâmetro externo do alinhador, matriz do alinhador, material do alinhador, velocidade da linha do alinhador, tipo

de trançado do fio, tensão do entrelaçamento, diâmetro do fio, tensão do alinhador, temperatura do alinhador, material de revestimento, tipo de matriz de revestimento, temperatura de fundição, *screen pack*, método de resfriamento e velocidade da linha. Suspeitava-se que apenas alguns desses fatores fossem importantes, de modo que foi realizado um experimento de seleção na forma de um experimento fatorial  $2^{15-11}$  (uma fração  $1/2^{11}$  de um experimento fatorial  $2^{15}$ ). A estrutura análoga resultante é bem complicada; em particular, todo efeito principal é confundido com interações de dois fatores. A variável de resposta foi a porcentagem de encolhimento de uma amostra de cabo produzida em níveis específicos dos fatores.

A Figura 11.12 exibe um gráfico de probabilidade normal dos contrastes do efeito. Com exceção de dois pontos, todos os outros estão bem próximos de uma linha reta. Os pontos discrepantes correspondem a efeitos  $E$  = tipo de entrelaçamento do fio e  $G$  = diâmetro do fio, o que sugere que esses dois fatores são os únicos que influenciam o grau de encolhimento.

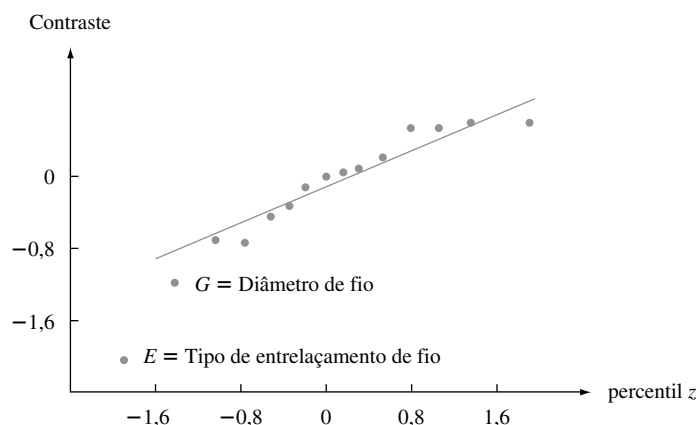


Figura 11.12 Gráfico de probabilidade normal de contrastes do Exemplo 11.19

Os tópicos do experimento fatorial, da confusão e da replicação fracionária englobam vários modelos e técnicas não discutidos aqui. Para obter mais informações, é aconselhável consultar as referências do capítulo.

## Exercícios | Seção 11.4 (38–49)

38. Os dados a seguir provêm de um experimento realizado para estudar a natureza da dependência da corrente de solda elétrica de três fatores: voltagem de solda elétrica, velocidade de alimentação do fio e distância entre a ponta da solda e a peça. Havia dois níveis de cada fator (um experimento  $2^3$ ) com duas replicações por combinação de níveis (as médias das replicações estão de acordo com os valores fornecidos no artigo “A Study on Prediction of Welding Current in Gas Metal Arc Welding”, *J. Engr. Manuf.*, 1991, p. 64-69). Os dois primeiros números dados são do tratamento (1), os dois seguintes são para  $a$ , assim por diante, na ordem-padrão: 200,0, 204,2, 215,5, 219,5, 272,7, 276,9, 299,5, 302,7, 166,6, 172,6, 186,4, 192,0, 232,6, 240,8, 253,4, 261,6.

- Verifique se as somas dos quadrados são como as fornecidas na tabela ANOVA do MINITAB a seguir.
- Que efeitos parecem importantes e por quê?

### Análise de Variância da Corrente

Fonte	DF	SS	MS	F	P
Voltagem	1	1685,1	1685,1	102,38	0,000
Velocidade	1	21272,2	21272,2	1292,37	0,000
Distância	1	5076,6	5076,6	308,42	0,000
Voltagem*veloc.	1	36,6	36,6	2,22	0,174
Voltagem*dist.	1	0,4	0,4	0,03	0,877
Velocidade*dist.	1	109,2	109,2	6,63	0,033
Volt.*veloc.*dist.	1	23,5	23,5	1,43	0,266
Erro	8	131,7	16,5		
Total	15	28335,3			

39. Os dados a seguir resultam de um experimento  $2^3$  com três replicações por combinação de tratamentos cujo objetivo era estudar os efeitos do detergente ( $A$ ), da concentração de carbonato de sódio ( $B$ ) e da concentração de carboximetil celulose de sódio ( $C$ ) na capacidade de limpeza de uma solução em testes de lavagem (um número maior indica maior capacidade de limpeza do que um número menor).



**Níveis do Fator**

A	B	C	Condição	Observações
1	1	1	(1)	106, 93, 116
2	1	1	a	198, 200, 214
1	2	1	b	197, 202, 185
2	2	1	ab	329, 331, 307
1	1	2	c	149, 169, 135
2	1	2	ac	243, 247, 220
1	2	2	bc	255, 230, 252
2	2	2	abc	383, 360, 364

- a. Depois de obter os totais de célula  $x_{ijk..}$ , calcule as estimativas de  $\beta_1$ ,  $\gamma_{11}^{AC}$  e  $\gamma_{21}^{AC}$ .
- b. Use os totais de célula e o método de Yates para calcular os contrastes do efeito e as somas dos quadrados. Em seguida, elabore uma tabela ANOVA e teste todas as hipóteses apropriadas usando  $\alpha = 0,05$ .
40. Em um estudo sobre os processos usados para remover impurezas de produtos de celulose (“Optimization of RopeRange Bleaching of Cellulosic Fabrics”, *Textile Research J.*, 1976, p. 493-496), foram obtidos os dados a seguir de um experimento  $2^4$  envolvendo o processo de limpeza. Os quatro fatores foram concentração de enzima (A), pH (B), temperatura (C) e tempo (D).

Tratamento	Enzima (g/L)	pH	Temp. (°C)	Tempo (h)	% de Goma Peso	
					1ª Replic.	2ª Replic.
(1)	0,50	6,0	60,0	6	9,72	13,50
a	0,75	6,0	60,0	6	9,80	14,04
b	0,50	7,0	60,0	6	10,13	11,27
ab	0,75	7,0	60,0	6	11,80	11,30
c	0,50	6,0	70,0	6	12,70	11,37
ac	0,75	6,0	70,0	6	11,96	12,05
bc	0,50	7,0	70,0	6	11,38	9,92
abc	0,75	7,0	70,0	6	11,80	11,10
d	0,50	6,0	60,0	8	13,15	13,00
ad	0,75	6,0	60,0	8	10,60	12,37
bd	0,50	7,0	60,0	8	10,37	12,00
abd	0,75	7,0	60,0	8	11,30	11,64
cd	0,50	6,0	70,0	8	13,05	14,55
acd	0,75	6,0	70,0	8	11,15	15,00
bcd	0,50	7,0	70,0	8	12,70	14,10
abcd	0,75	7,0	70,0	8	13,20	16,12

- a. Use o algoritmo de Yates para obter as somas dos quadrados e a tabela ANOVA.
- b. Parece haver algum efeito de interação de segunda, terceira ou quarta ordem? Explique seu raciocínio. Que efeitos principais parecem significativos?
41. No Exercício 39, suponha que, para obter os dados, tenha-se usado uma temperatura baixa de água. Todo o experimento é em seguida repetido com uma temperatura mais alta para obter os dados a seguir. Use o algoritmo de Yates em todo o conjunto de 48 observações para obter as somas dos quadrados e a tabela ANOVA e, em seguida, teste as hipóteses apropriadas no nível 0,05.

**Condição Observações**

d	144, 154, 158
ad	239, 227, 244
bd	232, 242, 246
abd	364, 362, 346
cd	194, 162, 203
acd	284, 295, 291
bcd	291, 287, 297
abcd	411, 406, 395

42. Os dados a seguir sobre o consumo de energia de altas temperaturas em fornalhas elétricas (kW consumido por tonelada de produto fundido) resultaram de um experimento fatorial  $2^4$  com três replicações (“Studies on a 10-cwt Arc Furnace”, *J. Iron and Steel Institute*, 1956, p. 22). Os fatores foram: tipo do telhado A (baixo, alto), ajuste de energia B (baixo, alto), sobras usadas C (tubos, chapas) e carga D (700 lb, 1000 lb).

Tratamento	$x_{ijklm}$	Tratamento	$x_{ijklm}$
a	946, 800, 840	ad	966, 976, 876
b	774, 834, 746	bd	702, 658, 650
ab	709, 789, 646	abd	784, 700, 596
c	1017, 990, 954	cd	922, 808, 868
ac	1028, 906, 977	acd	1056, 870, 908
bc	817, 783, 771	bcd	798, 726, 700
abc	829, 806, 691	abcd	752, 714, 714

Monte a tabela ANOVA e teste todas as hipóteses de interesse usando  $\alpha = 0,01$ .

43. O artigo “Statistical Design and Analysis of Qualification Test Program for a Small Rocket Engine” (*Industrial Quality Control*, 1964, p. 14-18) apresenta os dados de um experimento para avaliar os efeitos da vibração (A), da ciclagem da temperatura (B), ciclagem da altitude (C) e temperatura de ciclagem da altitude e da detonação (D) sobre a duração do empuxo. Um subconjunto de dados é fornecido aqui. (No artigo, havia quatro níveis de D, em vez de apenas dois.) Use o método de Yates para obter as somas dos quadrados e a tabela ANOVA. Em seguida, suponha que as interações de três e quatro fatores não existam, combine as somas dos quadrados correspondentes para obter uma estimativa de  $\sigma^2$  e teste todas as hipóteses apropriadas no nível 0,05.

		$D_1$		$D_2$	
		$C_1$	$C_2$	$C_1$	$C_2$
$A_1$	$B_1$	21,60	21,60	11,54	11,50
	$B_2$	21,09	22,17	11,14	11,32
$A_2$	$B_1$	21,60	21,86	11,75	9,82
	$B_2$	19,57	21,85	11,69	11,18

44. a. Em um experimento  $2^4$ , suponha que dois blocos serão usados e que se decida confundir a interação  $ABCD$  com o efeito do bloco. Que tratamentos devem ser formados no primeiro bloco [aquele contendo o tratamento (1)] e que tratamentos devem ser designados ao segundo bloco?
- b. Em um experimento para investigar a retenção de niacina em vegetais, como uma função da temperatura de cozimento ( $A$ ), do tamanho do crivo ( $B$ ), do tipo de processamento ( $C$ ) e do tempo de cozimento ( $D$ ), cada fator foi mantido em dois níveis. Foram usados dois blocos com a designação de blocos apresentada no item (a) para confundir somente a interação  $ABCD$  com os blocos. Use o procedimento de Yates para obter a tabela ANOVA para os dados a seguir.

Tratamento	$x_{ijkl}$	Tratamento	$x_{ijkl}$
(1)	91	$d$	72
$a$	85	$ad$	78
$b$	92	$bd$	68
$ab$	94	$abd$	79
$c$	86	$cd$	69
$ac$	83	$acd$	75
$bc$	85	$bcd$	72
$abc$	90	$abcd$	71

- c. Suponha que todos os efeitos de interação tripla não existam, de modo que as somas de quadrados correspondentes possam ser combinadas para gerar uma estimativa de  $\sigma^2$ , realize todos os testes necessários no nível 0,05.
45. a. Um experimento foi realizado para investigar os efeitos sobre a sensibilidade de áudio de uma resistência variável ( $A$ ), duas capacitâncias ( $B$ ,  $C$ ) e a indutância de uma bobina ( $D$ ) em parte de um circuito de televisão. Se fossem usados quatro blocos com quatro tratamentos por bloco e os efeitos de confusão determinantes fossem  $AB$  e  $CD$ , que tratamentos apareceriam em cada bloco?
- b. Suponha que fossem realizadas duas replicações do experimento descrito no item (a), resultando nos dados a seguir. Obtenha a tabela ANOVA e teste todas as hipóteses relevantes, usando o nível 0,01.

Tratamento	$x_{ijkl}$		Tratamento	$x_{ijkl}$	
	$x_{ijk1}$	$x_{ijk2}$		$x_{ijk1}$	$x_{ijk2}$
(1)	618	598	$d$	598	585
$a$	583	560	$ad$	587	541
$b$	477	525	$bd$	480	508
$ab$	421	462	$abd$	462	449
$c$	601	595	$cd$	603	577
$ac$	550	589	$acd$	571	552
$bc$	505	484	$bcd$	502	508
$abc$	452	451	$abcd$	449	455

46. Em um experimento envolvendo quatro fatores ( $A$ ,  $B$ ,  $C$  e  $D$ ) e quatro blocos, mostre que pelo menos um efeito principal ou um efeito de interação de dois fatores deve ser confundido com o efeito do bloco.

47. a. Em um experimento de sete fatores ( $A, \dots, G$ ), suponha que uma replicação pelo quarto seja de fato realizada. Se os efeitos determinantes forem  $ABCDE$  e  $CDEFG$ , qual será o terceiro efeito não-estimável e quais tratamentos estarão no grupo que contém (1)? Quais são os grupos análogos dos sete efeitos principais?

- b. Se a replicação pelo quarto tiver de ser realizada usando-se quatro blocos (com oito tratamentos por bloco), quais serão os blocos se os efeitos de confusão escolhidos forem  $ACF$  e  $BDG$ ?

48. Suponha que no problema de empuxo do foguete do Exercício 43 houvesse recursos suficientes apenas para uma replicação pela metade do experimento  $2^4$ .

- a. Se o efeito  $ABCD$  for escolhido como determinante para a replicação e o grupo de oito tratamentos para o qual são obtidos os dados incluir o tratamento (1), que outros tratamentos existirão no grupo observado e quais serão os pares análogos?

- b. Suponha que, ao realizar o experimento da maneira como foi descrita no item (a), sejam obtidos os resultados registrados aqui (fornecidos na ordem-padrão depois de excluir a metade não observada). Supondo que as interações de dois e três fatores são insignificantes, verifique se existem efeitos principais, usando o nível 0,05. Além disso, construa um gráfico de probabilidade normal.

19,09 20,11 21,66 20,44  
13,72 11,26 11,72 12,29

49. Uma replicação pela metade de um experimento  $2^5$  para investigar os efeitos do tempo de aquecimento ( $A$ ), do tempo de esfriamento ( $B$ ), do tempo de trefilação ( $C$ ), da posição das serpentinas de aquecimento ( $D$ ) e a posição de medição ( $E$ ) sobre a dureza das fundições de aço produziu os dados a seguir. Elabore a tabela ANOVA e (supondo que as interações de segunda ordem e de ordem superior são insignificantes) verifique se existem efeitos principais no nível 0,01. Além disso, construa um gráfico de probabilidade normal.

Tratamento	Observação	Tratamento	Observação
$a$	70,4	$acd$	66,6
$b$	72,1	$ace$	67,5
$c$	70,4	$ade$	64,0
$d$	67,4	$bcd$	66,8
$e$	68,0	$bce$	70,3
$abc$	73,8	$bde$	67,9
$abd$	67,0	$cde$	65,9
$abe$	67,8	$abcde$	68,0

## Exercícios Suplementares (50–61)

50. Os resultados de um estudo que verificou a eficácia da secagem no varal para a maciez do tecido foram resumidos no artigo "Line-Dried vs. Machine-Dried Fabrics: Comparison of Appearance, Hand, and Consumer Acceptance" (*Home Econ. Research J.*, 1984, p. 27-35). A avaliação dos níveis de maciez foi feita para nove tipos diferentes de tecido e cinco diferentes métodos de secagem: (1) secagem a máquina, (2) secagem no varal, (3) secagem no varal seguida de 15 min. na secadora, (4) secagem no varal com amaciante e (5) secagem no varal com ventilação. Considerando os diferentes tipos de tecido como blocos, construa uma tabela ANOVA. Usando um nível de significância 0,05, verifique se existe alguma diferença para a pontuação na média real da maciez dos métodos de secagem.

	Método de Secagem				
	1	2	3	4	5
<b>Tecido</b>					
<b>Crepe</b>	3,3	2,5	2,8	2,5	1,9
<b>Malha</b>	3,6	2,0	3,6	2,4	2,3
<b>Sarja</b>	4,2	3,4	3,8	3,1	3,1
<b>Sarja</b>	3,4	2,4	2,9	1,6	1,7
<b>Veludo</b>	3,8	1,3	2,8	2,0	1,6
<b>Casimira</b>	2,2	1,5	2,7	1,5	1,9
<b>Pano de</b>	3,5	2,1	2,8	2,1	2,2
<b>Veludo</b>	3,6	1,3	2,8	1,7	1,8
<b>Brim</b>	2,6	1,4	2,4	1,3	1,6

51. A absorção de água de dois tipos de argamassa usados para reparar cimentos danificados foi discutida no artigo "Polymer Mortar Composite Matrices for Maintenance-Free, Highly Durable Ferrocement" (*J. Ferrocement*, 1984, p. 337-345). Amostras de argamassa de cimento comum (ACC) e argamassa composta de cimento e polímero (ACP) foram submersas por períodos de tempo variáveis (5, 9, 24 ou 48 horas), registrando-se a absorção de água (% por peso). Sendo o fator *A* (com dois níveis) o tipo de argamassa e o fator *B* (com quatro níveis) o período de submersão, foram feitas três observações para cada combinação de níveis do fator. Os dados fornecidos no artigo foram usados para calcular as somas dos quadrados:  $SQA = 322,667$ ,  $SQB = 35,623$ ,  $SQAB = 8,557$  e  $SQT = 372,113$ . Use essas informações para construir uma tabela ANOVA. Teste as hipóteses apropriadas no nível de significância 0,05.

52. Havia quatro canteiros disponíveis para um experimento de comparação da quantidade de trevos para quatro diferentes taxas de semeadura ("Performance of Overdrilled Red Clover with Different Sowing Rates and Initial Grazing Managements", *N. Zeal. J. Exp. Ag.*, 1984, p. 71-81). Visto que os canteiros foram adubados diferentemente antes do experimento e que se avaliou que isso podia influenciar na quantidade de trevos, foi

usado um experimento em bloco aleatorizado, e todas as quatro taxas de semeadura foram experimentadas em uma parte de cada canteiro. Use os dados fornecidos para testar a hipótese nula de que não existe nenhuma diferença na média real da acumulação de trevos (kg dm/ha) para as diferentes taxas de semeadura.

	Taxa de Semeadura (kg/ha)			
	3,6	6,6	10,2	13,5
<b>Canteiro</b>				
<b>1</b>	1155	2255	3505	4632
<b>2</b>	123	406	564	416
<b>3</b>	68	416	662	379
<b>4</b>	62	75	362	564

53. Em um processo de revestimento químico automatizado, a velocidade de acordo com a qual os objetos em uma esteira rolante são passados por um *spray* químico (velocidade da esteira), a quantidade de substância química pulverizada (volume do *spray*) e a marca da substância usada (marca) são fatores que podem influenciar na uniformidade do revestimento aplicado. Um experimento de 2<sup>3</sup> replicado foi conduzido na tentativa de aumentar a uniformidade do revestimento. Na tabela a seguir, os valores mais altos da variável de resposta estão associados com uma maior uniformidade da superfície:

Seqüência	Volume de Spray	Velocidade da Esteira	Marca	Uniformidade da Superfície	
				Replição 1	Replição 2
1	-	-	-	40	36
2	+	-	-	25	28
3	-	+	-	30	32
4	+	+	-	50	48
5	-	-	+	45	43
6	+	-	+	25	30
7	-	+	+	30	29
8	+	+	+	52	49

Analise esses dados e expresse suas conclusões.

54. As usinas elétricas abastecidas a carvão usadas no setor elétrico têm recebido atenção pública crescente em decorrência dos problemas ambientais associados a lixos sólidos gerados pela combustão em larga escala ("Fly Ash Binders in Stabilization of FGD Wastes", *J. of Environmental Engineering*, 1998, p. 43-49). Um estudo foi conduzido para analisar a influência de três fatores – tipo de aglutinante (*A*), quantidade de água (*B*) e método de descarte (*C*) – que afetam determinadas características de lixiviação de lixos sólidos da

combustão. Cada um dos fatores foi estudado em dois níveis. Um experimento  $2^3$  não-replicável foi realizado, e um valor de resposta CE50 (de concentração efetiva em mg/L, que diminui 50% da luz em um ensaio de bioluminescência) foi medido para cada combinação de níveis do fator. Os dados experimentais são fornecidos na tabela a seguir:

Seqüência	Fator			Resposta EC50
	A	B	C	
1	-1	-1	-1	23.100
2	1	-1	-1	43.000
3	-1	1	-1	71.400
4	1	1	-1	76.000
5	-1	-1	1	37.000
6	1	-1	1	33.200
7	-1	1	1	17.000
8	1	1	1	16.500

Construa uma ANOVA apropriada e expresse suas conclusões.

55. Impurezas na forma de óxidos de ferro diminuem o valor econômico e a utilidade dos minerais industriais, como o caulim, em indústrias cerâmicas e de papel. Um experimento  $2^4$  foi conduzido para avaliar os efeitos de quatro fatores sobre a porcentagem de ferro extraída de amostras de caulim (“Factorial Experiments in the Development of a Kaolin Bleaching Process Using Thiourea in Sulphuric Acid Solutions,” *Hydrometallurgy*, 1997, p. 181-197). Os fatores e respectivos níveis estão relacionados na tabela a seguir:

Fator	Descrição	Unidades	Nível	Nível
			Baixo	Alto
A	H <sub>2</sub> SO <sub>4</sub>	M	0,10	0,25
B	Tiouréia	g/L	0,0	5,0
C	Temperatura	°C	70	90
D	Tempo	min.	30	150

Os dados de um experimento  $2^4$  não-replicado são apresentados na tabela a seguir.

Seqüência do Teste	Extração de Ferro		Seqüência do Teste	Extração de Ferro (%)
	Ferro (%)			
(1)	7		<i>d</i>	28
<i>a</i>	11		<i>ad</i>	51
<i>b</i>	7		<i>bd</i>	33
<i>ab</i>	12		<i>abd</i>	57
<i>c</i>	21		<i>cd</i>	70
<i>ac</i>	41		<i>acd</i>	95
<i>bc</i>	27		<i>bcd</i>	77
<i>abc</i>	48		<i>abcd</i>	99

- a. Calcule a estimativa de todos os efeitos principais e dos efeitos das interações de dois fatores para esse experimento.

- b. Crie um gráfico de probabilidade desses efeitos. Que efeitos parecem importantes?

56. Modelos fatoriais foram usados na silvicultura para avaliar os efeitos de vários fatores sobre o comportamento do crescimento das árvores. Em um experimento desse tipo, os pesquisadores acreditavam que mudas saudáveis de abetos provavelmente germinariam mais rapidamente que mudas doentes (“Practical Analysis of Factorial Experiments in Forestry”, *Canadian J. of Forestry*, 1995, p. 446-461). Além disso, antes do plantio, as mudas foram também expostas a três níveis de pH para verificar se esse fator tem um efeito sobre a passagem de vírus para as raízes. A tabela a seguir mostra os dados de um experimento  $2 \times 3$  para estudar ambos os fatores:

		pH		
		3	5,5	7
Saúde	Doente	1,2, 1,4,	0,8, 0,6,	1,0, 1,0,
		1,0, 1,2,	0,8, 1,0,	1,2, 1,4,
		1,4	0,8	1,2
	Saudável	1,4, 1,6,	1,0, 1,2,	1,2, 1,4,
		1,6, 1,6,	1,2, 1,4,	1,2, 1,2,
		1,4	1,4	1,4

A variável de resposta é uma avaliação média dos cinco brotos de uma muda. Os índices são 0 (broto não-quebrado), 1 (broto parcialmente desenvolvido) e 2 (broto totalmente desenvolvido). Analise esses dados.

57. Uma propriedade dos *air bags* de automóvel que contribui para sua capacidade de absorver energia é a permeabilidade (pés<sup>3</sup>/pés<sup>2</sup>/min.) do material entrelaçado usado para fabricar os *air bags*. Entender como a permeabilidade é influenciada por vários fatores é fundamental para a eficácia dos *air bags*. Em um determinado estudo, foram analisados os efeitos de três fatores, cada um em três níveis (“Analysis of Fabrics used in Passive Restraint Systems-Airbags”, *J. of the Textile Institute*, 1996, p. 554-571):

A (Temperatura): 8°C, 50°C, 75°C

B (Espessura do tecido, em *denier*): 420-D, 630-D, 840-D

C (Pressão do ar): 17,2 kPa, 34,4 kPa, 103,4 kPa

Temperatura 8°

Denier	Pressão		
	17,2	34,4	103,4
420-D	73	157	332
	80	155	322
630-D	35	91	288
	433	98	271
840-D	125	234	477
	111	233	464

**Temperatura 50°**

Denier	Pressão		
	17,2	34,4	103,4
420-D	52	125	281
	51	118	264
630-D	16	72	169
	12	78	173
840-D	96	149	338
	100	155	350

**Temperatura 75°**

Denier	Pressão		
	17,2	34,4	103,4
420-D	37	95	276
	31	106	281
630-D	30	91	213
	41	100	211
840-D	102	170	307
	98	160	311

Análise esses dados e expresse suas conclusões (suponha que todos os fatores são fixos).

- 58. Um engenheiro químico realizou um experimento para estudar os efeitos dos fatores fixos da pressão do silo (*A*), tempo de cozimento da polpa (*B*) e concentração de madeira (*C*) na resistência do papel. Esse experimento usou duas pressões, quatro tempos de cozimento, três concentrações e duas observações em cada combinação desses níveis. As somas dos quadrados calculadas são  $SQA = 6,94$ ,  $SQB = 5,61$ ,  $SQC = 12,33$ ,  $SQAB = 4,05$ ,  $SQAC = 7,32$ ,  $SQBC = 15,80$ ,  $SQE = 14,40$  e  $SQT = 70,82$ . Construa a tabela ANOVA e realize os testes apropriados no nível de significância 0,05.
- 59. A resistência de contato na montagem de um circuito integrado em um substrato de vidro metalizado foi estudada como uma função do fator *A* = tipo de adesivo, do fator *B* = tempo de cura, e do fator *C* = material condutor (cobre e níquel). Os dados são apresentados a seguir, bem como uma tabela ANOVA do MINITAB. Que conclusões você pode tirar desses dados?

Cobre		Tempo de Cura		
		1	2	3
1	1	72,7	74,6	80,0
	2	80,0	77,5	82,7
Adesivo	2	77,8	78,5	84,6
	3	75,3	81,1	78,3
3	1	77,3	80,9	83,9
	2	76,5	82,6	85,0
Níquel		1	2	3
		1	74,7	75,7
Adesivo	2	77,4	78,2	74,6
	3	79,3	78,8	83,0
3	1	77,8	75,4	83,9
	2	77,2	84,5	89,4
		78,4	77,5	81,2

Análise de Variância da Resistência

Fonte	GL	SQ	QM	F	P
Adesivo	2	101,317	50,659	6,54	0,007
Tempo de cura	2	151,317	75,659	9,76	0,001
Conmater	1	0,722	0,722	0,09	0,764
Adesivo*Tempo de cura	4	30,526	7,632	0,98	0,441
Adesivo*conmater	2	8,015	4,008	0,52	0,605
Tempo de cura*conmater	2	5,952	2,976	0,38	0,687
Adesivo*tempo de cura*conmater	4	33,298	8,325	1,07	0,398
Erro	18	139,515	7,751		
Total	35	470,663			

- 60. O artigo “Food Consumption and Energy Requirements of Captive Bald Eagles” (*J. Wildlife Mgmt.*, 1982, p. 646-654) investigou a ingestão média bruta de calorias diárias (a variável de resposta) de diferentes tipos de dieta (fator *A*, com três níveis) e temperatura (fator *B*, com três níveis). As quantidades resumidas fornecidas no artigo foram usadas para gerar os dados, resultando em  $SQA = 18.138$ ,  $SQB = 5182$ ,  $SQAB = 1737$ ,  $SQT = 36.348$  e gl do erro = 36. Construa uma tabela ANOVA e teste as hipóteses relevantes.
- 61. Analogamente ao quadrado latino, um modelo em quadrado greco-latino pode ser usado quando existe a suspeita de que três fatores externos podem influenciar a variável de resposta e os quatro fatores (os três externos e o de interesse) têm o mesmo número de níveis. No quadrado latino, cada nível do fator de interesse (*C*) aparece uma vez em cada linha (com cada nível de *A*) e uma vez em cada coluna (com cada nível de *B*). No quadrado greco-latino, cada nível do fator *D* aparece uma vez em cada linha, em cada coluna e também com cada nível do terceiro fator externo *C*. Alternativamente, esse modelo pode ser usado quando os quatro fatores têm o mesmo grau de interesse, o número de níveis de cada um é *N* e os recursos são suficientes apenas para  $N^2$  observações. Um quadrado  $5 \times 5$  é ilustrado em (a), no qual (*k*, *l*) em cada célula representa o *k*-ésimo nível de *C* e o *l*-ésimo de *D*. Em (b), apresentamos os dados da perda de peso em barras de silício usadas no material semicondutor como uma função do volume de solução ácida (*A*), cor do ácido nítrico na solução ácida (*B*), tamanho das barras (*C*) e tempo dentro da solução ácida (*D*) (com base em “Applications of Analytic Techniques to the Semiconductor Industry”, Fourteenth Midwest Quality Control Conference, 1959).

Considere  $X_{ij(kl)}$  a perda de peso observada quando o fator *A* está no nível *i*, *B* está no nível *j*, *C* está no nível *k* e *D* está no nível *l*. Supondo que não há nenhuma interação entre os fatores, a soma de quadrados total  $SQT$  (com  $N^2 - 1$  gl) pode ser dividida em  $SQA$ ,  $SQB$ ,  $SQC$ ,  $SQD$  e  $SQE$ . Apresente expressões para essas somas de quadrados, incluindo fórmulas de cálculo, obtenha a tabela ANOVA para os dados especificados e teste cada uma das hipóteses dos quatro efeitos principais, usando  $\alpha = 0,05$ .

(C, D)		B				
		1	2	3	4	5
A	1	(1, 1)	(2, 3)	(3, 5)	(4, 2)	(5, 4)
	2	(2, 2)	(3, 4)	(4, 1)	(5, 3)	(1, 5)
	3	(3, 3)	(4, 5)	(5, 2)	(1, 4)	(2, 1)
	4	(4, 4)	(5, 1)	(1, 3)	(2, 5)	(3, 2)
	5	(5, 5)	(1, 2)	(2, 4)	(3, 1)	(4, 3)

(a)

65	82	108	101	126
84	109	73	97	83
105	129	89	89	52
119	72	76	117	84
97	59	94	78	106

(b)

---

## Bibliografia

- BOX, G., HUNTER, W. e HUNTER, S. *Statistics for Experimenters*. Nova York: Wiley, 1978. Contém uma profusão de sugestões e idéias sobre análise de dados fundamentadas na extensa experiência dos autores em consultoria.
- DEVOR, R., CHANG, T. e SUTHERLAND, J. W. *Statistical Quality Design and Control*. Nova York: Macmillan, 1992. Inclui uma moderna pesquisa sobre experimentos fatoriais e fracionais, utilizando o mínimo possível de processos matemáticos.
- HOCKING, R. *The Analysis of Linear Models*. Pacific Grove, CA: Brooks/Cole, 1985. Um tratado bem geral de análise de variância escrito por uma das maiores autoridades nesse campo.
- KLEINBAUM, D., KUPPER, L., MULLER, K. e NIZAM, A. *Applied Regression Analysis and Other Multivariable Methods*. Pacific Grove: Duxbury Press, 1998. 3. ed. Contém uma discussão especialmente exaustiva sobre problemas associados à análise de “dados desequilibrados” isto é,  $K_{ijs}$  desiguais.
- KUEHL, R. O. *Statistical Principles of Research Design and Analysis*. Belmont, CA: Wadsworth, 1994. Um tratamento atualizado e abrangente sobre experimentos projetados e análises dos dados resultantes.
- MONTGOMERY, D. *Design and Analysis of Experiments*. Nova York: Wiley, 2001. 5. ed. Consulte a bibliografia do Capítulo 10.
- NETER, J., WASSERMAN, W. e KUTNER, M. *Applied Linear Statistical Models*. Homewood, IL: Irwin, 1996. 4. ed. Consulte a bibliografia do Capítulo 10.
- VARDEMAN, S. *Statistics for Engineering Problem Solving*, PWS. Boston: 1994. Introdução geral para engenheiros, com muita metodologia descritiva e inferencial para dados de experimentos projetados.

# Regressão Linear Simples e Correlação

## Introdução

Nos problemas de duas amostras discutidos no Capítulo 9, concentramo-nos na comparação de valores de parâmetros da distribuição  $x$  e da distribuição  $y$ . Mesmo quando as observações foram pareadas, não tentamos usar informações sobre uma das variáveis ao estudar a outra variável. Este é precisamente o objetivo da análise de regressão: explorar a relação entre duas (ou mais) variáveis, de modo que possamos obter informações sobre uma delas, por meio dos valores conhecidos da(s) outra(s).

O objetivo de grande parte dos cálculos é investigar as variáveis que estão relacionadas deterministicamente. Afirmar que  $x$  e  $y$  estão relacionadas dessa maneira significa dizer que o conhecimento do valor de  $x$  implica o conhecimento exato do valor de  $y$ . Por exemplo: suponha que decidamos alugar uma van por um dia e que o valor do aluguel seja \$ 25,00 mais \$ 0,30 por quilômetro rodado. Sendo  $x$  = o número de quilômetros rodados e  $y$  = a despesa de aluguel, então  $y = 25 + 0,3x$ . Se percorrermos 100 quilômetros com a van ( $x = 100$ ), então  $y = 25 + 0,3(100) = 55$ . Como outro exemplo, se a velocidade inicial de uma partícula for  $v_0$  e ela sofrer uma aceleração constante  $a$ , a distância percorrida  $y = v_0x + \frac{1}{2}ax^2$ , onde  $x$  = tempo.

Muitas variáveis  $x$  e  $y$  podem até parecer relacionadas uma com a outra, mas não de maneira determinística. Um exemplo comum para vários alunos é dado pelas variáveis  $x$  = média total geral de notas escolares do ensino médio (Grade Point Average – GPA) e  $y$  = GPA da faculdade. O valor de  $y$  não pode ser determinado apenas com base no valor conhecido de  $x$  e dois diferentes alunos podem ter o mesmo valor de  $x$ , mas com valores de  $y$  bem diferentes. Contudo, existe uma tendência de alunos com GPAs altos (baixos) no ensino médio terem também GPAs altos (baixos) na faculdade. Conhecer o GPA do ensino médio de um aluno pode ser muito útil para nos ajudar a prever seu desempenho na faculdade.

Outros exemplos de variáveis relacionadas de maneira não-determinística incluem  $x$  = idade de uma criança e  $y$  = tamanho do vocabulário dessa criança,  $x$  = tamanho de um motor em centímetros cúbicos e  $y$  = rendimento do combustível de um automóvel

equipado com esse motor e  $x$  = força de tração aplicada e  $y$  = quantidade de alongamento de uma tira de metal.

A **análise de regressão** é a parte da estatística que investiga a relação entre duas ou mais variáveis relacionadas de maneira não-determinística. Neste capítulo, estendemos (generalizamos) a relação linear determinística  $y = \beta_0 + \beta_1 x$  para uma relação linear probabilística; desenvolvemos procedimentos para fazer inferências sobre os parâmetros do modelo e obter uma medida quantitativa (o coeficiente de correlação) sobre até que ponto as duas variáveis estão relacionadas. No capítulo seguinte, abordaremos as técnicas de validação de um determinado modelo e investigaremos as relações não-lineares e as relações com mais de duas variáveis.

## 12.1 O Modelo de Regressão Linear Simples

A relação matemática determinística mais simples entre duas variáveis  $x$  e  $y$  é uma relação linear  $y = \beta_0 + \beta_1 x$ . O conjunto de pares  $(x, y)$  para o qual  $y = \beta_0 + \beta_1 x$  determina uma reta com coeficiente angular  $\beta_1$  e termo constante  $\beta_0$ .<sup>1</sup> O objetivo desta seção é desenvolver um modelo probabilístico linear.

Se as duas variáveis não estiverem relacionadas deterministicamente, então, para um valor fixo de  $x$ , o valor da segunda variável será aleatório. Por exemplo: se estivermos investigando a relação entre a idade de uma criança e o tamanho do seu vocabulário e decidirmos selecionar uma criança de idade  $x = 5,0$  anos, então, antes que a seleção seja feita, o tamanho do vocabulário será uma variável aleatória  $Y$ . Depois que uma determinada criança de 5 anos tiver sido selecionada e testada, o resultado será, por exemplo, um vocabulário de 2000 palavras. Portanto, podemos dizer que o valor observado de  $Y$ , associado com a fixação de  $x = 5,0$ , foi  $y = 2000$ .

Na maioria das vezes, a variável cujo valor é fixado pelo pesquisador será representada por  $x$  e será chamada de variável **independente, previsão** ou **explicativa**. Para  $x$  fixo, a segunda variável será aleatória; representamos essa variável aleatória e seu valor observado por  $Y$  e  $y$ , respectivamente, e as chamamos de variável **dependente** ou **resposta**.

Em geral, as observações serão feitas para inúmeros conjuntos da variável independente. Sejam  $x_1, x_2, \dots, x_n$  os valores da variável independente para as quais são feitas as observações e sejam  $Y$  e  $y$ , respectivamente, a variável aleatória e o valor observado associados a  $x_i$ . Portanto, os dados bivariados fornecidos consistem nos  $n$  pares  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ . O primeiro passo na análise de regressão com duas variáveis é elaborar um gráfico de dispersão dos dados observados. Em um gráfico desse tipo, cada  $(x_i, y_i)$  é representado como um ponto representado graficamente em um sistema de coordenadas bidimensional.

### Exemplo 12.1

Problemas visuais e musculoesqueléticos associados com o uso de terminais de exibição visual (Visual Display Terminals – VDTs) tornaram-se muito comuns nos últimos anos. Alguns pesquisadores concentraram-se na direção vertical do olhar como uma fonte de esforço e irritação ocular. Acredita-se que essa direção esteja intimamente relacionada com a Área da Superfície Ocular (ASO), de modo que é necessário um método para mediá-la. Os dados representativos a seguir sobre  $y = \text{ASO (cm}^2)$  e  $x = \text{largura da fissura da pálpebra (isto é, a largura horizontal da abertura do olho, em cm)}$  foram reproduzidos do artigo "Analysis of Ocular Surface Area for Comfortable VDT Workstation Layout" (*Ergonomics*, 1996, p. 877–884). A ordem na qual as observações foram obtidas não foi dada, de maneira que, por conveniência, elas são relacionadas na ordem crescente dos valores  $x$ .

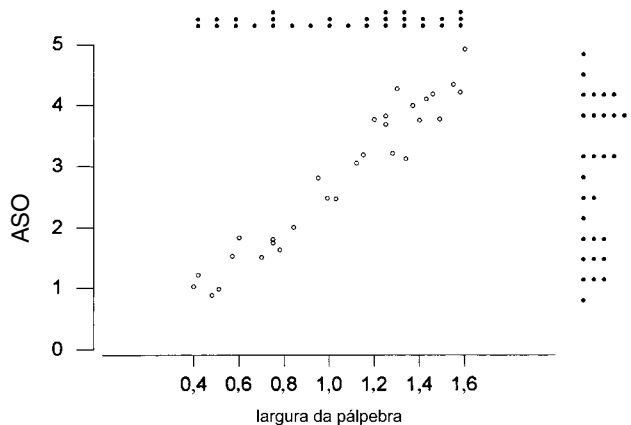
<sup>1</sup> O coeficiente angular de uma reta é a mudança em  $y$  decorrente do aumento de 1 unidade em  $x$ . Por exemplo: se  $y = -3x + 10$ ,  $y$  diminui em 3 quando  $x$  cresce em 1, de modo que o coeficiente angular é  $-3$ . O termo constante é a altura em que a reta cruza o eixo vertical e é obtido estabelecendo-se  $x = 0$  na equação.



$i$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$x_i$	0,40	0,42	0,48	0,51	0,57	0,60	0,70	0,75	0,75	0,78	0,84	0,95	0,99	1,03	1,12
$y_i$	1,02	1,21	0,88	0,98	1,52	1,83	1,50	1,80	1,74	1,63	2,00	2,80	2,48	2,47	3,05
$i$	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
$x_i$	1,15	1,20	1,25	1,25	1,28	1,30	1,34	1,37	1,40	1,43	1,46	1,49	1,55	1,58	1,60
$y_i$	3,18	3,76	3,68	3,82	3,21	4,27	3,12	3,99	3,75	4,10	4,18	3,77	4,34	4,21	4,92

Portanto,  $(x_1, y_1) = (0,40, 1,02)$ ,  $(x_5, y_5) = (0,57, 1,52)$ , e assim por diante. Um gráfico de dispersão do MINITAB é mostrado na Figura 12.1; usamos uma opção que produziu um gráfico de pontos tanto dos valores  $x$  como dos valores  $y$  individualmente ao longo das margens direita e superior do gráfico, o que facilita a visualização das distribuições das variáveis individuais (outras alternativas são os histogramas ou *bloxplots*). Aqui, há alguns detalhes a serem observados nos dados e no gráfico:

- Várias observações têm valores  $x$  idênticos, mas diferentes valores  $y$  (por exemplo,  $x_8 = x_9 = 0,75$ , mas  $y_8 = 1,80$  e  $y_9 = 1,74$ ). Portanto, o valor de  $y$  não é determinado apenas por  $x$ , como também por vários outros fatores.
- Existe uma grande tendência de  $y$  aumentar à medida que  $x$  aumenta. Ou seja, valores maiores da ASO tendem a estar associados com valores maiores da largura da fissura – uma relação positiva entre as variáveis.



**Figura 12.1** Gráfico de dispersão do MINITAB para os dados do Exemplo 12.1, bem como gráfico de pontos dos valores  $x$  e  $y$

- Aparentemente o valor de  $y$  pode ser previsto com base no valor de  $x$ , encontrando-se uma reta razoavelmente próxima dos pontos no gráfico (os autores do artigo citado fizeram isso). Em outras palavras, há evidência de uma relação linear considerável (embora não perfeita) entre as duas variáveis. ■

Os eixos horizontal e vertical no gráfico de dispersão da Figura 12.1 cruzam no ponto  $(0, 0)$ . Em vários conjuntos de dados, os valores  $x$  ou  $y$  ou os valores de ambas as variáveis diferem consideravelmente de zero em relação à(s) amplitude(s) dos valores. Por exemplo: um estudo sobre como a eficiência do ar-condicionado está relacionada com a temperatura máxima diária externa pode envolver observações para temperaturas no intervalo de  $80^{\circ}\text{F}$  a  $100^{\circ}\text{F}$ . Quando esse for o caso, um gráfico mais informativo mostraria o cruzamento dos eixos adequadamente em algum ponto diferente de  $(0, 0)$ .

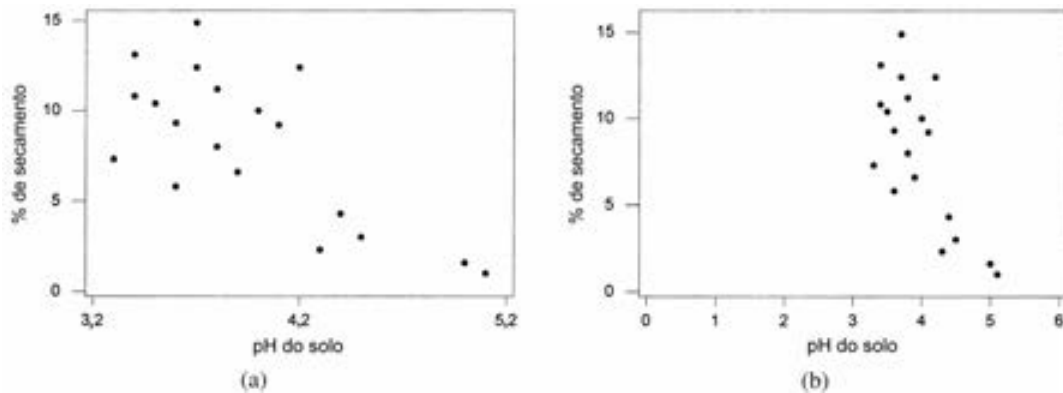
**Exemplo 12.2**

Os fenômenos mundiais de crescimento e declínio florestal atraíram grande interesse público e científico. O artigo “Relationships Among Crown Condition, Growth, and Stand Nutrition in Seven Northern Vermont Sugarbushes” (*Canad. J. of Forest Res.*, 1995, p. 386–397) incorporou um gráfico de dispersão de  $y =$  secamento médio dos ramos (%), um indicador de retardamento no crescimento, e  $x =$  pH do solo (um pH mais alto corresponde a um solo mais ácido), do qual foram tiradas as observações a seguir:

$x$	3,3	3,4	3,4	3,5	3,6	3,6	3,7	3,7	3,8	3,8
$y$	7,3	10,8	13,1	10,4	5,8	9,3	12,4	14,9	11,2	8,0
$x$	3,9	4,0	4,1	4,2	4,3	4,4	4,5	5,0	5,1	
$y$	6,6	10,0	9,2	12,4	2,3	4,3	3,0	1,6	1,0	

A Figura 12.2 a seguir mostra dois gráficos de dispersão desses dados feitos no MINITAB. Na Figura 12.2(a), o MINITAB selecionou a escala para ambos os eixos. Obtivemos a Figura 12.2(b) especificando valores mínimos e máximos de  $x$  e  $y$  de modo que os eixos cruzassem aproximadamente no ponto  $(0, 0)$ . O segundo gráfico é mais compacto que o primeiro; essa compactação pode dificultar a especificação da característica geral de qualquer relação. Por exemplo: pode ser mais difícil perceber uma curvatura em um gráfico compacto.

Porcentagens altas de secamento dos ramos tendem a estar associadas a solos com níveis baixos de pH, uma relação negativa ou inversa. Além do mais, as duas variáveis parecem estar relacionadas pelo menos quase linearmente, embora reconhecendo que os pontos estariam dispersos em qualquer reta traçada sobre eles.



**Figura 12.2** Gráficos de dispersão do MINITAB dos dados do Exemplo 12.2

**Modelo Probabilístico Linear**

Para o modelo determinístico  $y = \beta_0 + \beta_1x$ , o valor real observado de  $y$  é uma função linear de  $x$ . A generalização apropriada dessa característica para um modelo probabilístico pressupõe que o valor esperado de  $Y$  é uma função linear de  $x$ , mas que, para um  $x$  fixo, a variável  $Y$  difere de seu valor esperado de uma quantidade aleatória.

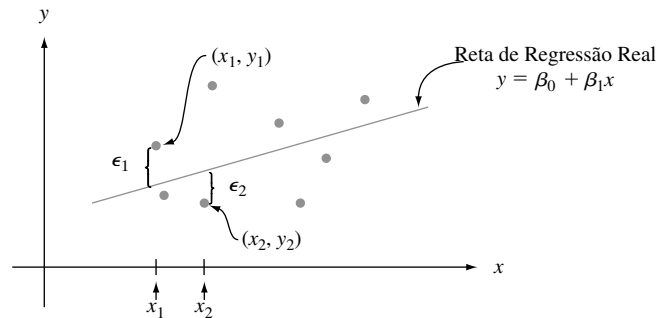
**O Modelo de Regressão Linear Simples**

Existem parâmetros  $\beta_0, \beta_1,$  e  $\sigma^2$  tais que, para qualquer valor fixo da variável independente  $x$ , a variável dependente está relacionada a  $x$  por meio da **equação do modelo**.

$$Y = \beta_0 + \beta_1x + \epsilon \tag{12.1}$$

A quantidade  $\epsilon$  na equação do modelo é uma variável aleatória, considerada normalmente distribuída com  $E(\epsilon) = 0$  e  $V(\epsilon) = \sigma^2$ .

A variável  $\epsilon$  normalmente é chamada de **desvio aleatório** ou **erro aleatório** do modelo. Sem  $\epsilon$ , qualquer par observado  $(x, y)$  corresponderia a um ponto disposto exatamente na reta  $y = \beta_0 + \beta_1 x$ , denominada **reta de regressão real** (ou da **população**). A inclusão do termo “erro aleatório” permite que  $(x, y)$  fique acima da reta de regressão real (quando  $\epsilon > 0$ ) ou abaixo da reta (quando  $\epsilon < 0$ ). Os pontos  $(x_1, y_1), \dots, (x_n, y_n)$  resultantes de  $n$  observações independentes serão então dispersos próximos da reta de regressão, como ilustrado na Figura 12.3. Ocasionalmente, a aplicabilidade do modelo de regressão linear simples pode ser sugerida por considerações teóricas (por exemplo: existe uma relação linear exata entre as duas variáveis, em que  $\epsilon$  representa o erro de medição). Contudo, bem mais freqüentemente, a plausibilidade do modelo é indicada por um gráfico de dispersão exibindo um padrão linear significativo (como na Figura 12.1).



**Figura 12.3** Pontos correspondentes a observações do modelo de regressão linear simples

As implicações da equação do modelo (12.1) podem ser melhor compreendidas com a ajuda da notação a seguir. Seja  $x^*$  um valor particular da variável independente  $x$  e

$\mu_{Y,x^*}$  = o valor esperado (ou médio) de  $Y$  quando  $x = x^*$

$\sigma_{Y,x^*}^2$  = a variância de  $Y$  quando  $x = x^*$

Uma notação alternativa é  $E(Y|x^*)$  e  $V(Y|x^*)$ . Por exemplo: se  $x$  = força aplicada (kg/mm<sup>2</sup>) e  $y$  = tempo até a ruptura ( $h$ ), então  $\mu_{Y,20}$  representaria o valor esperado do tempo até a ruptura, quando a força aplicada fosse 20 kg/mm<sup>2</sup>. Se pensarmos em uma população completa de pares  $(x, y)$ , então  $\mu_{Y,x^*}$  será a média de todos os valores  $y$  para os quais  $x = x^*$ , e  $\sigma_{Y,x^*}^2$  é a medida de dispersão dos valores em torno do valor médio. Se, por exemplo,  $x$  = idade da criança e  $y$  = tamanho do vocabulário, então  $\mu_{Y,5}$  é o tamanho médio do vocabulário de todas as crianças de 5 anos na população e  $\sigma_{Y,5}^2$  descreve a variabilidade no tamanho do vocabulário para essa parte da população. Assim que  $x$  é fixado, a única aleatoriedade no lado direito da equação do modelo (12.1) está no erro aleatório  $\epsilon$ , e seu valor médio e variância são 0 e  $\sigma^2$ , respectivamente, seja qual for o valor de  $x$ . Isso significa que

$$\mu_{Y,x^*} = E(\beta_0 + \beta_1 x^* + \epsilon) = \beta_0 + \beta_1 x^* + E(\epsilon) = \beta_0 + \beta_1 x^*$$

$$\sigma_{Y,x^*}^2 = V(\beta_0 + \beta_1 x^* + \epsilon) = V(\beta_0 + \beta_1 x^*) + V(\epsilon) = 0 + \sigma^2 = \sigma^2$$

Substituindo  $x^*$  em  $\mu_{Y,x^*}$  por  $x$ , obtém-se a relação  $\mu_{Y,x} = \beta_0 + \beta_1 x$ , que estabelece que o valor médio de  $Y$ , em vez de o próprio  $Y$ , é uma função linear de  $x$ . A reta de regressão real  $y = \beta_0 + \beta_1 x$  é, portanto, a reta de valores médios; sua altura acima de qualquer valor particular de  $x$  é o valor esperado de  $Y$  para esse valor de  $x$ . O coeficiente angular  $\beta_1$  da reta de regressão real é interpretado como a mudança esperada em  $Y$  associada com um aumento de 1 unidade no valor de  $x$ . A segunda relação estabelece que a variabilidade na distribuição de valores  $Y$  é a mesma em cada valor diferente de  $x$  (homogeneidade da variância). No exemplo que envolve a idade de uma criança e o tamanho do vocabulário, o modelo significa que o tamanho médio do vocabulário muda linearmente com a idade (a expectativa é que  $\beta_1$  é positivo) e que a quantidade de variabilidade do vocabulário em qualquer idade em particular é a mesma que em qualquer outra idade. Finalmente, para  $x$  fixo,  $Y$  é a soma de um  $\beta_0 + \beta_1 x$  constante e a  $\epsilon$  normalmente distribuída, implicando que  $Y$  tem uma distribuição normal. Essas

propriedades são ilustradas na Figura 12.4. O parâmetro de variância  $\sigma^2$  determina até que ponto cada curva normal se dispersa ao redor de seu valor médio (a altura da reta). Quando  $\sigma^2$  é pequeno, um ponto observado  $(x, y)$  quase sempre ficará bem próximo da reta de regressão real, ao passo que as observações podem desviar consideravelmente de seus valores esperados (correspondendo a pontos distantes da reta) quando  $\sigma^2$  for grande.

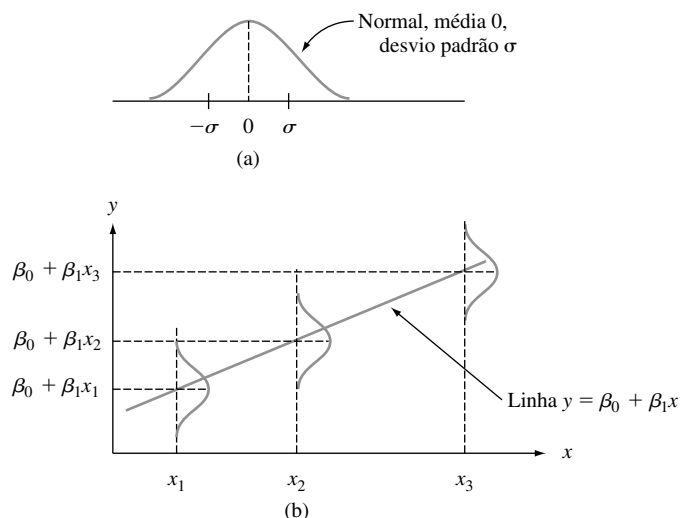


Figura 12.4 (a) Distribuição de  $\epsilon$ ; (b) distribuição de Y para diferentes valores de x

**Exemplo 12.3**

Suponha que a relação entre a força aplicada  $x$  e o tempo até a falha  $y$  seja descrita por um modelo de regressão linear simples com uma reta de regressão real  $y = 65 - 1,2x$  e  $\sigma = 8$ . Portanto, para qualquer valor fixo  $x^*$  de força, o tempo até a falha tem uma distribuição normal com valor médio  $65 - 1,2x^*$  e desvio padrão 8. *Grosso modo*, na população que compreende todos os pontos  $(x, y)$ , a magnitude de um desvio comum da reta de regressão real é aproximadamente 8. Para  $x = 20$ , o valor médio de  $Y$  é  $\mu_{Y,20} = 65 - 1,2(20) = 41$ , de modo que

$$P(Y > 50 \text{ quando } x = 20) = P\left(Z > \frac{50 - 41}{8}\right) = 1 - \Phi(1,13) = 0,1292$$

A probabilidade de que o tempo até a falha ultrapasse 50 quando a força aplicada for 25 é, uma vez que  $\mu_{Y,25} = 35$ ,

$$P(Y > 50 \text{ quando } x = 25) = P\left(Z > \frac{50 - 35}{8}\right) = 1 - \Phi(1,88) = 0,0301$$

Essas probabilidades são representadas pelas áreas sombreadas na Figura 12.5.

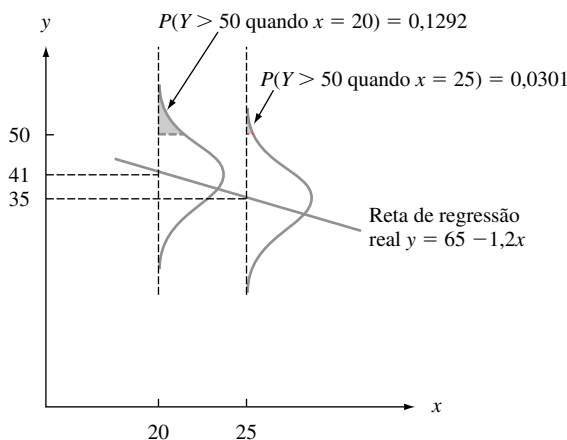


Figura 12.5 Probabilidades baseadas no modelo de regressão linear simples

Suponha que  $Y_1$  represente uma observação do tempo até a falha feita com  $x = 25$  e  $Y_2$  represente uma observação independente feita com  $x = 24$ . Logo,  $Y_1 - Y_2$  tem distribuição normal com valor médio  $E(Y_1 - Y_2) = \beta_1 = -1,2$ , variância  $V(Y_1 - Y_2) = \sigma^2 + \sigma^2 = 128$  e desvio padrão  $\sqrt{128} = 11,314$ . A probabilidade de  $Y_1$  ultrapassar  $Y_2$  é

$$P(Y_1 - Y_2 > 0) = P\left(Z > \frac{0 - (-1,2)}{11,314}\right) = P(Z > 0,11) = 0,4562$$

Ou seja, mesmo que nossa expectativa fosse de que  $Y$  diminuísse quando  $x$  aumentasse em 1 unidade, não é improvável que o  $Y$  observado em  $x + 1$  seja maior que o  $Y$  observado em  $x$ . ■

**Exercícios** | Seção 12.1 (1–11)

1. A taxa de eficiência de uma amostra de aço imersa em um tanque de fosfatação é o peso do revestimento de fosfato dividido pela perda do metal (ambos em mg/pés<sup>2</sup>). O artigo “Statistical Process Control of a Phosphate Coating Line” (*Wire J. Intl.*, maio de 1997, p. 78–81) forneceu os dados a seguir sobre a temperatura do tanque ( $x$ ) e a taxa de eficiência ( $y$ ).

<b>Temp.</b>	170	172	173	174	174	175	176
<b>Taxa</b>	0,84	1,31	1,42	1,03	1,07	1,08	1,04
<b>Temp.</b>	177	180	180	180	180	180	181
<b>Taxa</b>	1,80	1,45	1,60	1,61	2,13	2,15	0,84
<b>Temp.</b>	181	182	182	182	182	184	184
<b>Taxa</b>	1,43	0,90	1,81	1,94	2,68	1,49	2,52
<b>Temp.</b>	185	186	188				
<b>Taxa</b>	3,00	1,87	3,08				

- a. Elabore um diagrama de caule e folha da temperatura e da taxa de eficiência; comente os aspectos de interesse.
- b. O valor da taxa de eficiência é determinado completa e exclusivamente pela temperatura do tanque? Explique seu raciocínio.
- c. Elabore um gráfico de dispersão dos dados. Há algum indício de que a taxa de eficiência pode ser prevista de maneira altamente satisfatória pelo valor da temperatura? Explique seu raciocínio.

2. O artigo “Exhaust Emissions from Four-Stroke Lawn Mower Engines” (*J. of the Air and Water Mgmt. Assoc.*, 1997, p. 945–952) apresentou dados de um estudo no qual foram usadas uma mistura de gasolina com baseline e uma gasolina reformulada. Considere as observações a seguir, sobre o período (ano) e a emissão de  $\text{NO}_x$  (g/kWh):

<b>Motor</b>	1	2	3	4	5
<b>Período</b>	0	0	2	11	7
<b>Baseline</b>	1,72	4,38	4,06	1,26	5,31
<b>Reformulada</b>	1,88	5,93	5,54	2,67	6,53
<b>Motor</b>	6	7	8	9	10
<b>Período</b>	16	9	0	12	4
<b>Baseline</b>	0,57	3,37	3,44	0,74	1,24
<b>Reformulada</b>	0,74	4,94	4,89	0,69	1,42

Elabore gráficos de dispersão das emissões de  $\text{NO}_x$  versus período. Qual parece ser a característica da relação entre essas duas variáveis (*Observação*: os autores do artigo citado fizeram comentários sobre a relação.)

3. Os dados bivariados freqüentemente provêm do uso de duas diferentes técnicas para medir uma mesma quantidade. Como exemplo, as observações a seguir sobre  $x$  = concentração de hidrogênio (ppm) usando um método cromatográfico de gás e  $y$  = concentração, usando um novo método sensor, foram lidas de um gráfico no artigo “A New Method to Measure the Diffusible Hydrogen Content in Steel Weldments Using a Polymer Electrolyte-Based Hydrogen Sensor” (*Welding Res.*, julho de 1997, p. 251s–256s).

$x$	47	62	65	70	70	78	95	100	114	118
$y$	38	62	53	67	84	79	93	106	117	116
$x$	124	127	140	140	140	150	152	164	198	221
$y$	127	114	134	139	142	170	149	154	200	215

Elabore um gráfico de dispersão. Parece haver uma relação extremamente sólida entre os dois tipos de medida de concentração? Os dois métodos parecem estar medindo aproximadamente a mesma quantidade? Explique seu raciocínio.

4. Um estudo para avaliar a capacidade dos sistemas de zonas alagadiças (*wetlands*) de fluxo de subsuperfície, para remover a demanda bioquímica de oxigênio (DBO) e vários outros componentes químicos produziu os dados a seguir sobre  $x$  = aumento do volume de DBO (kg/ha/d) e  $y$  = remoção do volume de DBO (kg/ha/d) (“Subsurface Flow Wetlands — A Performance Evaluation”, *Water Envir. Res.*, 1995, p. 244–247).

$x$	3	8	10	11	13	16	27	30	35	37	38	44	103	142
$y$	4	7	8	8	10	11	16	26	21	9	31	30	75	90

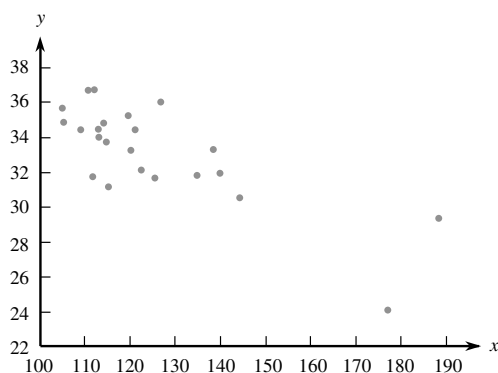
- a. Elabore um *boxplot* tanto para o aumento do volume como para sua remoção e comente qualquer característica interessante.

b. Elabore um gráfico de dispersão dos dados e comente qualquer característica interessante.

5. O artigo “Objective Measurement of the Stretchability of Mozzarella Cheese” (*J. of Texture Studies*, 1992, p. 185-194) descreveu um experimento realizado para investigar como o comportamento do queijo mussarela variou com a temperatura. Considere os dados a seguir sobre  $x$  = temperatura e  $y$  = elasticidade (%) na ruptura do queijo. (Observação: os pesquisadores eram italianos e usaram um queijo mussarela autêntico, não o “primo pobre” que pode ser obtido em quase qualquer parte dos Estados Unidos.)

$x$	59	63	68	72	74	78	83
$y$	118	182	247	208	197	135	132

- a. Elabore um gráfico de dispersão em que os eixos se cruzam em  $(0, 0)$ . Marque 0, 20, 40, 60, 80 e 100 no eixo horizontal e 0, 50, 100, 150, 200 e 250 no eixo vertical.
- b. Elabore um gráfico de dispersão em que os eixos se cruzam em  $(55, 100)$ , como foi feito no artigo citado. Será que esse gráfico parece preferível ao da parte (a)? Explique seu raciocínio.
- c. O que os gráficos das partes (a) e (b) sugerem sobre a natureza da relação entre as duas variáveis?
6. Um dos fatores do desenvolvimento do “cotovelo de tenista” (*tennis elbow*), doença que amedronta todos os tenistas praticantes, é a vibração induzida pelo impacto da raquete e do braço no contato com a bola. É de conhecimento geral que a probabilidade de o esportista ser acometido por essa doença depende de várias propriedades no uso da raquete. Considere o gráfico de dispersão de  $x$  = frequência de ressonância da raquete (Hz) e  $y$  = soma da aceleração pico a pico (característica da vibração do braço, em m/s/s) para  $n = 23$  raquetes diferentes (“Transfer of Tennis Racket Vibrations into the Human Forearm”, *Medicine and Science in Sports and Exercise*, 1992, p. 1134–1140). Comente as características interessantes dos dados e do gráfico de dispersão.



7. O artigo "Some Field Experience in the Use of an Accelerated Method in Estimating 28-Day Strength of Concrete" (*J. Amer. Concrete Institute*, 1969, p. 895)

experimentou regredir a resistência ao tratamento padrão aos  $y = 28$  dias em relação a  $x =$  resistência acelerada (psi). Suponha que a equação da reta de regressão real seja  $y = 1800 + 1,3x$ .

- a. Qual é o valor esperado da resistência aos 28 dias quando a resistência acelerada = 2500?
- b. Até que ponto podemos esperar que a resistência aos 28 dias mude quando a resistência acelerada é aumentada em 1 psi?
- c. Responda à parte (b) para um aumento de 100 psi.
- d. Responda à parte (b) para uma diminuição de 100 psi.
8. Em relação ao Exercício 7, suponha que o desvio padrão do desvio aleatório  $\epsilon$  seja 350 psi.
- a. Qual é a probabilidade de o valor observado de resistência aos 28 dias ultrapassar 5000 psi, quando o valor da resistência acelerada for 2000?
- b. Repita a parte (a), substituindo 2000 por 2500.
- c. Experimente fazer duas observações independentes sobre a resistência aos 28 dias, a primeira para uma resistência acelerada de 2000 e a segunda para  $x = 2500$ . Qual é a probabilidade de a segunda observação ultrapassar a primeira em mais de 1000 psi?
- d. Sejam  $Y_1$  e  $Y_2$  as observações sobre a resistência aos 28 dias quando  $x = x_1$  e  $x = x_2$ , respectivamente. Quanto  $x_2$  tem que exceder  $x_1$  para que  $P(Y_2 > Y_1) = 0,95$ ?
9. A taxa de fluxo  $y$  ( $m^2/min.$ ) em um dispositivo usado para avaliar a qualidade do ar depende da queda de pressão  $x$  (polegadas de água) por meio do filtro do dispositivo. Suponha que, para  $x$  valores entre 5 e 20, as duas variáveis sejam relacionadas de acordo com o modelo de regressão linear simples com uma reta de regressão real  $y = -0,12 + 0,095x$ .
- a. Qual é a mudança esperada na taxa de fluxo associada com um aumento de 1 pol. na queda da pressão? Explique.
- b. Que mudança na taxa de fluxo pode ser esperada quando a queda da pressão sofre uma diminuição de 5 pol.?
- c. Qual é a taxa de fluxo esperada para uma queda de pressão de 10 pol.? E uma queda de 15 pol.?
- d. Suponha  $\sigma = 0,025$  e considere uma queda de pressão de 10 pol. Qual é a probabilidade de o valor observado da taxa de fluxo ultrapassar 0,835? E de que a taxa de fluxo ultrapasse 0,840?
- e. Qual é a probabilidade de uma observação sobre a taxa de fluxo, com uma queda de pressão de 10 pol., ultrapassar uma observação sobre a taxa de fluxo feita quando a queda de pressão for 11 pol.?
10. Suponha que o custo esperado de uma operação de produção esteja relacionado com o tamanho da operação por meio da equação  $y = 4000 + 10x$ . Seja  $Y$  uma observação sobre o custo de uma operação. Se as variáveis *tamanho* e *custo* estiverem relacionadas de acordo com o modelo de regressão linear simples, seria possível que  $P(Y > 5500 \text{ quando } x = 100) = 0,05$  e  $P(Y > 6500 \text{ quando } x = 200) = 0,10$ ? Explique.

11. Suponha que, em um determinado processo químico, o tempo de reação  $y$  (h) esteja relacionado com a temperatura ( $^{\circ}\text{F}$ ) na câmara em que a reação ocorre, de acordo com o modelo de regressão linear simples, com a equação  $y = 5,00 - 0,01x$  e  $\sigma = 0,075$ .
- Qual é a mudança esperada no tempo de reação para um aumento de temperatura de  $1^{\circ}\text{F}$ ? Para um aumento de temperatura de  $10^{\circ}\text{F}$ ?
  - Qual é o tempo de reação esperado quando a temperatura é  $200^{\circ}\text{F}$ ? Quando a temperatura é  $250^{\circ}\text{F}$ ?
- Suponha que cinco observações sejam feitas independentemente do tempo de reação, cada uma para uma temperatura de  $250^{\circ}\text{F}$ . Qual é a probabilidade de os cinco tempos estarem entre 2,4 e 2,6 h?
  - Qual é a probabilidade de que dois tempos de reação observados independentemente para temperaturas com  $1^{\circ}$  de diferença sejam tais que o tempo na temperatura mais alta ultrapasse o tempo na temperatura mais baixa?

## 12.2 | Estimando Parâmetros de Modelo

Nesta e em várias outras seções a seguir, assumiremos que as variáveis  $x$  e  $y$  estão relacionadas de acordo com o modelo de regressão linear simples. Os valores de  $\beta_0$ ,  $\beta_1$  e  $\sigma^2$  praticamente nunca serão conhecidos pelo investigador. Em vez disso, serão fornecidos dados amostrais, compreendendo  $n$  pares observados  $(x_1, y_1)$ , ...,  $(x_n, y_n)$ , com base nos quais os parâmetros de modelo e a própria reta de regressão real podem ser estimados. Supõe-se que essas observações tenham sido obtidas independentemente uma da outra. Ou seja,  $y_i$  é o valor observado de uma  $Y_i$ , onde  $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$  e os  $n$  desvios  $\epsilon_1, \epsilon_2, \dots, \epsilon_n$  são  $n$   $v$ as independentes. A independência de  $Y_1, Y_2, \dots, Y_n$  depende da independência dos  $\epsilon_i$ s.

De acordo com o modelo, os pontos observados serão distribuídos nas imediações da reta de regressão real de maneira aleatória. A Figura 12.6 mostra um gráfico típico de pares observados com duas candidatas à reta de regressão estimada,  $y = a_0 + a_1 x$  e  $y = b_0 + b_1 x$ . Intuitivamente, a reta  $y = a_0 + a_1 x$  não é uma estimativa razoável da reta real  $y = \beta_0 + \beta_1 x$  porque, se  $y = a_0 + a_1 x$  fosse a reta real, os pontos observados quase certamente ficariam mais próximos dessa reta. A reta  $y = b_0 + b_1 x$  é uma estimativa mais plausível porque, em torno dessa reta, os pontos observados estão dispersos, e não próximos.

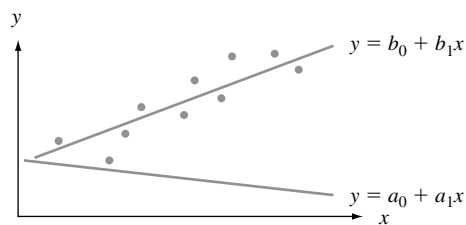


Figura 12.6 Duas diferentes estimativas da reta de regressão real

A Figura 12.6 e a discussão precedente sugerem que nossa estimativa de  $y = \beta_0 + \beta_1 x$  deve ser uma reta que, em certo sentido, ofereça a melhor aderência aos pontos dos dados observados. Isso é o que induz o princípio dos mínimos quadrados, que remonta ao matemático alemão Gauss (1777–1855). De acordo com esse princípio, uma reta oferece uma boa aderência aos dados, se as distâncias verticais (desvios) dos pontos observados em relação à reta são pequenos (veja a Figura 12.7). A medida da aderência é a soma dos quadrados desses desvios. A reta de melhor aderência é, portanto, aquela que tem a menor soma possível de desvios quadrados.

### Princípio dos Mínimos Quadrados

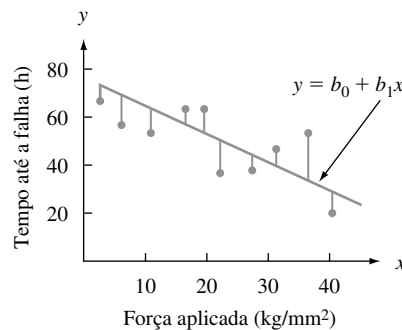
O desvio vertical do ponto  $(x_i, y_i)$  da reta  $y = b_0 + b_1 x$  é

$$\text{altura do ponto} - \text{altura da reta} = y_i - (b_0 + b_1 x_i)$$

A soma dos desvios quadrados verticais dos pontos  $(x_1, y_1), \dots, (x_n, y_n)$  à reta é, portanto,

$$f(b_0, b_1) = \sum_{i=1}^n [y_i - (b_0 + b_1 x_i)]^2$$

As estimativas pontuais de  $\beta_0$  e  $\beta_1$ , representadas por  $\hat{\beta}_0$  e  $\hat{\beta}_1$  e denominadas **estimativas dos mínimos quadrados**, são aqueles valores que minimizam  $f(b_0, b_1)$ . Ou seja,  $\hat{\beta}_0$  e  $\hat{\beta}_1$  são tais que  $f(\hat{\beta}_0, \hat{\beta}_1) \leq f(b_0, b_1)$  para qualquer  $b_0$  e  $b_1$ . A **reta de regressão estimada** ou a **reta dos mínimos quadrados** é, portanto, a reta cuja equação é  $y = \hat{\beta}_0 + \hat{\beta}_1 x$ .



**Figura 12.7** Desvios dos dados observados da reta  $y = b_0 + b_1 x$

Os valores de minimização de  $b_0$  e  $b_1$  são identificados tomando-se derivadas parciais de  $f(b_0, b_1)$  em relação a  $b_0$  e  $b_1$ , igualando-as a zero [analogamente a  $f'(b) = 0$  em cálculo univariado] e resolvendo as equações

$$\frac{\partial f(b_0, b_1)}{\partial b_0} = \sum 2(y_i - b_0 - b_1 x_i)(-1) = 0$$

$$\frac{\partial f(b_0, b_1)}{\partial b_1} = \sum 2(y_i - b_0 - b_1 x_i)(-x_i) = 0$$

Cancelando-se o fator  $-2$  e reorganizando as equações, obtemos o sistema a seguir, denominado **equações normais**:

$$\begin{aligned} nb_0 + (\sum x_i)b_1 &= \sum y_i \\ (\sum x_i)b_0 + (\sum x_i^2)b_1 &= \sum x_i y_i \end{aligned}$$

As equações normais são lineares nas duas incógnitas  $b_0$  e  $b_1$ . Contanto que pelo menos dois dos  $x_i$ s sejam diferentes, as estimativas dos mínimos quadrados são a única solução para esse sistema.

A estimativa dos mínimos quadrados do coeficiente angular  $\beta_1$  da reta de regressão real é

$$b_1 = \hat{\beta}_1 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = \frac{S_{xy}}{S_{xx}} \quad (12.2)$$

As fórmulas de cálculo do numerador e denominador de  $\hat{\beta}_1$  são

$$S_{xy} = \sum x_i y_i - (\sum x_i)(\sum y_i)/n \quad S_{xx} = \sum x_i^2 - (\sum x_i)^2/n$$

A estimativa dos mínimos quadrados do termo constante  $\beta_0$  da reta de regressão real é

$$b_0 = \hat{\beta}_0 = \frac{\sum y_i - \hat{\beta}_1 \sum x_i}{n} = \bar{y} - \hat{\beta}_1 \bar{x} \quad (12.3)$$



As fórmulas de cálculo de  $S_{xy}$  e  $S_{xx}$  exigem apenas as estatísticas  $\sum x_i, \sum y_i, \sum x_i^2, \sum x_i y_i$  ( $\sum y_i^2$  será necessário em breve) e que se minimizem os efeitos do arredondamento. Ao calcular  $\hat{\beta}_0$ , use números suplementares em  $\hat{\beta}_1$  porque, se a magnitude de  $\bar{x}$  for grande, o arredondamento influenciará a resposta final. Enfatizamos que, *antes de  $\beta_1$  e  $\hat{\beta}_0$  serem calculados, deve-se examinar um gráfico de dispersão para verificar se um modelo probabilístico linear é plausível*. Se os pontos não tenderem a se agrupar ao redor de uma reta com aproximadamente o mesmo grau de dispersão para todo  $x$ , devem ser investigados outros modelos. Na prática, os gráficos e os cálculos de regressão normalmente são feitos usando-se um pacote de *software* estatístico.

#### Exemplo 12.4

O concreto sem finos, fabricado com um agregado rústico nivelado de maneira uniforme e uma pasta de cimento-água, é benéfico em áreas propensas a muita chuva por causa de suas excelentes propriedades de drenagem. O artigo “Pavement Thickness Design for No-Fines Concrete Parking Lots” (*J. of Transportation Engr.*, 1995, p. 476–484) empregou uma análise de mínimos quadrados ao estudar como  $y$  = porosidade (%) está relacionada com  $x$  = peso unitário (pcf) em amostras de concreto. Considere os dados representativos a seguir, exibidos em um formato tabular conveniente para calcular os valores das estatísticas:

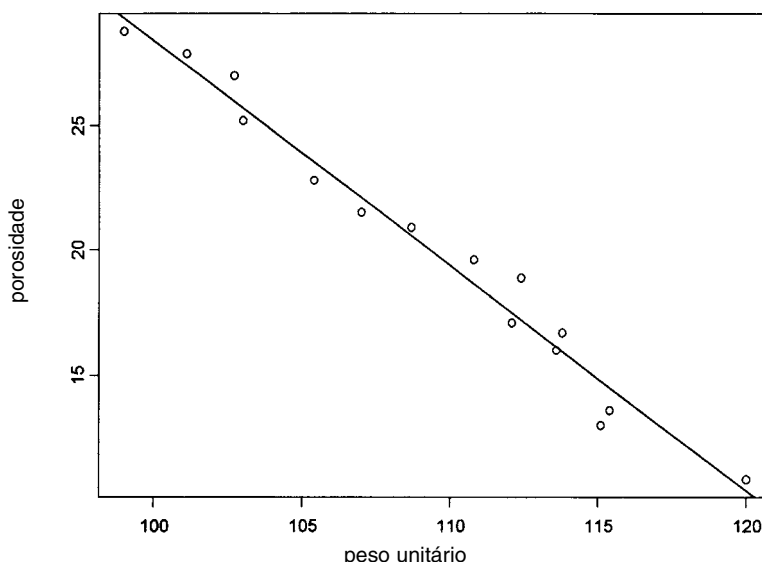
Obs	$x$	$y$	$x^2$	$xy$	$y^2$
1	99,0	28,8	9.801,00	2.851,20	829,44
2	101,1	27,9	10.221,21	2.820,69	778,41
3	102,7	27,0	10.547,29	2.772,90	729,00
4	103,0	25,2	10.609,00	2.595,60	635,04
5	105,4	22,8	11.109,16	2.403,12	519,84
6	107,0	21,5	11.449,00	2.300,50	462,25
7	108,7	20,9	11.815,69	2.271,83	436,81
8	110,8	19,6	12.276,64	2.171,68	384,16
9	112,1	17,1	12.566,41	1.916,91	292,41
10	112,4	18,9	12.633,76	2.124,36	357,21
11	113,6	16,0	12.904,96	1.817,60	256,00
12	113,8	16,7	12.950,44	1.900,46	278,89
13	115,1	13,0	13.248,01	1.496,30	169,00
14	115,4	13,6	13.317,16	1.569,44	184,96
15	120,0	10,8	14.400,00	1.296,00	116,64
Soma	1640,1	299,8	179.849,73	32.308,59	6.430,06

Portanto  $\bar{x} = 109,34$ ,  $\bar{y} = 19,986.667$ , e

$$\begin{aligned}\hat{\beta}_1 &= \frac{S_{xy}}{S_{xx}} = \frac{32.308,59 - (1640,1)(299,8)/15}{179.849,73 - (1640,1)^2/15} \\ &= \frac{-471,542}{521,196} = -0,90473066 \approx -0,905\end{aligned}$$

$$\hat{\beta}_0 = 19,986667 - (-0,90473066)(109,34) = 118,909917 \approx 118,91$$

De acordo com nossa estimativa, a mudança esperada na porosidade associada com um aumento de 1 pcf no peso unitário é  $-0,905\%$  (um decréscimo de  $0,905\%$ ). A equação da reta de regressão estimada (reta dos mínimos quadrados) é, portanto,  $y = 118,91 - 0,905x$ . A Figura 12.8, gerada pelo pacote de *software* estatístico S-Plus, mostra que a reta dos mínimos quadrados oferece uma excelente aproximação da relação entre as duas variáveis.



**Figura 12.8** Gráfico de dispersão dos dados do Exemplo 12.4 com a reta dos mínimos quadrados sobreposta, feito do S-Plus

A reta de regressão estimada pode ser usada imediatamente para dois diferentes propósitos. Para um valor fixo  $x^*$ ,  $\hat{\beta}_0 + \hat{\beta}_1 x^*$  (a altura da reta acima de  $x^*$ ) fornece tanto (1) uma estimativa pontual do valor esperado de  $Y$  quando  $x = x^*$  como (2) uma previsão pontual do valor  $Y$  que resultará de uma única nova observação feita em  $x = x^*$ .

**Exemplo 12.5**

Refira-se aos dados sobre peso unitário–porosidade no exemplo anterior. Uma estimativa pontual da porosidade média real de todas as amostras cujo peso unitário é 110 é

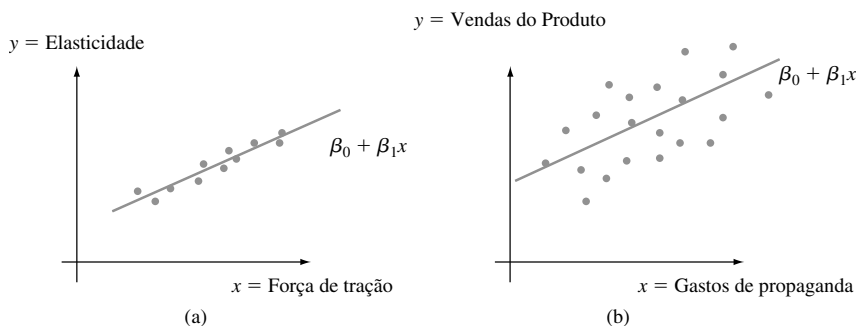
$$\hat{\mu}_{y \cdot 110} = \hat{\beta}_0 + \hat{\beta}_1(110) = 118,91 - 0,905(110) = 19,4\%$$

Se for selecionada uma única amostra cujo peso unitário é 110 pcf, 19,4% será também uma previsão pontual da porosidade dessa amostra.

A reta dos mínimos quadrados não deve ser usada para estimar um valor  $x$  que esteja muito além da amplitude dos dados, como  $x = 90$  ou  $x = 135$  no Exemplo 12.4. O risco de extrapolação é o de que a relação ajustada (aqui, uma reta) possa não ser válida para tais valores  $x$ . (No exemplo precedente,  $x = 135$  dá  $\hat{y} = -3,3$ , um valor nitidamente ridículo de porosidade, mas a extrapolação nem sempre resultará nessas inconsistências.)

**Estimando  $\sigma^2$  e  $\sigma$**

O parâmetro  $\sigma^2$  determina a variabilidade inerente no modelo de regressão. Um valor grande de  $\sigma^2$  fará com que os  $(x_i, y_i)$ s observados fiquem muito dispersos em relação à reta de regressão real, ao passo que, quando  $\sigma^2$  for pequeno, os pontos observados tenderão a ficar bem próximos da reta real (veja a Figura 12.9).



**Figura 12.9** Amostra típica de  $\sigma^2$ : (a) pequena; (b) grande

Uma estimativa de  $\sigma^2$  será usada nas fórmulas do Intervalo de Confiança (IC) e procedimentos de teste de hipótese que serão apresentados nas duas seções seguintes. Pelo fato de não se conhecer a equação da reta real, a estimativa se baseia em até que ponto as observações amostrais se desviam da reta estimada. Muitos desvios grandes (residuais) sugerem um valor grande de  $\sigma^2$ , ao passo que todos os desvios de pequena magnitude sugerem que  $\sigma^2$  é pequeno.

## DEFINIÇÃO

Os **valores ajustados** (ou **previstos**)  $\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n$  são obtidos substituindo-se sucessivamente  $x_1, \dots, x_n$  na equação da reta de regressão estimada:  $\hat{y}_1 = \hat{\beta}_0 + \hat{\beta}_1 x_1, \hat{y}_2 = \hat{\beta}_0 + \hat{\beta}_1 x_2, \dots, \hat{y}_n = \hat{\beta}_0 + \hat{\beta}_1 x_n$ . Os **resíduos** são os desvios verticais  $y_1 - \hat{y}_1, y_2 - \hat{y}_2, \dots, y_n - \hat{y}_n$  da reta estimada.

Traduzindo em palavras, o valor previsto  $\hat{y}_i$  é o valor de  $y$  que previmos ou esperávamos ao usar a reta de regressão estimada com  $x = x_i$ ;  $\hat{y}_i$  é a altura da reta de regressão estimada sobre o valor  $x_i$  para o qual a  $i$ -ésima observação foi feita. O resíduo  $y_i - \hat{y}_i$  é a diferença entre  $y_i$  observado e  $\hat{y}_i$  previsto. Se os resíduos forem todos pequenos em magnitude, a maior parte da variabilidade nos valores  $y$  observados parece decorrer da relação linear entre  $x$  e  $y$ , ao passo que muitos resíduos grandes sugerem grande variabilidade inerente em  $y$ , em relação à quantidade decorrente da relação linear. Supondo que a reta na Figura 12.7 seja a reta de mínimos quadrados, os resíduos são identificados pelos segmentos de linha verticais a partir dos pontos observados até a reta. Quando a reta de regressão estimada é obtida por meio do princípio dos mínimos quadrados, teoricamente a soma dos resíduos deve ser zero. Na prática, a soma pode desviar um pouco de zero em decorrência do arredondamento.

## Exemplo 12.6

A alta densidade populacional do Japão provocou milhares de problemas relacionados ao uso de seus recursos. Uma dificuldade especialmente importante diz respeito à remoção de lixo. O artigo “Innovative Sludge Handling Through Pelletization Thickening” (*Water Research*, 1999, p. 3245–3252) descreveu o desenvolvimento de uma nova máquina de compressão para processamento do lodo de esgoto. Uma parte importante da investigação relacionou o controle da umidade de grânulos comprimidos ( $y$ , em %) com a taxa de filtragem da máquina ( $x$ , em kg–DS/m/h). Os dados a seguir foram reproduzidos de um gráfico desse artigo:

$x$	125,3	98,2	201,4	147,3	145,9	124,7	112,2	120,2	161,2	178,9
$y$	77,9	76,8	81,5	79,8	78,2	78,3	77,5	77,0	80,1	80,2
$x$	159,5	145,8	75,1	151,4	144,2	125,0	198,8	132,5	159,6	110,7
$y$	79,9	79,0	76,7	78,2	79,5	78,1	81,5	77,0	79,0	78,6

As estatísticas pertinentes são:  $\sum x_i = 2817,9$ ,  $\sum y_i = 1574,8$ ,  $\sum x_i^2 = 415.949,85$ ,  $\sum x_i y_i = 222.657,88$ , e  $\sum y_i^2 = 124.039,58$ , pelas quais  $\bar{x} = 140,895$ ,  $\bar{y} = 78,74$ ,  $S_{xx} = 18.921,8295$ , e  $S_{xy} = 776,434$ . Portanto,

$$\hat{\beta}_1 = \frac{776,434}{18.921,8295} = 0,04103377 \approx 0,041$$

$$\hat{\beta}_0 = 78,74 - (0,04103377)(140,895) = 72,958547 \approx 72,96$$

com base nas quais a equação da reta dos mínimos quadrados é  $\hat{y} = 72,96 + 0,041x$ . Para obter precisão numérica, os valores ajustados são calculados pela equação  $\hat{y}_i = 72,958547 + 0,04103377x_i$ :

$$\hat{y}_1 = 72,958547 + 0,04103377(125,3) \approx 78,100, \quad y_1 - \hat{y}_1 \approx -0,200 \text{ etc.}$$

Um resíduo positivo corresponde a um ponto no gráfico de dispersão disposto acima do gráfico da reta dos mínimos quadrados, ao passo que um resíduo negativo resulta de um ponto disposto abaixo da reta. Todos os valores previstos (ajustes) e resíduos são exibidos na tabela a seguir.

Obs	Líquido Filtrado	Concentração de Umidade	Ajuste	Resíduo
1	125,3	77,9	78,100	-0,200
2	98,2	76,8	76,988	-0,188
3	201,4	81,5	81,223	0,277
4	147,3	79,8	79,003	0,797
5	145,9	78,2	78,945	-0,745
6	124,7	78,3	78,075	0,225
7	112,2	77,5	77,563	-0,063
8	120,2	77,0	77,891	-0,891
9	161,2	80,1	79,573	0,527
10	178,9	80,2	80,299	-0,099
11	159,5	79,9	79,503	0,397
12	145,8	79,0	78,941	0,059
13	75,1	76,7	76,040	0,660
14	151,4	78,2	79,171	-0,971
15	144,2	79,5	78,876	0,624
16	125,0	78,1	78,088	0,012
17	198,8	81,5	81,116	0,384
18	132,5	77,0	78,396	-1,396
19	159,6	79,0	79,508	-0,508
20	110,7	78,6	77,501	1,099

Da mesma forma que os desvios da média no caso de uma única amostra foram somados para obter a estimativa  $s^2 = \sum(x_i - \bar{x})^2/(n - 1)$ , a estimativa de  $\sigma^2$  na análise de regressão tem por base elevar ao quadrado e somar os resíduos. Continuaremos a usar o símbolo  $s^2$  para essa variância estimada. Portanto, não o confunda com o  $s^2$  previamente usado.

**DEFINIÇÃO**

A **soma dos quadrados dos erros** (que equivale à soma dos quadrados dos resíduos), representada por SQE, é

$$\text{SQE} = \sum(y_i - \hat{y}_i)^2 = \sum [y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)]^2$$

e a estimativa de  $\sigma^2$  é

$$\hat{\sigma}^2 = s^2 = \frac{\text{SQE}}{n - 2} = \frac{\sum(y_i - \hat{y}_i)^2}{n - 2}$$

O divisor  $n - 2$  em  $s^2$  é o número de graus de liberdade (gl) associados com a estimativa (ou, de modo equivalente, com a soma dos quadrados dos erros). Isso porque, para obter  $s^2$ , os dois parâmetros  $\beta_0$  e  $\beta_1$  devem ser estimados primeiro, o que resulta em uma perda de 2 gl (assim como  $\mu$  teve de ser estimado nos problemas de uma única amostra, gerando uma variância estimada com base em  $n - 1$  gl). Substituindo cada  $y_i$  pela  $Y$  na fórmula para  $s^2$ , obtém-se o estimador  $S^2$ . É possível mostrar que  $S^2$  é um estimador não-tendencioso de  $\sigma^2$  (embora o estimador  $S$  seja tendencioso para  $\sigma$ ).

**Exemplo 12.7** (Continuação do Exemplo 12.6)

Os resíduos dos dados sobre a taxa de filtragem–quantidade de umidade foram calculados previamente. A soma dos quadrados dos erros correspondente é

$$\text{SQE} = (-0,200)^2 + (-0,188)^2 + \dots + (1,099)^2 = 7,968$$

A estimativa de  $\sigma^2$  é, portanto,  $\hat{\sigma}^2 = s^2 = 7,968/(20 - 2) = 0,4427$ , e o desvio padrão estimado é  $\hat{\sigma} = s = \sqrt{0,4427} = 0,665$ . *Grosso modo*, 0,665 é a magnitude de um desvio típico da reta de regressão estimada. ■

O cálculo de SQE com base na fórmula determinante requer muitos cálculos tediosos porque tanto os valores previstos como os resíduos devem primeiro ser calculados. O uso da fórmula de cálculo a seguir não requer essas quantidades.

$$\text{SQE} = \sum y_i^2 - \hat{\beta}_0 \sum y_i - \hat{\beta}_1 \sum x_i y_i$$

Essa expressão resulta da substituição de  $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$  em  $\sum (y_i - \hat{y}_i)^2$ , elevando ao quadrado o adendo, fazendo a soma aos três termos resultantes e simplificando. Essa fórmula de cálculo é especialmente sensível aos efeitos do arredondamento em  $\hat{\beta}_0$  e  $\hat{\beta}_1$ , de modo que usar o maior número de dígitos possível nos cálculos intermediários impedirá erros de arredondamento.

### Exemplo 12.8

O artigo “Promising Quantitative Nondestructive Evaluation Techniques for Composite Materials” (*Materials Evaluation*, 1985, p. 561–565) relata um estudo para investigar como a propagação de uma onda de tensão ultra-sônica por meio de uma substância depende das propriedades dela. Os dados a seguir sobre resistência à ruptura ( $x$ , como uma porcentagem da resistência à tração extrema) e atenuação ( $y$ , em neper/cm, o decréscimo de amplitude da onda de tensão) em compostos de poliéster reforçados por fibra de vidro foram reproduzidos de um gráfico apresentado nesse artigo. O modelo de regressão linear simples é sugerido pelo padrão linear significativo do gráfico de dispersão.

$x$	12	30	36	40	45	57	62	67	71	78	93	94	100	105
$y$	3,3	3,2	3,4	3,0	2,8	2,9	2,7	2,6	2,5	2,6	2,2	2,0	2,3	2,1

As quantidades resumidas necessárias são  $n = 14$ ,  $\sum x_i = 890$ ,  $\sum x_i^2 = 67.182$ ,  $\sum y_i = 37,6$ ,  $\sum y_i^2 = 103,54$  e  $\sum x_i y_i = 2234,30$ , das quais  $S_{xx} = 10.603,4285714$ ,  $S_{xy} = -155,98571429$ ,  $\hat{\beta}_1 = -0,0147109$  e  $\hat{\beta}_0 = 3,6209072$ . A fórmula de cálculo de SQE fornece

$$\text{SQE} = 103,54 - (3,6209072)(37,6) - (-0,0147109)(2234,30) = 0,2624532$$

de modo que  $s^2 = 0,2624532/12 = 0,0218711$  e  $s = 0,1479$ . Quando  $\hat{\beta}_0$  e  $\hat{\beta}_1$  são arredondados para três casas decimais na fórmula de cálculo de SQE, o resultado é

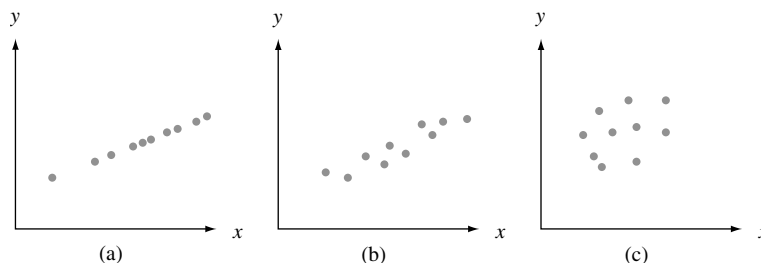
$$\text{SQE} = 103,54 - (3,621)(37,6) - (-0,015)(2234,30) = 0,905$$

que é mais de três vezes o valor correto. ■

## O Coeficiente de Determinação

A Figura 12.10 mostra três diferentes gráficos de dispersão de dados bivariados. Em todos eles, as alturas dos diferentes pontos variam substancialmente, indicando que há muita variabilidade nos valores  $y$  observados. Todos os pontos no primeiro gráfico dispõem-se exatamente em uma reta. Nesse caso, toda a variação amostral (100%) em  $y$  pode ser atribuída ao fato de  $x$  e  $y$  estarem relacionados linearmente. Os pontos na Figura 12.10(b) não se dispõem exatamente em uma reta, mas, em comparação com a variabilidade global de  $y$ , os desvios da

reta dos mínimos quadrados são pequenos. É razoável concluir, nesse caso, que grande parte da variação observada de  $y$  pode ser atribuída à relação linear aproximada entre as variáveis postuladas pelo modelo de regressão linear simples. Quando o gráfico de dispersão fica parecido com o da Figura 12.10(c), há uma variação significativa ao redor da reta dos mínimos quadrados em relação à variação global de  $y$ , de modo que o modelo de regressão linear simples não consegue explicar a variação em  $y$  relacionando-o a  $x$ .

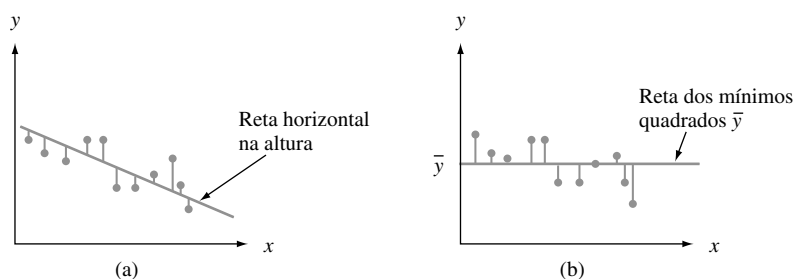


**Figura 12.10** Usando o modelo para explicar a variação de  $y$ : (a) dados para os quais toda a variação é justificada; (b) dados para os quais grande parte da variação é justificada; (c) dados para os quais pouca variação é justificada

A soma dos quadrados dos erros SQE pode ser interpretada como uma medida da quantidade de variação em  $y$  deixada inexplicada pelo modelo – ou seja, quanto não pode ser atribuído a uma relação linear. Na Figura 12.10(a),  $SQE = 0$ , e não há qualquer variação inexplicada, ao passo que a variação inexplicada é pequena para os dados da Figura 12.10(b) e muito maior na Figura 12.10(c). Uma medida quantitativa da quantidade total de variação nos valores observados de  $y$  é dada pela soma total dos quadrados.

$$SQT = S_{yy} = \sum (y_i - \bar{y})^2 = \sum y_i^2 - (\sum y_i)^2/n$$

A soma total dos quadrados é a soma dos desvios quadrados ao redor da média amostral dos valores observados de  $y$ . Portanto, o mesmo número  $\bar{y}$  é subtraído de cada  $y_i$  em SQT, ao passo que SQE envolve a subtração de cada diferente valor previsto  $\hat{y}_i$  do  $y_i$  observado correspondente. Assim como SQE é a soma dos desvios quadrados em torno da reta de mínimos quadrados  $y = \hat{\beta}_0 + \hat{\beta}_1 x$ , SQT é a soma dos desvios quadrados em torno da reta horizontal na altura  $\bar{y}$  (assim, os desvios verticais são  $y_i - \bar{y}$ ), como ilustrado na Figura 12.11. Além disso, pelo fato de a soma dos desvios quadrados em torno da reta dos mínimos quadrados ser menor do que a soma dos desvios quadrados em torno de qualquer outra reta,  $SQE < SQT$  a menos que a própria reta horizontal seja a reta dos mínimos quadrados. A razão  $SQE/SQT$  é a proporção da variação total que não pode ser explicada pelo modelo de regressão linear simples e  $1 - SQE/SQT$  (um número entre 0 e 1) é a proporção da variação  $y$  observada explicada pelo modelo.



**Figura 12.11** Somas dos quadrados ilustradas: (a) SQE = soma dos desvios quadrados em torno da reta dos mínimos quadrados; (b) SQT = soma dos quadrados total em torno da reta horizontal

**DEFINIÇÃO**

O **coeficiente de determinação**, representado por  $r^2$ , é dado por

$$r^2 = 1 - \frac{SQE}{SQT}$$

Esse coeficiente é interpretado como a proporção da variação de  $y$  observada que pode ser explicada pelo modelo de regressão linear simples (atribuída a uma relação linear aproximada entre  $y$  e  $x$ ).

Quanto mais alto o valor de  $r^2$ , mais o modelo de regressão linear simples consegue explicar a variação  $y$ . Quando a análise de regressão é feita por um pacote de *software* estatístico, tanto  $r^2$  quanto  $100r^2$  (a porcentagem de variação explicada pelo modelo) são uma parte importante do resultado. Se  $r^2$  for pequeno, em geral o analista vai querer procurar um modelo alternativo (um modelo não-linear ou um modelo de regressão múltipla com mais de uma variável independente) que possa explicar mais eficientemente a variação  $y$ .

**Exemplo 12.9** (Continuação do Exemplo 12.4)

O gráfico de dispersão dos dados sobre o concreto sem finos na Figura 12.8 certamente prognostica um valor  $r^2$  alto. Com

$$\hat{\beta}_0 = 118,909917 \quad \hat{\beta}_1 = -0,90473066 \quad \sum y_i = 299,8$$

$$\sum x_i y_i = 32.308,59 \quad \sum y_i^2 = 6430,06$$

temos

$$SQT = 6430,06 - \frac{299,8^2}{15} = 438,057333 \approx 438,06$$

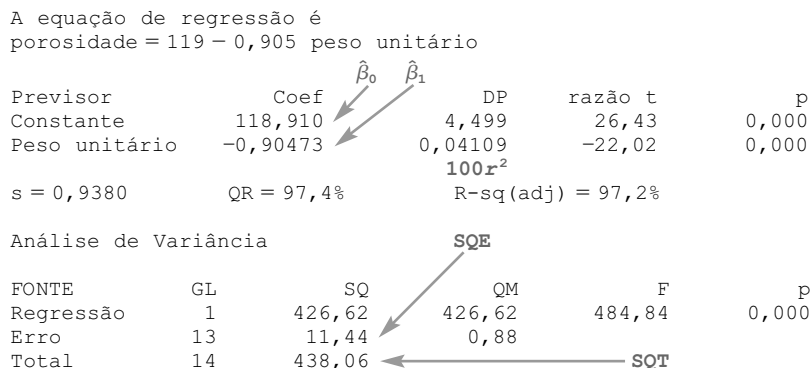
$$SQE = 6430,06 - (118,909917)(299,8) - (-0,90473066)(32.308,59) = 11,4388 \approx 11,44$$

O coeficiente de determinação, portanto,

$$r^2 = 1 - \frac{11,44}{438,06} = 1 - 0,026 = 0,974$$

Ou seja, 97,4% da variação observada na porosidade é atribuída à (pode ser explicada por) relação linear aproximada entre porosidade e peso unitário, um resultado notável. (Vários cientistas sociais morreriam por um valor  $r^2$  bem superior a 0,5!)

A Figura 12.12 mostra o resultado parcial do MINITAB para os dados sobre porosidade–peso unitário dos Exemplos 12.4 e 12.9; o *software* também fornecerá os valores previstos e os resíduos quando solicitados, bem como outras informações. Os formatos usados por outros *softwares* diferem ligeiramente do formato do MINITAB, mas o conteúdo das informações é bastante semelhante. Quantidades como desvios padrão, razões  $t$  e tabela ANOVA são discutidas na Seção 12.3.



**Figura 12.12** Resultado do MINITAB para a regressão dos Exemplos 12.4 e 12.9

O coeficiente de determinação pode ser escrito de maneira ligeiramente diferente, introduzindo-se uma terceira soma de quadrados: **a soma dos quadrados da regressão, SQR** – dada por  $SQR = SQT - SQE$ . A soma dos quadrados da regressão é interpretada como a quantidade de variação total explicada pelo modelo. Portanto, temos

$$r^2 = 1 - SQE/SQT = (SQT - SQE)/SQT = SQR/SQT$$

a razão entre a variação explicada pela variação total. A tabela ANOVA na Figura 12.12 mostra que  $SQR = 426,62$ , de onde  $r^2 = 426,62/438,06 = 0,974$ .

## Terminologia e Escopo da Análise de Regressão

O termo *análise de regressão* foi usado pela primeira vez por Francis Galton, no final do século XIX, relativamente à sua obra sobre a relação entre a altura do pai  $x$  e a altura do filho  $y$ . Depois de reunir inúmeros pares  $(x_i, y_i)$ , Galton usou o princípio dos mínimos quadrados para obter a equação da reta de regressão estimada, com o objetivo de usá-la para prever a altura do filho com base na altura do pai. Ao usar a reta derivada, Galton descobriu que, se a altura do pai estivesse acima da média, a altura do filho também apresentaria a probabilidade de estar acima da média, *mas não tanto quanto a do pai*. De modo semelhante, a altura do filho cujo pai tem altura abaixo da média também apresentava a probabilidade de estar abaixo da média, mas não tanto quanto a do pai. Portanto, a altura prevista de um filho sofreu um "retrocesso" em direção à média; pelo fato de regressão significar volta ou retorno, Galton adotou a terminologia *reta de regressão*. Esse fenômeno de sofrer um retrocesso em relação à média foi observado em várias outras situações (por exemplo, médias de rebatidas de ano para ano no beisebol) e é denominado **efeito de regressão**.

Nossa discussão até agora presumiu que a variável independente é controlada pelo pesquisador, de modo que apenas a variável dependente  $Y$  é aleatória. Entretanto, isso não ocorreu com o experimento de Galton: a altura dos pais não foi previamente selecionada. Em vez disso,  $X$  e  $Y$  eram aleatórios. Métodos e conclusões da análise de regressão podem ser aplicados tanto quando os valores da variável independente são fixos como quando são aleatórios, mas, pelo fato de as derivações e interpretações serem mais diretas no primeiro caso, continuaremos a trabalhar com ele sem reservas. Para obter mais explicações, veja o excelente livro de John Neter et al., relacionado na bibliografia deste capítulo.

### Exercícios | Seção 12.2 (12–29)

12. O Exercício 4 forneceu dados sobre  $x$  = o aumento do volume de DBO e  $y$  = remoção do volume de DBO. Os valores pertinentes são

$$\begin{array}{ll} n = 14 & \sum x_i = 517 \\ \sum y_i = 346 & \sum x_i^2 = 39,095 \\ \sum y_i^2 = 17,454 & \sum x_i y_i = 25,825 \end{array}$$

- Obtenha a equação da reta dos mínimos quadrados.
- Preveja o valor da remoção do volume de DBO para uma única observação, feita quando o aumento do volume de DBO é 35, e calcule o valor do resíduo correspondente.
- Calcule a SQE e, em seguida, uma estimativa pontual de  $\sigma$ .
- Que proporção da variação observada na remoção pode ser explicada pela relação linear aproximada entre as duas variáveis?
- Os dois últimos valores  $x$ , 103 e 142, são bem maiores que os outros. Como a equação da reta dos mínimos quadrados e o valor de  $r^2$  são afetados pela

exclusão das duas observações correspondentes da amostra? Ajuste os valores fornecidos e use o fato de que o novo valor de SQE é 311,79.

13. Os dados a seguir sobre  $x$  = densidade da corrente ( $\text{mA}/\text{cm}^2$ ) e  $y$  = taxa de deposição ( $(\mu\text{m})/\text{min.}$ ) foram tirados do artigo "Plating of 60/40 Tin/Lead Solder for Head Termination Metallurgy" (*Plating and Surface Finishing*, janeiro de 1997, p. 38–40). Você concorda com a afirmação do autor desse artigo de que "uma relação linear foi obtida a partir da taxa de deposição de estanho–chumbo como uma função da densidade da corrente"? Explique seu raciocínio.

$x$	20	40	60	80
$y$	0,24	1,20	1,71	2,22

14. Refira-se aos dados sobre temperatura do tanque–taxa de eficiência do Exercício 1.
- Determine a equação da reta de regressão estimada.
  - Calcule uma estimativa pontual da taxa de eficiência média real quando a temperatura do tanque é 182.



- c. Calcule os valores dos resíduos da reta de mínimos quadrados das quatro observações para as quais a temperatura é 182. Por que nem todas têm o mesmo sinal?
- d. Que proporção da variação observada na taxa de eficiência pode ser atribuída à relação de regressão linear simples entre as duas variáveis?

15. Os valores do módulo de elasticidade (MOE, a razão da força, isto é, força por área unitária, para o escoamento, ou seja, deformação por comprimento unitário, em GPa) e a resistência à flexão (uma medida da capacidade de resistência a falhas decorrentes de dobramentos, em MPa) foram determinados para um tipo de amostra de vigas de concreto, gerando os dados a seguir (reproduzidos de um gráfico do artigo “Effects of Aggregates and Microfillers on the Flexural Properties of Concrete”, *Magazine of Concrete Research*, 1997, p. 81–98):

MOE	29,8	33,2	33,7	35,3	35,5	36,1	36,2
Resistência	5,9	7,2	7,3	6,3	8,1	6,8	7,0
MOE	36,3	37,5	37,7	38,7	38,8	39,6	41,0
Resistência	7,6	6,8	6,5	7,0	6,3	7,9	9,0
MOE	42,8	42,8	43,5	45,6	46,0	46,9	48,0
Resistência	8,2	8,7	7,8	9,7	7,4	7,7	9,7
MOE	49,3	51,7	62,6	69,8	79,5	80,0	
Resistência	7,8	7,7	11,6	11,3	11,8	10,7	

- a. Elabore um diagrama caule e folha dos valores do MOE e comente qualquer característica interessante.
- b. O valor da resistência é determinado exclusiva e unicamente pelo valor do MOE? Explique.
- c. Use o resultado do MINITAB a seguir, a fim de obter a equação da reta dos mínimos quadrados para prever a resistência do módulo de elasticidade e, em seguida, prever a resistência de uma viga cujo módulo de elasticidade é 40. Você se sentiria à vontade usando a reta dos mínimos quadrados para prever a resistência se o módulo de elasticidade fosse 100? Explique.

Preditor	Coef	DP	razão t	p
Constante	3,2925	0,6008	5,48	0,000
mód. elast.	0,10748	0,01280	8,40	0,000

s = 0,8657 R-sq = 73,8% R-sq (ajuste) = 72,8%

Análise de Variância

FONTE	GL	SQ	QM	F	P
Regressão	1	52,870	52,870	70,55	0,000
Erro	25	18,736	0,749		
Total	26	71,605			

- d. Quais são os valores de SSE, SST e o coeficiente de determinação? Esses valores sugerem que o modelo de regressão linear simples descreve com eficiência a relação entre as duas variáveis? Explique.

16. O artigo “Characterization of Highway Runoff in Austin, Texas, Area” (*J. of Envir. Engr.*, 1998, p. 131–137) apresentou um gráfico de dispersão, com a reta dos mínimos

quadrados, de  $x$  = volume de chuva ( $m^3$ ) e  $y$  = volume de *runoff* [chuva não absorvida pelo solo] ( $m^3$ ) de um determinado lugar. Os valores a seguir foram reproduzidos do gráfico.

x	5	12	14	17	23	30	40	47
y	4	10	13	15	15	25	27	46
x	55	67	72	81	96	112	127	
y	38	46	53	70	82	99	100	

- a. Um gráfico de dispersão dos dados apóia o uso do modelo de regressão linear simples?
  - b. Calcule estimativas pontuais do coeficiente angular e do termo constante da reta de regressão populacional.
  - c. Calcule uma estimativa pontual do volume de *runoff* médio real quando o volume de chuva for 50.
  - d. Calcule uma estimativa pontual do desvio padrão  $\sigma$ .
  - e. Que proporção da variação observada no volume de *runoff* pode ser atribuída à relação de regressão linear simples entre o volume de *runoff* e o volume de chuva?
17. Uma regressão de  $y$  = volume de cálcio (g/L) em  $x$  = material dissolvido ( $mg/cm^2$ ) foi descrita no artigo “Use of Fly Ash or Silica Fume to Increase the Resistance of Concrete to Feed Acids” (*Magazine of Concrete Research*, 1997, p. 337–344). A equação da reta de regressão estimada era  $y = 3,678 + 0,144x$ , onde  $r^2 = 0,860$ , com base em  $n = 23$ .
- a. Interprete o coeficiente angular estimado 0,144 e o coeficiente de determinação 0,860.
  - b. Calcule uma estimativa pontual do volume médio real de cálcio quando a quantidade de material dissolvido for igual a 50  $mg/cm^2$ .
  - c. O valor total da soma dos quadrados era  $SQT = 320,398$ . Calcule uma estimativa do desvio padrão do erro  $\sigma$  no modelo de regressão linear simples.
18. As estatísticas resumidas a seguir foram obtidas de um estudo que usou a análise de regressão para investigar a relação entre a deflexão de pavimento e a temperatura da superfície do pavimento de vários lugares em uma rodovia estadual. Aqui,  $x$  = temperatura ( $^{\circ}F$ ) e  $y$  = fator de ajuste de deflexão ( $y \geq 0$ ):
- $$n = 15 \quad \sum x_i = 1425 \quad \sum y_i = 10,68$$
- $$\sum x_i^2 = 139.037,25 \quad \sum x_i y_i = 987,645$$
- $$\sum y_i^2 = 7,8518$$
- (Foram feitas bem mais de 15 observações nesse estudo; a referência é “Flexible Pavement Evaluation and Rehabilitation”, *Transportation Eng. J.*, 1977, p. 75–85.)
- a. Calcule  $\hat{\beta}_1$ ,  $\hat{\beta}_0$  e a equação da reta de regressão estimada. Represente em um gráfico a reta estimada.
  - b. Qual é a estimativa de mudança esperada no fator de ajuste de deflexão quando a temperatura sofre um aumento de 1  $^{\circ}F$ ?

- c. Suponha que a temperatura fosse medida em °C em vez de °F. Qual seria a reta de regressão estimada? Responda à parte (b) para um aumento de 1°C. (Sugestão:  $^{\circ}\text{F} = (9/5)^{\circ}\text{C} + 32$ ; substitua então o “x antigo” em função do “novo x”).
- d. Se uma temperatura de superfície de 200°F estivesse dentro do domínio das possibilidades, você usaria a reta estimada da parte (a) para prever o fator de deflexão para essa temperatura? Por quê?

19. Os dados a seguir são representativos dos dados relatados no artigo “An Experimental Correlation of Oxides of Nitrogen Emissions from Power Boilers Based on Field Data” (*J. Eng. for Power*, julho de 1973, p. 165–170), onde  $x$  = taxa de liberação na área do queimador (MBtu/h–pés<sup>2</sup>) e  $y$  = taxa de emissão de NO<sub>x</sub> (ppm):

$x$	100	125	125	150	150	200	200
$y$	150	140	180	210	190	320	280
$x$	250	250	300	300	350	400	400
$y$	400	430	440	390	600	610	670

- a. Supondo que o modelo de regressão linear simples seja válido, obtenha a estimativa dos mínimos quadrados da reta de regressão real.
- b. Qual é a estimativa da taxa de emissão de NO<sub>x</sub>, quando a taxa de liberação na área do queimador é igual a 225?
- c. Calcule a mudança esperada na taxa de emissão de NO<sub>x</sub> quando a taxa de liberação na área do queimador for diminuída em 50.
- d. Você usaria a reta de regressão estimada visando prever a taxa de emissão para uma taxa de liberação de 500? Por quê?

20. Inúmeros estudos já mostraram que os líquens (determinadas plantas compostas de algas e fungos) são excelentes bioindicadores de poluição do ar. O artigo “The Epiphytic Lichen Hypogymnia Physodes as a Biomonitor of Atmospheric Nitrogen and Sulphur Deposition in Norway” (*Envir. Monitoring and Assessment*, 1993, p. 27–47) apresenta os dados a seguir (reproduzidos de um gráfico) sobre  $x$  = deposição de umidade de NO<sub>3</sub><sup>-</sup> (g N/m<sup>2</sup>) e  $y$  = líquens N (% peso seco):

$x$	0,05	0,10	0,11	0,12	0,31	0,37	0,42
$y$	0,48	0,55	0,48	0,50	0,58	0,52	1,02
$x$	0,58	0,68	0,68	0,73	0,85	0,92	
$y$	0,86	0,86	1,00	0,88	1,04	1,70	

O autor usou regressão linear simples para analisar os dados. Use o resultado do MINITAB a seguir para responder às seguintes perguntas:

- a. Quais são as estimativas por mínimos quadrados de  $\beta_0$  e  $\beta_1$ ?

- b. Preveja os líquens N para um valor de deposição de NO<sub>3</sub><sup>-</sup> de 0,5.
- c. Qual é a estimativa de  $\sigma$ ?
- d. Qual é o valor da variação total e quanto ela pode ser explicada pela relação do modelo?

A equação de regressão é líquens  
 $N = 0,365 + 0,967 \text{ no}^3 \text{ depos}$

Preditor	Coef	DP	razão t	p
Constante	0,36510	0,09904	3,69	0,004
no <sup>3</sup> depos,	0,9668	0,1829	5,29	0,000

$s = 0,1932$     QR = 71,7%    QR (ajuste) = 69,2%

Análise de Variância

FONTE	GL	SQ	QM	F	P
Regressão	1	1,0427	1,0427	27,94	0,000
Erro	11	0,4106	0,0373		
Total	12	1,4533			

21. O artigo “Effects of Bike Lanes on Driver and Bicyclist Behavior” (*ASCE Transportation Eng. J.*, 1977, p. 243–256) relata os resultados de uma análise de regressão com  $x$  = espaço percorrido disponível em pés (uma medida conveniente da largura da pista, definida como a distância entre um ciclista e a linha central da pista) e distância de separação  $y$  entre uma bicicleta e um carro em ultrapassagem (determinado por fotografia). Aqui seguem os dados para 10 ruas com pistas de ciclismo:

$x$	12,8	12,9	12,9	13,6	14,5
$y$	5,5	6,2	6,3	7,0	7,8
$x$	14,6	15,1	17,5	19,5	20,8
$y$	8,3	7,1	10,0	10,8	11,0

- a. Demonstre que  $\sum x_i = 154,20$ ,  $\sum y_i = 80$ ,  $\sum x_i^2 = 2452,18$ ,  $\sum x_i y_i = 1282,74$  e  $\sum y_i^2 = 675,16$ .
- b. Deduza a equação da reta de regressão estimada.
- c. Que distância de separação você estimaria para outra pista com 15,0 de espaço disponível para trafegar?
- d. Qual seria a estimativa da distância de separação esperada de todas as ruas com um valor 15,0 de espaço percorrido utilizável?

- 22. a. Use os dados e resultados do Exercício 18 para estimar o desvio padrão do erro aleatório e no modelo de regressão linear simples.
- b. Com base nos dados e resultados do Exercício 18, que proporção de variação no fator de ajuste de deflexão pode ser explicada pela relação de regressão linear simples entre o fator de ajuste e a temperatura?

23. a. Obtenha a SQE para os dados do Exercício 19 por meio da fórmula de definição  $[SQE = \sum (y_i - \hat{y}_i)^2]$  e

compare com o valor calculado por meio da fórmula de cálculo.

- b. Calcule o valor da soma dos quadrados total. O modelo de regressão linear simples parece ser eficaz para explicar uma variação na taxa de emissão? Justifique a sua afirmação.

24. Os dados a seguir foram reproduzidos de um gráfico apresentado no artigo “Reactions on Painted Steel Under the Influence of Sodium Chloride, and Combinations Thereof” (*Ind. Engr. Chem. Prod. Res. Dev.*, 1985, p. 375–378). A variável independente é a taxa de deposição (mg/m<sup>2</sup>/d) de SO<sub>2</sub> e a variável dependente é a perda de peso do aço (g/m<sup>2</sup>).

x	14	18	40	43	45	112
y	280	350	470	500	560	1200

- a. Elabore um gráfico de dispersão. O modelo de regressão linear simples parece razoável nessa circunstância?
- b. Calcule a equação da reta de regressão estimada.
- c. Que porcentagem da variação observada na perda de peso do aço pode ser atribuída à relação do modelo com a variação na taxa de deposição?
- d. Pelo fato de o maior valor x na amostra ser sensivelmente maior do que os outros, essa observação pode ter influenciado muito na determinação da equação da reta estimada. Elimine essa observação e recalcule a equação. A nova equação parece diferir substancialmente da original (os valores previstos podem ser considerados)?

25. Mostre que  $b_1$  e  $b_0$  das expressões (12.2) e (12.3) satisfazem as equações normais.

26. Mostre que o “ponto das médias”  $(\bar{x}, \bar{y})$  cai sobre a reta de regressão estimada.

27. Suponha que um investigador tenha dados sobre a quantidade de espaço na prateleira  $x$  reservado para expor um determinado produto e a receita de vendas  $y$  desse produto. O pesquisador talvez queira ajustar um modelo no qual a reta de regressão real passe por  $(0, 0)$ . O modelo apropriado é  $Y = \beta_1 x + \epsilon$ . Suponha que  $(x_1, y_1), \dots, (x_n, y_n)$  sejam pares observados por meio desse modelo e de-

rive o estimador dos mínimos quadrados de  $\beta_1$ . (*Sugestão:* escreva a soma dos desvios quadrados como uma função de  $b_1$ , um valor experimental, e use cálculo para determinar o valor de  $b_1$  que minimize essa soma.)

28. a. Considere os dados do Exercício 20. Suponha que, em vez de a reta dos mínimos quadrados passar pelos pontos  $(x_1, y_1), \dots, (x_n, y_n)$ , queiramos que ela passe por  $(x_1 - \bar{x}, y_1), \dots, (x_n - \bar{x}, y_n)$ . Elabore um gráfico de dispersão dos pontos  $(x_i, y_i)$  e, em seguida, dos pontos  $(x_i - \bar{x}, y_i)$ . Use os gráficos para explicar intuitivamente como as duas retas dos mínimos quadrados estão relacionadas uma com a outra.

b. Suponha que, em vez do modelo  $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$  ( $i = 1, \dots, n$ ), queiramos ajustar um modelo da forma  $Y_i = \beta_0^* + \beta_1^*(x_i - \bar{x}) + \epsilon_i$  ( $i = 1, \dots, n$ ). Quais são os estimadores dos mínimos quadrados de  $\beta_0^*$  e  $\beta_1^*$  e como eles estão relacionados com  $\hat{\beta}_0$  e  $\hat{\beta}_1$ ?

29. Considere os três conjuntos de dados a seguir, nos quais as variáveis de interesse são  $x =$  distância de comutação e  $y =$  tempo de comutação. Com base em um gráfico de dispersão e nos valores de  $s$  e  $r^2$ , em que situação a regressão linear simples seria a mais (menos) eficaz e por quê?

Conjunto de Dados	1		2		3	
	x	y	x	y	x	y
	15	42	5	16	5	8
	16	35	10	32	10	16
	17	45	15	44	15	22
	18	42	20	45	20	23
	19	49	25	63	25	31
	20	46	50	115	50	60

$S_{xx}$	17,50	1270,8333	1270,8333
$S_{yy}$	29,50	2722,5	1431,6667
$\hat{\beta}_1$	1,685714	2,142295	1,126557
$\hat{\beta}_0$	13,666672	7,868852	3,196729
SST	114,83	5897,5	1627,33
SSE	65,10	65,10	14,48

## 12.3 Inferências sobre o Parâmetro do Coeficiente Angular $\beta_1$

Em quase todos os nossos trabalhos inferenciais até agora, a noção de variabilidade da amostragem tem sido muito difundida. Particularmente as propriedades das distribuições amostrais de várias estatísticas têm servido de base para o desenvolvimento de fórmulas de intervalo de confiança e métodos em testes de hipóteses. A principal idéia aqui é que o valor de praticamente qualquer quantidade calculada dos dados amostrais – o valor de praticamente qualquer estatística – varia de uma amostra para a outra.

**Exemplo 12.10**

Reconsidere os dados sobre  $x$  = taxa de liberação na área do queimador e  $y$  = taxa de emissão de  $\text{NO}_x$  do Exercício 12.19, na seção anterior. Há 14 observações feitas nos valores  $x$  100, 125, 125, 150, 150, 200, 200, 250, 250, 300, 300, 350, 400 e 400, respectivamente. Suponha que o coeficiente angular e o termo constante da reta de regressão real sejam  $\beta_1 = 1,70$  e  $\beta_0 = -50$ , com  $\sigma = 35$  (consistente com os valores  $\hat{\beta}_1 = 1,7114$ ,  $\hat{\beta}_0 = -45,55$ ,  $s = 36,75$ ). Prosseguimos então para gerar uma amostra de desvios aleatórios  $\tilde{\epsilon}_1, \dots, \tilde{\epsilon}_{14}$  de uma distribuição normal com média 0 e desvio padrão 35, e adicionamos  $\tilde{\epsilon}_i$  to  $\beta_0 + \beta_1 x_i$  para obter 14 valores  $y$  correspondentes. Os cálculos de regressão foram então realizados para obter o coeficiente angular estimado, o termo constante e o desvio padrão. Esse processo foi repetido 20 vezes, no total, gerando os valores apresentados na Tabela 12.1.

**Tabela 12.1** Resultados da simulação do Exemplo 12.10

	$\hat{\beta}_1$	$\hat{\beta}_0$	$s$		$\hat{\beta}_1$	$\hat{\beta}_0$	$s$
1.	1,7559	-60,62	43,23	11.	1,7843	-67,36	41,80
2.	1,6400	-49,40	30,69	12.	1,5822	-28,64	32,46
3.	1,4699	-4,80	36,26	13.	1,8194	-83,99	40,80
4.	1,6944	-41,95	22,89	14.	1,6469	-32,03	28,11
5.	1,4497	5,80	36,84	15.	1,7712	-52,66	33,04
6.	1,7309	-70,01	39,56	16.	1,7004	-58,06	43,44
7.	1,8890	-95,01	42,37	17.	1,6103	-27,89	25,60
8.	1,6471	-40,30	43,71	18.	1,6396	-24,89	40,78
9.	1,7216	-42,68	23,68	19.	1,7857	-77,31	32,38
10.	1,7058	-63,31	31,58	20.	1,6342	-17,00	30,93

Sem dúvida há uma variação nos valores do coeficiente angular e do termo constante estimados, bem como do desvio padrão estimado. A equação da reta dos mínimos quadrados varia, portanto, de uma amostra para outra. A Figura 12.13 mostra um *dotplot* dos coeficientes angulares estimados, bem como os gráficos da reta de regressão real e as 20 retas de regressão da amostra.

O coeficiente angular  $\beta_1$  da reta de regressão populacional é a mudança média real na variável dependente  $y$  associada com um aumento de 1 unidade na variável independente  $x$ . O coeficiente angular da reta dos mínimos quadrados,  $\hat{\beta}_1$ , fornece uma estimativa pontual de  $\beta_1$ . Da mesma forma que um intervalo de confiança de  $\mu$  e os procedimentos de teste de hipóteses sobre  $\mu$  basearam-se em propriedades da distribuição amostral de  $\bar{X}$ , as inferências adicionais sobre  $\beta_1$  têm por base considerar  $\hat{\beta}_1$  como uma estatística e investigar sua distribuição amostral.

Supõe-se que os valores dos  $x_i$ s sejam escolhidos antes da execução do experimento, de modo que apenas os  $Y_i$ s são aleatórios. Os estimadores (estatísticas e, portanto, variáveis aleatórias) de  $\beta_0$  e  $\beta_1$  são obtidos substituindo-se  $y_i$  por  $Y_i$  em (12.2) e (12.3):

$$\hat{\beta}_1 = \frac{\sum(x_i - \bar{x})(Y_i - \bar{Y})}{\sum(x_i - \bar{x})^2}$$

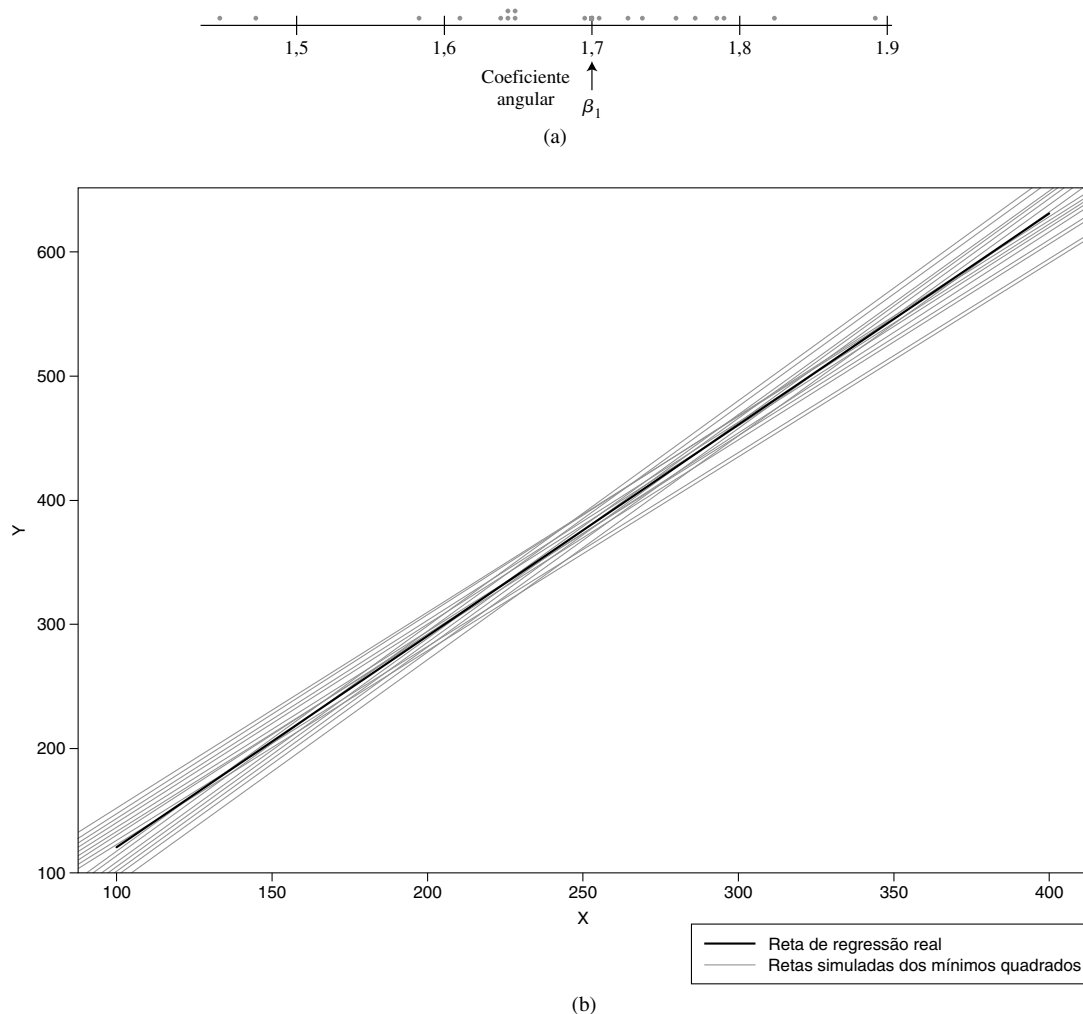
$$\hat{\beta}_0 = \frac{\sum Y_i - \hat{\beta}_1 \sum x_i}{n}$$

De modo semelhante, o estimador de  $\sigma^2$  resulta da substituição de cada  $y_i$ , na fórmula  $s^2$  pela  $Y_i$ :

$$\hat{\sigma}^2 = S^2 = \frac{\sum Y_i^2 - \hat{\beta}_0 \sum Y_i - \hat{\beta}_1 \sum x_i Y_i}{n - 2}$$

O denominador de  $\hat{\beta}_1$ ,  $S_{xx} = \sum(x_i - \bar{x})^2$  depende somente dos  $x_i$ s e não dos  $Y_i$ s, de modo que ele é uma constante. Portanto, pelo fato de  $\sum(x_i - \bar{x})\bar{Y} = \bar{Y}\sum(x_i - \bar{x}) = \bar{Y} \cdot 0 = 0$ , o estimador do coeficiente angular pode ser escrito como

$$\hat{\beta}_1 = \frac{\sum(x_i - \bar{x})Y_i}{S_{xx}} = \sum c_i Y_i \quad \text{onde } c_i = (x_i - \bar{x})/S_{xx}$$



**Figura 12.13** Resultados da simulação do Exemplo 12.10: (a) *dotplot* dos coeficientes angulares estimados; (b) gráficos da reta de regressão real e das 20 retas de mínimos quadrados (a partir do S-Plus)

Ou seja,  $\hat{\beta}_1$  é uma função linear das  $n$  variáveis independentes  $Y_1, Y_2, \dots, Y_n$ , sendo cada uma normalmente distribuída. Recorrer a propriedades de uma função linear das variáveis aleatórias discutidas na Seção 5.5 nos conduz aos resultados a seguir.

1. O valor médio de  $\hat{\beta}_1$  é  $E(\hat{\beta}_1) = \mu_{\hat{\beta}_1} = \beta_1$ , de modo que  $\hat{\beta}_1$  é um estimador não-tendencioso de  $\beta_1$  (a distribuição de  $\hat{\beta}_1$  é sempre centrada no valor de  $\beta_1$ ).
2. A variância e o desvio padrão de  $\hat{\beta}_1$  são

$$V(\hat{\beta}_1) = \sigma_{\hat{\beta}_1}^2 = \frac{\sigma^2}{S_{xx}} \quad \sigma_{\hat{\beta}_1} = \frac{\sigma}{\sqrt{S_{xx}}} \quad (12.4)$$

onde  $S_{xx} = \sum(x_i - \bar{x})^2 = \sum x_i^2 - (\sum x_i)^2/n$ . Substituindo  $\sigma$  por sua estimativa  $s$ , obtemos uma estimativa de  $\sigma_{\hat{\beta}_1}$  (o desvio padrão estimado, isto é, o erro-padrão estimado, de  $\hat{\beta}_1$ ):

$$s_{\hat{\beta}_1} = \frac{s}{\sqrt{S_{xx}}}$$

(Essa estimativa pode ainda ser representada por  $\hat{\sigma}_{\hat{\beta}_1}$ .)

3. O estimador  $\hat{\beta}_1$  tem uma distribuição normal (porque é uma função linear das  $n$  variáveis normais independentes).

De acordo com (12.4), a variância de  $\hat{\beta}_1$  é igual à variância de  $\sigma^2$  do termo do erro aleatório – ou, de modo equivalente, de qualquer  $Y_i$  – dividida por  $\sum(x_i - \bar{x})^2$ . Pelo fato de  $\sum(x_i - \bar{x})^2$  ser uma medida da dispersão dos  $x_i$ s em torno de  $\bar{x}$ , concluímos que fazer observações em valores  $x_i$  muito dispersos produz um estimador mais preciso do parâmetro do coeficiente angular (variância menor de  $\hat{\beta}_1$ ), ao passo que, quando todos os valores  $x_i$ s estão muito próximos uns dos outros, o estimador é altamente variável. Obviamente, se os  $x_i$ s ficarem muito dispersos, um modelo linear talvez não seja apropriado em todos os âmbitos de observação.

Vários procedimentos inferenciais foram discutidos antes para padronizar um estimador, primeiramente subtraindo-se seu valor médio e em seguida dividindo-se por seu desvio padrão estimado. Em particular, os procedimentos de teste e um IC para a média  $\mu$  de uma população normal utilizaram-se do fato de a variável padronizada  $(\bar{X} - \mu)/(S/\sqrt{n})$  – isto é,  $(\bar{X} - \mu)/S_{\bar{x}}$  – ter uma distribuição  $t$  com  $n - 1$  gl. Um resultado semelhante oferece aqui a solução para inferências posteriores concernentes a  $\beta_1$ .

**TEOREMA**

As suposições do modelo de regressão linear simples significam que a variável padronizada

$$T = \frac{\hat{\beta}_1 - \beta_1}{S/\sqrt{S_{xx}}} = \frac{\hat{\beta}_1 - \beta_1}{S_{\hat{\beta}_1}}$$

tem uma distribuição  $t$  com  $n - 2$  gl.

**Intervalo de Confiança para  $\beta_1$** 

Como na derivação dos ICs anteriores, começamos com uma declaração de probabilidade:

$$P\left(-t_{\alpha/2, n-2} < \frac{\hat{\beta}_1 - \beta_1}{S_{\hat{\beta}_1}} < t_{\alpha/2, n-2}\right) = 1 - \alpha$$

A manipulação das desigualdades entre parênteses para isolar  $\beta_1$  e a substituição a estimativas pelos estimadores resultam na fórmula do IC.

Um IC de  $100(1 - \alpha)\%$  para o coeficiente angular  $\beta_1$  da reta de regressão real é

$$\hat{\beta}_1 \pm t_{\alpha/2, n-2} \cdot s_{\hat{\beta}_1}$$

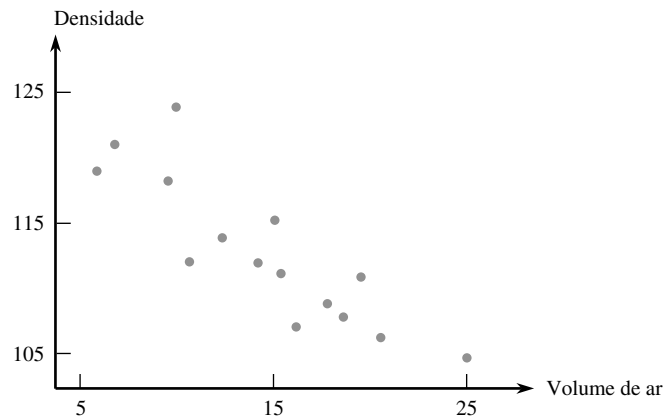
Esse intervalo tem a mesma forma geral que a de vários de nossos intervalos anteriores. Ele está centralizado na estimativa pontual do parâmetro e o quanto ele se estende em cada lado da estimativa depende do nível de confiança desejado (por meio do valor crítico  $t$ ) e da quantidade de variabilidade no estimador  $\hat{\beta}_1$  (por meio de  $s_{\hat{\beta}_1}$ , que tende a ser pequeno, quando houver pouca variabilidade na distribuição de  $\hat{\beta}_1$ , e grande, na situação inversa).

**Exemplo 12.11**

Variações no peso da alvenaria de tijolos de argila têm implicações não apenas no projeto estrutural e acústico, como também no projeto dos sistemas térmico, de ventilação e de ar-condicionado. O artigo “Clay Brick Masonry Weight Variation” (*J. of Architectural Engr.*, 1996, p. 135–137) fornece um gráfico de dispersão de  $y =$  densidade da argamassa seca (lb/pés<sup>3</sup>) versus  $x =$  volume de ar da argamassa (%) para uma amostra de espécimes de argamassa, de onde foram reproduzidos os seguintes dados representativos:

$x$	5,7	6,8	9,6	10,0	10,7	12,6	14,4	15,0	15,3
$y$	119,0	121,3	118,2	124,0	112,3	114,1	112,2	115,1	111,3
$x$	16,2	17,8	18,7	19,7	20,6	25,0			
$y$	107,2	108,9	107,8	111,0	106,2	105,0			

O gráfico de dispersão desses dados na Figura 12.14 certamente sugere a aplicabilidade do modelo de regressão linear simples; parece haver uma relação linear negativa significativa entre o volume de ar e a densidade, relação em que a densidade tende a diminuir à medida que o volume de ar aumenta.



**Figura 12.14** Gráfico de dispersão dos dados do Exemplo 12.11

Os valores das estatísticas requeridas para calcular as estimativas dos mínimos quadrados são

$$\begin{aligned}\sum x_i &= 218,1 & \sum y_i &= 1693,6 & \sum x_i^2 &= 3577,01 \\ \sum x_i y_i &= 24.252,54 & \sum y_i^2 &= 191.672,90\end{aligned}$$

dos quais  $S_{xy} = -372,404$ ,  $S_{xx} = 405,836$ ,  $\hat{\beta}_1 = -0,917622$ ,  $\hat{\beta}_0 = 126,248889$ ,  $SQT = 454,1693$ ,  $SQE = 112,4432$  e  $r^2 = 1 - 112,4432/454,1693 = 0,752$ . Aproximadamente 75% da variação observada na densidade pode ser atribuída à relação do modelo de regressão linear simples entre a densidade e o volume de ar. O gl do erro é  $15 - 2 = 13$ , que dá  $s^2 = 112,4432/13 = 8,6495$  e  $s = 2,941$ .

O desvio padrão estimado de  $\hat{\beta}_1$  é

$$s_{\hat{\beta}_1} = \frac{s}{\sqrt{S_{xx}}} = \frac{2,941}{\sqrt{405,836}} = 0,1460$$

O valor crítico  $t$  para um nível de confiança de 95% é  $t = 2,160$ . O intervalo de confiança é

$$-0,918 \pm (2,160)(0,1460) = -0,918 \pm 0,315 = (-1,233, -0,603)$$

Com alto grau de confiança, estimamos que uma diminuição média na densidade, entre 0,603 lb/pés<sup>3</sup> e 1,233 lb/pés<sup>3</sup>, está associada com um aumento de 1% do volume de ar (pelo menos para valores de volume de ar entre 5% e 25%, aproximadamente, que correspondem aos valores  $x$  em nossa amostra). O intervalo é razoavelmente estreito, o que indica que o coeficiente angular da reta populacional foi estimado com precisão. Observe que o intervalo compreende apenas valores negativos, de maneira que podemos ficar bem seguros em relação à tendência de a densidade diminuir à medida que o conteúdo de ar aumenta.

Ao examinar o resultado do SAS na Figura 12.15 encontramos o valor de  $s_{\hat{\beta}_1}$  sob o título Estimativas do Parâmetro tal como o segundo número na coluna Erro-Padrão. Todos os *softwares* estatísticos amplamente usados incluem esse erro-padrão estimado no resultado. Há também um erro-padrão estimado para a estatística  $\hat{\beta}_0$ , com base na qual é possível calcular um intervalo de confiança para o termo constante  $\hat{\beta}_0$  da reta de regressão populacional.

Variável Dependente: DENSIDADE

Análise de Variância					
Fonte	GL	Soma dos Quadrados	Quadrado Médio	Valor F	Prob > F
Modelo	1	341,72606	341,72606	39,508	0,0001
Erro	13	112,44327	8,64948		
Total C	14	454,16933			
	Raíz	2,94100	Quadrado Médio	0,7524	
	Média Dep.	112,90667	Ajuste do Quadrado	0,7334	
	CV	2,60481			
Estimativa do Parâmetro					
Variável	GL	Parâmetro Estimado	ERRO PADRÃO	T para H0: Parâmetro = 0	Prob >  T
TERMO CONST.	1	126,248889	2,25441683	56,001	0,0001
VOLUME DE AR	1	-0,917622	0,14598888	-6,286	0,0001
	Obs	Var, Dep, DENSIDADE	VALOR PREVISTO	Resíduo	
	1	119,0	121,0	-2,0184	
	2	121,3	120,0	1,2909	
	3	118,2	117,4	0,7603	
	4	124,0	117,1	6,9273	
	5	112,3	116,4	-4,1303	
	6	114,1	114,7	-0,5869	
	7	112,2	113,0	-0,8351	
	8	115,1	112,5	2,6154	
	9	111,3	112,2	-0,9093	
	10	107,2	111,4	-4,1834	
	11	108,9	109,9	-1,0152	
	12	107,8	109,1	-1,2894	
	13	111,0	108,2	2,8283	
	14	106,2	107,3	-1,1459	
	15	105,0	103,3	1,6917	
		Soma dos Resíduos		0	
		Soma dos Resíduos Quadrados		112,4433	
		SQ dos Resíduos Previstos (Press)		146,4144	

Figura 12.15 Resultado do SQA dos dados do Exemplo 12.11

## Procedimentos de Teste de Hipóteses

Como antes, a hipótese nula em um teste sobre  $\beta_1$  será uma declaração de igualdade. O valor nulo (valor de  $\beta_1$  declarado real pela hipótese nula) será representado por  $\beta_{10}$  (leia-se “beta um zero”, não “beta dez”). A estatística de teste resulta da substituição de  $\beta_1$  na variável padronizada  $T$  pelo valor nulo  $\beta_{10}$  – ou seja, da padronização do estimador de  $\beta_1$  segundo a suposição de que  $H_0$  é verdadeira. A estatística de teste, portanto, tem uma distribuição  $t$  com  $n - 2$  gl quando  $H_0$  é verdadeira, de maneira que a probabilidade de erro tipo I é controlada no nível desejado  $\alpha$ , usando-se um valor crítico  $t$  apropriado.

O par de hipóteses sobre  $\beta_1$  mais comumente encontrado é  $H_0: \beta_1 = 0$  versus  $H_a: \beta_1 \neq 0$ . Quando essa hipótese nula é verdadeira,  $\mu_{y,x} = \beta_0$  independentemente de  $x$ , então o conhecimento de  $x$  não oferece nenhuma informação sobre o valor da variável dependente. O teste dessas duas hipóteses em geral é chamado de *teste de utilidade do modelo* na regressão linear simples. A menos que  $n$  seja muito pequeno,  $H_0$  será rejeitada e a utilidade do modelo será confirmada precisamente quando  $r^2$  for razoavelmente grande. O modelo de regressão linear simples não deve ser usado para inferências adicionais (estimativas do valor médio ou previsões de futuros valores), a menos que o resultado do teste de utilidade do modelo rejeite  $H_0$  para um  $\alpha$  adequadamente pequeno.



Hipótese nula:  $H_0: \beta_1 = \beta_{10}$

Valor da estatística de teste:  $t = \frac{\hat{\beta}_1 - \beta_{10}}{s_{\hat{\beta}_1}}$

**Hipótese Alternativa      Região de Rejeição para Teste de Nível  $\alpha$**

$H_a: \beta_1 > \beta_{10} \quad t \geq t_{\alpha, n-2}$   
 $H_a: \beta_1 < \beta_{10} \quad t \leq -t_{\alpha, n-2}$   
 $H_a: \beta_1 \neq \beta_{10} \quad \text{ou } t \geq t_{\alpha/2, n-2} \quad \text{ou } t \leq -t_{\alpha/2, n-2}$

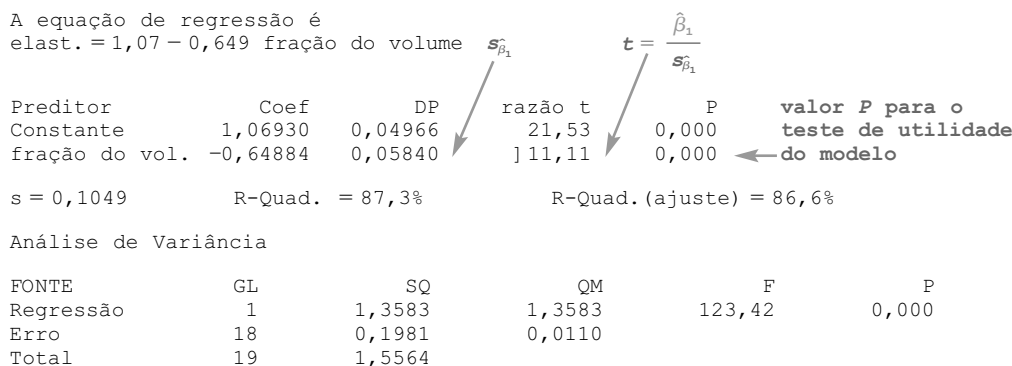
Um valor  $P$  baseado em  $n - 2$  gl pode ser calculado como anteriormente para os testes  $t$ , nos capítulos 8 e 9.

O teste de utilidade do modelo é o teste de  $H_0: \beta_1 = 0$  versus  $H_a: \beta_1 \neq 0$ , caso em que o valor da estatística de teste é a razão  $t = \hat{\beta}_1/s_{\hat{\beta}_1}$ .

**Exemplo 12.12**

A limpeza de metal ou liga de alumínio antes da sua fundição é determinada principalmente pelo volume de hidrogênio e de inclusão da fundição. O artigo "Effect of Melt Cleanliness on the Properties of an Al-10 Wt Pct Si-10 Vol Pct SiC(p) Composite" (*Metallurgical Trans.*, 1993, p. 1631-1645) relata um estudo em que várias propriedades de tensão foram relacionadas com  $x$  = fração do volume de óxidos/inclusões (%). Aqui apresentamos os dados (reproduzidos de um gráfico) sobre  $y$  = elasticidade (%) das barras de teste. Os autores afirmam que o gráfico de dispersão mostra uma relação linear e fornecem a equação da reta dos mínimos quadrados. Vamos usar o resultado do MINITAB da Figura 12.16 para realizar o teste de utilidade do modelo no nível de significância  $\alpha = 0,01$ .

$x$	0,10	0,16	0,31	0,37	0,37	0,46	0,50	0,50	0,60	0,70
$y$	0,96	1,10	0,80	0,84	0,77	0,87	0,60	0,87	0,60	0,61
$x$	0,75	0,80	0,90	1,00	1,07	1,08	1,11	1,30	1,37	1,54
$y$	0,70	0,41	0,40	0,41	0,45	0,59	0,25	0,25	0,08	0,10



**Figura 12.16** Resultado do MINITAB para o Exemplo 12.12

O parâmetro de interesse é  $\beta_1$ , a mudança esperada na porcentagem de elasticidade associada com 1% de aumento na fração do volume de óxidos/inclusões.  $H_0: \beta_1 = 0$  será rejeitada em favor de  $H_a: \beta_1 \neq 0$ , se a razão  $t = \hat{\beta}_1/s_{\hat{\beta}_1}$  satisfizer  $t \geq t_{\alpha/2, n-2} = t_{0,005, 18} = 2,878$  ou  $t \leq -2,878$ . Com base na Figura 12.15,  $\hat{\beta}_1 = -0,64884$ ,  $s_{\hat{\beta}_1} = 0,05840$ , e

$$t = \frac{-0,64884}{0,05840} = -11,11 \quad (\text{Também no resultado})$$

Obviamente,  $-11,11 \leq -2,878$ , então  $H_0$  é rejeitada de modo retumbante. De modo alternativo, o valor  $P$  é duas vezes a área incluída sob a curva  $t$  de 18 gl, à esquerda de  $-11,11$ . O MINITAB fornece o valor  $P = 0,000$ , de forma que  $H_0$  deve ser rejeitada em qualquer  $\alpha$  moderado. Essa confirmação da utilidade do modelo de regressão linear simples nos permite calcular várias estimativas e previsões, como descrito na Seção 12.4. ■

## Regressão e ANOVA

A divisão total da soma dos quadrados  $\sum(y_i - \bar{y})^2$  em uma parte SQE, que mede uma variação inexplicada, e uma parte SQR, que mede uma variação explicada por meio da relação linear, é bastante similar à ANOVA unidirecional. Na verdade, a hipótese nula  $H_0: \beta_1 = 0$  pode ser testada em relação a  $H_a: \beta_1 \neq 0$  elaborando-se uma tabela ANOVA (Tabela 12.2) e rejeitando-se  $H_0$  se  $f \geq F_{\alpha,1,n-2}$ .

O teste  $F$  fornece exatamente o mesmo resultado que o teste  $t$  de utilidade do modelo porque  $t^2 = f$  e  $t_{\alpha/2,n-2}^2 = F_{\alpha,1,n-2}$ . Praticamente todos os softwares de computador que dispõem de opções de regressão incluem no resultado uma tabela ANOVA desse tipo. Por exemplo: a Figura 12.15 mostra o resultado do SAS dos dados sobre argamassa do Exemplo 12.11. A tabela ANOVA na parte superior do resultado tem  $f = 39,508$ , com um valor  $P$  de 0,0001, para o teste de utilidade do modelo. A tabela de estimativas de parâmetro fornece  $t = -6,286$ , novamente com  $P = 0,0001$  e  $(-6,286)^2 = 39,51$ .

**Tabela 12.2** Tabela ANOVA da regressão linear simples

Fonte de variação	gl	Soma dos Quadrados	Quadrado Médio	$f$
Regressão	1	SQR	SQR	$\frac{\text{SQR}}{\text{SQE}/(n-2)}$
Erro	$n - 2$	SQE	$s^2 = \frac{\text{SQE}}{n-2}$	
Total	$n - 1$	SQT		

### Exercícios | Seção 12.3 (30–43)

- 30.** Reconsidere a situação descrita no Exercício 7, no qual  $x$  = resistência acelerada do concreto e  $y$  = resistência ao tratamento de 28 dias. Suponha que o modelo de regressão linear simples seja válido para  $x$  entre 1000 e 4000 e que  $\beta_1 = 1,25$  e  $\sigma = 350$ . Suponha um experimento no qual  $n = 7$  e os valores  $x$ , nos quais as observações são feitas, são  $x_1 = 1000, x_2 = 1500, x_3 = 2000, x_4 = 2500, x_5 = 3000, x_6 = 3500$  e  $x_7 = 4000$ .
- Calcule  $\sigma_{\hat{\beta}_1}$ , o desvio padrão de  $\hat{\beta}_1$ .
  - Qual é a probabilidade de o coeficiente angular estimado baseado nessas observações ficar entre 1,00 e 1,50?
  - Suponha que também seja possível fazer uma única observação em cada um dos  $n = 11$  valores  $x_1 = 2000, x_2 = 2100, \dots, x_{11} = 3000$ . Se o objetivo principal for estimar  $\beta_1$  com o máximo de precisão possível, o experimento com  $n = 11$  seria preferível ao experimento com  $n = 7$ ?
- 31.** Reconsidere as quantidades fornecidas no Exercício 18 para a regressão de  $y$  = fator de deflexão sobre  $x$  = temperatura.
- Calcule o desvio padrão estimado  $s_{\hat{\beta}_1}$ .
  - Calcule um IC de 95% para  $\beta_1$ , a mudança esperada no fator de deflexão associado com um aumento de temperatura de 1°F.
- 32.** O Exercício 16 da Seção 12.2 forneceu dados sobre  $x$  = volume de chuva e  $y$  = volume de runoff (ambos em m<sup>3</sup>). Use o resultado do MINITAB a seguir para determinar se existe uma relação linear útil entre o volume de chuva e o volume de runoff, e em seguida calcule um intervalo de confiança para a mudança média real do volume de runoff associado com um aumento de 1 m<sup>3</sup> no volume de chuva.
- A equação de regressão é  
 $runoff = -1,13 + 0,827 \text{ chuva}$
- | Previsor  | Coef    | DP      | razão t | p     |
|-----------|---------|---------|---------|-------|
| Constante | -1,128  | 2,368   | -0,48   | 0,642 |
| chuva     | 0,82697 | 0,03652 | 22,64   | 0,000 |
- $s = 5,240$     QR = 97,5%    QR (ajuste) = 97,3%
- 33.** O Exercício 15 da Seção 12.2 incluiu o resultado do MINITAB para uma regressão da resistência à flexão de vigas de concreto no módulo de elasticidade.
- Use esse resultado para calcular um intervalo de confiança com um nível de confiança de 95% para o coeficiente angular  $\beta_1$  da reta de regressão populacional e interprete o intervalo resultante.
  - Suponha que tivéssemos acreditado que, quando o módulo de elasticidade aumenta 1 GPa, a mudança média real associada na resistência à flexão é de no máximo 0,1 MPa. Os dados amostrais contradizem essa afirmação? Expresse e teste as hipóteses relevantes.
- 34.** Consulte o resultado do MINITAB do Exercício 20, no qual  $x$  = deposição de umidade de  $x = NO_3^-$  e  $y$  = líquens N (%).

- a. Execute o teste de utilidade do modelo no nível 0,01, usando a abordagem da região de rejeição.
- b. Repita a parte (a), usando a abordagem do valor  $P$ .
- c. Suponha que antes tivéssemos acreditado que, quando a deposição de umidade de  $\text{NO}_3^-$  aumenta em torno de  $0,1 \text{ g N/m}^2$ , a mudança associada no líquen  $N$  esperado é de pelo menos 0,15%. Teste as hipóteses no nível 0,01 para determinar se os dados contradizem esse ponto de vista prévio.

35. Como a aceleração lateral – forças laterais experimentadas em viagens que em grande medida podem ser controladas pelo motorista – influencia na náusea sentida por passageiros de ônibus? O artigo "Motion Sickness in Public Road Transport: The Effect of Driver, Route, and Vehicle" (*Ergonomics*, 1999, p. 1646–1664) apresentou dados sobre  $x$  = quantidade de náusea de movimento (calculada de acordo com um padrão britânico para avaliar movimentos semelhantes no mar) e  $y$  = náusea relatada (%). As quantidades pertinentes são

$$n = 17, \sum x_i = 222,1, \sum y_i = 193, \sum x_i^2 = 3056,69, \sum x_i y_i = 2759,6, \sum y_i^2 = 2975$$

Os valores da quantidade na amostra estão entre 6,0 e 17,6.

- a. Supondo que o modelo de regressão linear simples seja válido para relacionar essas duas variáveis (isso é confirmado pelos dados brutos), calcule e interprete uma estimativa do parâmetro do coeficiente angular que transmita informações sobre a precisão e a confiabilidade da estimativa.
- b. Parece haver uma relação linear útil entre essas duas variáveis? Responda à pergunta, empregando a abordagem do valor  $P$ .
- c. Seria sensato usar o modelo de regressão linear simples como base para prever a porcentagem de náusea quando a quantidade = 5,0? Explique seu raciocínio.
- d. Quando o MINITAB foi usado para ajustar o modelo de regressão linear simples aos dados brutos, foi indicado que a observação (6,0, 2,50) possivelmente tivesse um impacto significativo sobre o ajuste. Elimine essa observação da amostra e recalcule a estimativas da parte (a). Com base nisso, a observação parece estar exercendo uma influência indevida?

36. Uma névoa (gotículas transportadas pelo ar ou aerossóis) é gerada quando fluidos para remoção de metal são usados em operações de máquina para resfriar e lubrificar a ferramenta e a peça. A produção de névoa é uma preocupação para a Occupational Safety and Health Administration (OSHA), que recentemente diminuiu substancialmente o nível nos locais de trabalho. O artigo "Variables Affecting Mist Generation from Metal Removal Fluids" (*Lubrication Engr.*, 2002, p. 10–17) apresentou os dados a seguir sobre  $x$  = velocidade do fluxo do fluido para um óleo 5% solúvel (cm/s) e  $y$  = a quantidade de gotículas de névoa com diâmetro menor que  $10 \mu\text{m}$  ( $\text{mg/m}^3$ ):

$x$	89	177	189	354	362	442	965
$y$	0,40	0,60	0,48	0,66	0,61	0,69	0,99

- a. Os pesquisadores executaram uma análise de regressão linear simples para relacionar as duas variáveis. O gráfico de dispersão dos dados apóia essa estratégia?
  - b. Que proporção da variação observada no volume de névoa pode ser atribuída à relação de regressão linear simples entre velocidade e névoa?
  - c. Os investigadores estavam particularmente interessados no impacto sobre a névoa ao aumentar a velocidade de 100 para 1000 (um fator de 10 correspondente à diferença entre o menor e o maior valores  $x$  na amostra). Quando  $x$  aumenta dessa forma, há alguma evidência significativa de que o aumento médio real em  $y$  é menor que 0,6?
  - d. Estime a mudança média real na névoa associada com um aumento de velocidade de 1 cm/s. Faça isso de maneira que transmita informações sobre precisão e confiança.
37. O artigo "Hydrogen, Oxygen, and Nitrogen in Cobalt Metal" (*Metallurgia*, 1969, p. 121–127) contém um gráfico dos pares de dados a seguir, onde  $x$  = pressão do gás extraído (mícrons) e  $y$  = tempo de extração (min.):

$x$	40	130	155	160	260	275	325	370	420	480
$y$	2,5	3,0	3,1	3,3	3,7	4,1	4,3	4,8	5,0	5,4

- a. Estime  $\sigma$  e o desvio padrão de  $\hat{\beta}_1$ .
  - b. Suponha que os investigadores antes do experimento acreditassem que haveria em média um aumento de 0,006 min. no tempo de extração, associado com um aumento de 1 mícron na pressão. Use a abordagem do valor  $P$ , com nível de significância 0,10, para determinar se os dados contradizem esse ponto de vista prévio.
38. Consulte os dados sobre  $x$  = taxa de liberação e  $y$  = taxa de emissão de  $\text{NO}_x$  fornecidos no Exercício 19.
- a. O modelo de regressão linear simples especifica uma relação útil entre as duas taxas? Use o procedimento de teste apropriado para obter informações sobre o valor  $P$  e, em seguida, tire uma conclusão usando o nível de significância 0,01.
  - b. Calcule um IC de 95% para a mudança esperada na taxa de emissão associada com um aumento de 10  $\text{MBtu/h-pés}^2$  na taxa de liberação.
39. Realize o teste de utilidade do modelo, usando a abordagem ANOVA para os dados sobre taxa de filtração–quantidade de umidade do Exemplo 12.6. Verifique se isso oferece um resultado equivalente ao do teste  $t$ .
40. Use as regras do valor esperado para mostrar que  $\hat{\beta}_0$  é um estimador não-tendencioso de  $\beta_0$  (supondo que  $\hat{\beta}_1$  não seja tendencioso para  $\beta_1$ ).
41. a. Verifique que  $E(\hat{\beta}_1) = \beta_1$  usando as regras do valor esperado do Capítulo 5.
- b. Use as regras de variância do Capítulo 5 para verificar a expressão para  $V(\hat{\beta}_1)$  fornecida nesta seção.

42. Demonstre que, se cada  $x_i$  for multiplicado por uma constante positiva  $c$  e cada  $y_i$  for multiplicado por outra constante positiva  $d$ , a estatística  $t$  para testar  $H_0: \beta_1 = 0$  versus  $H_a: \beta_1 \neq 0$  não muda em valor (o valor de  $\hat{\beta}_1$  mudará, o que mostra que sua magnitude não é por si só uma indicação da utilidade do modelo).
43. A probabilidade de um erro tipo II para o teste  $t$  de  $H_0: \beta_1 = \beta_{10}$  pode ser calculada da mesma maneira que para os testes  $t$  do Capítulo 8. Se o valor alternativo de  $\beta_1$  for representado por  $\beta'_1$ , o valor de

$$d = \frac{|\beta_{10} - \beta'_1|}{\sigma \sqrt{\frac{n-1}{\sum x_i^2 - (\sum x_i)^2/n}}}$$

é calculado primeiro; depois, o conjunto de curvas apropriado na Tabela A.17 do apêndice será inserido no eixo horizontal, no valor de  $d$ , e o  $\beta$  será reproduzido da curva para  $n - 2$  gl. Um artigo publicado no *Journal of Public Health Engineering* apresenta os resultados de uma análise de regressão baseada em  $n = 15$  observações, nas quais  $x$  = temperatura de aplicação do filtro ( $^{\circ}\text{C}$ ) e  $y$  = % eficiência da remoção de DBO. As quantidades calculadas incluem  $\sum x_i = 402$ ,  $\sum x_i^2 = 11.098$ ,  $s = 3,725$  e  $\hat{\beta}_1 = 1,7035$ . Experimente testar, no nível 0,01,  $H_0: \beta_1 = 1$ , que expressa que o aumento esperado na % de remoção de DBO é 1, quando a temperatura de aplicação do filtro aumenta  $1^{\circ}\text{C}$ , comparando-a com a alternativa  $H_a: \beta_1 > 1$ . Determine  $P$ (erro tipo II) quando  $\beta'_1 = 2$ ,  $\sigma = 4$ .

## 12.4 Inferências sobre $\mu_{Y \cdot X^*}$ e a Previsão de Valores Futuros de $Y$

Seja  $x^*$  um valor especificado da variável independente  $x$ . Uma vez que as estimativas  $\hat{\beta}_0$  e  $\hat{\beta}_1$  tenham sido calculadas,  $\hat{\beta}_0 + \hat{\beta}_1 x^*$  pode ser considerada uma estimativa pontual de  $\mu_{Y \cdot X^*}$  (o valor esperado ou valor médio real de  $Y$  quando  $x = x^*$ ) ou uma previsão do valor  $Y$  que resultará de uma única observação feita quando  $x = x^*$ . A estimativa ou previsão pontual por si só não fornece qualquer informação sobre a precisão da estimativa de  $\mu_{Y \cdot X^*}$ . Isso pode ser remediado desenvolvendo-se um IC para  $\mu_{Y \cdot X^*}$  e um intervalo de previsão (IP) para um único valor  $Y$ .

Antes de obtermos dados amostrais, tanto  $\hat{\beta}_0$  como  $\hat{\beta}_1$  estão sujeitos à variabilidade amostral – ou seja, ambos são estatísticas cujos valores variam de uma amostra para outra. Suponha, por exemplo, que  $\beta_0 = 50$  e  $\beta_1 = 2$ . Então, uma primeira amostra de pares  $(x, y)$  pode fornecer  $\hat{\beta}_0 = 52,35$ ,  $\hat{\beta}_1 = 1,895$ , uma segunda amostra pode resultar em  $\hat{\beta}_0 = 46,52$ ,  $\hat{\beta}_1 = 2,056$ , e assim por diante. Segue-se a isso que  $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 x^*$ , por si só, varia em valor de amostra para amostra, de modo que é uma estatística. Se o termo constante e o coeficiente angular da reta populacional forem os valores antes mencionados, 50 e 2, respectivamente, e  $x^* = 10$ , então essa estatística está tentando estimar o valor  $50 + 2(10) = 70$ . A estimativa da primeira amostra acima será  $52,35 + 1,895(10) = 71,30$ , da segunda amostra será  $46,52 + 2,056(10) = 67,08$ , e assim por diante. Da mesma forma que um intervalo de confiança para  $\beta_1$  foi fundamentado em propriedades da distribuição amostral de  $\hat{\beta}_1$ , um intervalo de confiança para um valor  $y$  médio em regressão baseia-se nas propriedades da distribuição amostral da estatística  $\hat{\beta}_0 + \hat{\beta}_1 x^*$ .

A substituição das expressões de  $\beta_0$  e  $\hat{\beta}_1$  em  $\hat{\beta}_0 + \hat{\beta}_1 x^*$ , seguida de algumas manipulações algébricas, conduz à representação de  $\hat{\beta}_0 + \hat{\beta}_1 x^*$  como uma função linear dos  $Y_i$ s:

$$\hat{\beta}_0 + \hat{\beta}_1 x^* = \sum_{i=1}^n \left[ \frac{1}{n} + \frac{(x^* - \bar{x})(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \right] Y_i = \sum_{i=1}^n d_i Y_i$$

Os coeficientes  $d_1, d_2, \dots, d_n$  nessa função linear envolvem os  $x_i$ s e  $x^*$ , todos fixos. A aplicação das regras da Seção 5.5 a essa função linear fornece as seguintes propriedades:

Seja  $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 x^*$ , onde  $x^*$  é algum valor fixo de  $x$ . Então,

1. O valor médio de  $\hat{Y}$  é

$$E(\hat{Y}) = E(\hat{\beta}_0 + \hat{\beta}_1 x^*) = \mu_{\hat{\beta}_0 + \hat{\beta}_1 x^*} = \beta_0 + \beta_1 x^*$$

Portanto  $\hat{\beta}_0 + \hat{\beta}_1 x^*$  é um estimador não-tendencioso de  $\beta_0 + \beta_1 x^*$  (isto é, para  $\mu_{Y, x^*}$ ).

2. A variância de  $\hat{Y}$  é

$$V(\hat{Y}) = \sigma_{\hat{Y}}^2 = \sigma^2 \left[ \frac{1}{n} + \frac{(x^* - \bar{x})^2}{\sum x_i^2 - (\sum x_i)^2/n} \right] = \sigma^2 \left[ \frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}} \right]$$

e o desvio padrão  $\sigma_{\hat{Y}}$  é a raiz quadrada dessa expressão. O desvio padrão estimado de  $\hat{\beta}_0 + \hat{\beta}_1 x^*$ , representado por  $s_{\hat{Y}}$  ou  $s_{\hat{\beta}_0 + \hat{\beta}_1 x^*}$ , resulta da substituição de  $\sigma$  por sua estimativa  $s$ :

$$s_{\hat{Y}} = s_{\hat{\beta}_0 + \hat{\beta}_1 x^*} = s \sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}}$$

3.  $\hat{Y}$  tem uma distribuição normal.

A variância de  $\hat{\beta}_0 + \hat{\beta}_1 x^*$  é mínima quando  $x^* = \bar{x}$  e aumenta, à medida que  $x^*$  distancia-se de  $\bar{x}$  em ambas as direções. Logo, o estimador de  $\mu_{Y, x^*}$  é mais preciso quando  $x^*$  está próximo do centro dos  $x_i$ s do que quando está distante dos valores  $x$  amostrados. Isso significa que tanto o IC como o IP são mais estreitos para um  $x^*$  próximo de  $\bar{x}$  do que para um  $x^*$  distante de  $\bar{x}$ . A maioria dos *softwares* estatísticos fornecerá tanto  $\hat{\beta}_0 + \hat{\beta}_1 x^*$  como  $s_{\hat{\beta}_0 + \hat{\beta}_1 x^*}$ , para qualquer  $x^*$  especificado, quando solicitados.

### Inferências sobre $\mu_{Y, x^*}$

Assim como os procedimentos inferenciais para  $\beta_1$  basearam-se na variável  $t$  obtida por meio da padronização de  $\beta_1$ , aqui, a variável  $t$  obtida pela padronização de  $\hat{\beta}_0 + \hat{\beta}_1 x^*$  conduz a um IC e a procedimentos de teste.

#### TEOREMA

A variável

$$T = \frac{\hat{\beta}_0 + \hat{\beta}_1 x^* - (\beta_0 + \beta_1 x^*)}{S_{\hat{\beta}_0 + \hat{\beta}_1 x^*}} = \frac{\hat{Y} - (\beta_0 + \beta_1 x^*)}{S_{\hat{Y}}} \quad (12.5)$$

tem uma distribuição  $t$  com  $n - 2$  gl.

Assim como para  $\beta_1$  na seção anterior, as probabilidades compreendendo essa variável padronizada podem ser manuseadas para produzir um intervalo de confiança para  $\mu_{Y, x^*}$ .

Um IC de  $100(1 - \alpha)\%$  para  $\mu_{Y, x^*}$ , o valor esperado de  $Y$  quando  $x = x^*$ , é

$$\hat{\beta}_0 + \hat{\beta}_1 x^* \pm t_{\alpha/2, n-2} \cdot s_{\hat{\beta}_0 + \hat{\beta}_1 x^*} = \hat{y} \pm t_{\alpha/2, n-2} \cdot s_{\hat{Y}} \quad (12.6)$$

Esse IC está centrado na estimativa pontual para  $\mu_{Y, x^*}$  e estende-se para ambos os lados, segundo um valor que depende do nível de confiança e da extensão da variabilidade no estimador em que a estimativa pontual está baseada.

**Exemplo 12.13**

A corrosão de barras de aço de reforço é o problema de durabilidade mais importante das estruturas de concreto reforçadas. A carbonação do concreto resulta de uma reação química que diminui suficientemente o nível de pH para iniciar a corrosão do concreto armado. São apresentados a seguir dados representativos sobre  $x$  = nível de carbonação (mm) e  $y$  = resistência (MPa) de uma amostra de espécimes retirados de um determinado prédio (reproduzidos de um gráfico do artigo “The Carbonation of Concrete Structures in the Tropical Environment of Singapore”, *Magazine of Concrete Res.*, 1996, p. 293–300).

$x$	8,0	15,0	16,5	20,0	20,0	27,5	30,0	30,0	35,0
$y$	22,8	27,2	23,7	17,1	21,5	18,6	16,1	23,4	13,4
$x$	38,0	40,0	45,0	50,0	50,0	55,0	55,0	59,0	65,0
$y$	19,5	12,4	13,2	11,4	10,3	14,1	9,7	12,0	6,8

O gráfico de dispersão dos dados (veja a Figura 12.17) apóia em grande medida o uso do modelo de regressão linear simples. As quantidades pertinentes são as seguintes:

$$\begin{array}{llll} \sum x_i = 659,0 & \sum x_i^2 = 28.967,50 & \bar{x} = 36,6111 & S_{xx} = 4840,7778 \\ \sum y_i = 293,2 & \sum x_i y_i = 9.293,95 & \sum y_i^2 = 5335,76 & \\ \hat{\beta}_1 = -0,297561 & \beta_0 = 27,182936 & \text{SQE} = 131,2402 & \\ r^2 = 0,766 & s = 2,8640 & & \end{array}$$

Calculemos agora um intervalo de confiança, usando um nível de confiança de 95%, para a resistência média de todos os espécimes com um nível de carbonação de 45 mm – ou seja, um intervalo de confiança para  $\beta_0 + \beta_1(45)$ . O intervalo está centralizado em

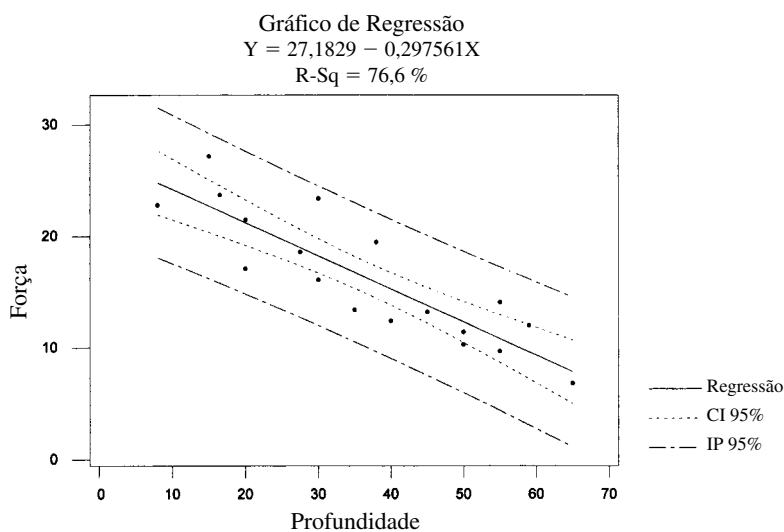
$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1(45) = 27,18 - 0,2976(45) = 13,79$$

O desvio padrão estimado da estatística  $\hat{Y}$  é

$$s_{\hat{y}} = 2,8640 \sqrt{\frac{1}{18} + \frac{(45 - 36,6111)^2}{4840,7778}} = 0,7582$$

O valor crítico  $t$  com 16 gl para um nível de confiança de 95% é 2,120, com base no qual determinamos o intervalo desejado como sendo

$$13,79 \pm (2,120)(0,7582) = 13,79 \pm 1,61 = (12,18, 15,40)$$



**Figura 12.17** Gráfico de dispersão do MINITAB com intervalos de confiança e intervalos de previsão para os dados do Exemplo 12.13

A estreiteza desse intervalo sugere que dispomos de informações razoavelmente precisas sobre o valor médio estimado. Lembre-se de que, se recalcularmos esse intervalo amostra após amostra, a longo prazo, aproximadamente 95% dos intervalos calculados incluirão  $\beta_0 + \beta_1(45)$ . Apenas esperamos que esse valor médio fique no único intervalo que calculamos.

A Figura 12.18 mostra um resultado do MINITAB proveniente de um ajuste do modelo de regressão linear simples e de cálculo de intervalos de confiança para o valor médio da resistência nos níveis de 45 mm e 35 mm. Os intervalos estão na parte inferior do resultado; observe que o segundo intervalo é mais estreito que o primeiro, porque 35 está mais próximo de  $\bar{x}$  do que 45. A Figura 12.17 mostra (1) as curvas correspondentes aos limites de confiança para cada valor diferente de  $x$  e (2) os limites de previsão, a serem discutidos brevemente. Observe que as curvas ficam cada vez mais distantes uma da outra à medida que  $x$  se afasta de  $\bar{x}$ .

A equação de regressão é  $\text{resistência} = 27,2 - (\text{nível})0,298$

Previsor	Coef	DP	razão t	p
Constante	27,183	1,651	16,46	0,000
nível	-0,29756	0,04116	-7,23	0,000

s = 2,864    RQ = 76,6%    RQ(ajuste) = 75,1%

Análise de Variância

FONTE	GL	SQ	QM	F	p
Regressão	1	428,62	428,62	52,25	0,000
Erro	16	131,24	8,20		
Total	17	559,86			

Ajuste	Ajuste DP	IC 95,0%	IP 95,0%
13,793	0,758	(12,185, 15,401)	(7,510, 20,075)
Ajuste	Ajuste DP	IC 95,0%	IP 95,0%
16,768	0,678	(15,330, 18,207)	(10,527, 23,009)

Figura 12.18 Resultado da regressão no MINITAB para os dados do Exemplo 12.13

Em algumas situações, um IC é desejável não apenas para um único  $x$ , como para dois ou mais valores de  $x$ . Suponha que um pesquisador queira um IC tanto para  $\mu_{y_v}$  como para  $\mu_{y_w}$ , onde  $v$  e  $w$  são dois diferentes valores da variável independente. É tentador calcular o intervalo (12.6) primeiro para  $x = v$  e, em seguida,  $x = w$ . Suponha que usemos  $\alpha = 0,05$  em cada cálculo para obter dois intervalos de 95%. Logo, se as variáveis envolvidas no cálculo dos dois intervalos fossem independentes entre si, o coeficiente de confiança conjunto seria  $(0,95) \cdot (0,95) \approx 0,90$ .

Entretanto, os intervalos não são independentes porque os mesmos  $\hat{\beta}_0$ ,  $\hat{\beta}_1$  e  $S$  são usados em cada um. Portanto, não podemos afirmar que o nível de confiança conjunto dos dois intervalos seja exatamente de 90%. É possível mostrar, entretanto, que, se o IC de  $100(1 - \alpha)\%$  de (12.6) for calculado tanto para  $x = v$  como para  $x = w$  para obter ICs conjuntos para  $\mu_{y_v}$  e  $\mu_{y_w}$  então, o nível de confiança conjunto no par de intervalos resultante é de pelo menos  $100(1 - 2\alpha)\%$ . Trocando em miúdos, se usarmos  $\alpha = 0,05$ , obteremos um nível de confiança conjunto de pelo menos 90%, ao passo que, se usarmos  $\alpha = 0,01$ , obteremos um nível de confiança de pelo menos 98%. Para ilustrar, no Exemplo 12.13, o IC de 95% para  $\mu_{y_{45}}$  foi (12,185, 15,401) e o IC de 95% para  $\mu_{y_{35}}$  foi (15,330, 18,207). O nível de confiança simultâneo ou conjunto para as duas declarações  $12,185 < \mu_{y_{45}} < 15,401$  e  $15,330 < \mu_{y_{35}} < 18,207$  é de pelo menos 90%.

A validade desses ICs conjuntos ou simultâneos depende de um resultado probabilístico denominado **Desigualdade de Bonferroni**, de modo que os ICs conjuntos são chamados de **Intervalos de Bonferroni**. O método é facilmente generalizado para obter intervalos conjuntos para  $k$  diferentes  $\mu_{y_{x^*}}$ . Usando o intervalo (12.6), separadamente, primeiro para  $x = x_1^*$ , em seguida para  $x = x_2^*$ , . . . e, finalmente, para  $x = x_k^*$ , obtemos um conjunto de  $k$  ICs para o qual o nível de confiança conjunto ou simultâneo é, seguramente, de pelo menos  $100(1 - k\alpha)\%$ .

Os testes das hipóteses sobre  $\beta_0 + \beta_1 x^*$  baseiam-se na estatística de teste  $T$  obtida pela substituição de  $\beta_0 + \beta_1 x^*$  no numerador de (12.5) pelo valor nulo de  $\mu_0$ . Por exemplo:  $H_0: \beta_0 + \beta_1(45) = 15$  no Exemplo 12.13 expressa que, quando o nível de carbonação é 45, a resistência esperada (isto é, média real) é 15. O valor da estatística de teste é, portanto,  $t = [\hat{\beta}_0 + \hat{\beta}_1(45) - 15]/s_{\hat{\beta}_0 + \hat{\beta}_1(45)}$ , e o teste é de cauda superior, de cauda inferior ou bicaudal, de acordo com a desigualdade em  $H_a$ .

## Intervalo de Previsão para um Valor de $Y$ Futuro

Analogamente ao IC (12.6) para  $\mu_{Y,x^*}$ , geralmente se deseja obter um intervalo de valores plausíveis para o valor de  $Y$  associado com alguma informação futura, quando a variável independente tem o valor  $x^*$ . A título de ilustração, no exemplo em que o tamanho do vocabulário  $y$  está relacionado com a idade  $x$  de uma criança, para  $x = 6$  anos (12.6) forneceria um IC para o tamanho médio real do vocabulário de todas as crianças com 6 anos de idade. De modo alternativo, podemos querer um intervalo de valores plausíveis para o tamanho do vocabulário de uma determinada criança da mesma faixa etária.

Um IC se refere a um parâmetro, ou a uma característica da população, cujo valor é fixo mas desconhecido para nós. Em contraposição, um valor futuro de  $Y$  não é um parâmetro, mas uma variável aleatória; por esse motivo, referimo-nos a um intervalo de valores plausíveis para um  $Y$  futuro como um **intervalo de previsão**, em vez de um intervalo de confiança. O erro da estimativa é  $\beta_0 + \beta_1 x^* - (\hat{\beta}_0 + \hat{\beta}_1 x^*)$ , uma diferença entre uma quantidade fixa (mas desconhecida) e uma variável aleatória. O erro da previsão é  $Y - (\hat{\beta}_0 + \hat{\beta}_1 x^*)$ , uma diferença entre duas variáveis aleatórias. Portanto, a incerteza na previsão é maior que na estimativa, de modo que IP será maior que um IC. Pelo fato de o valor de  $Y$  futuro ser independente dos  $Y_i$ s observados,

$$\begin{aligned} V[Y - (\hat{\beta}_0 + \hat{\beta}_1 x^*)] &= \text{variância do erro de previsão} \\ &= V(Y) + V(\hat{\beta}_0 + \hat{\beta}_1 x^*) \\ &= \sigma^2 + \sigma^2 \left[ \frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}} \right] \\ &= \sigma^2 \left[ 1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}} \right] \end{aligned}$$

Além disso, pelo fato de  $E(Y) = \beta_0 + \beta_1 x^*$  e  $E(\hat{\beta}_0 + \hat{\beta}_1 x^*) = \beta_0 + \beta_1 x^*$ , o valor esperado do erro de previsão é  $E(Y - (\hat{\beta}_0 + \hat{\beta}_1 x^*)) = 0$ . Portanto, é possível mostrar que a variável padronizada

$$T = \frac{Y - (\hat{\beta}_0 + \hat{\beta}_1 x^*)}{S \sqrt{1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}}}$$

tem uma distribuição  $t$  com  $n - 2$  gl. Substituindo esse  $T$  na probabilidade  $P(-t_{\alpha/2, n-2} < T < t_{\alpha/2, n-2}) = 1 - \alpha$  e manuseando para isolar  $Y$  entre as duas desigualdades, obteremos o intervalo a seguir.

Um IP de  $100(1 - \alpha)\%$  para uma observação futura de  $Y$  a ser feita quando  $x = x^*$  é

$$\begin{aligned} &\hat{\beta}_0 + \hat{\beta}_1 x^* \pm t_{\alpha/2, n-2} \cdot s \sqrt{1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}} && (12.7) \\ &= \hat{\beta}_0 + \hat{\beta}_1 x^* \pm t_{\alpha/2, n-2} \cdot \sqrt{s^2 + s_{\hat{\beta}_0 + \hat{\beta}_1 x^*}^2} \\ &= \hat{y} \pm t_{\alpha/2, n-2} \cdot \sqrt{s^2 + s_{\hat{y}}^2} \end{aligned}$$

A interpretação do nível de previsão  $100(1 - \alpha)\%$  é idêntica ao da interpretação dos níveis de confiança anteriores – se (12.7) for usado repetidamente, a longo prazo os intervalos resultantes conterão de fato os valores  $y$  observados  $100(1 - \alpha)\%$  do tempo. Observe que o 1 dentro do símbolo inicial de raiz quadrada torna o IP (12.7) maior que o IC (12.6), embora os intervalos estejam ambos centralizados em  $\hat{\beta}_0 + \hat{\beta}_1 x^*$ . Além disso, à medida que  $n \rightarrow \infty$ , a largura de IC aproxima-se de 0, ao passo que a largura de IP, não (porque, mesmo se conheçêssemos perfeitamente  $\beta_0$  e  $\beta_1$ , ainda assim haveria incerteza na previsão).



**Exemplo 12.14**

Retornemos aos dados sobre o nível de carbonação–resistência do Exemplo 12.13 e calculemos um intervalo de confiança de 95% para o valor de resistência que pode resultar da seleção de um único espécime com um nível de carbonação de 45 mm. As quantidades pertinentes do exemplo são

$$\hat{y} = 13,79 \quad s_{\hat{y}} = 0,7582 \quad s = 2,8640$$

Para um nível de previsão de 95%, baseado em  $n - 2 = 16$  gl, o valor crítico  $t$  é 2,120, exatamente o que usamos previamente para um nível de confiança de 95%. O intervalo de previsão é, portanto,

$$\begin{aligned} 13,79 \pm (2,120)\sqrt{(2,8640)^2 + (0,7582)^2} &= 13,79 \pm (2,120)(2,963) \\ &= 13,79 \pm 6,28 = (7,51, 20,07) \end{aligned}$$

Os valores plausíveis para uma única observação sobre a resistência quando o nível é de 45 mm estão (no nível de previsão de 95%) entre 7,51 MPa e 20,07 MPa. O intervalo de confiança de 95% para uma resistência média, com um nível de 45, foi (12,18, 15,40). O intervalo de previsão é bem maior que esse por causa do  $(2,8640)^2$  suplementar sob a raiz quadrada. A Figura 12.18, o resultado do MINITAB no Exemplo 12.13, mostra esse intervalo e, igualmente, o intervalo de confiança. ■

A técnica de Bonferroni pode ser empregada como no caso dos intervalos de confiança. Se um IP de  $100(1 - \alpha)\%$  for calculado para cada  $k$  diferentes valores de  $x$ , o nível de previsão simultâneo ou conjunto para todos os  $k$  intervalos será de pelo menos  $100(1 - k\alpha)\%$ .

**Exercícios** | Seção 12.4 (44–56)

- 44. Ao ajustar o modelo de regressão linear simples para as  $n = 27$  observações sobre  $x =$  módulo de elasticidade e  $y =$  resistência à flexão, no Exercício 15 da Seção 12.2, obtivemos  $\hat{y} = 7,592$ ,  $s_{\hat{y}} = 0,179$  quando  $x = 40$  e  $\hat{y} = 9,741$ ,  $s_{\hat{y}} = 0,253$  para  $x = 60$ .
  - a. Explique por que  $s_{\hat{y}}$  é maior quando  $x = 60$  do que quando  $x = 40$ .
  - b. Calcule um intervalo de confiança com um nível de confiança de 95% para a resistência média real de todas as vigas cujo módulo de elasticidade é 40.
  - c. Calcule um intervalo de previsão com um nível de previsão de 95% para a resistência de uma única viga cujo módulo de elasticidade é 40.
  - d. Se for calculado um IC de 95% para a resistência média real quando o módulo de elasticidade é 60, qual será o nível de confiança simultâneo para ambos os intervalos calculados na parte (b)?
- 45. Reconsidere os dados sobre taxa de filtragem–quantidade de umidade introduzidos no Exemplo 12.6 (veja também o Exemplo 12.7).
  - a. Calcule um IC de 90% para  $\beta_0 + 125\beta_1$ , a quantidade média real de umidade quando a taxa de filtragem é 125.
  - b. Preveja o valor da quantidade de umidade para uma única operação experimental na qual a taxa de filtragem é 125 usando um nível de previsão de 90%. Como esse intervalo se compara ao intervalo da parte (a)? Por que isso ocorre?

- c. Em que os intervalos das partes (a) e (b) se comparam com o IC e o IP quando a taxa de filtragem é 115? Responda sem de fato calcular esses novos intervalos.
  - d. Interprete as hipóteses  $H_0: \beta_0 + 125\beta_1 = 80$  e  $H_0: \beta_0 + 125\beta_1 < 80$ , e, em seguida, realize um teste ao nível de significância 0,01.
46. O artigo “The Incorporation of Uranium and Silver by Hydrothermally Synthesized Galena” (*Econ. Geology*, 1964, p. 1003–1024) descreve a determinação do volume de prata nos cristais de galena formados em um sistema hidrotérmico acima de uma amplitude de temperatura. Sendo  $x =$  temperatura de cristalização em °C e  $y =$  Ag<sub>2</sub>S em %mol, os dados são:

$x$	398	292	352	575	568	450	550	408	484	350	503	600	600
$y$	0,15	0,05	0,23	0,43	0,23	0,40	0,44	0,44	0,45	0,09	0,59	0,63	0,60

- de onde  $\sum x_i = 6130$ ,  $\sum x_i^2 = 3.022.050$ ,  $\sum y_i = 4,73$ ,  $\sum y_i^2 = 2,1785$ ,  $\sum x_i y_i = 2418,74$ ,  $\hat{\beta}_1 = 0,00143$ ,  $\hat{\beta}_0 = -0,311$ , e  $s = 0,131$ .
- a. Estime o volume médio real de prata quando a temperatura é de 500°C usando um intervalo de confiança de 95%.
  - b. Como a amplitude de um IC de 95% para o volume médio real de prata (quando a temperatura é de 400°C) pode se comparar com a amplitude do intervalo na parte (a)? Responda sem calcular esse novo intervalo.

- c. Calcule um IC de 95% para uma mudança média real no volume de prata associada com um aumento de temperatura de 1°C.
- d. Suponha que se tenha acreditado antes que, quando a temperatura de cristalização fosse de 400°C, o volume médio real de prata seria de 0,25. Realize um teste ao nível de significância 0,05 para verificar se os dados amostrais contradizem esse ponto de vista prévio.

47. O modelo de regressão linear simples oferece um ótimo ajuste para os dados sobre volume de chuva e de *runoff* fornecidos no Exercício 16 da Seção 12.2. A equação da reta dos mínimos quadrados é  $\hat{y} = -1,128 + 0,82697x$ ,  $r^2 = 0,975$  e  $s = 5,24$ .

- a. Use a constatação de que  $s_{\hat{y}} = 1,44$  quando o volume de chuva é 40 m<sup>3</sup> para prever o volume de *runoff*, de maneira que transmita informações sobre confiabilidade e precisão. O intervalo resultante sugere que existem informações precisas sobre o valor de *runoff* para essa observação futura? Explique seu raciocínio.
- b. Calcule um IP para o volume de *runoff* quando o volume de chuva é 50, usando o mesmo nível de previsão da parte (a). O que se pode dizer sobre o nível de previsão simultâneo para os dois intervalos que você calculou?

48. Um determinado estudo relatado no artigo “The Effects of Water Vapor Concentration on the Rate of Combustion of an Artificial Graphite in Humid Air Flow” (*Combustion and Flame*, 1983, p. 107–118) forneceu dados sobre  $x$  = temperatura de um composto de nitrogênio–oxigênio (1000°F) sob circunstâncias específicas e  $y$  = difusividade do oxigênio. As quantidades são  $n = 9$ ,  $\sum x_i = 12,6$ ,  $\sum y_i = 27,68$ ,  $\sum x_i^2 = 18,24$ ,  $\sum x_i y_i = 40,968$ , e  $\sum y_i^2 = 93,3448$ . Suponha que as duas variáveis estejam relacionadas de acordo com um modelo de regressão linear simples.

- a. Estime a difusividade média real de oxigênio quando a temperatura é de 1500°F. Execute essa tarefa de maneira que transmita informações sobre confiabilidade e precisão.
- b. Presuma a difusividade de oxigênio para uma única observação feita quando a temperatura é de 1500°F, de forma que transmita informações sobre confiabilidade e precisão. Como a sua previsão calculada aqui se compara com a estimativa que você calculou na parte (a)?
- c. Um intervalo de previsão da difusividade a uma temperatura de 1200°F, usando o mesmo nível de previsão da parte (b), seria mais largo ou mais estreito que o intervalo calculado na parte (b)? Responda, sem calcular esse segundo intervalo.

49. Você é informado de que um IC de 95% do volume de chumbo esperado quando o fluxo do trânsito é 15, com base em uma amostra de  $n = 10$  observações, é (462,1, 597,7). Calcule um IC com um nível de confiança de

99% para um volume de chumbo esperado em que o fluxo do trânsito seja 15.

50. Um experimento para medir o tempo de relaxamento magnético macroscópico dos cristais ( $\mu s$ ) como uma função da resistência do campo magnético de polarização externa (kg) produziu os dados a seguir (“An Optical Faraday Rotation Technique for the Determination of Magnetic Relaxation Times”, *IEEE Trans. Magnetics*, junho de 1968, p. 175–178), os quais foram reproduzidos de um gráfico nesse artigo:

$x$	11,0	12,5	15,2	17,2	19,0	20,8
$y$	187	225	305	318	367	365

$x$	22,0	24,2	25,3	27,0	29,0
$y$	400	435	450	506	558

As estatísticas são  $\sum x_i = 223,2$ ,  $\sum y_i = 4116$ ,  $\sum x_i^2 = 4877,50$ ,  $\sum x_i y_i = 90.096,1$ , e  $\sum y_i^2 = 1.666.782$ . Calcule o seguinte:

- a. IC de 95% do tempo de relaxamento esperado quando a resistência do campo é igual a 18.
- b. IP de 95% do tempo de relaxamento futuro quando a resistência do campo é igual a 18.
- c. ICs simultâneos do tempo de relaxamento esperado quando a resistência do campo é igual a 15, 18 e 20; seu coeficiente de confiança conjunto deve ser de pelo menos 97%.

51. Reporte-se ao Exemplo 12.12, no qual  $x$  = fração do volume de óxidos/inclusões e  $y$  = % de elasticidade.

- a. O MINITAB fornece  $s_{\hat{\beta}_0 + \hat{\beta}_1(0,40)} = 0,0311$  e  $s_{\hat{\beta}_0 + \hat{\beta}_1(1,20)} = 0,0352$ . Por que o primeiro desvio padrão é menor que o segundo?
- b. Use o resultado do MINITAB do exemplo para calcular um IC de 95% de uma % de elasticidade esperada quando a fração do volume = 0,40.
- c. Use o resultado do MINITAB para calcular um IP de 95% de um único valor da % de elasticidade a se observar quando a fração do volume = 1,20.

52. A corrosão (ataque) por plasma (*plasma etching*) é essencial para a transferência de padrão de linha fina nos atuais processos dos semicondutores. O artigo “Ion Beam–Assisted Etching of Aluminum with Chlorine” (*J. Electrochem. Soc.*, 1985, p. 2010–2012) apresenta os dados a seguir (reproduzidos de um gráfico) sobre o fluxo de cloro ( $x$ , em SCCM) por meio de um bocal usado no mecanismo de deposição e corrosão e na taxa de ataque ( $y$ , em 100A/min.).

$x$	1,5	1,5	2,0	2,5	2,5	3,0	3,5	3,5	4,0
$y$	23,0	24,5	25,0	30,0	33,5	40,0	40,5	47,0	49,0

As estatísticas são  $\sum x_i = 24,0$ ,  $\sum y_i = 312,5$ ,  $\sum x_i^2 = 70,50$ ,  $\sum x_i y_i = 902,25$ ,  $\sum y_i^2 = 11.626,75$ ,  $\hat{\beta}_0 = 6,448718$ ,  $\hat{\beta}_1 = 10,602564$ .

- a. O modelo de regressão linear simples especifica uma relação útil entre o fluxo de cloro e a taxa de ataque?

- b. Estime a mudança média real na taxa de ataque associada com 1 SCCM de aumento na taxa de fluxo, usando um intervalo de confiança de 95% e interprete esse intervalo.
- c. Calcule um IC de 95% para  $\mu_{y,3,0}$ , a taxa média real de ataque quando o fluxo = 3,0. Essa média foi estimada de maneira precisa?
- d. Calcule um IP de 95% para uma única observação futura sobre a taxa de ataque, a ser feita quando o fluxo = 3,0. Existe alguma hipótese de a previsão ser precisa?
- e. O IC e o IP de 95%, quando o fluxo = 2,5, seriam mais amplos ou mais estreitos que os intervalos correspondentes das partes (c) e (d)? Responda, sem, de fato, calcular os intervalos.
- f. Você recomendaria que se calculasse um IP de 95% para um fluxo de 6,0? Explique.
53. Considere os quatro intervalos a seguir com base nos dados do Exemplo 12.4 (Seção 12.2):
- Um IC de 95% para uma porosidade média quando o peso unitário é 110.
  - Um IP de 95% para uma porosidade em que o peso unitário é 110.
  - Um IC de 95% para uma porosidade média quando o peso unitário é 115.
  - Um IP de 95% para uma porosidade em que o peso unitário é 115.
- Sem calcular nenhum desses intervalos, o que se pode dizer sobre as amplitudes relativas entre um e outro?
54. O declínio no fornecimento de água em determinadas áreas dos Estados Unidos criou a necessidade de se compreender melhor as relações entre fatores econômicos, como o volume da colheita, fatores hidrológicos e do solo. O artigo “Variability of Soil Water Properties and Crop Yield in a Sloped Watershed” (*Water Resources Bull.*, 1988, p. 281–288) apresenta dados sobre a produção de sorgo ( $y$ , em g/m linear) e distância ascendente ( $x$ , em m) em uma bacia inclinada. Observações selecionadas são apresentadas na tabela a seguir.

$x$	0	10	20	30	45	50	70
$y$	500	590	410	470	450	480	510

$x$	80	100	120	140	160	170	190
$y$	450	360	400	300	410	280	350

- Elabore um gráfico de dispersão. Um modelo de regressão linear simples parece plausível?
  - Realize um teste de utilidade do modelo.
  - Estime a produção média real quando a distância ascendente é 75 dando um intervalo de valores plausíveis.
55. A infestação de insetos na colheita há muito tem sido motivo de grande preocupação de fazendeiros e cientistas agrícolas. O artigo “Cotton Square Damage by the Plant Bug, *Lygus hesperus*, and Abscission Rates” (*J. of Econ. Entom.*, 1988, p. 1328–1337) apresenta dados sobre  $x$  = período de vida de uma plantaçao de algodão (dias) e  $y$  = % de áreas prejudicadas. Considere as  $n = 12$  observações a seguir (reproduzidas de um gráfico do artigo).
- |     |    |    |    |    |    |    |
|-----|----|----|----|----|----|----|
| $x$ | 9  | 12 | 12 | 15 | 18 | 18 |
| $y$ | 11 | 12 | 23 | 30 | 29 | 52 |
- |     |    |    |    |    |    |    |
|-----|----|----|----|----|----|----|
| $x$ | 21 | 21 | 27 | 30 | 30 | 33 |
| $y$ | 41 | 65 | 60 | 72 | 84 | 93 |
- Por que a relação entre  $x$  e  $y$  não é determinística?
  - O gráfico de dispersão indica que o modelo de regressão linear simples descreverá a relação entre as duas variáveis?
  - As estatísticas são  $\sum x_i = 246$ ,  $\sum x_i^2 = 5742$ ,  $\sum y_i = 572$ ,  $\sum y_i^2 = 35.634$ , e  $\sum x_i y_i = 14.022$ . Determine a equação da reta dos mínimos quadrados.
  - Preveja a porcentagem de quadras prejudicadas para um período de vida de 20 dias, fornecendo um intervalo de valores plausíveis.
56. Verifique se  $V(\hat{\beta}_0 + \hat{\beta}_1 x)$  é de fato fornecida pela expressão no texto. [Sugestão:  $V(\sum d_i Y_i) = \sum d_i^2 \cdot V(Y_i)$ .]

## 12.5 | Correlação

Há várias situações em que o objetivo de estudar o comportamento conjunto de duas variáveis é verificar se elas estão relacionadas, e não usar uma para prever o valor da outra. Nesta seção, primeiro desenvolvemos o coeficiente de correlação amostral  $r$  como uma medida de quão fortemente as duas variáveis  $x$  e  $y$  estão relacionadas em uma amostra e, em seguida, relacionamos  $r$  ao coeficiente de correlação  $\rho$  definido no Capítulo 5.

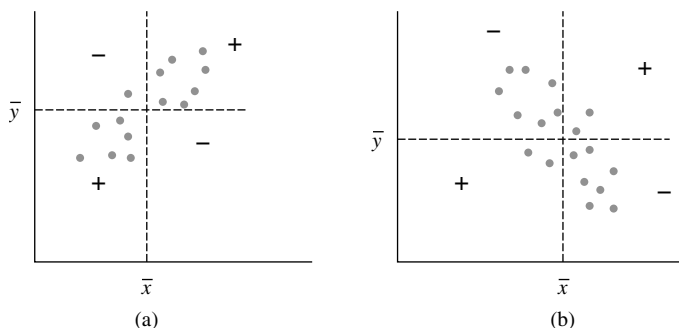
### O Coeficiente de Correlação Amostral $r$

Dados os  $n$  pares de observações  $(x_1, y_1)$ ,  $(x_2, y_2)$ , ...,  $(x_n, y_n)$ , é natural que se fale de  $x$  e  $y$  como tendo uma relação positiva, se  $x$ s grandes estiverem pareados com  $y$ s grandes e  $x$ s pequenos com  $y$ s pequenos. De modo

semelhante, se  $x$ s grandes estiverem pareados com  $y$ s pequenos e  $x$ s pequenos com  $y$ s grandes, então está implícita uma relação negativa entre as variáveis. Considere a quantidade

$$S_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^n x_i y_i - \left( \sum_{i=1}^n x_i \right) \left( \sum_{i=1}^n y_i \right) / n$$

Logo, se a relação for fortemente positiva, um  $x_i$  acima da média  $\bar{x}$  tenderá a estar pareado com um  $y$ , acima da média  $\bar{y}$ , de modo que  $(x_i - \bar{x})(y_i - \bar{y}) > 0$ , e esse produto será também positivo quando ambos,  $x_i$  e  $y_i$ , estiverem abaixo das respectivas médias. Portanto, uma relação positiva significa que  $S_{xy}$  será positiva. Um argumento análogo mostra que, quando a relação for negativa,  $S_{xy}$  será negativa, visto que a maioria dos produtos  $(x_i - \bar{x})(y_i - \bar{y})$  será negativa. Isso está ilustrado na Figura 12.19.



**Figura 12.19** (a) Gráfico de dispersão com  $S_{xy}$  positiva; (b) gráfico de dispersão com  $S_{xy}$  negativa [+ médias  $(x_i - \bar{x})(y_i - \bar{y}) > 0$ , e - médias  $(x_i - \bar{x})(y_i - \bar{y}) < 0$ ]

Embora  $S_{xy}$  pareça uma medida plausível da força de uma relação, ainda não temos nenhuma idéia sobre quão positiva ou negativa ela pode ser. Infelizmente,  $S_{xy}$  tem um sério defeito: ao mudar as unidades de medida de  $x$  ou  $y$ ,  $S_{xy}$  pode ser arbitrariamente grande em magnitude ou arbitrariamente próxima de zero. Por exemplo: se  $S_{xy} = 25$  quando  $x$  é medido em metros,  $S_{xy} = 25.000$  quando  $x$  é medido em milímetros e  $0,025$  quando  $x$  é expresso em quilômetros. Uma condição razoável a ser imposta a qualquer medida sobre quão forte  $x$  e  $y$  estão relacionados não deve depender das unidades particulares usadas para medi-los. Essa condição é alcançada modificando-se  $S_{xy}$  para obter o coeficiente de correlação amostral.

**DEFINIÇÃO**

O coeficiente de correlação amostral de  $n$  pares  $(x_1, y_1), \dots, (x_n, y_n)$  é

$$r = \frac{S_{xy}}{\sqrt{\sum(x_i - \bar{x})^2} \sqrt{\sum(y_i - \bar{y})^2}} = \frac{S_{xy}}{\sqrt{S_{xx}} \sqrt{S_{yy}}} \tag{12.8}$$

**Exemplo 12.15**

Uma avaliação precisa sobre a produtividade do solo é fundamental para o planejamento racional do uso da terra. Infelizmente, como afirma o autor do artigo “Productivity Ratings Based on Soil Series” (*Prof. Geographer*, 1980, p. 158–163), não é tão fácil de estabelecer um índice de produtividade do solo aceitável. Uma das dificuldades é que a produtividade é determinada parcialmente pela cultura plantada, e a relação entre a produção de duas diferentes culturas plantadas no mesmo solo pode não ser muito forte. Para ilustrar, esse artigo apresenta os dados a seguir sobre a produção de milho  $x$  e a produção de amendoim  $y$  (mT/Ha) de oito diferentes tipos de solo.

$x$	2,4	3,4	4,6	3,7	2,2	3,3	4,0	2,1
$y$	1,33	2,12	1,80	1,65	2,00	1,76	2,11	1,63

Com  $\sum x_i = 25,7$ ,  $\sum y_i = 14,40$ ,  $\sum x_i^2 = 88,31$ ,  $\sum x_i y_i = 46,856$ , e  $\sum y_i^2 = 26,4324$ ,

$$S_{xx} = 88,31 - \frac{(25,7)^2}{8} = 88,31 - 82,56 = 5,75$$

$$S_{yy} = 26,4324 - \frac{(14,40)^2}{8} = 0,5124$$

$$S_{xy} = 46,856 - \frac{(25,7)(14,40)}{8} = 0,5960$$

de onde

$$r = \frac{0,5960}{\sqrt{5,75}\sqrt{0,5124}} = 0,347$$

## Propriedades de $r$

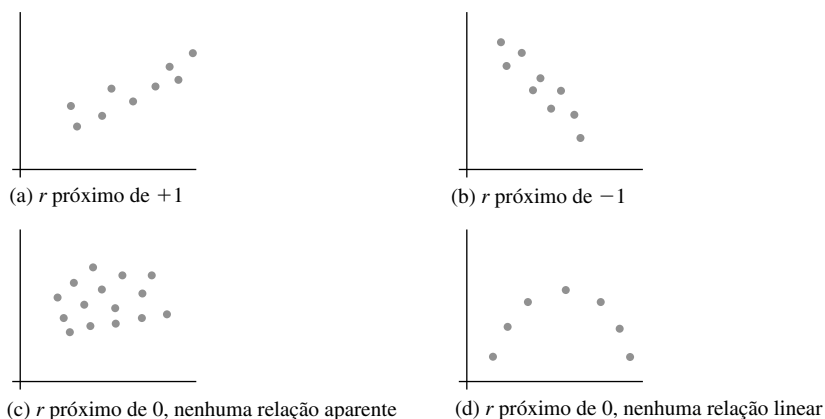
As propriedades mais importantes de  $r$  são as seguintes:

1. O valor de  $r$  não depende de qual das duas variáveis em estudo é chamada de  $x$  e qual é chamada de  $y$ .
2. O valor de  $r$  independe das unidades com as quais  $x$  e  $y$  são medidos.
3.  $-1 \leq r \leq 1$
4.  $r = 1$  se, e somente se, todos os pares  $(x_i, y_i)$  estiverem alinhados em linha reta com um coeficiente angular positivo e  $r = -1$  se, e somente se, todos os pares  $(x_i, y_i)$  estiverem alinhados com um coeficiente angular negativo.
5. O quadrado do coeficiente de correlação amostral fornece o valor do coeficiente de determinação que resultaria de um ajuste do modelo de regressão linear simples – em símbolos,  $(r)^2 = r^2$ .

A discrepância da Propriedade 1 é marcante em relação ao que ocorre na análise de regressão, na qual praticamente todas as quantidades de interesse (o coeficiente angular estimado, o termo constante  $y$  estimado,  $s^2$  etc.) dependem de qual das duas variáveis é considerada a variável dependente. Contudo, a Propriedade 5 mostra que a proporção de variação na variável dependente explicada pelo ajuste do modelo de regressão linear simples não depende de qual variável desempenha esse papel.

A Propriedade 2 equivale a dizer que  $r$  não muda se cada  $x_i$  for substituído por  $cx_i$  e cada  $y_i$  for substituído por  $dy_i$  (uma mudança na escala de medida), tanto como se cada  $x_i$  for substituído por  $x_i - a$  e  $y_i$  por  $y_i - b$  (o que muda o lugar de zero no eixo de medida). Isso significa, por exemplo, que  $r$  é o mesmo, independentemente de a temperatura ser medida em °F ou °C.

A Propriedade 3 expressa que o valor máximo de  $r$ , correspondente ao maior grau possível de relação positiva, é  $r = 1$ , ao passo que a relação mais negativa é identificada com  $r = -1$ . De acordo com a Propriedade 4, as maiores correlações positivas e negativas são alcançadas somente quando todos os pontos estendem-se sobre uma linha reta. Qualquer outra configuração de pontos, mesmo se a configuração sugerir uma relação determinística entre as variáveis, produzirá um valor  $r$  menor que 1 em magnitude absoluta. Portanto,  $r$  mede o grau de relação linear entre variáveis. Um valor de  $r$  próximo de 0 não é uma evidência de que não existe uma relação forte, mas apenas de que falta uma relação linear, de modo que esse valor de  $r$  deve ser interpretado com cuidado. A Figura 12.20 ilustra várias configurações de pontos associados a diferentes valores de  $r$ .



**Figura 12.20** Gráficos de dados para diferentes valores de  $r$

Uma pergunta freqüente é: “Quando é possível dizer que existe uma correlação forte entre as variáveis e quando a correlação é fraca?”. Uma regra prática razoável é afirmar que a correlação é fraca se  $0 \leq |r| \leq 0,5$ , forte se  $0,8 \leq |r| \leq 1$  e moderada em caso contrário. Talvez surpreenda o fato de que  $r = 0,5$  é considerado fraco, mas  $r^2 = 0,25$  significa que, em uma regressão de  $y$  sobre  $x$ , apenas 25% da variação  $y$  observada seria explicada pelo modelo. No Exemplo 12.15, a correlação entre a produção de milho e a produção de amendoim seria considerada fraca.

## O Coeficiente de Correção Populacional $\rho$ e Inferências sobre Correlação

O coeficiente de correlação  $r$  é uma medida de quão fortemente  $x$  e  $y$  estão relacionados na amostra observada. Podemos imaginar os pares  $(x_i, y_i)$  como sendo tirados de uma população de pares bivariados, em que  $(X, Y)$  tem uma distribuição de probabilidades conjunta  $f(x, y)$ . No Capítulo 5, definimos o coeficiente de correlação  $\rho(X, Y)$  como

$$\rho = \rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_x \cdot \sigma_y}$$

onde

$$\text{Cov}(X, Y) = \begin{cases} \sum_x \sum_y (x - \mu_x)(y - \mu_y)f(x, y) & (X, Y) \text{ discreto} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_x)(y - \mu_y)f(x, y) dx dy & (X, Y) \text{ contínuo} \end{cases}$$

Se considerarmos  $f(x, y)$  como a representação da distribuição de pares de valores dentro da população global,  $\rho$  torna-se uma medida de quão  $x$  e  $y$  estão fortemente relacionados nessa população. As propriedades de  $\rho$  análogas às de  $r$  foram apresentadas no Capítulo 5.

O coeficiente de correlação populacional  $\rho$  é um parâmetro ou uma característica da população, assim como  $\mu_x$ ,  $\mu_y$ ,  $\sigma_x$  e  $\sigma_y$ , de modo que podemos usar o coeficiente de correlação amostral para fazer várias inferências sobre  $\rho$ . Em particular,  $r$  é uma estimativa pontual para  $\rho$  e o estimador correspondente é

$$\hat{\rho} = R = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2} \sqrt{\sum(Y_i - \bar{Y})^2}}$$

### Exemplo 12.16

Em alguns lugares, existe uma associação significativa entre a concentração de dois diferentes poluentes. O artigo “The Carbon Component of the Los Angeles Aerosol: Source Apportionment and Contributions to the Visibility Budget” (*J. Air Pollution Control Fed.*, 1984, p. 643–650) apresenta os dados a seguir sobre a concentração de ozônio  $x$  (ppm) e a concentração secundária de carbono  $y$  ( $\mu\text{g}/\text{m}^3$ ).

$x$	0,066	0,088	0,120	0,050	0,162	0,186	0,057	0,100
$y$	4,6	11,6	9,5	6,3	13,8	15,4	2,5	11,8
$x$	0,112	0,055	0,154	0,074	0,111	0,140	0,071	0,110
$y$	8,0	7,0	20,6	16,6	9,2	17,9	2,8	13,0

As quantidades são  $n = 16$ ,  $\sum x_i = 1,656$ ,  $\sum y_i = 170,6$ ,  $\sum x_i^2 = 0,196912$ ,  $\sum x_i y_i = 20,0397$  e  $\sum y_i^2 = 2253,56$ , de onde

$$\begin{aligned} r &= \frac{20,0397 - (1,656)(170,6)/16}{\sqrt{0,196912 - (1,656)^2/16} \sqrt{2253,56 - (170,6)^2/16}} \\ &= \frac{2,3826}{(0,1597)(20,8456)} = 0,716 \end{aligned}$$

A estimativa pontual do coeficiente de correlação populacional  $\rho$  entre a concentração de ozônio e a concentração secundária de carbono é  $\hat{\rho} = r = 0,716$ . ■

Os intervalos para amostras pequenas e os procedimentos de teste apresentados nos capítulos 7–9 basearam-se na suposição de normalidade da população. Para testar hipóteses sobre  $\rho$ , devemos fazer uma suposição análoga sobre a distribuição de pares de valores  $(x, y)$  na população. No momento, estamos supondo que tanto  $X$  como  $Y$  são aleatórios, embora grande parte de nosso trabalho de regressão concentre-se no  $x$  fixado pelo pesquisador.

#### SUPOSIÇÃO

A distribuição de probabilidades conjunta de  $(X, Y)$  é definida por

$$\begin{aligned} f(x, y) &= \frac{1}{2\pi \cdot \sigma_1 \sigma_2 \sqrt{1 - \rho^2}} e^{-\frac{1}{2(1 - \rho^2)} \left[ \frac{(x - \mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x - \mu_1)(y - \mu_2)}{\sigma_1 \sigma_2} + \frac{(y - \mu_2)^2}{\sigma_2^2} \right]} \\ &\quad -\infty < x < \infty \\ &\quad -\infty < y < \infty \end{aligned} \tag{12.9}$$

onde  $\mu_1$  e  $\sigma_1$  são a média e o desvio padrão de  $X$  e  $\mu_2$  e  $\sigma_2$  são a média e o desvio padrão de  $Y$ ;  $f(x, y)$  é denominado **distribuição de probabilidades normal bivariada**.

A distribuição normal bivariada é, obviamente, muito complexa, mas, para os nossos propósitos, precisamos nos familiarizar apenas ligeiramente com suas várias propriedades. A superfície determinada por  $f(x, y)$  localiza-se totalmente acima do plano  $x, y$  [ $f(x, y) \geq 0$ ] e tem uma aparência tridimensional em forma de sino ou de montanha, como ilustrado na Figura 12.21. Se fatiarmos a superfície, cortando-a com qualquer plano perpendicular ao plano  $x, y$ , e observarmos a forma da curva traçada no “plano de corte”, o resultado será uma curva normal. Mais precisamente, se  $X = x$ , é possível mostrar que a distribuição (condicional) de  $Y$  é normal com média  $\mu_{Y|x} = \mu_2 - \rho\mu_1\sigma_2/\sigma_1 + \rho\sigma_2x/\sigma_1$  e variância  $(1 - \rho^2)\sigma_2^2$ . Esse é exatamente o modelo usado na regressão linear simples com  $\beta_0 = \mu_2 - \rho\mu_1\sigma_2/\sigma_1$ ,  $\beta_1 = \rho\sigma_2/\sigma_1$ , e  $\sigma^2 = (1 - \rho^2)\sigma_2^2$  independentes de  $x$ . Isso significa que, se os pares  $(x_i, y_i)$  observados forem selecionados de fato de uma distribuição normal bivariada, então o modelo de regressão linear simples é uma forma apropriada de estudar o comportamento de  $Y$  para um  $x$  fixo. Se  $\rho = 0$ , então  $\mu_{Y|x} = \mu_2$  é independente de  $x$ ; na verdade, quando  $\rho = 0$ , a função densidade de probabilidade conjunta  $f(x, y)$  de (12.9) pode ser fatorada em uma parte envolvendo apenas  $x$  e uma parte envolvendo apenas  $y$ , o que significa que  $X$  e  $Y$  são variáveis independentes.

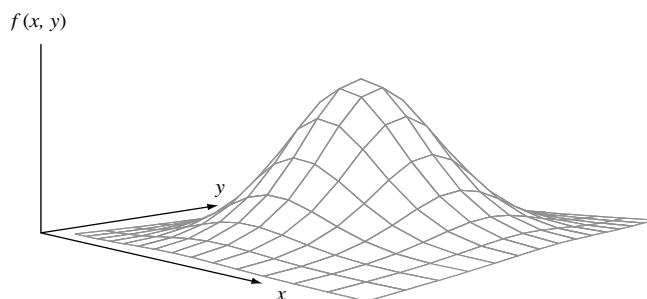


Figura 12.21 Um gráfico da fdp normal

Supondo que os pares são traçados de uma distribuição normal bivariada permite-nos testar as hipóteses sobre  $\rho$  e construir um IC. Não há nenhuma maneira plenamente satisfatória de verificar a plausibilidade da suposição de normalidade bivariada. Em uma verificação parcial, consiste em construir dois gráficos de probabilidade distintos, um para a amostra dos  $x_i$ s e outro para a amostra dos  $y_i$ s, visto que a normalidade bivariada significa que as distribuições marginais tanto de  $X$  como de  $Y$  são normais. Se ambos os gráficos desviarem substancialmente de um padrão de linha reta, os procedimentos inferenciais a seguir não devem ser usados quando o tamanho da amostra  $n$  for pequeno.

#### Verificação da ausência de correlação

Quando  $H_0: \rho = 0$  é verdadeira, a estatística de teste

$$T = \frac{R\sqrt{n-2}}{\sqrt{1-R^2}}$$

tem uma distribuição  $t$  com  $n - 2$  gl.

#### Hipótese Alternativa

$$H_a: \rho > 0$$

$$H_a: \rho < 0$$

$$H_a: \rho \neq 0$$

#### Região de Rejeição para um Teste de Nível $\alpha$

$$t \geq t_{\alpha, n-2}$$

$$t \leq -t_{\alpha, n-2}$$

$$\text{ou } t \geq t_{\alpha/2, n-2} \text{ ou } t \leq -t_{\alpha/2, n-2}$$

Um valor  $P$  baseado em  $n - 2$  gl pode ser calculado como descrito anteriormente.

#### Exemplo 12.17

Os efeitos neurotóxicos do manganês são bem conhecidos e em geral decorrem de um alto nível de exposição ocupacional durante longos períodos de tempo. Na área de higiene do trabalho e higiene ambiental, a relação entre a peroxidação lipídica, que é responsável pela deterioração de alimentos e danos a tecidos vivos, e a exposição ocupacional ainda não foram relatadas. O artigo “Lipid Peroxidation in Workers Exposed to Manganese” (*Scand. J. Work and Environ. Health*, 1996, p. 381–386) forneceu dados sobre  $x$  = concentração de manganês no sangue (ppb) e  $y$  = concentração ( $\mu\text{mol/L}$ ) de malondialdeído, um produto estável de peroxidação lipídica, ambos para uma amostra de 22 trabalhadores expostos ao manganês e para uma amostra de controle de 45 indivíduos. O valor de  $r$  da amostra de controle era 0,29, de onde

$$t = \frac{(0,29)\sqrt{45-2}}{\sqrt{1-(0,29)^2}} \approx 2,0$$

O valor  $P$  correspondente de um teste  $t$  bicaudal baseado em 43 gl é aproximadamente 0,052 (o artigo mencionado relatou apenas que o valor  $P > 0,05$ ). Não queremos rejeitar a alegação de que  $\rho = 0$  tanto no nível 0,01 como no nível 0,05. Em relação à amostra de trabalhadores expostos,  $r = 0,83$  e  $t \approx 6,7$ , uma clara evidência de que existe uma relação linear na população global de trabalhadores expostos da qual foi selecionada a amostra. ■

Pelo fato de  $\rho$  medir a extensão até o ponto em que existe uma relação linear entre as duas variáveis na população, a hipótese nula  $H_0: \rho = 0$  expressa que não há relação alguma com a população. Na Seção 12.3, usamos a razão  $t\hat{\beta}_1/s_{\hat{\beta}_1}$  para verificar se existia uma relação linear entre as duas variáveis no contexto da análise de regressão. Verifica-se que os dois procedimentos de teste são completamente equivalentes porque  $r\sqrt{n-2}/\sqrt{1-r^2} = \hat{\beta}_1/s_{\hat{\beta}_1}$ . Quando o que se quer é apenas avaliar a intensidade de qualquer relação linear, em vez de ajustar um modelo e usá-lo para fazer uma estimativa ou previsão, a fórmula da estatística de teste que acabamos de apresentar exige uma quantidade menor de cálculos do que a razão  $t$ .



## Outras Inferências sobre $\rho$

O procedimento para testar  $H_0: \rho = \rho_0$  quando  $\rho_0 \neq 0$  não é equivalente a qualquer procedimento da análise de regressão. A estatística de teste baseia-se em uma transformação de  $R$  denominada transformação de Fisher.

PROPOSIÇÃO

Quando  $(X_1, Y_1), \dots, (X_n, Y_n)$  é uma amostra de uma distribuição normal bivariada, a va

$$V = \frac{1}{2} \ln\left(\frac{1+R}{1-R}\right) \quad (12.10)$$

tem aproximadamente uma distribuição normal com média e variância

$$\mu_V = \frac{1}{2} \ln\left(\frac{1+\rho}{1-\rho}\right) \quad \sigma_V^2 = \frac{1}{n-3}$$

O raciocínio lógico para a transformação é obter uma função de  $R$  cuja variância seja independente de  $\rho$ ; esse não seria o caso de  $R$  propriamente. Além disso, a transformação não deve ser usada se  $n$  for pequeno, visto que a aproximação não será válida.

A estatística de teste para  $H_0: \rho = \rho_0$  é

$$Z = \frac{V - \frac{1}{2} \ln[(1+\rho_0)/(1-\rho_0)]}{1/\sqrt{n-3}}$$

Hipótese Alternativa

Região de Rejeição para um Teste de Nível  $\alpha$

$$H_a: \rho > \rho_0$$

$$z \geq z_\alpha$$

$$H_a: \rho < \rho_0$$

$$z \leq -z_\alpha$$

$$H_a: \rho \neq \rho_0$$

$$\text{ou } z \geq z_{\alpha/2} \text{ ou } z \leq -z_{\alpha/2}$$

Um valor  $P$  pode ser calculado da mesma maneira que para os testes  $z$  anteriores.

### Exemplo 12.18

O artigo “A Study of a Partial Nutrient Removal System for Wastewater Treatment Plants” (*Water Research*, 1972, p. 1389–1397) relatou um método de remoção de nitrogênio que compreende o tratamento do sobrenadante de um digestor aeróbico. Tanto o nitrogênio influente total  $x$  (mg/L) como a porcentagem  $y$  de nitrogênio removida foram registrados durante 20 dias, sendo as estatísticas  $\sum x_i = 285,90$ ,  $\sum x_i^2 = 4409,55$ ,  $\sum y_i = 690,30$ ,  $\sum y_i^2 = 29.040,29$  e  $\sum x_i y_i = 10.818,56$ . Os dados indicam que o nitrogênio influente total e a porcentagem de nitrogênio removida estão correlacionados de maneira positiva, pelo menos moderadamente?

Nossa interpretação anterior da correlação positiva moderada foi  $0,5 < \rho < 0,8$ , de modo que desejamos testar  $H_0: \rho = 0,5$  versus  $H_a: \rho > 0,5$ . O valor calculado de  $r$  é 0,733, de modo que

$$\frac{1}{2} \ln\left(\frac{1+0,733}{1-0,733}\right) = 0,935 \quad \frac{1}{2} \ln\left(\frac{1+0,5}{1-0,5}\right) = 0,549$$

Isso resulta em  $z = (0,935 - 0,549)\sqrt{17} = 1,59$ . Visto que  $1,59 < 1,645$ , no nível 0,05, não podemos concluir que  $\rho > 0,5$ , de maneira que a relação não se mostrou nem mesmo moderadamente forte (uma conclusão até certo ponto surpreendente, visto que  $r = 0,73$ ; porém, quando  $n$  é pequeno, é possível produzir um  $r$  grande mesmo quando  $\rho$  é pequeno). ■

Para obter um IC para  $\rho$ , primeiro derivamos um intervalo para  $\mu_v = \frac{1}{2} \ln[(1 + \rho)/(1 - \rho)]$ . Padronizando  $V$ , escrevendo uma declaração de probabilidade e manipulando as desigualdades resultantes, obtemos

$$\left( v - \frac{z_{\alpha/2}}{\sqrt{n-3}}, v + \frac{z_{\alpha/2}}{\sqrt{n-3}} \right) \tag{12.11}$$

como um intervalo  $100(1 - \alpha)\%$  para  $\mu_v$ , onde  $v = \frac{1}{2} \ln[(1 + r)/(1 - r)]$ . Esse intervalo pode ser manipulado para produzir um IC para  $\rho$ .

Um intervalo de confiança  $100(1 - \alpha)\%$  para  $\rho$  é

$$\left( \frac{e^{2c_1} - 1}{e^{2c_1} + 1}, \frac{e^{2c_2} - 1}{e^{2c_2} + 1} \right)$$

onde  $c_1$  e  $c_2$  são os pontos finais esquerdo e direito, respectivamente, do intervalo (12.11).

**Exemplo 12.19** (Continuação do Exemplo 12.18)

O coeficiente de correlação amostral entre o nitrogênio influente e a porcentagem de nitrogênio removida foi  $r = 0,733$ , o que resulta em  $v = 0,935$ . Com  $n = 20$ , um intervalo de confiança de 95% para  $\mu_v$  é  $(0,935 - 1,96/\sqrt{17}, 0,935 + 1,96/\sqrt{17}) = (0,460, 1,410) = (c_1, c_2)$ . O intervalo de 95% para  $\rho$  é

$$\left[ \frac{e^{2(0,46)} - 1}{e^{2(0,46)} + 1}, \frac{e^{2(1,41)} - 1}{e^{2(1,41)} + 1} \right] = (0,43, 0,89) \quad \blacksquare$$

No Capítulo 5, advertimos que um coeficiente de correlação grande (próximo de 1 ou -1) significa apenas associação, e não relação de causa e efeito. Isso se aplica tanto a  $\rho$  quanto a  $r$ .

---

**Exercícios** | Seção 12.5 (57–67)

- 57.** O artigo “Behavioral Effects of Mobile Telephone Use During Simulated Driving” (*Ergonomics*, 1995, p. 2536–2562) relatou que, para uma amostra de 20 pessoas, o coeficiente de correlação amostral de  $x =$  idade e  $y =$  tempo decorrido desde que o indivíduo tirou sua carteira de habilitação (ano) foi 0,97. Por que acreditamos que o valor de  $r$  é tão próximo de 1? (O autor desse artigo apresentou seu argumento.)
- 58.** O Teste de Oxidação do Óleo de Turbina (TOOT) e o Teste de Oxidação da Bomba de Rotação (TOBR) são dois diferentes procedimentos de avaliação da estabilidade da oxidação do óleo de turbina a vapor. O artigo “Dependence of Oxidation Stability of Steam Turbine Oil on Base Oil Composition” (*J. of the Society of Tribologists and Lubrication Engrs.*, outubro de 1997, p. 19–24) relatou as observações a seguir sobre  $x =$  tempo do TOOT (h) e  $y =$  tempo do TOBR (min.) de 12 amostras de óleo.

<b>TOOT</b>	4200	3600	3750	3675	4050	2770
<b>TOBR</b>	370	340	375	310	350	200
<b>TOOT</b>	4870	4500	3450	2700	3750	3300
<b>TOBR</b>	400	375	285	225	345	285

- Calcule e interprete o valor do coeficiente de correlação amostral (como o fizeram os autores do artigo).
  - Como o valor de  $r$  seria influenciado se tivéssemos representado  $x =$  tempo do TOBR e  $y =$  tempo do TOOT?
  - Como o valor de  $r$  seria influenciado se o tempo TOBR fosse expresso em horas?
  - Elabore gráficos de probabilidade normal e comente-os.
  - Teste as hipóteses para verificar se o tempo do TOBR e o tempo do TOOT estão relacionados linearmente.
- 59.** A rigidez e a fibrosidade do aspargo são os principais determinantes de sua qualidade. Esse foi o objeto de estudo relatado em “Post-Harvest Glyphosphate Applica-

tion Reduces Toughening, Fiber Content, and Lignification of Stored Asparagus Spears” (*J. of the Amer. Soc. of Horticultural Science*, 1988, p. 569–572). Esse artigo relatou os dados a seguir (reproduzidos de um gráfico) sobre  $x$  = força de corte (kg) e  $y$  = porcentagem do peso seco da fibra.

$x$	46	48	55	57	60	72	81	85	94
$y$	2,18	2,10	2,13	2,28	2,34	2,53	2,28	2,62	2,63
$x$	109	121	132	137	148	149	184	185	187
$y$	2,50	2,66	2,79	2,80	3,01	2,98	3,34	3,49	3,26

$$n = 18 \quad \sum x_i = 1950 \quad \sum x_i^2 = 251,970$$

$$\sum y_i = 47,92 \quad \sum y_i^2 = 130,6074 \quad \sum x_i y_i = 5530,92$$

- a. Calcule o valor do coeficiente de correlação amostral. Com base nesse valor, como você descreveria a natureza da relação entre as duas variáveis?
  - b. Se uma primeira amostra tiver um valor maior de força de corte do que uma segunda amostra, que porcentagem do peso seco da fibra tende a ser real para as duas amostras?
  - c. Se a força de corte for expressa em libras, o que acontecerá ao valor de  $r$ ? Por quê?
  - d. Se o modelo de regressão linear simples fosse ajustado a esses dados, que proporção da variação observada na porcentagem do peso seco da fibra pode ser explicada pela relação do modelo?
  - e. Faça um teste no nível de significância 0,01 para verificar se existe uma associação linear positiva entre as duas variáveis.
60. O artigo “A Dual–Buffer Titration Method for Lime Requirement of Acid Mine–soils” (*J. of Environ. Qual.*, 1988, p. 452–456) relata os resultados de um estudo relacionado com a revegetação do solo em áreas de recuperação dos mananciais. Com  $x$  = alumínio extraível de KCl e  $y$  = quantidade de cal necessária para tornar o nível de pH do solo igual a 7,0, os dados do artigo produziram as seguintes estatísticas:  $n = 24$ ,  $\sum x = 48,15$ ,  $\sum x^2 = 155,4685$ ,  $\sum y = 263,5$ ,  $\sum y^2 = 3750,53$  e  $\sum xy = 658,455$ . Realize um teste no nível de significância 0,01 para verificar se o coeficiente de correlação populacional é de alguma maneira diferente de 0.
61. Os autores do artigo “Objective Effects of a Six Months' Endurance and Strength Training Program in Outpatients with Congestive Heart Failure” (*Medicine and Science in Sports and Exercise*, 1999, p. 1102–1107) apresentaram uma análise de correlação para investigar a relação entre o nível máximo de lactato  $x$  e a resistência muscular  $y$ . Os dados a seguir foram reproduzidos de um gráfico desse artigo.

$x$	400	750	770	800	850	1025	1200
$y$	3,80	4,00	4,90	5,20	4,00	3,50	6,30

$x$	1250	1300	1400	1475	1480	1505	2200
$y$	6,88	7,55	4,95	7,80	4,45	6,60	8,90

$S_{xx} = 36,9839$ ,  $S_{yy} = 2.628.930,357$ ,  $S_{xy} = 7377,704$ . O gráfico de dispersão mostra um padrão linear.

- a. Verifique se existe uma correlação positiva entre o nível máximo de lactato e a resistência muscular na população da qual esses dados foram selecionados.
  - b. Se uma análise de regressão tivesse de ser realizada para prever a resistência muscular com base no nível de lactato, que proporção de variação observada na resistência podia ser atribuída à relação linear aproximada? Responda a essa mesma pergunta se a regressão for usada para prever o nível de lactato com base na resistência – e responda a ambas as perguntas sem fazer nenhum cálculo de regressão.
62. Supõe-se que o conteúdo de hidrogênio seja um fator importante na porosidade de fundições de liga de alumínio. O artigo “The Reduced Pressure Test as a Measuring Tool in the Evaluation of Porosity/Hydrogen Content in Al–7 Wt Pct Si–10 Vol Pct SiC(p) Metal Matrix Composite” (*Metallurgical Tran.*, 1993, p. 1857–1868) apresenta os dados a seguir sobre  $x$  = conteúdo e  $y$  = porosidade do gás para uma determinada técnica de medição.

$x$	0,18	0,20	0,21	0,21	0,21	0,22	0,23
$y$	0,46	0,70	0,41	0,45	0,55	0,44	0,24
$x$	0,23	0,24	0,24	0,25	0,28	0,30	0,37
$y$	0,47	0,22	0,80	0,88	0,70	0,72	0,75

O MINITAB apresenta o resultado a seguir em resposta ao comando CORRELATION (correlação):

Correlação do Volume de Hidrogênio e Porosidade = 0,449

- a. Realize um teste, usando o nível 0,05 para verificar se o coeficiente de correlação populacional é diferente de 0.
  - b. Se fosse realizada uma análise de regressão linear simples, qual porcentagem de variação observada na porosidade pode ser atribuída à relação do modelo?
63. As propriedades físicas de seis amostras de tecido antichama foram investigadas no artigo “Sensory and Physical Properties of Inherently Flame–Retardant Fabrics” (*Textile Research*, 1984, p. 61–68). Use os dados a seguir e um nível de significância 0,05 para determinar se existe uma relação linear entre rigidez  $x$  (mg–cm) e espessura  $y$  (mm). O resultado do teste é surpreendente, levando em consideração o valor de  $r$ ?

$x$	7,98	24,52	12,47	6,92	24,11	35,71
$y$	0,28	0,65	0,32	0,27	0,81	0,57

64. O artigo “Increases in Steroid Binding Globulins Induced by Tamoxifen in Patients with Carcinoma of the Breast” (*J. Endocrinology*, 1978, p. 219–226) apresenta dados sobre os efeitos da droga tamoxifen na alteração do nível de globulina ligadora de cortisol (*cortisol-binding globulin* – CBG) dos pacientes durante o tratamento. Com idade =  $x$  e  $\Delta$ CBG =  $y$ , os valores resumidos são  $n = 26$ ,  $\sum x_i = 1613$ ,  $\sum (x_i - \bar{x})^2 = 3756,96$ ,  $\sum y_i = 281,9$ ,  $\sum (y_i - \bar{y})^2 = 465,34$  e  $\sum x_i y_i = 16,731$ .

- a. Calcule um IC de 90% para o coeficiente de correlação real  $\rho$ .
- b. Teste  $H_0: \rho = -0,5$  versus  $H_a: \rho < -0,5$  no nível 0,05.
- c. Em uma análise de regressão de  $y$  sobre  $x$ , que proporção da variação na alteração do nível de globulina ligadora de cortisol pode ser explicada pela variação na idade do paciente dentro da amostra?
- d. Se você decidir realizar uma análise de regressão usando a idade como variável dependente, que proporção da variação na idade é explicada pela variação em  $\Delta$ CBG?

65. O artigo “Chronological Trend in Blood Lead Levels” (*N. Engl. J. Med.*, 1983, p. 1373–1377) apresenta os dados a seguir sobre  $y$  = nível médio de chumbo no sangue de crianças brancas entre 6 meses e 5 anos de idade e  $x$  = quantidade de chumbo usada na produção de gasolina (em 1000 t) em dez períodos de seis meses:

$x$	48	59	79	80	95
$y$	9,3	11,0	12,8	14,1	13,6
$x$	95	97	102	102	107
$y$	13,8	14,6	14,6	16,0	18,2

- a. Construa gráficos distintos de probabilidade normal para  $x$  e  $y$ . Você acha razoável supor que os pares  $(x, y)$  são de uma população normal bivariada?
- b. Os dados fornecem evidência suficiente para concluir que existe uma relação linear entre o nível de chumbo no sangue e a quantidade de chumbo usada na produção de gasolina? Use  $\alpha = 0,01$ .

66. Considere uma série temporal – isto é, uma seqüência de observações  $X_1, X_2, \dots$  obtidas no decorrer do tempo – com valores observados  $x_1, x_2, \dots, x_n$ . Suponha que

essa série não mostre nenhuma tendência ascendente ou descendente no decorrer do tempo. Um pesquisador freqüentemente deseja saber exatamente até que ponto os valores na série separados por um número específico de unidades de tempo estão fortemente relacionados. O primeiro coeficiente de autocorrelação amostral de atraso  $r_1$  é exatamente o valor do coeficiente de correlação amostral  $r$  dos pares  $(x_1, x_2), (x_2, x_3), \dots, (x_{n-1}, x_n)$ , ou seja, pares de valores separados por uma unidade de tempo. De modo semelhante, o segundo coeficiente de autocorrelação amostral de atraso  $r_2$  é  $r$  para  $n - 2$  pares  $(x_1, x_3), (x_2, x_4), \dots, (x_{n-2}, x_n)$ .

- a. Calcule os valores de  $r_1, r_2$  e  $r_3$  para os dados de temperatura do Exercício 82 do Capítulo 1 e comente.
- b. De maneira semelhante ao coeficiente de correlação populacional  $\rho, \rho_1, \rho_2, \dots$  denotam os coeficientes de autocorrelação teóricos ou de longo prazo nos vários atrasos. Se todos esses  $\rho$ s forem 0, não haverá relação (linear) em qualquer atraso. Nesse caso, se  $n$  for grande, cada  $R_i$  terá aproximadamente uma distribuição normal com média 0 e desvio padrão  $1/\sqrt{n}$  e diferentes  $R_i$ s serão quase independentes. Portanto,  $H_0: \rho_i = 0$  pode ser rejeitada no nível de significância de aproximadamente 0,05 se  $r_i \geq 2/\sqrt{n}$  ou  $r_i \leq -2/\sqrt{n}$ . Se  $n = 100$  e  $r_1 = 0,16, r_2 = -0,09$  e  $r_3 = -0,15$ , há alguma evidência de autocorrelação teórica nos primeiros três atrasos?
- c. Se você estivesse testando simultaneamente as hipóteses nulas do item (b) para mais de um atraso, por que aumentar a constante de corte 2 na região de rejeição?

67. Uma amostra de  $n = 500$  pares  $(x, y)$  foi selecionada e um teste de  $H_0: \rho = 0$  versus  $H_a: \rho \neq 0$  foi realizado. O valor  $P$  resultante foi calculado como sendo 0,00032.

- a. Que conclusão seria apropriada no nível de significância 0,001?
- b. Esse pequeno valor  $P$  indica que existe uma relação fortemente linear entre  $x$  e  $y$  (um valor de  $\rho$  consideravelmente diferente de 0)? Explique.
- c. Agora, forme uma amostra de  $n = 10.000$  pares  $(x, y)$  que resultaram em  $r = 0,022$ . Teste  $H_0: \rho = 0$  versus  $H_a: \rho \neq 0$  no nível 0,05. O resultado é estatisticamente significativo? Comente o significado prático de sua análise.

Exercícios Suplementares (68–85)

68. O artigo “Refuse-Derived Fuel Evaluation in an Industrial Spreader-Stoker Boiler” (*J. Engr. for Gas Turbines and Power*, 1984, p. 782–788) apresenta os dados a seguir sobre  $x$  = % de entrada térmica de combustível derivado rejeitado (RDF) e  $y$  = % eficiência de uma determinada caldeira.

$x$	37	30	48	29	27	16	0	20
$y$	78,0	77,2	74,4	77,7	76,9	79,0	82,1	76,5

- a. Obtenha a equação da reta de regressão estimada.
- b. O modelo de regressão linear simples especifica uma relação útil entre a % de entrada térmica de

RDF e a % de eficiência? Expresse e teste as hipóteses apropriadas.

- c. Para obter uma estimativa precisa de  $\beta_1$ , teria sido preferível fazer quatro observações em  $x = 0$  e quatro observações em  $x = 50$  (supondo que o modelo seja válido para  $x$  entre 0 e 50)? E no caso de três observações em  $x = 0$  e três em  $x = 50$ ? Explique.
- d. Estime a % média real de eficiência quando a % de entrada térmica de RDF é 25, usando um IC de 95%. A % média real de eficiência parece ter sido estimada de maneira precisa? Explique.

69. Os dados a seguir sobre  $x =$  taxa de consumo de óleo diesel pelo método de peso líquido e  $y =$  taxa medida pelo método de traçado de IC, ambos em g/h, foram reproduzidos de um gráfico apresentado no artigo “A New Measurement Method of Diesel Engine Oil Consumption Rate” (*J. Society Auto Engr.*, 1985, p. 28–33).

$x$	4	5	8	11	12	16	17	20	22	28	30	31	39
$y$	5	7	10	10	14	15	13	25	20	24	31	28	39

- a. Supondo que  $x$  e  $y$  estejam relacionados pelo modelo de regressão linear simples, teste  $H_0: \beta_1 = 1$  versus  $H_a: \beta_1 \neq 1$ , usando um nível de significância de 0,05.
- b. Calcule o valor do coeficiente de correlação amostral para esses dados.

70. O resultado do SAS apresentado a seguir baseia-se em dados extraídos do artigo “Evidence for and the Rate of Denitrification in the Arabian Sea” (*Deep Sea Research*, 1978, p. 431–435). As variáveis em estudo são  $x =$  nível de salinidade (%) e  $y =$  nível de nitrato ( $\mu\text{M/L}$ ).

- a. Qual é o tamanho da amostra  $n$ ? [Sugestão: procure graus de liberdade para SQE.]
- b. Calcule uma estimativa pontual do nível de nitrato esperado quando o nível de salinidade é 35,5.
- c. Parece haver alguma relação linear útil entre as duas variáveis?
- d. Qual é o valor do coeficiente de correlação amostral?
- e. Você usaria o modelo de regressão linear simples para tirar conclusões quando o nível de salinidade fosse 40?

71. A presença de carbeto de liga dura em ligas de ferro de cromo branco alto produz uma excelente resistência à abrasão, o que as torna adequadas à manipulação de materiais na mineração e às indústrias de processamento de materiais. Os dados a seguir sobre  $x =$  volume de austenita  $x$  retida (%) e  $y =$  perda de resistência abrasiva ( $\text{mm}^3$ ) em testes de desgaste de pinos usando a granada como abrasivo foram reproduzidos de um gráfico apresentado no artigo “Microstructure Property Relationships in High Chromium White Iron Alloys” (*Intl. Materials Reviews*, 1996, p. 59–82).

$x$	4,6	17,0	17,4	18,0	18,5	22,4	26,5	30,0	34,0
$y$	0,66	0,92	1,45	1,03	0,70	0,73	1,20	0,80	0,91

$x$	38,8	48,2	63,5	65,8	73,9	77,2	79,8	84,0
$y$	1,19	1,15	1,12	1,37	1,45	1,50	1,36	1,29

Use os dados e o resultado SAS a seguir para responder às perguntas da próxima página.

Resultado do SAS para o Exercício 70

Variável Dependente:NÍVEL NITRATO

Análise de Variância

Fonte	GL	Soma dos Quadrados	Quadrado Médio	SQ Valor F	Prob > F
Modelo	1	64,49622	64,49622	63,309	0,0002
Erro	6	6,11253	1,01875		
Total C	7	70,60875			
	MQE Raiz	1,00933	RQ	0,9134	
	Média Dep.	26,91250	RQ ajustada	0,8990	
	C,V,	3,75043			

Estimativas de Parâmetro

Variável	GL	Estimativa de Parâmetro	Erro Padrão	T para H0: Parâmetro = 0	Prob > :T:
TERMO CONSTANTE	1	326,976038	37,71380243	8,670	0,0001
SALINIDADE	1	- 8,403964	1,05621381	- 7,957	0,0002

## Resultado do SAS para o Exercício 71

Variável Dependente: PERDA ABRASIVA

## Análise de Variância

Fonte	GL	Soma dos Quadrados	Quadrado Médio	Valor F	Prob > F
Modelo	1	0,63690	0,63690	15,444	0,0013
Erro	15	0,61860	0,04124		
Total C	16	1,25551			
	MQE Raiz	0,20308	RQ	0,5073	
	Média Dep.	1,10765	RQ ajustada	0,4744	
	CV	18,33410			

## Estimativas de Parâmetro

Variável	GL	Estimativa de Parâmetro	Erro-Padrão	T para H0: Parâmetro = 0	Prob >  T
Termo Constante	1	0,787218	0,09525879	8,264	0,0001
VOLAUSTENTITA	1	0,007570	0,00192626	3,930	0,0013

- a. Que proporção de variação observada na perda de resistência abrasiva pode ser atribuída à relação do modelo de regressão linear simples?
- b. Qual é o valor do coeficiente de correlação amostral?
- c. Teste a utilidade do modelo de regressão linear simples usando  $\alpha = 0,01$ .
- d. Estime a perda média real por desgaste quando o volume é 50%. Faça os cálculos de maneira que transmitam informações sobre confiabilidade e precisão.
- e. Que valor de perda por desgaste você preveria com um volume de 30% e qual é o valor do resíduo correspondente?
72. Os dados a seguir foram reproduzidos de um gráfico de dispersão apresentado no artigo “Urban Emissions Measured with Aircraft” (*J. of the Air and Waste Mgmt. Assoc.*, 1998, p. 16–25). A variável de resposta é  $\Delta\text{NO}_y$  e a variável explicativa é  $\Delta\text{CO}$ .
- |                     |     |     |      |      |     |
|---------------------|-----|-----|------|------|-----|
| $\Delta\text{CO}$   | 50  | 60  | 95   | 108  | 135 |
| $\Delta\text{NO}_y$ | 2,3 | 4,5 | 4,0  | 3,7  | 8,2 |
| $\Delta\text{CO}$   | 210 | 214 | 315  | 720  |     |
| $\Delta\text{NO}_y$ | 5,4 | 7,2 | 13,8 | 32,1 |     |
- a. Ajuste um modelo apropriado aos dados e avalie a sua utilidade.
- b. Preveja o valor de  $\Delta\text{NO}_y$  que resultaria de uma observação a mais quando  $\Delta\text{CO}$  for 400. Faça os cálculos de maneira que transmita informações sobre precisão e confiabilidade. Será que  $\Delta\text{NO}_y$  pode ser previsto com precisão? Explique.
- c. O maior valor de  $\Delta\text{CO}$  é bem maior que os outros valores. Essa observação parece ter tido um impacto significativo sobre a equação ajustada?
73. Os dados a seguir são um subconjunto dos dados apresentados no artigo “Radial Tension Strength of Pipe and Other Curved Flexural Members” (*J. Amer. Concrete Inst.*, 1980, p. 33–39). As variáveis são tempo de vida de uma amostra de tubo ( $x$ , em dias) e carga necessária para obter a primeira rachadura ( $y$ , em 1000 lb/pés).
- |     |       |       |       |       |       |
|-----|-------|-------|-------|-------|-------|
| $x$ | 20    | 20    | 20    | 25    | 25    |
| $y$ | 11,45 | 10,42 | 11,14 | 10,84 | 11,17 |
| $x$ | 25    | 31    | 31    | 31    |       |
| $y$ | 10,54 | 9,47  | 9,19  | 9,54  |       |
- a. Calcule a equação da reta de regressão estimada.
- b. Suponha que um modelo teórico sugira que a queda esperada na carga associada com um dia a mais no tempo de vida seja no máximo 0,10. Os dados por acaso contradizem essa afirmação? Expresse e teste as hipóteses de interesse usando um nível de significância de 0,05.
- c. Para estimar o coeficiente angular da reta de regressão real com a maior precisão possível, teria sido preferível fazer uma única observação em cada um dos tempos de vida 20, 21, 22, ..., 30, e 31? Explique.
- d. Calcule uma estimativa da carga média real para a primeira rachadura com o tempo de vida de 28 dias. A estimativa deve transmitir informações sobre a precisão da estimativa.
74. Uma investigação foi realizada para estudar a relação entre a velocidade (pés/s) e a taxa de passadas largas (número de passos por segundo) entre os corredores de

maratona do sexo feminino. As quantidades resultantes incluem  $n = 11$ ,  $\sum(\text{velocidade}) = 205,4$ ,  $\sum(\text{velocidade})^2 = 3880,08$ ,  $\sum(\text{taxa}) = 35,16$ ,  $\sum(\text{taxa})^2 = 112,681$  e  $\sum(\text{velocidade})(\text{taxa}) = 660,130$ .

- a. Calcule a equação da reta dos mínimos quadrados que você usaria para prever a taxa de passadas largas com base na velocidade.
- b. Calcule a equação da reta dos mínimos quadrados que você usaria para prever a velocidade com base na taxa de passadas largas.
- c. Calcule o coeficiente de determinação da regressão da taxa de passadas largas sobre a velocidade do item (a) e o da regressão da velocidade sobre a taxa de passadas largas do item (b). Como eles estão relacionados?

75. O artigo “Photocharge Effects in Dye Sensitized Ag(Br,I) Emulsions at Millisecond Range Exposures” (*Photographic Sci. and Engr.*, 1981, p. 138–144) apresenta os dados a seguir sobre  $x = \%$  de absorção de luz em 5800 Å e  $y =$  fotovoltagem de pico.

$x$	4,0	8,7	12,7	19,1	21,4
$y$	0,12	0,28	0,55	0,68	0,85

$x$	24,6	28,9	29,8	30,5
$y$	1,02	1,15	1,34	1,29

- a. Construa um gráfico de dispersão desses dados. O que ele sugere?
- b. Supondo que o modelo de regressão linear simples seja apropriado, obtenha a equação da reta de regressão estimada.
- c. Que proporção da variação observada na fotovoltagem de pico pode ser explicada pela relação do modelo?
- d. Preveja a fotovoltagem de pico quando a  $\%$  de absorção for 19,1 e calcule o valor do resíduo correspondente.
- e. Os autores desse artigo afirmam que existe uma relação linear útil entre a  $\%$  de absorção e a fotovoltagem de pico. Você concorda? Realize um teste formal.
- f. Apresente uma estimativa da mudança na fotovoltagem de pico esperada, associada com um aumento de 1% na absorção de luz. Sua estimativa deve transmitir informações sobre a precisão da estimação.
- g. Repita o item (f) para o valor esperado de fotovoltagem de pico quando a  $\%$  de absorção for 20.

76. Na Seção 12.4, apresentamos uma fórmula para  $V(\hat{\beta}_0 + \hat{\beta}_1 x^*)$  e um IC para  $\beta_0 + \beta_1 x^*$ . Tomando  $x^* = 0$ , obtemos  $\sigma_{\hat{\beta}_0}^2$  e um IC para  $\beta_0$ . Use os dados do Exemplo 12.11 para calcular o desvio padrão estimado de  $\hat{\beta}_0$  e um IC de 95% para o termo constante  $y$  da reta de regressão real.

77. Mostre que  $SQE = S_{yy} - \hat{\beta}_1 S_{xy}$ , que oferece uma fórmula de cálculo alternativa para SQE.

78. Suponha que  $x$  e  $y$  sejam variáveis positivas e que uma amostra de  $n$  pares resulte em  $r \approx 1$ . Se o coeficiente de correlação amostral for calculado para os pares  $(x, y^2)$ , o valor resultante também será aproximadamente 1? Explique.

79. Sejam  $s_x$  e  $s_y$  os desvios padrão amostrais dos  $x$ s e  $y$ s observados, respectivamente [portanto,  $s_x^2 = \sum(x_i - \bar{x})^2 / (n - 1)$  e algo semelhante pode-se afirmar em relação a  $s_y^2$ ].

- a. Mostre que uma expressão alternativa para a reta de regressão estimada  $y = \hat{\beta}_0 + \hat{\beta}_1 x$  é

$$y = \bar{y} + r \cdot \frac{s_y}{s_x} (x - \bar{x})$$

- b. Essa expressão para a reta de regressão linear pode ser interpretada como a seguir. Suponha que  $r = 0,5$ . Qual será, portanto, o  $y$  previsto para um  $x$  que esteja a 1 DP ( $s_x$  unidades) acima da média dos  $x_i$ s? Se  $r$  fosse 1, segundo a previsão,  $y$  ficaria 1 DP acima de sua média  $\bar{y}$ . Porém, visto que  $r = 0,5$ , prevemos um  $y$  que está apenas 0,5 DP ( $0,5 s_y$  unidade) acima de  $\bar{y}$ . Usando os dados do Exercício 64 para um paciente cuja idade está a 1 DP abaixo da idade média na amostra, quantos desvios padrão o  $\Delta$ CBG previsto do paciente está acima ou abaixo da média  $\Delta$ CBG da amostra?

80. Demonstre que a estatística  $t$  para testar  $H_0: \beta_1 = 0$  na Seção 12.3 é idêntica à estatística  $t$  na Seção 12.5 para testar  $H_0: \rho = 0$ .

81. Use a fórmula para calcular SQE para demonstrar que  $r^2 = 1 - SQE/SQT$ .

82. Na biofiltração de esgoto, o ar descarregado de uma ins-talação de tratamento passa através de uma membrana porosa úmida que faz com que os contaminantes dissolvam-se na água e sejam transformados em produtos inofensivos. Os dados a seguir sobre  $x =$  temperatura inicial ( $^{\circ}\text{C}$ ) e  $y =$  eficiência da remoção ( $\%$ ) serviram de base para a elaboração de um gráfico de dispersão apresentado no artigo “Treatment of Mixed Hydrogen Sulfide and Organic Vapors in a Rock Medium Biofilter” (*Water Environment Research*, 2001, p. 426–435).

Obs	Remoção		Obs	Remoção	
	Temp	%		Temp	%
1	7,68	98,09	17	8,55	98,27
2	6,51	98,25	18	7,57	98,00
3	6,43	97,82	19	6,94	98,09
4	5,48	97,82	20	8,32	98,25
5	6,57	97,82	21	10,50	98,41
6	10,22	97,93	22	16,02	98,51
7	15,69	98,38	23	17,83	98,71
8	16,77	98,89	24	17,03	98,79
9	17,13	98,96	25	16,18	98,87
10	17,63	98,90	26	16,26	98,76
11	16,72	98,68	27	14,44	98,58
12	15,45	98,69	28	12,78	98,73
13	12,06	98,51	29	12,25	98,45
14	11,44	98,09	30	11,69	98,37
15	10,17	98,25	31	11,34	98,36
16	9,64	98,36	32	10,97	98,45

As quantidades calculadas são  $\sum x_i = 384,26$ ,  $\sum y_i = 3149,04$ ,  $\sum x_i^2 = 5099,2412$ ,  $\sum x_i y_i = 37.850,7762$  e  $\sum y_i^2 = 309.892,6548$ .

- a. Um gráfico de dispersão dos dados sugere que o modelo de regressão linear simples é apropriado?
- b. Ajuste o modelo de regressão linear simples, obtenha uma previsão pontual da eficiência da remoção quando a temperatura = 10,50 e calcule o valor do resíduo correspondente.
- c. Qual é o tamanho aproximado de um desvio típico de pontos no gráfico de dispersão em relação à reta dos mínimos quadrados?
- d. Que proporção da variação observada na eficiência da remoção pode ser atribuída à relação do modelo?
- e. Estime o coeficiente angular de maneira que transmita informações sobre a confiabilidade e precisão, e interprete sua estimativa.
- f. Numa conversa particular com os autores do artigo, eles revelaram que uma observação adicional não foi incluída no respectivo gráfico de dispersão: (6,53, 96,55). Que impacto tem essa observação adicional sobre a equação da reta dos mínimos quadrados e os valores de  $s$  e  $r^2$ ?

83. Processos normais de criação de peixes em aquaculturas inevitavelmente deixam os peixes estressados, o que pode influenciar negativamente no crescimento, na reprodução, na qualidade da carne e na suscetibilidade a doenças. Esse estresse manifesta-se em níveis corticosteróides elevados e retidos. O artigo “Evaluation of Simple Instruments for the Measurement of Blood Glucose and Lactate, and Plasma Protein as Stress Indicators in Fish” (*J. of the World Aquaculture Society*, 1999, p. 276–284) apresenta um experimento no qual os peixes foram submetidos a um protocolo de avaliação de estresse e, em seguida, removidos e testados em vários horários após a aplicação do protocolo. Os dados a seguir sobre  $x$  = tempo (min.) e  $y$  = nível de glicose no sangue (mmol/L) foram reproduzidos de um gráfico.

$x$	2	2	5	7	12	13	17	18	23	24	26	28
$y$	4,0	3,6	3,7	4,0	3,8	4,0	5,1	3,9	4,4	4,3	4,3	4,4
$x$	29	30	34	36	40	41	44	56	56	57	60	60
$y$	5,8	4,3	5,5	5,6	5,1	5,7	6,1	5,1	5,9	6,8	4,9	5,7

Use os métodos desenvolvidos neste capítulo para analisar os dados e escreva um relatório sucinto, resumindo suas conclusões (suponha que os pesquisadores estejam particularmente interessados no nível de glicose 30 min. após o estresse).

84. O artigo “Evaluating the BOD POD for Assessing Body Fat in Collegiate Football Players” (*Medicine and Science in Sports and Exercise*, 1999, p. 1350–1356) relata um novo dispositivo de deslocamento de ar para medir a gordura do corpo. O procedimento habitual utiliza o dispositivo de pesagem hidrostática, que mede a porcentagem de gordura do corpo por meio do deslocamento de água. Aqui são apresentados dados representativos reproduzidos de um gráfico desse artigo.

BOD	2,5	4,0	4,1	6,2	7,1	7,0	8,3	9,2	9,3	12,0	12,2
HW	8,0	6,2	9,2	6,4	8,6	12,2	7,2	12,0	14,9	12,1	15,3
BOD	12,6	14,2	14,4	15,1	15,2	16,3	17,1	17,9			17,9
HW	14,8	14,3	16,3	17,9	19,5	17,5	14,3	18,3			16,2

- a. Use vários métodos para determinar se é plausível que as duas técnicas meçam em média a mesma quantidade de gordura.
  - b. Use os dados para desenvolver uma forma de prever uma medida HW com base em uma medida BOD POD e investigue a eficácia dessas previsões.
85. Reconsidere a situação do Exercício 68, no qual  $x$  = % de entrada térmica de RDF e  $y$  = % de eficiência de uma determinada caldeira foram relacionadas por meio de um modelo de regressão linear simples  $Y = \beta_0 + \beta_1 x + \epsilon$ . Suponha que, para uma segunda caldeira, essas variáveis sejam também relacionadas por meio do modelo de regressão linear simples  $Y = \gamma_0 + \gamma_1 x + \epsilon$  e que  $V(\epsilon) = \sigma^2$  para ambas as caldeiras. Se o conjunto de dados consistir em  $n_1$  observações para a primeira caldeira e  $n_2$  para a segunda e se  $SQE_1$  e  $SQE_2$  representarem as duas somas de quadrados do erro, então uma estimativa conjunta de  $\sigma^2$  e  $\hat{\sigma}^2 = (SQE_1 + SQE_2) / (n_1 + n_2 - 4)$ . Representemos  $SS_{x1}$  e  $SS_{x2}$  por  $\sum (x_i - \bar{x})^2$  para os dados na primeira e na segunda caldeiras, respectivamente. Um teste de  $H_0: \beta_1 - \gamma_1 = 0$  (coeficientes angulares iguais) baseia-se na estatística

$$T = \frac{\hat{\beta}_1 - \hat{\gamma}_1}{\hat{\sigma} \sqrt{\frac{1}{SS_{x1}} + \frac{1}{SS_{x2}}}}$$

Quando  $H_0$  é verdadeira,  $T$  tem uma distribuição  $t$  com  $n_1 + n_2 - 4$  gl. Suponha que as seis observações sobre a segunda caldeira sejam (0, 81,3), (10, 78,4), (20, 78,2), (25, 79,1), (30, 77,6) e (40, 77,4). Usando esses valores com os dados do Exercício 68, realize um teste no nível 0,05 para verificar se a mudança esperada na % de eficiência associada com um aumento de 1% na entrada térmica de RDF é idêntica para as duas caldeiras.



---

## Bibliografia

DRAPER, N. e HARRY S. *Applied Regression Analysis*.  
Nova York: Wiley, 1999. 3. ed. É o livro mais abrangente  
e competente já publicado.

NETER, J., KUTNER, M., NACHTSHEIM, C. e  
WASSERMAN, W. *Applied Linear Statistical Models*.

Homewood, IL: Irwin, 1996. 4. ed. Os primeiros 15  
capítulos constituem uma pesquisa extremamente  
agradável e informativa sobre análise de regressão.

# Regressão Não-Linear e Múltipla

## Introdução

Segundo o modelo probabilístico estudado no Capítulo 12, o valor observado da variável dependente  $Y$  distingue-se da função de regressão linear  $\mu_{Y,x} = \beta_0 + \beta_1 x$  por um valor aleatório. Aqui, avaliamos duas formas de generalizar o modelo de regressão linear simples. A primeira é substituir  $\beta_0 + \beta_1 x$  por uma função não-linear de  $x$  e a segunda é usar uma função de regressão compreendendo mais de uma única variável independente. Depois de ajustar a função de regressão da forma escolhida aos dados fornecidos, obviamente é importante ter à disposição métodos para fazer inferências sobre os parâmetros do modelo escolhido. Antes de usar esses métodos, entretanto, o analista de dados primeiro verifica a validade do modelo escolhido. Na Seção 13.1, discutimos métodos, com base principalmente em uma análise gráfica dos resíduos ( $y$ s observados menos  $y$ s previstos), para checar a aplicabilidade de um modelo ajustado.

Na Seção 13.2 avaliamos as funções de regressão não-linear de uma única variável independente  $x$  que são "intrinsecamente lineares". Com isso, queremos dizer que é possível transformar uma ou ambas as variáveis de modo que a relação entre as novas variáveis seja linear. Uma classe alternativa de relações não-lineares é obtida por meio de funções de regressão polinomial da forma  $\mu_{Y,x} = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_k x^k$ ; tais modelos polinomiais são o assunto da Seção 13.3. A análise da regressão múltipla requer a elaboração de modelos para relacionar  $y$  a duas ou mais variáveis independentes. O foco da Seção 13.4 está em interpretar vários modelos de regressão múltipla e em compreender e usar os resultados de regressão de diversos softwares estatísticos. A última seção deste capítulo examina algumas extensões e armadilhas da modelagem por regressão múltipla.

## 13.1 | Aplicabilidade do Modelo e Verificação do Modelo

Um gráfico dos pares observados  $(x_i, y_i)$  é o primeiro passo necessário para determinar a forma de uma relação matemática entre  $x$  e  $y$ . É possível ajustar aos dados várias funções diferentes da função linear ( $y = b_0 + b_1x$ ), usando o princípio dos mínimos quadrados ou outro método de aderência. Uma vez que uma função da forma escolhida é ajustada ou aderida, é fundamental verificar a aderência do modelo para confirmar se é de fato apropriado. Uma maneira de avaliar a calibração é sobrepor um gráfico da função mais bem aderida ou ajustada no gráfico de dispersão dos dados. Entretanto, qualquer inclinação ou curvatura da função mais aderida pode obscurecer alguns aspectos da aderência que devem ser investigados. Além disso, a escala no eixo vertical pode dificultar a avaliação sobre até que ponto os valores observados desviam das funções mais bem aderidas.

### Resíduos e Resíduos Padronizados

Uma abordagem mais eficaz de avaliar a adequação do modelo é calcular os valores  $\hat{y}_i$  ajustados ou previstos e os resíduos  $e_i = y_i - \hat{y}_i$  e, em seguida, dispor graficamente várias funções dessas quantidades calculadas. Examinamos então os gráficos para confirmar nossa opção de modelo ou para obter indicações de que o modelo não é apropriado. Suponha que o modelo de regressão linear simples seja correto e considere  $y = \hat{\beta}_0 + \hat{\beta}_1x$  a equação da reta de regressão estimada. Portanto, o  $i$ -ésimo resíduo é  $e_i = y_i - (\hat{\beta}_0 + \hat{\beta}_1x_i)$ . Para derivar propriedades dos resíduos, considere  $e_i = Y_i - \hat{Y}_i$  o  $i$ -ésimo resíduo como uma variável aleatória (va) (antes de as observações serem de fato feitas). Então

$$E(Y_i - \hat{Y}_i) = E(Y_i) - E(\hat{\beta}_0 + \hat{\beta}_1x_i) = \beta_0 + \beta_1x_i - (\beta_0 + \beta_1x_i) = 0 \quad (13.1)$$

Como  $\hat{Y}_i (= \hat{\beta}_0 + \hat{\beta}_1x_i)$  é uma função linear dos  $Y_j$ s,  $Y_i - \hat{Y}_i$  (os coeficientes dependem dos  $x_j$ s) também o é. Portanto, a normalidade dos  $Y_j$ s implica que cada resíduo é normalmente distribuído. Pode-se mostrar também que

$$V(Y_i - \hat{Y}_i) = \sigma^2 \cdot \left[ 1 - \frac{1}{n} - \frac{(x_i - \bar{x})^2}{S_{xx}} \right] \quad (13.2)$$

Substituindo  $\sigma^2$  por  $s^2$  e extraindo a raiz quadrada da Equação (13.2), obtemos o desvio padrão estimado de um resíduo.

Padronizemos agora cada resíduo subtraindo o valor médio (zero) e, em seguida, dividindo-o pelo desvio padrão estimado.

Os **resíduos padronizados** são dados por

$$e_i^* = \frac{y_i - \hat{y}_i}{s \sqrt{1 - \frac{1}{n} - \frac{(x_i - \bar{x})^2}{S_{xx}}}} \quad i = 1, \dots, n \quad (13.3)$$

Se, por exemplo, um determinado resíduo padronizado for 1,5, o resíduo em si será 1,5 desvio padrão (estimado) maior do que o que seria esperado da aderência do modelo correto. Observe que as variâncias dos resíduos diferem entre si. Contudo, se  $n$  for moderadamente grande, o termo entre colchetes em (13.2) será aproximadamente 1, de modo que algumas fontes usam  $e_i/s$  como o resíduo padronizado. Os cálculos dos  $e_i^*$ s podem ser tediosos, mas os softwares estatísticos mais amplamente usados fornecem esses valores de maneira automática e, (por solicitação), podem elaborar vários gráficos compreendendo esses valores.

**Exemplo 13.1**

O Exercício 19 no Capítulo 12 apresentou dados sobre  $x$  = taxa de liberação na área do queimador e  $y$  = emissões de  $\text{NO}_x$ . Aqui, reproduzimos esses dados e fornecemos os valores ajustados, os resíduos e os resíduos padronizados. A reta de regressão estimada é  $y = -45,55 + 1,71x$  e  $r^2 = 0,961$ . Observe que os resíduos padronizados não são um múltiplo constante dos resíduos (isto é,  $e_i^* \neq e_i/s$ ).

$x_i$	$y_i$	$\hat{y}_i$	$e_i$	$e_i^*$
100	150	125,6	24,4	0,75
125	140	168,4	-28,4	-0,84
125	180	168,4	11,6	0,35
150	210	211,1	-1,1	-0,03
150	190	211,1	-21,1	-0,62
200	320	296,7	23,3	0,66
200	280	296,7	-16,7	-0,47
250	400	382,3	17,7	0,50
250	430	382,3	47,7	1,35
300	440	467,9	-27,9	-0,80
300	390	467,9	-77,9	-2,24
350	600	553,4	46,6	1,39
400	610	639,0	-29,0	-0,92
400	670	639,0	31,0	0,99

**Gráficos de Diagnóstico**

Os gráficos básicos recomendados por vários estatísticos para verificar a validade e utilidade do modelo são os seguintes:

1.  $e_i^*$  (ou  $e_i$ ) no eixo vertical versus  $x_i$  no eixo horizontal.
2.  $e_i^*$  (ou  $e_i$ ) no eixo vertical versus  $\hat{y}_i$  no eixo horizontal.
3.  $\hat{y}_i$  no eixo vertical versus  $y_i$  no eixo horizontal.
4. Um gráfico de probabilidade normal dos resíduos padronizados.

Os Gráficos 1 e 2 são denominados **gráficos dos resíduos** (em relação à variável independente e aos valores ajustados, respectivamente), ao passo que o Gráfico 3 é de valores ajustados em relação aos valores observados.

Se o Gráfico 3 apresentar pontos próximos à reta de 45° [coeficiente angular +1 passando através da origem (0, 0)], então a função de regressão estimada fornecerá previsões precisas dos valores de fato observados. Portanto, o Gráfico 3 oferece uma avaliação visual sobre quão eficaz é o modelo para fazer previsões. Contudo que o modelo seja o correto, nenhum gráfico dos resíduos deve exibir padrões distintos. Os resíduos devem ser distribuídos aleatoriamente em torno de 0, de acordo com uma distribuição normal, de modo que todos, menos alguns poucos resíduos padronizados, devem ficar entre  $-2$  e  $+2$  (isto é, todos os resíduos, exceto alguns, devem ficar a 2 desvios padrão do valor esperado 0). O gráfico dos resíduos padronizados versus  $\hat{y}$  é, na verdade, uma combinação dos dois outros gráficos, mostrando implicitamente como os resíduos variam com  $x$  e como os valores ajustados se comparam com os valores observados. Este último gráfico é um dos mais frequentemente recomendados para a análise de regressão múltipla. O Gráfico 4 permite que o analista avalie a plausibilidade da suposição de que  $\epsilon$  tem uma distribuição normal.

**Exemplo 13.2** (Continuação do Exemplo 13.1)

A Figura 13.1 apresenta um gráfico de dispersão dos dados e os quatro gráficos que acabamos de recomendar. O gráfico de  $\hat{y}$  versus  $y$  confirma a impressão dada por  $r^2$  de que  $x$  é eficaz para prever  $y$  e, além disso, indica que não existe nenhum  $y$  observado para o qual o valor previsto esteja extremamente longe da marca. Nenhum

dos gráficos dos resíduos mostra um padrão incomum ou valores discrepantes. Há um resíduo padronizado ligeiramente fora do intervalo  $(-2, 2)$ , mas isso não é novidade em uma amostra de tamanho 14. O gráfico de probabilidade normal dos resíduos padronizados é razoavelmente reto. Em resumo, os gráficos não nos deixam dúvidas sobre a aplicabilidade de uma relação linear simples ou o ajuste dos dados fornecidos. ■

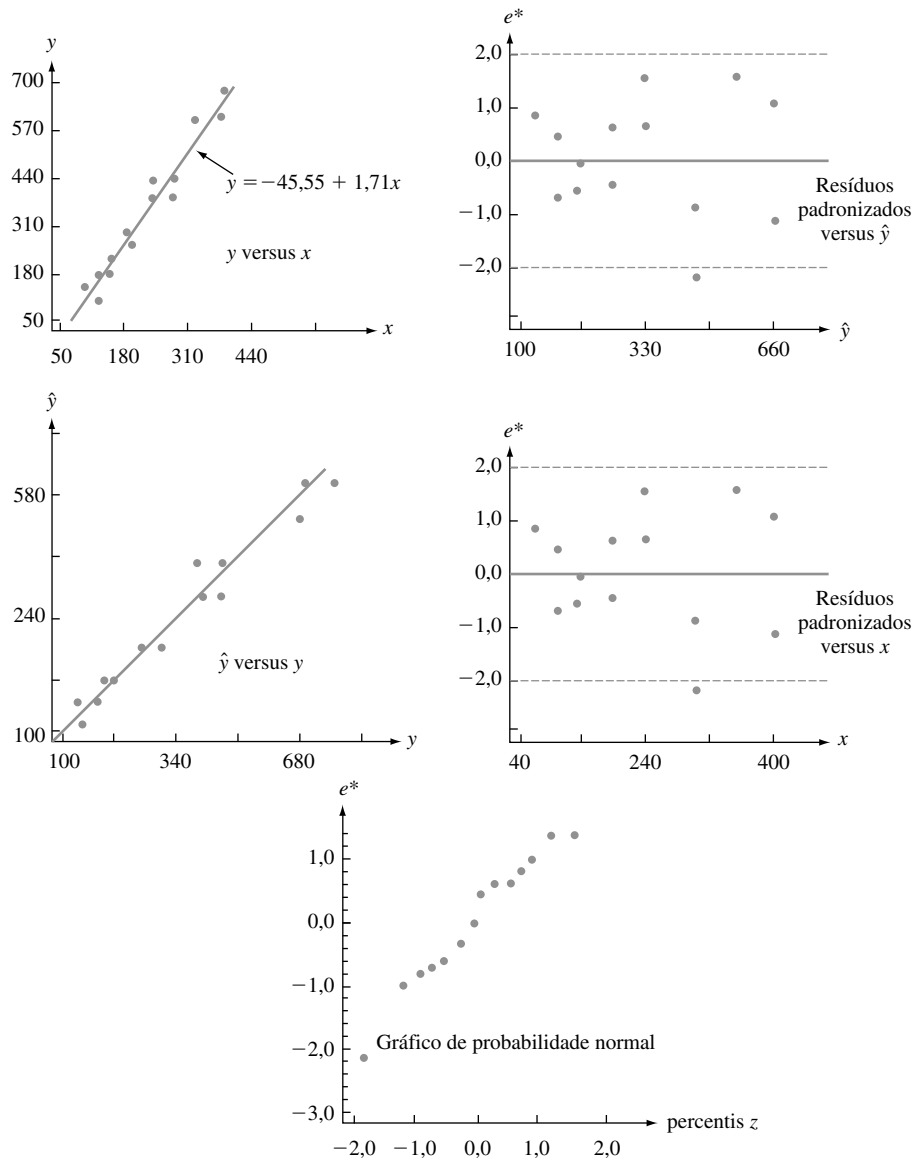


Figura 13.1 Gráficos dos dados do Exemplo 13.1

### Dificuldades e Soluções

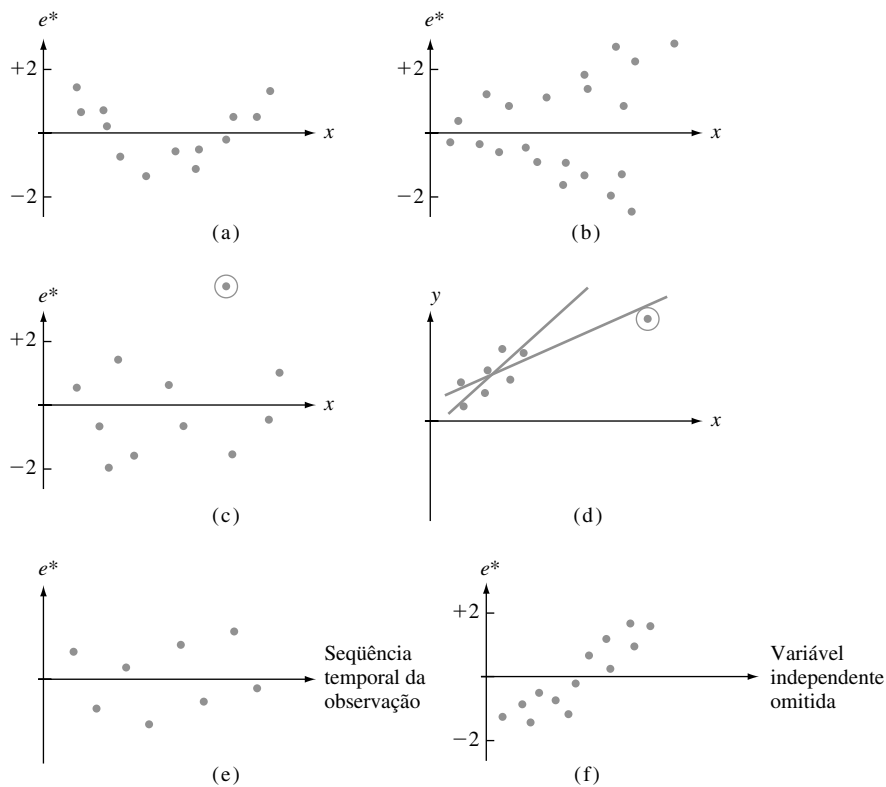
Embora nossa expectativa seja de que nossa análise gere gráficos como os da Figura 13.1, muito freqüentemente os gráficos vão apresentar uma ou mais das seguintes dificuldades:

1. Uma relação probabilística não-linear entre  $x$  e  $y$  é apropriada.
2. A variância de  $\epsilon$  (e de  $Y$ ) não é uma constante  $\sigma^2$ , mas depende de  $x$ .
3. O modelo selecionado se ajusta bem aos dados, exceto alguns poucos valores discrepantes ou aberrantes, os quais podem ter influenciado substancialmente a escolha da função mais ajustada.
4. O termo  $\epsilon$  do erro não tem uma distribuição normal.

5. Quando o subscrito  $i$  indica a seqüência temporal das observações, os  $\epsilon_i$ s exibem dependência ao longo do tempo.
6. Uma ou mais variáveis independentes de interesse foram omitidas do modelo.

A Figura 13.2 apresenta os gráficos dos resíduos correspondentes aos itens 1–3, 5 e 6. No Capítulo 4 falamos sobre os padrões nos gráficos de probabilidade normal que lançam dúvidas sobre a suposição de uma distribuição normal. Observe que os resíduos dos dados da Figura 13.2(d) que contêm o ponto circulado não necessariamente propõem, por si sós, análises adicionais; contudo, quando uma nova reta é ajustada a esse ponto excluído, essa nova reta difere consideravelmente da reta original. Esse tipo de comportamento é mais difícil de identificar na regressão múltipla. É mais provável surgir quando existe um único (ou muito poucos) ponto(s) de dados(s) com valor(es) de variável independente bem afastado(s) dos dados restantes.

Indicamos a seguir, resumidamente, as soluções existentes para dificuldades desse tipo. Para uma discussão mais abrangente, é necessário consultar uma ou mais referências sobre análise de regressão. Se o gráfico dos resíduos for de alguma forma parecido com o da Figura 13.2(a), exibindo um padrão curvo, uma função não-linear de  $x$  pode ser adequada.



**Figura 13.2** Gráficos que indicam anormalidade nos dados: (a) relação não-linear; (b) variância não-constante; (c) observação discrepante; (d) observação com ampla influência; (e) dependência em relação a erros; (f) variável omitida

O gráfico dos resíduos da Figura 13.2(b) sugere que, embora uma relação de linha reta possa ser razoável, a suposição de que  $V(Y_i) = \sigma^2$  para cada  $i$  tem validade duvidosa. Quando as hipóteses do Capítulo 12 são válidas, é possível mostrar que, entre todos os estimadores não-tendenciosos de  $\beta_0$  e  $\beta_1$ , os estimadores dos mínimos quadrados ordinários têm variância mínima. Esses estimadores dão o mesmo peso para cada  $(x_i, Y_i)$ . Se a variância de  $Y$  aumentar com  $x$ , os  $Y_i$ s para  $x_i$  grande devem receber um peso menor que aqueles com  $x_i$  pequeno. Isso sugere que  $\beta_0$  e  $\beta_1$  devem ser estimados minimizando-se

$$f_w(b_0, b_1) = \sum w_i [y_i - (b_0 + b_1 x_i)]^2 \quad (13.4)$$

onde os  $w_i$ s são pesos que diminuem à medida que  $x_i$  aumenta. Minimizando a Expressão (13.4), obtemos estimativas dos **mínimos quadrados ponderados**. Por exemplo, se o desvio padrão de  $Y$  for proporcional a  $x$  (para

$x > 0$ ) — ou seja,  $V(Y) = kx^2$  —, é possível mostrar que os pesos  $w_i = 1/x_i^2$  fornecem melhores estimadores de  $\beta_0$  e  $\beta_1$ . Os livros de John Neter et al. e S. Chatterjee e Bertram Price contêm informações mais detalhadas (veja a Bibliografia deste capítulo). Os mínimos quadrados ponderados são usados frequentemente por econometristas (economistas que usam métodos estatísticos) para estimar parâmetros.

Quando os gráficos ou outras evidências sugerem que o conjunto de dados contém *outliers* ou pontos de grande influência no ajuste resultante, uma abordagem possível é omitir esses pontos afastados e recalculá-la equação da regressão estimada. Isso com certeza seria correto se houvesse sido identificado que os *outliers* resultaram de erros durante o registro dos valores dos dados ou de erros experimentais. Se não for possível identificar nenhuma causa determinável para os *outliers*, continua sendo desejável descrever a equação estimada omitindo e não omitindo os *outliers*. Além disso, outra abordagem é manter possíveis *outliers*, mas usar um princípio de estimativa que atribui um peso relativamente menor aos valores afastados, do que o princípio dos mínimos quadrados. Um desses princípios é conhecido por MSDA (Mínima Soma dos Desvios Absolutos), que seleciona  $\hat{\beta}_0$  e  $\hat{\beta}_1$  para minimizar  $\sum |y_i - (b_0 + b_1x_i)|$ . Diferentemente das estimativas dos mínimos quadrados, não existe nenhuma fórmula ideal para as estimativas MSDA; os respectivos valores devem ser encontrados por meio de um procedimento de cálculo iterativo. Procedimentos desse tipo são também usados quando existe a suspeita de que os  $\epsilon_i$ s têm distribuição com “caudas pesadas”, em vez de normal (tornando bem mais provável do que na distribuição normal, em que valores discrepantes entrem na amostra); procedimentos de regressão robustos são aqueles que produzem estimativas confiáveis para uma ampla variedade de distribuições de erros subjacentes. Os estimadores dos mínimos quadrados não são robustos, da mesma forma que a média amostral  $\bar{X}$  não é um estimador robusto para  $\mu$ .

Quando um gráfico sugere dependência de tempo nos termos do erro, uma análise apropriada pode exigir a transformação dos  $y$ s ou então um modelo que inclua explicitamente uma variável temporal. Por fim, é necessário examinar um gráfico como o da Figura 13.2(f), que mostra um padrão nos resíduos quando ilustrados graficamente em relação a uma variável omitida, que sugere que seja considerado um modelo de regressão múltipla que inclua a variável omitida anteriormente.

## Exercícios | Seção 13.1 (1–14)

- Suponha que as variáveis  $x$  = distância de comutação e  $y$  = tempo de comutação estejam relacionadas de acordo com o modelo de regressão linear simples, com  $\sigma = 10$ .
  - Se  $n = 5$  observações forem feitas nos valores  $x_1 = 5$ ,  $x_2 = 10$ ,  $x_3 = 15$ ,  $x_4 = 20$  e  $x_5 = 25$ , calcule os desvios padrão dos cinco resíduos correspondentes.
  - Repita a parte (a) para  $x_1 = 5$ ,  $x_2 = 10$ ,  $x_3 = 15$ ,  $x_4 = 20$  e  $x_5 = 50$ .
  - Que significado têm os resultados das partes (a) e (b) para os desvios da reta estimada em relação à observação feita no maior valor  $x$  amostrado?
- Os valores  $x$  e os resíduos padronizados dos dados sobre o fluxo de cloro/taxa de ataque do Exercício 52 (Seção 12.4) são apresentados na tabela a seguir. Elabore um gráfico dos resíduos padronizados e fale sobre sua aparência.

$x$	1,50	1,50	2,00	2,50	2,50
$e^*$	0,31	1,02	-1,15	-1,23	0,23

$x$	3,00	3,50	3,50	4,00
$e^*$	0,73	-1,36	1,53	0,07

- O Exemplo 12.6 apresenta os resíduos de uma regressão linear simples da quantidade de umidade  $y$  sobre a taxa de filtragem  $x$ .
  - Represente graficamente os resíduos em relação a  $x$ . O gráfico resultante sugere que uma função de regressão de linha reta é uma opção de modelo razoável? Explique seu raciocínio.
  - Usando  $s = 0,665$ , calcule os valores dos resíduos padronizados. Será que  $e_i^* \approx e_i/s$  para  $i = 1, \dots, n$  ou os  $e_i^*$ s não são tão próximos de serem proporcionais aos  $e_i$ s?
  - Represente graficamente os resíduos padronizados em relação a  $x$ . Em sua aparência global, o gráfico difere significativamente do gráfico da parte (a)?
- A resistência ao desgaste de determinados componentes do reator nuclear feitos de Zircaloy-2 é parcialmente determinada pelas propriedades da camada de óxido. Os dados a seguir foram apresentados em um

artigo que propôs um novo método de teste não-destrutivo para monitorar a espessura da camada ("Monitoring of Oxide Layer Thickness on Zircaloy-2 by the Eddy Current Test Method", *J. of Testing and Eval.*, 1987, p. 333–336). As variáveis são  $x$  = espessura da camada de óxido ( $\mu\text{m}$ ) e  $y$  = resposta de corrente parasita ou de Foucault (unidades arbitrárias).

$x$	0	7	17	114	133
$y$	20,3	19,8	19,5	15,9	15,1

$x$	142	190	218	237	285
$y$	14,7	11,9	11,5	8,3	6,6

- a. Os autores resumiram a relação fornecendo a equação da reta dos mínimos quadrados como  $y = 20,6 - 0,047x$ . Calcule e represente graficamente os resíduos em relação a  $x$  e, em seguida, comente sobre a aplicabilidade do modelo de regressão linear simples.
- b. Use  $s = 0,7921$  para calcular os resíduos padronizados de uma regressão linear simples. Elabore um gráfico de resíduos padronizados e comente. Além disso, elabore um gráfico de probabilidade normal e comente.

5. À medida que a temperatura do ar cai, as águas dos rios ficam supercongeladas, formando cristais de gelo. Esse gelo pode influenciar significativamente a hidráulica de um rio. O artigo "Laboratory Study of Anchor Ice Growth" (*J. of Cold Regions Engr.*, 2001, p. 60–66) descreve um experimento no qual a espessura do gelo (mm) foi examinada como uma função do tempo decorrido (h) sob circunstâncias específicas. Os dados a seguir foram reproduzidos de um gráfico desse artigo:  $n = 33$ ;  $x = 0,17, 0,33, 0,50, 0,67, \dots, 5,50$ ;  $y = 0,50, 1,25, 1,50, 2,75, 3,50, 4,75, 5,75, 5,60, 7,00, 8,00, 8,25, 9,50, 10,50, 11,00, 10,75, 12,50, 12,25, 13,25, 15,50, 15,00, 15,25, 16,25, 17,25, 18,00, 18,25, 18,15, 20,25, 19,50, 20,00, 20,50, 20,60, 20,50, 19,80$ .

- a. O valor  $r^2$  resultante do ajuste dos mínimos quadrados é 0,977. Interprete esse valor e fale sobre a aplicabilidade da suposição de uma relação linear aproximada.
- b. Os resíduos, relacionados na mesma ordem dos valores  $x$ , são

-1,03	-0,92	-1,35	-0,78	-0,68	-0,11	0,21
-0,59	0,13	0,45	0,06	0,62	0,94	0,80
-0,14	0,93	0,04	0,36	1,92	0,78	0,35
0,67	1,02	1,09	0,66	-0,09	1,33	-0,10
-0,24	-0,43	-1,01	-1,75	-3,14		

Represente os resíduos graficamente em relação ao tempo decorrido. O que o gráfico sugere?

6. Os dados a seguir sobre  $x$  = densidade real ( $\text{kg}/\text{mm}^3$ ) e  $y$  = quantidade de umidade (% d.b.) foram reproduzidos de um gráfico do artigo "Physical Properties of Cumin Seed" (*J. Agric. Engr. Res.*, 1996, p. 93–98).

$x$	7,0	9,3	13,2	16,3	19,1	22,0
$y$	1.046	1.065	1.094	1.117	1.130	1.135

A equação da reta dos mínimos quadrados é  $y = 1.008,14 + 6,19268x$  (essa equação difere muito ligeiramente da equação apresentada no artigo);  $s = 7,265$  e  $r^2 = 0,968$ .

- a. Realize um teste de utilidade do modelo e comente.
- b. Calcule os valores dos resíduos e represente-os graficamente em relação a  $x$ . O gráfico sugere que uma função de regressão linear é imprópria?
- c. Calcule os valores dos resíduos padronizados e represente-os graficamente em relação a  $x$ . Existe algum resíduo padronizado extraordinariamente grande (positivo ou negativo)? Esse gráfico fornece a mesma mensagem que o gráfico da parte (b) em relação à aplicabilidade da função de regressão linear?

7. O artigo "Effects of Gamma Radiation on Juvenile and Mature Cuttings of Quaking Aspen" (*Forest Science*, 1967, p. 240–245) apresenta os dados a seguir sobre tempo de exposição à radiação ( $x$ , em  $\text{kr}/16 \text{ h}$ ) e o peso seco das raízes ( $y$ , em  $\text{mg} \times 10^{-1}$ ).

$x$	0	2	4	6	8
$y$	110	123	119	86	62

- a. Elabore um gráfico de dispersão. O gráfico sugere que uma relação probabilística linear é apropriada?
- b. Uma regressão linear resulta na reta dos mínimos quadrados  $y = 127 - 6,65x$ , com  $s = 16,94$ . Calcule os resíduos e os resíduos padronizados e, em seguida, elabore os respectivos gráficos. O que esses gráficos sugerem? Que tipo de função pode fornecer um melhor ajuste para os dados do que a linha reta?

8. O registro contínuo da taxa de batimentos cardíacos pode ser usado para obter informações sobre o nível de intensidade de exercício ou de esforço físico no trabalho, durante atividades esportivas ou durante outras atividades diárias. O artigo "The Relationship between Heart Rate and Oxygen Uptake During Non-Steady State Exercise" (*Ergonomics*, 2000, p. 1.578–1.592) relatou em um estudo que usou a resposta da taxa de batimentos cardíacos ( $x$ , como uma porcentagem da taxa máxima) para prever a absorção de oxigênio ( $y$ , como uma porcentagem da absorção máxima) durante exercícios físicos. Os dados a seguir foram reproduzidos de um gráfico desse artigo:

TBC	43,5	44,0	44,0	44,5	44,0	45,0	48,0	49,0
VO <sub>2</sub>	22,0	21,0	22,0	21,5	25,5	24,5	30,0	28,0

TBC	49,5	51,0	54,5	57,5	57,7	61,0	63,0	72,0
VO <sub>2</sub>	32,0	29,0	38,5	30,5	57,0	40,0	58,0	72,0

Use um software estatístico para executar uma análise de regressão linear simples, observando principal-



mente se existe alguma observação extraordinária ou influente.

9. Considere os quatro conjuntos de dados  $(x, y)$  a seguir; os três primeiros têm os mesmos valores  $x$ , de modo que esses valores são relacionados apenas uma única vez (Frank Anscombe, “Graphs in Statistical Analysis”, *Amer. Statistician*, 1973, p. 17–21):

Conjunto de dados	1–3	1	2	3	4	4
Variável	$x$	$y$	$y$	$y$	$x$	$y$
	10,0	8,04	9,14	7,46	8,0	6,58
	8,0	6,95	8,14	6,77	8,0	5,76
	13,0	7,58	8,74	12,74	8,0	7,71
	9,0	8,81	8,77	7,11	8,0	8,84
	11,0	8,33	9,26	7,81	8,0	8,47
	14,0	9,96	8,10	8,84	8,0	7,04
	6,0	7,24	6,13	6,08	8,0	5,25
	4,0	4,26	3,10	5,39	19,0	12,50
	12,0	10,84	9,13	8,15	8,0	5,56
	7,0	4,82	7,26	6,42	8,0	7,91
	5,0	5,68	4,74	5,73	8,0	6,89

Para cada um desses quatro conjuntos de dados, os valores das estatísticas  $\sum x_i$ ,  $\sum x_i^2$ ,  $\sum y_i$ ,  $\sum y_i^2$  e  $\sum x_i y_i$  são praticamente idênticos, de modo que todas as quantidades calculadas com base nesses cinco termos serão essencialmente idênticas para os quatro conjuntos — a reta dos mínimos quadrados ( $y = 3 + 0,5x$ ), SQE,  $s^2$ ,  $r^2$ , intervalos  $t$ , estatísticas  $t$ , e assim por diante. As estatísticas não apresentam nenhuma maneira de diferenciar os quatro conjuntos de dados. Com base no gráfico de dispersão e no gráfico dos resíduos de cada conjunto, fale sobre a aplicabilidade ou inadequação de ajustar um modelo de linha reta; inclua em seus comentários qualquer sugestão específica sobre como uma “análise de linha reta” deveria ser modificada ou qualificada.

10. a. Mostre que  $\sum_{i=1}^n e_i = 0$  quando os  $e_i$ s são os resíduos de uma regressão linear simples.  
 b. Os resíduos de uma regressão linear simples são independentes entre si, positivamente correlacionados ou negativamente correlacionados? Explique.  
 c. Mostre que  $\sum_{i=1}^n x_i e_i = 0$  para os resíduos de uma regressão linear simples. [Esse resultado, juntamente à parte (a), mostra que existem duas restrições lineares nos  $e_i$ s, o que resulta em uma perda de 2 gl quando os resíduos quadrados são usados para estimar  $\sigma^2$ .]  
 d. É verdade que  $\sum_{i=1}^n e_i^* = 0$ ? Forneça uma prova ou um exemplo contrário.
11. a. Expresse o  $i$ -ésimo resíduo  $Y_i - \hat{Y}_i$  (onde  $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ ) na forma  $\sum c_j Y_j$ , uma função linear dos  $Y_j$ s. Em seguida, use regras de variância para provar que  $V(Y_i - \hat{Y}_i)$  é dada pela Expressão (13.2).  
 b. É possível mostrar que  $\hat{Y}_i$  e  $Y_i - \hat{Y}_i$  (o  $i$ -ésimo valor previsto e resíduo) são independentes entre si. Use essa constatação, a relação  $Y_i = \hat{Y}_i + (Y_i - \hat{Y}_i)$  e a

expressão de  $V(\hat{Y})$  da Seção 12.4 para provar novamente a Expressão (13.2).

- c. Quando  $x_i$  afasta-se mais de  $\bar{x}$ , o que ocorre com  $V(\hat{Y}_i)$  e  $V(Y_i - \hat{Y}_i)$ ?
12. a. Uma regressão linear poderia resultar nos resíduos 23, -27, 5, 17, -8, 9 e 15? Por quê?  
 b. Uma regressão linear poderia resultar nos resíduos 23, -27, 5, 17, -8, -12 e 2 correspondentes aos valores  $x$  3, -4, 8, 12, -14, -20 e 25? Por quê? (*Sugestão*: Veja o Exercício 10.)
13. Lembre-se de que  $\hat{\beta}_0 + \hat{\beta}_1 x$  tem uma distribuição normal, com valor esperado  $\beta_0 + \beta_1 x$  e variância

$$\sigma^2 \left\{ \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right\}$$

de modo que

$$Z = \frac{\hat{\beta}_0 + \hat{\beta}_1 x - (\beta_0 + \beta_1 x)}{\sigma \left( \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right)^{1/2}}$$

tem uma distribuição normal padronizada. Se  $S = \sqrt{\text{SQE}/(n - 2)}$  for substituída por  $\sigma$ , a variável resultante terá uma distribuição  $t$  com  $n - 2$  gl. Por analogia, qual é a distribuição de qualquer resíduo padronizado? Se  $n = 25$ , qual é a probabilidade de um determinado resíduo padronizado ficar fora do intervalo  $(-2,50; 2,50)$ ?

14. Se houver pelo menos um valor  $x$  no qual foi feita mais de uma observação, existe um procedimento de teste formal para testar

$H_0: \mu_{y,x} = \beta_0 + \beta_1 x$  para alguns valores  $\beta_0, \beta_1$  (a função de regressão real é linear)

versus

$H_a: H_0$  não é verdadeira (a função de regressão real é não-linear)

Suponha que se façam observações em  $x_1, x_2, \dots, x_c$ . Sejam  $Y_{11}, Y_{12}, \dots, Y_{1n_1}$  as  $n_1$  observações quando  $x = x_1$ ; ...;  $Y_{c1}, Y_{c2}, \dots, Y_{cn_c}$  as  $n_c$  observações quando  $x = x_c$ . Com  $n = \sum n_i$  (número total de observações), SQE tem  $n - 2$  gl. Dividimos SQE em duas partes, SQEP (erro puro) e SQFA (falta de aderência), como a seguir:

$$\begin{aligned} \text{SQEP} &= \sum_i \sum_j (Y_{ij} - \bar{Y}_i)^2 \\ &= \sum \sum Y_{ij}^2 - \sum n_i \bar{Y}_i^2 \end{aligned}$$

$$\text{SQFA} = \text{SQE} - \text{SQEP}$$

As  $n_i$  observações em  $x_i$  contribuem com  $n_i - 1$  gl para SQEP, de modo que o número de graus de liberdade de SQEP é  $\sum_i (n_i - 1) = n - c$  e o de SQFA é  $n - 2 - (n - c) = c - 2$ . Seja  $\text{QMFP} = \text{SQEP}/(n - c)$  e  $\text{QMFA} = \text{SQFA}/(c - 2)$ . Logo, é possível mostrar que, embora  $E(\text{QMFP}) = \sigma^2$ , independentemente de  $H_0$  ser ou não verdadeira,  $E(\text{QMFA}) = \sigma^2$  se  $H_0$  for verdadeira e  $E(\text{QMFA}) > \sigma^2$  se  $H_0$  for falsa.

Estatística de teste:  $F = \frac{\text{MQFA}}{\text{QMFP}}$

Região de rejeição:  $f \geq F_{\alpha, c-2, n-c}$

Os dados a seguir foram apresentados no artigo “Changes in Growth Hormone Status Related to Body Weight of Growing Cattle” (*Growth*, 1977, p. 241-247), com  $x$  = peso corporal e  $y$  = taxa de depuração metabólica/peso corporal.

$x$	110	110	110	230	230	230	360
$y$	235	198	173	174	149	124	115

$x$	360	360	360	505	505	505	505
$y$	130	102	95	122	112	98	96

(Logo  $c = 4$ ,  $n_1 = n_2 = 3$ ,  $n_3 = n_4 = 4$ .)

- a. Teste  $H_0$  versus  $H_a$  ao nível 0,05 usando o teste de falta de aderência que acabamos de descrever.
- b. Um gráfico de dispersão dos dados sugere que a relação entre  $x$  e  $y$  é linear? Como isso se compara com o resultado da parte (a)? (Foi usada no artigo uma função de regressão não-linear.)

## 13.2 | Regressão com Variáveis Transformadas

A necessidade de um modelo alternativo ao modelo probabilístico linear  $Y = \beta_0 + \beta_1 x + \epsilon$  pode ser sugerida por um argumento teórico ou então examinando-se gráficos de diagnóstico de uma análise de regressão linear. Em ambos os casos, é desejável estabelecer um modelo cujos parâmetros possam ser facilmente estimados. Uma classe importante de modelos desse tipo é especificada por meio de funções “intrinsecamente lineares”.

### DEFINIÇÃO

Uma função que relaciona  $y$  a  $x$  é **intrinsecamente linear** se, por meio de uma transformação em  $x$  e/ou  $y$ , essa função puder ser expressa como  $y' = \beta_0 + \beta_1 x'$ , onde  $x'$  = a variável independente transformada e  $y'$  = a variável dependente transformada.

Quatro dentre as funções intrinsecamente lineares mais comuns são apresentadas na Tabela 13.1. Em cada caso, a transformação apropriada é uma transformação logarítmica — logaritmo base 10 ou logaritmo natural (base  $e$ ) — ou uma transformação recíproca. Gráficos representativos das quatro funções são exibidos na Figura 13.3.

**Tabela 13.1** Funções intrinsecamente lineares comuns<sup>1</sup>

Função	Transformação(ões) para Linearizar	Forma Linear
a. Exponencial: $y = \alpha e^{\beta x}$	$y' = \ln(y)$	$y' = \ln(\alpha) + \beta x$
b. Potência: $y = \alpha x^\beta$	$y' = \log(y)$ , $x' = \log(x)$	$y' = \log(\alpha) + \beta x'$
c. $y = \alpha + \beta \cdot \log(x)$	$x' = \log(x)$	$y = \alpha + \beta x'$
d. Recíproca: $y = \alpha + \beta \cdot \frac{1}{x}$	$x' = \frac{1}{x}$	$y = \alpha + \beta x'$

<sup>1</sup>Quando aparece  $\log(\cdot)$  é possível usar logaritmo base 10 ou logaritmo base  $e$ .

Para uma relação do tipo função exponencial, apenas  $y$  é transformado para conseguir linearidade, ao passo que, para uma relação do tipo função potencial, tanto  $x$  quanto  $y$  são transformados. Pelo fato de a variável  $x$  estar no expoente, em uma relação exponencial,  $y$  aumenta (se  $\beta > 0$ ) ou diminui (se  $\beta < 0$ ) bem mais rapidamente à medida que  $x$  aumenta, em comparação ao que ocorre na função potencial, embora em um curto intervalo de valores  $x$  seja difícil distinguir as duas funções. Exemplos de funções não-intrinsecamente lineares são  $y = \alpha + \gamma e^{\beta x}$  e  $y = \alpha + \gamma x^\beta$ .

As funções intrinsecamente lineares conduzem-nos diretamente a modelos probabilísticos que, embora não lineares em  $x$  como uma função, têm parâmetros cujos valores são facilmente estimados por meio de mínimos quadrados ordinários.

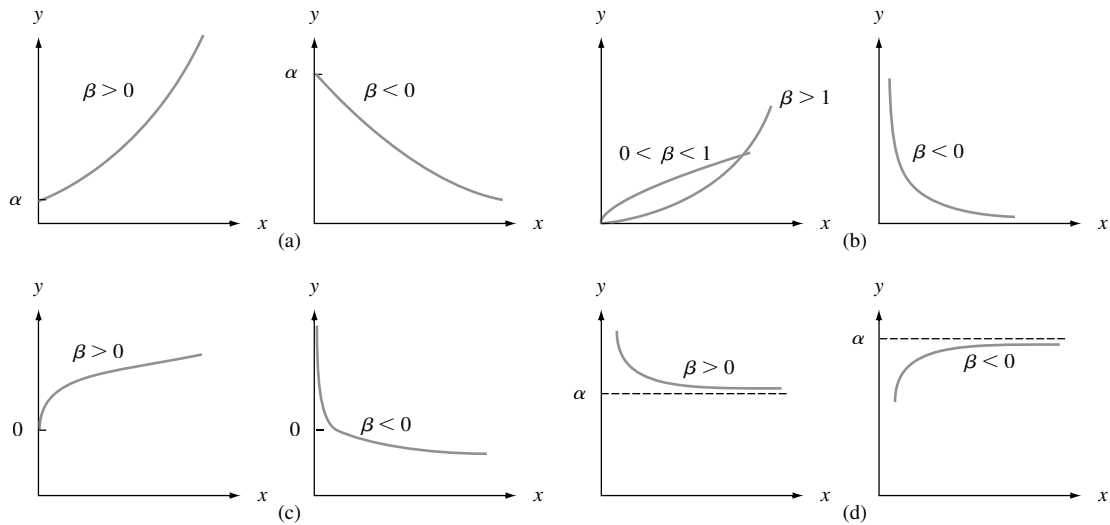


Figura 13.3 Gráficos das funções intrinsecamente lineares mostradas na Tabela 13.1

**DEFINIÇÃO**

Um modelo probabilístico que relacione  $Y$  a  $x$  é **intrinsecamente linear** se, por meio de uma transformação em  $Y$  e/ou em  $x$ , puder reduzir-se a um modelo probabilístico linear  $Y' = \beta_0 + \beta_1 x' + \epsilon'$ .

Os modelos probabilísticos intrinsecamente lineares que correspondem às quatro funções da Tabela 13.1 são os seguintes:

- a.  $Y = \alpha e^{\beta x} \cdot \epsilon$ , um modelo exponencial multiplicativo, de modo que  $\ln(Y) = Y' = \beta_0 + \beta_1 x' + \epsilon'$  com  $x' = x$ ,  $\beta_0 = \ln(\alpha)$ ,  $\beta_1 = \beta$  e  $\epsilon' = \ln(\epsilon)$ .
- b.  $Y = \alpha x^\beta \cdot \epsilon$ , um modelo multiplicativo de potência, de maneira que  $\log(Y) = Y' = \beta_0 + \beta_1 x' + \epsilon'$  com  $x' = \log(x)$ ,  $\beta_0 = \log(\alpha)$ ,  $\beta_1 = \beta$  e  $\epsilon' = \log(\epsilon)$ .
- c.  $Y = \alpha + \beta \log(x) + \epsilon$ , de modo que  $x' = \log(x)$  lineariza imediatamente o modelo.
- d.  $Y = \alpha + \beta \cdot 1/x + \epsilon$ , de maneira que  $x' = 1/x$  produz um modelo linear.

Os modelos aditivos exponencial e de potência,  $Y = \alpha e^{\beta x} + \epsilon$  e  $Y = \alpha x^\beta + \epsilon$ , não são intrinsecamente lineares. Observe que tanto (a) quanto (b) exigem uma transformação em  $Y$  e, conseqüentemente, uma transformação na variável de erro  $\epsilon$ . Na verdade, se  $\epsilon$  tiver uma distribuição lognormal (veja o Capítulo 4) com  $E(\epsilon) = e^{\sigma^2/2}$  e  $V(\epsilon) = \tau^2$  independente de  $x$ , então os modelos transformados tanto de (a) quanto de (b) vão satisfazer todas as hipóteses do Capítulo 12 sobre o modelo probabilístico linear; isso, por sua vez, significa que serão válidas todas as inferências para os parâmetros do modelo transformado que se baseiam nessas hipóteses. Se  $\sigma^2$  for pequeno,  $\mu_{y \cdot x} \approx \alpha e^{\beta x}$  em (a) ou  $\alpha x^\beta$  em (b).

A maior vantagem de um modelo intrinsecamente linear é que os parâmetros  $\beta_0$  e  $\beta_1$  do modelo transformado podem ser estimados de imediato por meio do princípio dos mínimos quadrados apenas substituindo  $x'$  e  $y'$  nas fórmulas de estimação:

$$\hat{\beta}_1 = \frac{\sum x'_i y'_i - \sum x'_i \sum y'_i / n}{\sum (x'_i)^2 - (\sum x'_i)^2 / n} \tag{13.5}$$

$$\hat{\beta}_0 = \frac{\sum y'_i - \hat{\beta}_1 \sum x'_i}{n} = \bar{y}' - \hat{\beta}_1 \bar{x}'$$

Portanto, os parâmetros do modelo não-linear original podem ser estimados transformando de volta  $\hat{\beta}_0$  e  $\hat{\beta}_1$ , se necessário. Uma vez calculado um intervalo de previsão para  $y'$  quando  $x' = x'^*$ , a reversão da transformação fornecerá um IP para o próprio  $y$ . Nos casos (a) e (b), quando  $\sigma^2$  é pequeno, obtém-se um IC aproximado para  $\mu_{y,x^*}$  calculando-se os antilogs dos limites no IC de  $\beta_0 + \beta_1 x'^*$  (estritamente falando, ao calcular os antilogs, obtém-se um IC para a mediana da distribuição  $Y$ , isto é, para  $\tilde{\mu}_{y,x^*}$ . Pelo fato de a distribuição lognormal ter inclinação positiva,  $\mu > \tilde{\mu}$ ; ambas são aproximadamente iguais se  $\sigma^2$  estiver próximo de 0).

### Exemplo 13.3

A equação de Taylor para o tempo de vida de uma ferramenta  $y$  como função do tempo de corte  $x$  expressa que  $xy^c = k$  ou, de modo equivalente, que  $y = \alpha x^\beta$ . O artigo “The Effect of Experimental Error on the Determination of Optimum Metal Cutting Conditions” (*J. Eng. for Industry*, 1967, p. 315–322) observa que a relação não é exata (determinística) e que os parâmetros  $\alpha$  e  $\beta$  devem ser estimados com base nos dados. Assim, um modelo apropriado é o modelo multiplicativo de potência  $Y = \alpha \cdot x^\beta \cdot \epsilon$ , o qual o autor ajusta aos dados a seguir, que consistem em 12 observações sobre o tempo de vida da ferramenta de carboneto (Tabela 13.2). Além dos valores  $x$ ,  $y$ ,  $x'$ , e  $y'$ , são fornecidos os valores transformados previstos ( $\hat{y}'$ ) e os valores previstos na escala original ( $\hat{y}$ , depois de transformados de volta).

**Tabela 13.2** Dados do Exemplo 13.3

	$x$	$y$	$x' = \ln(x)$	$y' = \ln(y)$	$\hat{y}'$	$\hat{y} = e^{\hat{y}'}$
1	600,	2,3500	6,39693	0,85442	1,12754	3,0881
2	600,	2,6500	6,39693	0,97456	1,12754	3,0881
3	600,	3,0000	6,39693	1,09861	1,12754	3,0881
4	600,	3,6000	6,39693	1,28093	1,12754	3,0881
5	500,	6,4000	6,21461	1,85630	2,11203	8,2650
6	500,	7,8000	6,21461	2,05412	2,11203	8,2650
7	500,	9,8000	6,21461	2,28238	2,11203	8,2650
8	500,	16,5000	6,21461	2,80336	2,11203	8,2650
9	400,	21,5000	5,99146	3,06805	3,31694	27,5760
10	400,	24,5000	5,99146	3,19867	3,31694	27,5760
11	400,	26,0000	5,99146	3,25810	3,31694	27,5760
12	400,	33,0000	5,99146	3,49651	3,31694	27,5760

As estatísticas para ajustar uma linha reta aos dados transformados são  $\sum x'_i = 74,41200$ ,  $\sum y'_i = 26,22601$ ,  $\sum x_i'^2 = 461,75874$ ,  $\sum y_i'^2 = 67,74609$  e  $\sum x'_i y'_i = 160,84601$ , de modo que

$$\hat{\beta}_1 = \frac{(160,84601) - (74,41200)(26,22601)/12}{461,75874 - (74,41200)^2/12} = -5,3996$$

$$\hat{\beta}_0 = \frac{26,22601 - (-5,3996)(74,41200)}{12} = 35,6684$$

Os valores estimados de  $\alpha$  e  $\beta$ , os parâmetros do modelo de função potencial, são  $\hat{\beta} = \hat{\beta}_1 = -5,3996$  e  $\hat{\alpha} = e^{\hat{\beta}_0} = 3,094491530 \cdot 10^{15}$ . Portanto, a função de regressão estimada é  $\hat{\mu}_{y,x} \approx 3,094491530 \cdot 10^{15} \cdot x^{-5,3996}$ . Para recapturar a equação (estimada) de Taylor, estabeleça  $y = 3,094491530 \cdot 10^{15} \cdot x^{-5,3996}$ , de onde  $xy^{185} = 740$ .

A Figura 13.4(a) a seguir fornece um gráfico dos resíduos padronizados da regressão linear usando variáveis transformadas (para as quais  $r^2 = 0,922$ ); não há nenhum padrão aparente no gráfico, embora um resíduo padronizado seja um pouco grande, e os resíduos parecem como os de uma regressão linear simples. A Figura 13.4(b) apresenta um gráfico de  $\hat{y}$  versus  $y$ , que indica previsões satisfatórias na escala original.

Para obter um intervalo de confiança para o tempo de vida mediano da ferramenta, com um tempo de corte de 500, transforme  $x = 500$  em  $x' = 6,21461$ . Portanto,  $\hat{\beta}_0 + \hat{\beta}_1 x' = 2,1120$  e um IC de 95% para  $\beta_0 + \beta_1(6,21461)$  é (com base na Seção 12.4)  $2,1120 \pm (2,228)(0,0824) = (1,928, 2,296)$ . O IC de 95% para  $\tilde{\mu}_{y,500}$  é portanto obtido calculando-se os antilogs:  $(e^{1,928}, e^{2,296}) = (6,876, 9,930)$ . É fácil verificar que, para os dados transformados  $s^2 = \hat{\sigma}^2 \approx 0,081$ . Pelo fato de esse valor ser muito pequeno,  $(6,876, 9,930)$  é um intervalo aproximado para  $\mu_{y,500}$ .

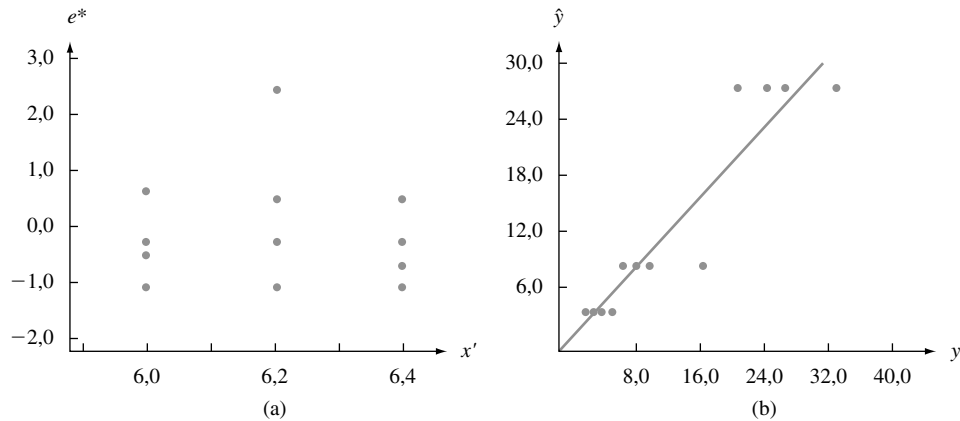


Figura 13.4 (a) Resíduos padronizados versus  $x'$  do Exemplo 13.3; (b)  $\hat{y}$  versus  $y$  do Exemplo 13.3

Exemplo 13.4

No artigo “Ethylene Synthesis in Lettuce Seeds: Its Physiological Significance” (*Plant Physiology*, 1972, p. 719–722), o volume de etileno em sementes de alface ( $y$ , em peso seco, nL/g) foi investigado como uma função do tempo de exposição ( $x$ , em min.) a um absorvente do etileno. A Figura 13.5 apresenta um gráfico de dispersão dos dados e um gráfico dos resíduos gerados de uma regressão linear de  $y$  sobre  $x$ . Ambos os gráficos mostram um padrão curvo significativo, sugerindo que, para conseguir linearidade, é apropriado realizar uma transformação. Além disso, uma regressão linear fornece previsões negativas para  $x = 90$  e  $x = 100$ .

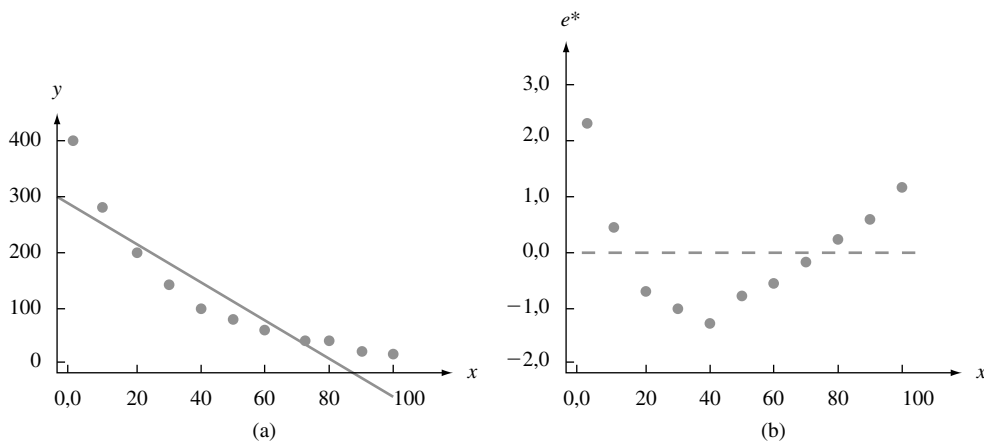


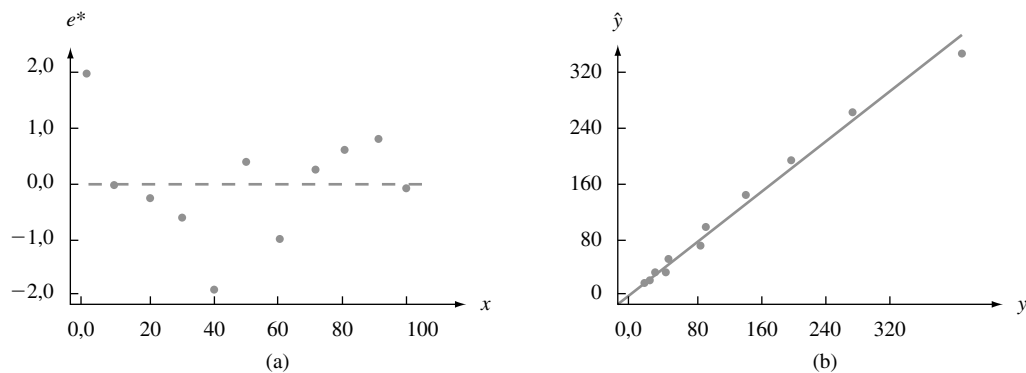
Figura 13.5 (a) Gráfico de dispersão; (b) gráfico dos resíduos de uma regressão linear dos dados do Exemplo 13.4

O autor não apresentou nenhum argumento para um modelo teórico, mas seu gráfico de  $y' = \ln(y)$  versus  $x$  mostra uma sólida relação linear, sugerindo que uma função exponencial fornecerá uma boa aderência aos dados. A Tabela 13.3 mostra os valores dos dados e outras informações da regressão linear de  $y'$  sobre  $x$ . As estimativas dos parâmetros do modelo linear são  $\beta_1 = -0,0323$  e  $\hat{\beta}_0 = 5,941$ , com  $r^2 = 0,995$ . A função de regressão estimada para o modelo exponencial é  $\hat{\mu}_{y,x} \approx e^{\hat{\beta}_0} \cdot e^{\hat{\beta}_1 x} = 380,32e^{-0,0323x}$ . Os valores previstos  $\hat{y}_i$  podem então ser obtidos substituindo-se  $x_i$  ( $i = 1, \dots, n$ ) em  $\hat{\mu}_{y,x}$  ou calculando-se  $\hat{y}_i = e^{\hat{y}'_i}$ , onde os  $\hat{y}'_i$ s são as previsões

do modelo de linha reta transformado. A Figura 13.6 apresenta um gráfico de  $e^{*}$  versus  $x$  (os resíduos padronizados de uma regressão linear) e um gráfico de  $\hat{y}$  versus  $y$ . Esses gráficos apóiam a escolha de um modelo exponencial.

**Tabela 13.3** Dados para o Exemplo 13.4

$x$	$y$	$y' = \ln(y)$	$\hat{y}'$	$\hat{y} = e^{\hat{y}'}$
2	408	6,01	5,876	353,32
10	274	5,61	5,617	275,12
20	196	5,28	5,294	199,12
30	137	4,92	4,971	144,18
40	90	4,50	4,647	104,31
50	78	4,36	4,324	75,50
60	51	3,93	4,001	54,64
70	40	3,69	3,677	39,55
80	30	3,40	3,354	28,62
90	22	3,09	3,031	20,72
100	15	2,71	2,708	15,00



**Figura 13.6** Gráfico dos (a) resíduos padronizados (após a transformação) versus  $x$ ; (b)  $\hat{y}$  versus  $y$  para os dados do Exemplo 13.4

Ao analisar os dados transformados, deve-se ter em mente os seguintes pontos:

1. Estimar  $\beta_1$  e  $\beta_0$  em (13.5) e, em seguida, transformá-los de volta para obter estimativas dos parâmetros originais não equivale a usar o princípio dos mínimos quadrados diretamente no modelo original. Logo, para o modelo exponencial, poderíamos estimar  $\alpha$  e  $\beta$  minimizando  $\sum (y_i - \alpha e^{\beta x_i})^2$ . Seria necessário realizar cálculos iterativos. As estimativas resultantes não seriam iguais:  $\hat{\alpha} \neq e^{\hat{\beta}_0}$  e  $\hat{\beta} \neq \hat{\beta}_1$ .
2. Se o modelo escolhido não for intrinsecamente linear, a abordagem resumida em (13.5) não poderá ser usada. Em vez disso, os mínimos quadrados (ou outro procedimento de calibração) teria de ser aplicado ao modelo não transformado. Desse modo, para o modelo exponencial aditivo  $Y = \alpha e^{\beta x} + \epsilon$ , os mínimos quadrados envolveriam a minimização de  $\sum (y_i - \alpha e^{\beta x_i})^2$ . Tomando-se as derivadas parciais em relação a  $\alpha$  e  $\beta$ , são obtidas duas equações normais não-lineares em  $\alpha$  e  $\beta$ ; portanto, essas equações devem ser solucionadas usando um procedimento iterativo.
3. Quando o modelo linear transformado satisfaz todas as hipóteses relacionadas no Capítulo 12, o método dos mínimos quadrados fornece as melhores estimativas dos parâmetros transformados. Entretanto, as estimativas dos parâmetros originais talvez não sejam de forma alguma as melhores, embora possam

ser razoáveis. Por exemplo, no modelo exponencial, o estimador  $\hat{\alpha} = e^{\hat{\beta}_0}$  não será não-tendencioso, apesar de que será o estimador de máxima verossimilhança de  $\alpha$  se a variável de erro  $\epsilon'$  for normalmente distribuída. Usar os mínimos quadrados diretamente (sem transformar) poderia oferecer estimativas melhores, embora os cálculos possam ser muito cansativos.

4. Se  $y$  tiver sido transformado e alguém desejar usar as fórmulas-padrão para testar hipóteses ou construir ICs,  $\epsilon'$  deve ter no mínimo uma distribuição normal aproximada. Para confirmar isso, os resíduos da regressão transformada devem ser examinados.
5. Quando  $y$  é transformado, o valor  $r^2$  da regressão resultante refere-se à variação nos  $y$ 's explicada pelo modelo de regressão transformado. Embora um valor alto de  $r^2$  indique aqui uma boa aderência do modelo não-linear original estimado aos  $y$ 's,  $r^2$  não se refere a essas observações originais. Talvez a melhor maneira de avaliar a qualidade da calibração seja calcular os valores previstos  $\hat{y}'_i$  usando o modelo transformado, transformá-los de volta à escala  $y$  original para obter  $\hat{y}_i$  e, em seguida, representar  $\hat{y}$  versus  $y$  graficamente. Uma boa aderência é, portanto, evidenciada por pontos próximos da reta de 45°. Poder-se-ia calcular  $SQE = \sum (y_i - \hat{y}_i)^2$  como uma medida numérica da aderência. Quando o modelo era linear, comparamos isso com  $SQT = \sum (y_i - \bar{y})^2$ , a variação total em torno da reta horizontal na altura  $\bar{y}$ ; isso levou ao  $r^2$ . No caso não-linear, contudo, não é necessariamente informativo medir a variação total dessa maneira. Portanto, um valor  $r^2$  não é tão útil quanto no caso linear.

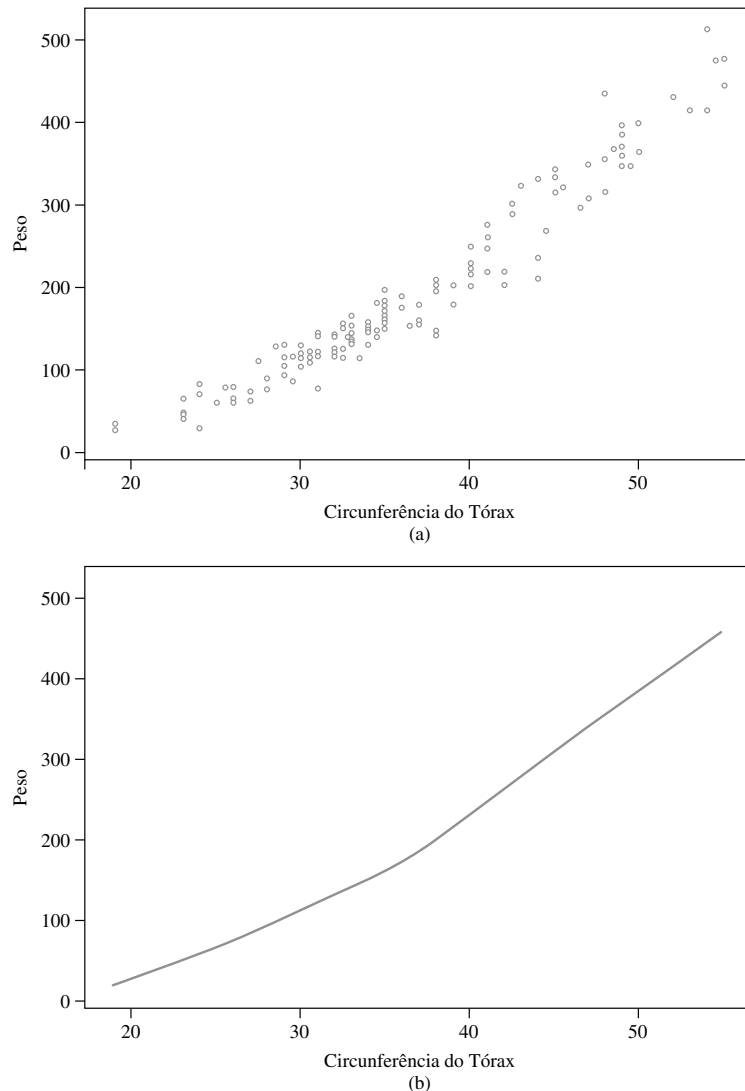
## Métodos de Regressão mais Gerais

Até agora, supusemos que  $Y = f(x) + \epsilon$  (um modelo aditivo) ou que  $Y = f(x) \cdot \epsilon$  (um modelo multiplicativo). No caso de um modelo aditivo,  $\mu_{y|x} = f(x)$ , de modo que estimar a função de regressão  $f(x)$  implica estimar a curva de valores  $y$  médios. Às vezes, um gráfico de dispersão dos dados sugere que não existe nenhuma expressão matemática simples para  $f(x)$ . Os estatísticos desenvolveram recentemente métodos mais flexíveis que permitem que inúmeros padrões sejam modelados por meio do mesmo procedimento de calibração. Um desses métodos é o **LOWESS** (ou LOESS), abreviatura de *locally weighted scatter plot smoother* (regressão localmente ponderada). Seja  $(x^*, y^*)$  um determinado par entre os  $n$  pares  $(x, y)$  na amostra. O valor  $\hat{y}$  correspondente a  $(x^*, y^*)$  é obtido ajustando-se a linha reta usando apenas uma porcentagem específica dos dados (por exemplo, 25%) cujos valores  $x$  estão mais próximos de  $x^*$ . Além disso, em vez de usar os mínimos quadrados “ordinários”, que fornecem o mesmo peso para todos os pontos, aqueles com valores  $x$  mais próximos de  $x^*$  recebem um peso bem maior que aqueles cujos valores  $x$  estão mais afastados. A altura da reta resultante acima de  $x^*$  é o valor ajustado  $\hat{y}^*$ . Esse processo é repetido para cada um dos  $n$  pontos, de modo que  $n$  retas diferentes são ajustadas (você certamente não vai querer fazer tudo isso à mão). Por fim, os pontos ajustados são conectados para gerar uma curva LOWESS.

### Exemplo 13.5

Pesar grandes animais mortos encontrados em áreas selvagens normalmente não é viável. Portanto, é desejável ter um método para estimar o peso de várias características de um animal que possam ser facilmente determinadas. O MINITAB dispõe de um conjunto de dados de várias características de uma amostra de  $n = 143$  ursos selvagens. A Figura 13.7(a) exibe um gráfico de dispersão de  $y =$  peso versus  $x =$  distância ao redor do tórax (circunferência do tórax). À primeira vista, parece que uma única reta obtida dos mínimos quadrados ordinários poderia resumir eficazmente o padrão. A Figura 13.7(b) mostra a curva LOWESS produzida pelo MINITAB usando uma distância de 50% (o ajuste em  $(x^*, y^*)$  é determinado pelos 50% mais próximos da amostra). A curva parece consistir em dois segmentos de linha unidos acima de  $x = 38$ , aproximadamente. A reta mais inclinada está à direita de 38, indicando que o peso tende a aumentar mais rapidamente que a circunferência, no caso de circunferências acima de 38 polegadas.

É complicado fazer outras inferências (por exemplo, obter um IC para um valor  $y$  médio) com base nesse tipo geral de modelo de regressão. A técnica bootstrap mencionada anteriormente pode ser usada para esse propósito.



**Figura 13.7** Gráfico de dispersão do MINITAB e curva LOWESS para dados de peso dos ursos

## Regressão Logística

O modelo de regressão linear simples é apropriado para relacionar uma variável de resposta quantitativa  $y$  com um previsor quantitativo  $x$ . Suponha que  $y$  seja uma variável dicotômica com os possíveis valores 1 e 0 correspondentes a sucesso (acerto) e fracasso (falha). Seja  $p = P(S) = P(y = 1)$ . Em geral o valor de  $p$  dependerá do valor de alguma variável quantitativa  $x$ . Por exemplo, a probabilidade de um carro precisar de determinados serviços cobertos pela garantia pode muito bem depender da quilometragem do carro ou a probabilidade de evitar uma infecção de um determinado tipo pode depender da dosagem na inoculação. Em vez de usar apenas o símbolo  $p$  para a probabilidade de sucesso, agora usamos  $p(x)$  para ressaltar a dependência dessa probabilidade em relação ao valor de  $x$ . A equação da regressão linear simples  $Y = \beta_0 + \beta_1 x + \epsilon$  não é mais apropriada, visto que, se tomarmos o valor médio em cada lado da equação, teremos

$$\mu_{y,x} = 1 \cdot p(x) + 0 \cdot (1 - p(x)) = p(x) = \beta_0 + \beta_1 x$$

Embora  $p(x)$  seja uma probabilidade e, portanto, deva estar entre 0 e 1,  $\beta_0 + \beta_1 x$  não precisam estar nesse intervalo.

Em vez de permitir que o valor médio de  $y$  seja uma função linear de  $x$ , consideramos agora um modelo no qual alguma função do valor médio de  $y$  é uma função linear de  $x$ . Em outras palavras, permitimos que  $p(x)$  seja



uma função de  $\beta_0 + \beta_1 x$ , em vez de  $\beta_0 + \beta_1 x$  propriamente dita. Uma função que se mostrou muito útil em várias aplicações é a **função logit**

$$p(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}$$

A Figura 13.8 mostra um gráfico de  $p(x)$  para determinados valores de  $\beta_0$  e  $\beta_1$  com  $\beta_1 > 0$ . À medida que  $x$  aumenta, a probabilidade de sucesso aumenta. Para  $\beta_1$  negativo, a probabilidade de sucesso seria uma função decrescente de  $x$ .

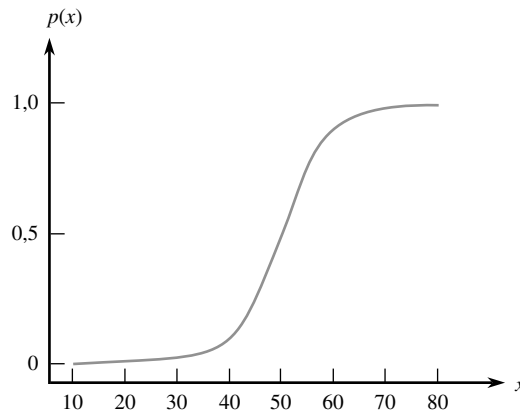


Figura 13.8 Gráfico de uma função logística

A *regressão logística* implica supor que  $p(x)$  está relacionada a  $x$  pela função logit. Uma álgebra simples mostra que

$$\frac{p(x)}{1 - p(x)} = e^{\beta_0 + \beta_1 x}$$

A expressão no lado esquerdo é denominada *razão das chances*. Se, por exemplo,  $\frac{p(60)}{1 - p(60)} = 3$ , quando  $x = 60$ , um sucesso é três vezes mais provável do que um fracasso. Percebemos agora que o logaritmo da razão das chances é uma função linear do predictor. Especificamente falando, o parâmetro do coeficiente angular  $\beta_1$  é a mudança nas chances logarítmicas associada ao aumento de uma unidade em  $x$ . Isso significa que a razão das chances em si muda segundo o fator multiplicativo  $e^{\beta_1}$  quando  $x$  aumenta em 1 unidade.

Ajustar a regressão logística aos dados amostrais exige que os parâmetros  $\beta_0$  e  $\beta_1$  sejam estimados. Normalmente isso é feito usando a técnica de máxima verossimilhança descrita no Capítulo 6. Os detalhes são muito complicados, mas, felizmente, os softwares estatísticos mais populares executarão essa tarefa quando solicitados e fornecerão indicações quantitativas e ilustrativas sobre a qualidade de aderência do modelo.

**Exemplo 13.6**

Aqui são apresentados os dados de temperatura de lançamento e a incidência de falhas nos anéis circulares (O-rings) em 24 lançamentos de naves espaciais antes do desastre da *Challenger* em janeiro de 1986.

Temperatura	Falha	Temperatura	Falha	Temperatura	Falha
53	S	68	N	75	N
56	S	69	N	75	S
57	S	70	N	76	N
63	N	70	S	76	N
66	N	70	S	78	N
67	N	70	S	79	N
67	N	72	N	80	N
67	N	73	N	81	N

A Figura 13.9 (apresentada logo a seguir) mostra o resultado do software estatístico JMP usado para uma análise de regressão logística. Optamos por supor que  $p$  seja a probabilidade de falha. As falhas tendem a ocorrer em temperaturas mais baixas e os acertos em temperaturas mais altas, de modo que o gráfico de  $\hat{p}$  diminui à medida que a temperatura aumenta. A estimativa de  $\beta_1$  é  $\hat{\beta}_1 = -0,1713$ , e o desvio padrão estimado de  $\hat{\beta}_1$  é  $s_{\hat{\beta}_1} = 0,08344$ . Contanto que  $n$  seja grande o suficiente e, nesse caso, supomos que esse seja o caso,  $\hat{\beta}_1$  terá aproximadamente uma distribuição normal. Se  $\beta_1 = 0$  (a temperatura não influencia na probabilidade de falha dos anéis circulares),  $z = \hat{\beta}_1/s_{\hat{\beta}_1}$  tem aproximadamente uma distribuição normal padronizada. O valor da razão  $z$  é  $-2,05$  e o valor  $P$  de um teste bicaudal é  $0,0404$  (duas vezes a área sob a curva  $z$  à esquerda de  $-2,05$ ). O JMP exibe o valor de uma estatística qui-quadrado, que é exatamente  $z^2$ , e o valor  $P$  qui-quadrado difere do valor de  $z$  somente por causa do arredondamento. Para cada 1 grau de aumento na temperatura, a chance de falha diminui por um fator ao redor de  $0,84$ . A temperatura de lançamento na missão da *Challenger* era de apenas  $31^\circ\text{F}$ . Pelo fato de esse valor ser bem menor do que qualquer temperatura em nossa amostra, é perigoso extrapolar a relação estimada. Entretanto, parece que para uma temperatura tão baixa, a falha no anel circular é quase certa.

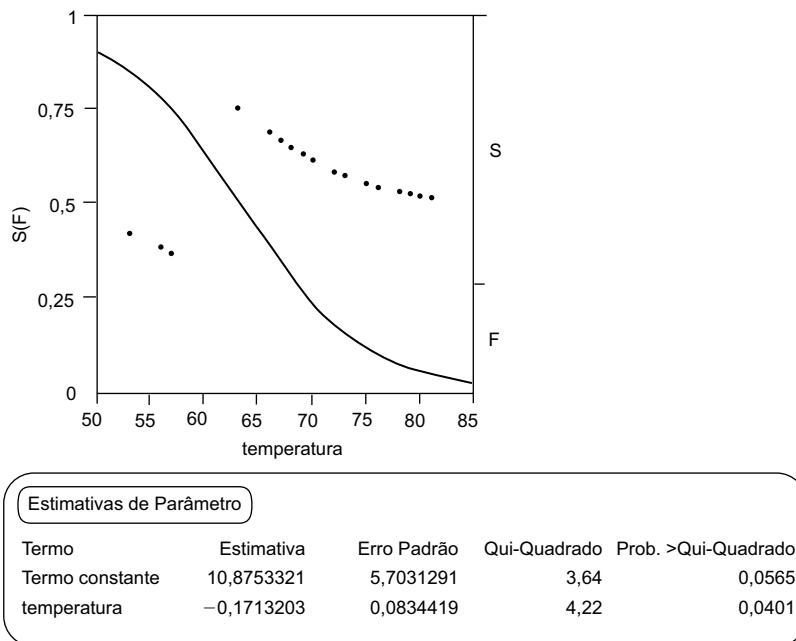


Figura 13.9 Resultado de uma regressão logística do JMP

**Exercícios** | Seção 13.2 (15–25)

15. Nenhum aficionado por chips de tortilla gosta de chips engordurados, de modo que é fundamental identificar características do processo de produção que resultem chips com uma textura atraente. Os dados a seguir sobre  $x =$  tempo de fritura (s) e  $y =$  quantidade de umidade (%) foram apresentados no artigo “Thermal and Physical Properties of Tortilla Chips as a Function of Frying Time” (*J. of Food Processing and Preservation*, 1995, p. 175–189).

$x$	5	10	15	20	25	30	45	60
$y$	16,3	9,7	8,1	4,2	3,4	2,9	1,9	1,3

a. Construa um gráfico de dispersão de  $y$  versus  $x$  e comente.

b. Construa um gráfico de dispersão de pares  $(\ln(x), \ln(y))$  e comente.

c. Que relação probabilística entre  $x$  e  $y$  é sugerida pelo padrão linear no gráfico do item (b)?

d. Presuma o valor da quantidade de umidade quando o tempo de fritura é 20, de uma maneira que transmita informações sobre confiabilidade e precisão.

e. Analise os resíduos do ajuste do modelo de regressão linear simples aos dados transformados e comente.

16. O artigo “The Luminosity–Spectral Index Relationship for Radio Galaxies” (*Nature*, 1972, p. 88–89) propôs que, para as galáxias da classe S,  $\ln(L_{178})$  está relacionada linearmente com o índice espectral, onde  $L_{178}$

representa luminosidade a 178 MHz. Aqui são apresentados os dados representativos:

<b>Índice Espectral</b>	0,59	0,67	0,72	0,80	0,85	0,90
<b>ln(L<sub>178</sub>)</b>	23,6	25,6	26,4	25,7	26,8	26,7
<b>Índice Espectral</b>	0,94	0,66	1,00	0,86	1,03	0,70
<b>ln(L<sub>178</sub>)</b>	27,0	24,9	27,1	27,2	26,9	25,2

- a. Estime os parâmetros do modelo exponencial sugeridos pela relação linear entre ln(L<sub>178</sub>) e o índice espectral.
  - b. Que valor de L<sub>178</sub> seria previsto para um índice espectral de 0,75?
  - c. Calcule um intervalo de previsão de 95% para a luminosidade quando o índice espectral é 0,95.
17. Os dados a seguir sobre a taxa de combustão de massa x e comprimento da chama y são representativos dos dados apresentados no artigo "Some Burning Characteristics of Filter Paper" (*Combustion Science and Technology*, 1971, p. 103–120):

x	1,7	2,2	2,3	2,6	2,7	3,0	3,2
y	1,3	1,8	1,6	2,0	2,1	2,2	3,0
x	3,3	4,1	4,3	4,6	5,7	6,1	
y	2,6	4,1	3,7	5,0	5,8	5,3	

- a. Estime os parâmetros de um modelo de função potencial.
  - b. Construa gráficos de diagnóstico para verificar se uma função potencial é uma opção de modelo apropriada.
  - c. Teste  $H_0: \beta = \frac{4}{3}$  versus  $H_a: \beta < \frac{4}{3}$ , usando um nível de teste 0,05.
  - d. Teste a hipótese nula que expressa que a mediana do comprimento da chama quando a taxa de combustão é 5,0 é duas vezes a mediana do comprimento da chama quando a taxa de combustão é 2,5 em relação à possibilidade de que isso não é verdadeiro.
18. As falhas decorrentes de fadiga de alto ciclo em motores de avião de turbina a gás são um problema que tende a se propagar. O artigo "Effect of Crystal Orientation on Fatigue Failure of Single Crystal Nickel Base Turbine Blade Superalloys" (*J. of Engineering for Gas Turbines and Power*, 2002, p. 161–176) apresentou os dados a seguir e ajustou um modelo de regressão não-linear a fim de prever a amplitude do esforço dos ciclos até a fratura. Ajuste um modelo apropriado, investigue a qualidade desse ajuste e preveja a amplitude quando o número de ciclos até a fratura = 5.000.

Obs	Ciclo de Falha	Ampl. do Esforço	Obs	Ciclo até a Falha	Ampl. do Esforço
1	1.326	0,01495	11	7.356	0,00576
2	1.593	0,01470	12	7.904	0,00580
3	4.414	0,01100	13	79	0,01212
4	5.673	0,01190	14	4.175	0,00782
5	29.516	0,00873	15	34.676	0,00596
6	26	0,01819	16	114.789	0,00600
7	843	0,00810	17	2.672	0,00880
8	1.016	0,00801	18	7.532	0,00883
9	3.410	0,00600	19	30.220	0,00676
10	7.101	0,00575			

19. Testes de resistência térmica foram executados para investigar a relação entre a temperatura e o tempo de vida de fios esmaltados de poliéster ("Thermal Endurance of Polyester Enameled Wires Using Twisted Wire Specimens", *IEEE Trans. Insulation*, 1965, p. 38–44), gerando os dados a seguir.

<b>Temp.</b>	200	200	200	200	200	200
<b>Tempo de vida</b>	5.933	5.404	4.947	4.963	3.358	3.878
<b>Temp.</b>	220	220	220	220	220	220
<b>Tempo de vida</b>	1.561	1.494	747	768	609	777
<b>Temp.</b>	240	240	240	240	240	240
<b>Tempo de vida</b>	258	299	209	144	180	184

- a. Um gráfico de dispersão dos dados sugere uma relação probabilística linear entre o tempo de vida e a temperatura?
  - b. Que modelo é sugerido por uma relação linear entre os valores esperados do ln(tempo de vida) e 1/temperatura? Um gráfico de dispersão dos dados transformados parece consistente com a relação?
  - c. Estime os parâmetros do modelo sugerido no item (b). Que tempo de vida seria previsto para uma temperatura de 220?
  - d. Pelo fato de haver múltiplas observações em cada valor x, o método do Exercício 14 pode ser usado para testar a hipótese nula que expressa que o modelo sugerido no item (b) é correto. Execute um teste no nível 0,01.
20. O Exercício 14 apresentou dados sobre o peso corporal x e a taxa de depuração metabólica/peso corporal y. Considere as seguintes funções intrinsecamente lineares para especificar a relação entre as duas variáveis: (a) ln(y) versus x, (b) ln(y) versus ln(x), (c) y versus ln(x), (d) y versus 1/x e (e) ln(y) versus 1/x. Use qualquer gráfico de diagnóstico e análise apropriados para determinar qual dessas

funções você selecionaria para especificar um modelo probabilístico. Explique seu raciocínio.

21. Um gráfico apresentado no artigo “Thermal Conductivity of Polyethylene: The Effects of Crystal Size, Density, and Orientation on the Thermal Conductivity” (*Polymer Eng. and Science*, 1972, p. 204–208) sugere que o valor esperado de condutividade térmica  $y$  é uma função linear de  $10^4 \cdot 1/x$ , onde  $x$  é a espessura lamelar.

$x$	240	410	460	490	520	590	745	8.300
$y$	12,0	14,7	14,7	15,2	15,2	15,6	16,0	18,1

- a. Estime os parâmetros da função de regressão e a própria função de regressão.  
 b. Preveja o valor da condutividade térmica quando a espessura lamelar é 500 Å.
22. Em cada um dos casos a seguir, determine se a função dada é intrinsecamente linear. Se sim, identifique  $x'$  e  $y'$  e, em seguida, explique como um termo de erro aleatório  $\epsilon$  pode ser introduzido para gerar um modelo probabilístico intrinsecamente linear.
- a.  $y = 1/(\alpha + \beta x)$   
 b.  $y = 1/(1 + e^{\alpha + \beta x})$   
 c.  $y = e^{\alpha + \beta x}$  (uma curva de Gompertz)  
 d.  $y = \alpha + \beta e^{\lambda x}$

23. Suponha que  $x$  e  $y$  estejam relacionados de acordo com um modelo probabilístico exponencial  $Y = \alpha e^{\beta x} \cdot \epsilon$ , com  $V(\epsilon)$  como uma constante independente de  $x$  (como no caso do modelo linear simples  $Y = \beta_0 + \beta_1 x + \epsilon$ ).  $V(Y)$  é uma constante independente de  $x$  [como no caso de  $Y = \beta_0 + \beta_1 x + \epsilon$ , onde  $V(Y) = \sigma^2$ ]? Explique seu raciocínio. Desenhe um gráfico de dispersão como um protótipo resultante desse modelo. Responda as mesmas perguntas para o modelo potencial  $Y = \alpha x^\beta \cdot \epsilon$ .

24. O termo cifose refere-se a uma grave curvatura protuberante da coluna vertebral que necessita de cirurgia corretiva. Um estudo realizado para determinar os fatores de risco da cifose relatou as idades a seguir (meses) para 40 indivíduos no momento da cirurgia; os primeiros 18 indivíduos tiveram cifose e os 22 restantes não.

<b>Com cifose</b>	12	15	42	52	59	73
	82	91	96	105	114	120
	121	128	130	139	139	157
<b>Sem cifose</b>	1	1	2	8	11	18
	22	31	37	61	72	81
	97	112	118	127	131	140
	151	159	177	206		

Use o resultado da regressão logística do MINITAB apresentado a seguir para determinar se a idade parece ter um impacto significativo sobre a existência da cifose.

25. Os dados a seguir resultaram de um estudo patrocinado por uma grande empresa de consultoria de gerenciamento para investigar a relação entre a quantidade de experiência profissional (meses) de um consultor júnior e a probabilidade de o consultor ser capaz de executar uma determinada tarefa complexa.

<b>Sucesso</b>	8	13	14	18	20	21	21	22	25
	26	28	29	30	32				
<b>Fracasso</b>	4	5	6	6	7	9	10	11	11
	13	15	18	19	20	23	27		

Interprete o resultado da regressão logística do MINITAB apresentado a seguir e desenhe um gráfico da probabilidade estimada da execução da tarefa como uma função da experiência.

Tabela de regressão logística do Exercício 24

Previsor	Coef	DP	Z	P	Razão das chances	95% Inferior	CI Superior
Constante	-0,5727	0,6024	-0,95	0,342			
idade	0,004296	0,005849	0,73	0,463	1,00	0,99	1,02

Tabela de regressão logística do Exercício 25

Previsor	Coef	DP	Z	P	Razão das chances	95% Inferior	CI superior
Constante	-3,211	1,235	-2,60	0,009			
idade	0,17772	0,06573	2,70	0,007	1,19	1,05	1,36

### 13.3 | Regressão Polinomial

Os modelos não-lineares, mas intrinsecamente lineares da Seção 13.2 compreendiam funções da variável independente  $x$  que eram rigorosamente crescentes ou rigorosamente decrescentes. Em várias situações, o raciocínio teórico, ou então um gráfico de dispersão dos dados, sugere que a função de regressão real  $\mu_{y,x}$  tem um ou mais picos ou vales — isto é, pelo menos um mínimo ou máximo relativo. Em casos desse tipo, uma função polinomial  $y = \beta_0 + \beta_1 x + \dots + \beta_k x^k$  pode fornecer uma aproximação satisfatória para a função de regressão real.

DEFINIÇÃO

A equação de modelo de regressão polinomial de  $k$ -ésimo grau é

$$Y = \beta_0 + \beta_1x + \beta_2x^2 + \dots + \beta_kx^k + \epsilon \tag{13.6}$$

onde  $\epsilon$  é uma variável aleatória normalmente distribuída com

$$\mu_\epsilon = 0 \quad \sigma_\epsilon^2 = \sigma^2 \tag{13.7}$$

Com base em (13.6) e (13.7), conclui-se imediatamente que

$$\mu_{Y \cdot x} = \beta_0 + \beta_1x + \dots + \beta_kx^k \quad \sigma_{Y \cdot x}^2 = \sigma^2 \tag{13.8}$$

Traduzindo em palavras, o valor esperado de  $Y$  é uma função polinomial de  $k$ -ésimo grau de  $x$ , ao passo que a variância de  $Y$ , que controla a dispersão dos valores observados ao redor da função de regressão, é a mesma para cada valor de  $x$ . Acredita-se que os pares  $(x_1, y_1), \dots, (x_n, y_n)$  foram gerados independentemente do modelo (13.6). A Figura 13.10 ilustra um modelo quadrático e um modelo cúbico.

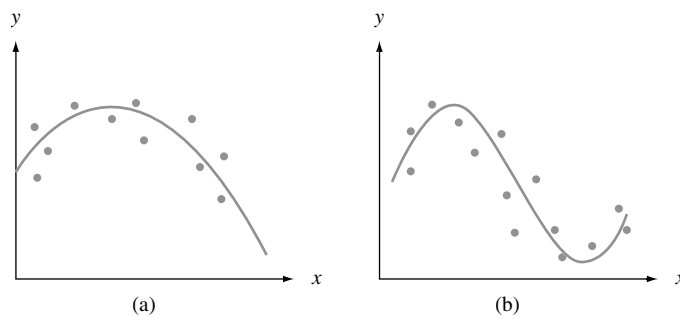


Figura 13.10 (a) Modelo de regressão quadrática; (b) modelo de regressão cúbica

### Estimativa de Parâmetros Usando os Mínimos Quadrados

Para estimar  $\beta_0, \beta_1, \dots, \beta_k$ , considere uma função de regressão de teste  $y = b_0 + b_1x + \dots + b_kx^k$ . Portanto, a qualidade da aderência dessa função aos dados observados pode ser avaliada calculando-se a soma dos desvios quadrados

$$f(b_0, b_1, \dots, b_k) = \sum_{i=1}^n [y_i - (b_0 + b_1x_i + b_2x_i^2 + \dots + b_kx_i^k)]^2 \tag{13.9}$$

De acordo com o princípio dos mínimos quadrados, as estimativas  $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$  são aqueles valores de  $b_0, b_1, \dots, b_k$  que minimizam a Expressão (13.9). É necessário observar que quando  $x_1, x_2, \dots, x_n$  são todos diferentes existe um polinômio de grau  $n - 1$  que se ajusta aos dados perfeitamente, de modo que o valor que minimiza (13.9) é 0 quando  $k = n - 1$ . Entretanto, em praticamente todas as aplicações, o modelo polinomial (13.6) com um  $k$  grande é muito improvável e, na maioria das aplicações,  $k = 2$  (quadrático) ou  $k = 3$  (cúbico) é apropriado.

Para encontrar os valores que minimizam (13.9), tomamos as  $k + 1$  derivadas parciais  $\partial f/\partial b_0, \partial f/\partial b_1, \dots, \partial f/\partial b_k$  e as igualamos a 0, gerando o sistema de equações normais para as estimativas. Pelo fato de a função experimental  $b_0 + b_1x + \dots + b_kx^k$  ser linear em  $b_0, \dots, b_k$  (embora não em  $x$ ), as  $k + 1$  equações normais são lineares para essas incógnitas:

$$\begin{aligned} b_0n + b_1\sum x_i + b_2\sum x_i^2 + \dots + b_k\sum x_i^k &= \sum y_i \\ b_0\sum x_i + b_1\sum x_i^2 + b_2\sum x_i^3 + \dots + b_k\sum x_i^{k+1} &= \sum x_iy_i \\ \vdots & \\ b_0\sum x_i^k + b_1\sum x_i^{k+1} + \dots + b_k\sum x_i^{2k} &= \sum x_i^ky_i \end{aligned} \tag{13.10}$$

Todos os softwares estatísticos padrão solucionarão automaticamente as equações em (13.10) e fornecerão as estimativas, bem como muitas outras informações.<sup>2</sup>

### Exemplo 13.7

O artigo “Determination of Biological Maturity and Effect of Harvesting and Drying Conditions on Milling Quality of Paddy” (*J. Agricultural Eng. Research*, 1975, p. 353-361) relata os dados a seguir sobre a data  $x$  da colheita (número de dias após a floração) e a produção  $y$  (kg/ha) de arroz com casca (paddy), grão cultivado na Índia:

$x$	16	18	20	22	24	26	28	30
$y$	2.508	2.518	3.304	3.423	3.057	3.190	3.500	3.883
$x$	32	34	36	38	40	42	44	46
$y$	3.823	3.646	3.708	3.333	3.517	3.241	3.103	2.776

O gráfico de dispersão na Figura 13.11(a) sugere que a escolha do autor de um modelo quadrático é razoável. A Figura 13.11(b) também fornece o resultado do MINITAB de uma regressão quadrática, do qual

$$\hat{\beta}_0 = -1070,4 \quad \hat{\beta}_1 = 293,48 \quad \hat{\beta}_2 = -4,5358$$

e a função de regressão estimada é

$$y = -1070,4 + 293,48x - 4,5358x^2$$

As funções de regressão estimadas podem ser usadas para calcular valores previstos (substituir  $x_1 = 16$ , em seguida  $x_2 = 18$  etc., para obter  $\hat{y}_1, \hat{y}_2, \dots$ ) e, portanto, os resíduos ( $y_i - \hat{y}_i$ ). Um gráfico de resíduos indica que uma função de regressão quadrática é apropriada, embora haja um pouco mais de dispersão nos resíduos para um  $x$  pequeno do que para um  $x$  grande. É possível usar mínimos quadrados ponderados para obter uma aderência mais satisfatória, mas os autores não consideraram este caso.

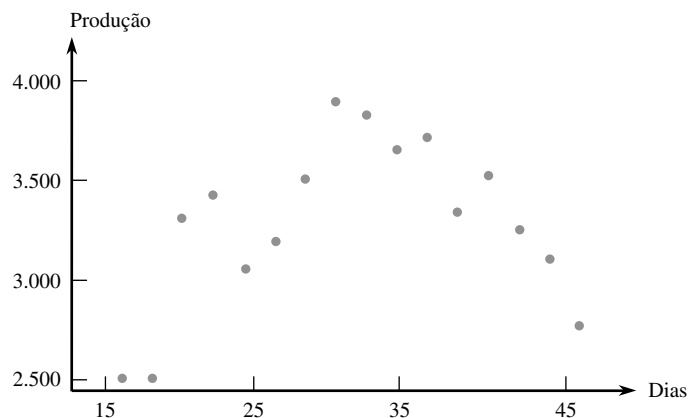


Figura 13.11(a) (a) Gráfico de dispersão do Exemplo 13.7

<sup>2</sup> Veremos na Seção 13.4 que a regressão polinomial é um caso especial de regressão múltipla, de modo que em geral se usa um comando apropriado para essa última tarefa.

A EQUAÇÃO DE REGRESSÃO É PRODUÇÃO = -1.070 + 293 DIAS - 4,54 DIAS QUADRADOS

Previsora	Coef	DP	razão t	p
Constante	-1070,4	617,3	-1,73	0,107
DIAS	293,48	42,18	6,96	0,000
DIAS QUADRADOS	-4,5358	0,6744	-6,73	0,000

s = 203,9 Raiz Quadrada = 79,4% R<sup>2</sup> (ajustado) = 76,2%

Análise de Variância

FONTE	GL	SQ	QM	F	p
Regressão	2	2084779	1042389	25,08	0,000
Erro	13	540388	41568		
Total	15	2625167			

Obs.	DIAS	PRODUÇÃO	Aderência	Aderência DP	Resíduo	Resíduo-padrão
1	16,0	2508,0	2464,2	135,6	43,8	0,29
2	18,0	2518,0	2742,7	104,8	-224,7	-1,28
3	20,0	3304,0	2984,9	83,0	319,1	1,71
4	22,0	3423,0	3190,9	71,3	232,1	1,22
5	24,0	3057,0	3360,6	68,4	-303,6	-1,58
6	26,0	3190,0	3494,0	70,7	-304,0	-1,59
7	28,0	3500,0	3591,1	74,2	-91,1	-0,48
8	30,0	3883,0	3651,9	76,4	231,1	1,22
9	32,0	3823,0	3676,4	76,4	146,6	0,78
10	34,0	3646,0	3664,6	74,2	-18,6	-0,10
11	36,0	3708,0	3616,6	70,7	91,4	0,48
12	38,0	3333,0	3532,3	68,4	-199,3	-1,04
13	40,0	3517,0	3411,6	71,3	105,4	0,55
14	42,0	3241,0	3254,7	83,0	-13,7	-0,07
15	44,0	3103,0	3061,5	104,8	41,5	0,24
16	46,0	2776,0	2832,1	135,6	-56,1	-0,37

Aderência	Aderência DP	IC 95%	95% IP
3.431,8	69,2	(3.282,2, 3.581,4)	(2.966,5, 3.897,1)

Figura 13.11(b) (b) Resultado do MINITAB para o Exemplo 13.7

## $\hat{\sigma}^2$ e $R^2$

Para fazer inferências adicionais sobre os parâmetros da função de regressão, a variância do erro  $\sigma^2$  deve ser estimada. Com  $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \dots + \hat{\beta}_k x_i^k$ , o  $i$ -ésimo resíduo é  $y_i - \hat{y}_i$ , e a soma dos resíduos quadrados (soma dos quadrados do erro) é  $SQE = \sum (y_i - \hat{y}_i)^2$ . A estimativa de  $\sigma^2$  é, portanto,

$$\hat{\sigma}^2 = s^2 = \frac{SQE}{n - (k + 1)} = QME \quad (13.11)$$

onde o denominador  $n - (k + 1)$  é usado porque  $k + 1$  gl são perdidos na estimativa de  $\beta_0, \beta_1, \dots, \beta_k$ .

Se, novamente, considerarmos  $SQT = \sum (y_i - \bar{y})^2$ , então  $SQE/SQT$  será a proporção da variação total dos  $y_i$  observados que não é explicada pelo modelo polinomial. A quantidade  $1 - SQE/SQT$ , a proporção da variação explicada pelo modelo, é denominada **coeficiente de determinação múltipla** e é representada por  $R^2$ .

Suponha que desejamos ajustar um modelo cúbico aos dados do Exemplo 13.7. Pelo fato de o modelo cúbico incluir o quadrático como um caso especial, o ajuste a um modelo cúbico será no mínimo tão adequado quanto o de um modelo quadrático. De forma geral, com  $SQE_k =$  soma dos quadrados do erro de um polinômio de  $k$ -ésimo grau,  $SQE_{k'} \leq SQE_k$  e  $R_k^2 \geq R_{k'}^2$  quando  $k' > k$ . Pelo fato de o objetivo da análise de regressão ser encontrar um modelo simples (relativamente poucos parâmetros) e que ao mesmo tempo ofereça aderência aos dados, um polinômio de grau mais alto talvez não especifique um modelo mais adequado do que um modelo de menor grau, não obstante o respectivo valor  $R^2$  mais alto. Para equilibrar o custo de usar mais parâmetros e o ganho em  $R^2$ , vários estatísticos usam o **coeficiente de determinação múltipla ajustado**

$$R^2 \text{ ajustado} = 1 - \frac{n - 1}{n - (k + 1)} \cdot \frac{SQE}{SQT} = \frac{(n - 1)R^2 - k}{n - 1 - k} \quad (13.12)$$

O  $R^2$  ajustado ajusta de modo crescente a proporção da variação não-explicada [visto que a razão  $(n - 1)/(n - k - 1)$  ultrapassa 1], o que resulta em  $R^2$  ajustado  $< R^2$ . Logo, se  $R_2^2 = 0,66$ ,  $R_3^2 = 0,70$  e  $n = 10$ , então

$$R_2^2 \text{ ajustado} = \frac{9(0,66) - 2}{10 - 3} = 0,563 \quad R_3^2 \text{ ajustado} = \frac{9(0,70) - 3}{10 - 4} = 0,550$$

de modo que o pequeno ganho em  $R^2$  de passar de um modelo quadrático para um modelo cúbico não é suficiente para compensar o custo de adicionar um parâmetro extra ao modelo.

**Exemplo 13.8** (Continuação do Exemplo 13.7)

SQE e SQT normalmente são encontradas em tabelas ANOVA produzidas por computador. A Figura 13.11 fornece SQE = 540.388 e SQT = 2.625.167 para os dados da produção de arroz com casca (paddy), donde

$$R^2 = 1 - \frac{\text{SQE}}{\text{SQT}} = 1 - \frac{540.388}{2.625.167} = 1 - 0,206 = 0,794$$

Portanto, 79,4% da variação observada na produção pode ser atribuída à relação do modelo.  $R^2$  ajustado = 0,762 é apenas uma pequena mudança decrescente em  $R^2$ . As estimativas de  $\sigma^2$  e  $\sigma$  são

$$s^2 = \frac{\text{SQE}}{n - (k + 1)} = \frac{540.388}{16 - (2 + 1)} = 41.568,3$$

$$s = \sqrt{s^2} = 203,9$$

Além de calcular  $R^2$  e  $R^2$  ajustado, deve-se examinar os gráficos de diagnóstico habituais para verificar se as hipóteses do modelo são válidas ou se a modificação pode ser apropriada.

## Intervalos Estatísticos e Procedimentos de Teste

Pelo fato de os  $y_i$ s aparecerem nas equações normais (13.10) apenas no lado direito e de maneira linear, as estimativas resultantes  $\hat{\beta}_0, \dots, \hat{\beta}_k$  são funções lineares dos  $y_i$ s. Logo, os estimadores são funções lineares dos  $Y_i$ s, de modo que cada  $\hat{\beta}_i$  tem uma distribuição normal. É possível mostrar que cada  $\hat{\beta}_i$  é um estimador não-tendencioso de  $\beta_i$ .

Seja  $\sigma_{\hat{\beta}_i}$  o desvio padrão do estimador  $\hat{\beta}_i$ . Esse desvio padrão tem a forma

$$\sigma_{\hat{\beta}_i} = \sigma \cdot \left\{ \begin{array}{l} \text{uma expressão complexa envolvendo todos os} \\ x_j, x_j^2, \dots, \text{ e } x_j^k \end{array} \right\}$$

Felizmente, a expressão entre chaves foi programada em todos os softwares estatísticos mais frequentemente usados. O desvio padrão estimado de  $\hat{\beta}_i$ ,  $s_{\hat{\beta}_i}$ , resulta da substituição de  $\sigma$  por  $s$  na expressão de  $\sigma_{\hat{\beta}_i}$ . Esses desvios padrão estimados,  $s_{\hat{\beta}_0}, s_{\hat{\beta}_1}, \dots, \text{ e } s_{\hat{\beta}_k}$ , aparecem no resultado de todos os pacotes estatísticos anteriormente mencionados. Seja  $S_{\hat{\beta}_i}$  o estimador de  $\sigma_{\hat{\beta}_i}$  — ou seja, a variável aleatória cujo valor observado é  $s_{\hat{\beta}_i}$ . Logo, é possível mostrar que a variável padronizada

$$T = \frac{\hat{\beta}_i - \beta_i}{S_{\hat{\beta}_i}} \quad (13.13)$$

tem uma distribuição  $t$  com base em  $n - (k + 1)$  gl. Isso conduz aos procedimentos inferenciais a seguir.



Um IC de  $100(1 - \alpha)\%$  para  $\beta_i$ , o coeficiente de  $x^i$  na função de regressão polinomial, é

$$\hat{\beta}_i \pm t_{\alpha/2, n-(k+1)} \cdot s_{\hat{\beta}_i}$$

Um teste de  $H_0: \beta_i = \beta_{i0}$  baseia-se no valor da estatística  $t$

$$t = \frac{\hat{\beta}_i - \beta_{i0}}{s_{\hat{\beta}_i}}$$

O teste tem por base  $n - (k + 1)$  gl e é de cauda superior, cauda inferior ou bicaudal, dependendo se a desigualdade em  $H_a$  for  $>$ ,  $<$  ou  $\neq$ .

Uma estimativa pontual de  $\mu_{y \cdot x}$  — ou seja, de  $\beta_0 + \beta_1 x + \dots + \beta_k x^k$  — é  $\hat{\mu}_{y \cdot x} = \hat{\beta}_0 + \hat{\beta}_1 x + \dots + \hat{\beta}_k x^k$ . O desvio padrão esperado do estimador correspondente é muito complexo. Vários softwares fornecerão esse desvio padrão estimado para qualquer valor  $x$  quando o usuário solicitar. Isso, com a variável  $t$  padronizada apropriada, pode ser usado para justificar os procedimentos a seguir.

Seja  $x^*$  um valor específico de  $x$ . Um IC de  $100(1 - \alpha)\%$  para  $\mu_{y \cdot x^*}$  é

$$\hat{\mu}_{y \cdot x^*} \pm t_{\alpha/2, n-(k+1)} \cdot \left\{ \begin{array}{l} \text{DP estimado de} \\ \hat{\mu}_{y \cdot x^*} \end{array} \right\}$$

Com  $Y = \hat{\beta}_0 + \hat{\beta}_1 x^* + \dots + \hat{\beta}_k (x^*)^k$ ,  $\hat{y}$  representando o valor calculado de  $\hat{Y}$  para os dados especificados e  $s_{\hat{y}}$  representando o desvio padrão estimado da estatística  $\hat{Y}$ , a fórmula para o IC é mais parecida com a da regressão linear simples:

$$\hat{y} \pm t_{\alpha/2, n-(k+1)} \cdot s_{\hat{y}}$$

Um IP de  $100(1 - \alpha)\%$  para um valor  $y$  futuro ser observado quando  $x = x^*$  é

$$\hat{\mu}_{y \cdot x^*} \pm t_{\alpha/2, n-(k+1)} \cdot \left\{ s^2 + \left( \text{DP estimado de} \right)^2 \right\}^{1/2} = \hat{y} \pm t_{\alpha/2, n-(k+1)} \cdot \sqrt{s^2 + s_{\hat{y}}^2}$$

### Exemplo 13.9

A Figura 13.11(b) mostra que  $\beta_2 = -4,5358$  e  $s_{\hat{\beta}_2} = 0,6744$  (da coluna do desvio padrão à parte superior do resultado). Para testar  $H_0: \beta_2 = 0$  versus  $H_a: \beta_2 \neq 0$  ( $H_0$  aqui expressa que o termo quadrático no modelo é desnecessário), a estatística de teste é  $T = \hat{\beta}_2 / s_{\hat{\beta}_2}$ , com o valor calculado  $-4,5358 / 0,6744 = -6,73$ . O teste baseia-se em  $n - (k + 1) = 16 - 3 = 13$  gl; com  $t_{0,025,13} = 2,160$ ,  $H_0$  é rejeitada no nível 0,05 se  $t \geq 2,160$  ou  $t \leq -2,160$ . Visto que  $-6,73 \leq -2,160$ ,  $H_0$  é rejeitada no nível 0,05, validando a inclusão do termo quadrático; o valor  $P$  correspondente é 0.

A última linha do resultado na Figura 13.11(b) contém informações sobre a estimativa e previsão de  $x = 25$ :

$$\begin{aligned} \hat{y} &= \hat{\beta}_0 + \hat{\beta}_1(25) + \hat{\beta}_2(25)^2 = \text{Aderência} = 3.431,8 \\ \left( \text{DP estimado de} \right) &= s_{\hat{y}} = \text{Aderência DP} = 69,2 \\ \left( \hat{\beta}_0 + \hat{\beta}_1(25) + \hat{\beta}_2(25)^2 \right) & \end{aligned}$$

Então, o IC de 95% para a produção média quando a data da colheita = 25 é

$$3.431,8 \pm (2,160)(69,2) = 3.431,8 \pm 149,5 = (3.282,3, 3.581,3)$$

ao passo que um IP de 95% para a produção quando  $x = 25$  é

$$3.431,8 \pm (2.160)[41.568,3 + (69,2)^2]^{1/2} = 3.431,8 \pm 465,1 = (2.966,7, 3896,9)$$

As pequenas discrepâncias entre esses limites de intervalo e aqueles mostrados no resultado podem ser atribuídas ao arredondamento. Observe novamente que o IP é bem mais amplo que o IC correspondente. Isso ocorre porque  $s^2$  é muito grande em comparação com a aderência do desvio padrão. ■

### Centralizando Valores $x$

Para o modelo quadrático com função de regressão  $\mu_{y,x} = \beta_0 + \beta_1x + \beta_2x^2$ , os parâmetros  $\beta_0$ ,  $\beta_1$  e  $\beta_2$  caracterizam o comportamento da função próximo de  $x = 0$ . Por exemplo,  $\beta_0$  é a altura em que a função de regressão cruza o eixo vertical  $x = 0$ , ao passo que  $\beta_1$  é a primeira derivada da função em  $x = 0$  (taxa instantânea da mudança de  $\mu_{y,x}$  em  $x = 0$ ). Se todos os  $x_i$ s ficarem distantes de 0, provavelmente não teremos informações precisas sobre os valores desses parâmetros. Considere  $\bar{x}$  = a média dos  $x_i$ s para os quais serão feitas observações e examine o modelo

$$Y = \beta_0^* + \beta_1^*(x - \bar{x}) + \beta_2^*(x - \bar{x})^2 + \epsilon \tag{13.14}$$

No modelo (13.14),  $\mu_{y,x} = \beta_0^* + \beta_1^*(x - \bar{x}) + \beta_2^*(x - \bar{x})^2$ , e os parâmetros agora descrevem o comportamento da função de regressão próximo ao centro  $\bar{x}$  dos dados.

Para estimar os parâmetros de (13.14), simplesmente subtraímos  $\bar{x}$  de cada  $x_i$  para obter  $x'_i = x_i - \bar{x}$  e, em seguida, usamos os  $x'_i$ s no lugar dos  $x_i$ s. Um importante benefício que disso traz é que os coeficientes de  $b_0, \dots, b_k$  nas equações normais (13.10) terão uma magnitude bem menor que no caso em que os  $x_i$ s originais foram usados. Quando o sistema é solucionado por computador, essa centralização evita qualquer possível erro de arredondamento.

#### Exemplo 13.10

O artigo “A Method for Improving the Accuracy of Polynomial Regression Analysis” (*J. Quality Technology*, 1971, p. 149–155) relata os dados a seguir sobre  $x$  = temperatura de tratamento (°F) e  $y$  = resistência máxima de cisalhamento de um composto de borracha (psi), com  $\bar{x} = 297,13$ :

$x$	280	284	292	295	298	305	308	315
$x'$	-17,13	-13,13	-5,13	-2,13	0,87	7,87	10,87	17,87
$y$	770	800	840	810	735	640	590	560

Uma análise por computador gerou os resultados mostrados na Tabela 13.4.

**Tabela 13.4** Coeficientes e desvios padrão estimados para o Exemplo 13.10

Parâmetro	Estimativa	DP Estimado	Parâmetro	Estimativa	DP Estimado
$\beta_0$	-26.219,64	11,912,78	$\beta_0^*$	759,36	23,20
$\beta_1$	189,21	80,25	$\beta_1^*$	-7,61	1,43
$\beta_2$	-0,3312	0,1350	$\beta_2^*$	-0,3312	0,1350

A função de regressão estimada usando o modelo original é  $y = -26.219,64 + 189,21x - 0,3312x^2$ , ao passo que, para o modelo centralizado, é  $y = 759,36 - 7,61(x - 297,13) - 0,3312(x - 297,13)^2$ . Essas funções estimadas são idênticas; a única distinção é que parâmetros diferentes foram estimados para os dois modelos. Os desvios padrão estimados indicam nitidamente que  $\beta_0^*$  e  $\beta_1^*$  foram estimados com maior precisão que  $\beta_0$  e  $\beta_1$ . Os parâmetros quadráticos são idênticos ( $\beta_2 = \beta_2^*$ ), como pode ser visto comparando-se o termo  $x^2$  em (13.14) com o modelo original. Enfatizamos novamente que o maior benefício da centralização é o ganho de precisão dos cálculos, não apenas nos modelos quadráticos, mas também em modelos de grau elevado. ■

O livro de Neter et al., relacionado na Bibliografia deste capítulo, é uma excelente fonte para informações adicionais sobre regressão polinomial.

**Exercícios** | Seção 13.3 (26–35)

**26.** Além da regressão linear da densidade real na quantidade de umidade, o artigo citado no Exercício 6 avaliou a regressão quadrática da densidade de massa versus quantidade de umidade. Os dados de um gráfico do artigo são apresentados a seguir, com o resultado do MINITAB de aderência quadrática.

A equação de regressão é  
 $\text{densdemassa} = 403 + 16,2 \text{ quantumidade} - 0,706 \text{ quantquad.}$

Previsor	Coef	Desv Padrão	T	P
Constante	403,24	36,45	11,06	0,002
quantumidade	16,164	5,451	2,97	0,059
quantquad.	-0,7063	0,1852	-3,81	0,032

S = 10,15 R quad. = 93,8% R quad. (ajustado) = 89,6%

Análise de Variância

Fonte	GL	SQ	QM	F	P
Regressão	2	4637,7	2318,9	22,51	0,016
Erro Residual	3	309,1	103,0		
Total	5	4946,8			

Obs	Quant umidade	densde- massanc	Ade- rência	DP	Resíduo	Resíduo Padrão
1	7,0	479,00	481,78	9,35	-2,78	-0,70
2	10,3	503,00	494,79	5,78	8,21	0,98
3	13,7	487,00	492,12	6,49	-5,12	-0,66
4	16,6	470,00	476,93	6,10	-6,93	-0,85
5	19,8	458,00	446,39	5,69	11,61	1,38
6	22,0	412,00	416,99	8,75	-4,99	-0,97

Aderência	Aderência	IC 95%	95% IP
491,10	6,52	(470,36, 511,83)	(452,71, 529,48)

- a. Um gráfico de dispersão dos dados parece consistente com o modelo de regressão quadrática?
- b. Que proporção da variação observada na densidade pode ser atribuída à relação do modelo?
- c. O modelo quadrático parece útil? Execute um teste ao nível de significância 0,05.
- d. A última linha do resultado provém da solicitação de informações sobre estimativa e previsão quando uma quantidade de umidade é 14. Calcule um PI de 99% para a densidade quando a quantidade de umidade é 14.
- e. A previsão quadrática parece fornecer informações úteis? Teste as hipóteses apropriadas ao nível de significância 0,05.

**27.** Os dados a seguir sobre  $y =$  concentração de glicose (g/L) e  $x =$  tempo de fermentação (dias) para uma determinada mistura de licor de malte foram reproduzidos de um gráfico do artigo “Improving Fermentation Productivity with Reverse Osmosis” (*Food Tech.*, 1984, p. 92–96):

x	1	2	3	4	5	6	7	8
y	74	54	52	51	52	53	58	71

- a. Verifique se o gráfico de dispersão dos dados é consistente com a escolha do modelo de regressão quadrática.
- b. A equação de regressão quadrática estimada é  $y = 84,482 - 15,875x + 1,7679x^2$ . Preveja o valor da concentração de glicose para um tempo de fermentação de 6 dias e calcule o resíduo correspondente.
- c. Usando  $SQE = 61,77$ , que proporção da variação observada pode ser atribuída à relação de regressão quadrática?
- d. Os  $n = 8$  resíduos padronizados, baseados no modelo quadrático, são 1,91, -1,95, -0,25, 0,58, 0,90, 0,04, -0,66 e 0,20. Construa um gráfico dos resíduos padronizados versus  $x$  e um gráfico de probabilidade normal. Os gráficos exibem alguma característica indesejável?
- e. O desvio padrão estimado de  $\hat{\mu}_{y=6}$  — isto é,  $\hat{\beta}_0 + \hat{\beta}_1(6) + \hat{\beta}_2(36)$  — é 1,69. Calcule um IC de 95% para  $\mu_{y=6}$ .
- f. Calcule um IP de 95% para uma observação sobre concentração de glicose feita seis dias depois do tempo de fermentação.

**28.** A viscosidade ( $y$ ) de um óleo foi medida por um viscosímetro de cone e chapa em seis diferentes velocidades de cone ( $x$ ). Supôs-se que um modelo de regressão quadrática era apropriado, e a função de regressão estimada com base em  $n = 6$  observações foi

$$y = -113,0937 + 3,3684x - 0,01780x^2$$

- a. Estime  $\mu_{y=75}$ , a viscosidade esperada quando a velocidade é 75 rpm.
- b. Que viscosidade é prevista para uma velocidade de cone de 60 rpm?
- c. Se  $\sum y_i^2 = 8.386,43$ ,  $\sum y_i = 210,70$ ,  $\sum x_i y_i = 17.002,00$  e  $\sum x_i^2 y_i = 1.419,780$ , calcule  $SQE [= \sum y_i^2 - \hat{\beta}_0 \sum y_i - \hat{\beta}_1 \sum x_i y_i - \hat{\beta}_2 \sum x_i^2 y_i]$ ,  $s^2$  e  $s$ .
- d. Com base no item (c),  $SQT = 8.386,43 - (210,70)^2/6 = 987,35$ . Usando a  $SQE$  calculada no item (c), qual é o valor calculado de  $R^2$ ?
- e. Se o desvio padrão estimado de  $\hat{\beta}_2$  for  $s_{\hat{\beta}_2} = 0,00226$ , teste  $H_0: \beta_2 = 0$  versus  $H_a: \beta_2 \neq 0$  ao nível de significância 0,01 e interprete o resultado.

**29.** O Exercício 7 apresentou os dados a seguir sobre o tempo de exposição à radiação  $x$  e o peso seco das raízes  $y$ :

x	0	2	4	6	8
y	110	123	119	86	62

A função de regressão quadrática estimada é  $y = 111,89 + 8,06x - 1,84x^2$ .

- a. Calcule os valores previstos e os resíduos. Em seguida, calcule SQE e  $s^2$ . [Veja o Exercício 28(c).]
- b. Calcule o coeficiente de determinação múltipla  $R^2$ .
- c. O desvio padrão estimado de  $\hat{\beta}_2$ , o estimador do coeficiente quadrático  $\beta_2$  é  $s_{\hat{\beta}_2} = 0,480$ . O termo quadrático pertence ao modelo? Expresse e teste as hipóteses apropriadas usando um nível de 0,05.
- d. O desvio padrão estimado de  $\hat{\beta}_1$  é  $s_{\hat{\beta}_1} = 4,01$ . Use este dado e as informações do item (c) para obter ICs conjuntos para  $\beta_1$  e  $\beta_2$  com um nível de confiança conjunto de (no mínimo) 95%.
- e. O desvio padrão estimado de  $\hat{\mu}_{y,4} (= \hat{\beta}_0 + 4\hat{\beta}_1 + 16\hat{\beta}_2)$  é 5,01. Calcule um IC de 90% para  $\mu_{y,4}$ .
- f. Estime o tempo de exposição que maximiza o peso seco esperado das raízes.

30. O artigo “A Simulation–Based Evaluation of Three Cropping Systems on Cracking–Clay Soils in a Summer Rainfall Environment” (*Agricultural Meteorology*, 1976, p. 211–229) propõe um modelo quadrático para a relação entre o índice de abastecimento de água ( $x$ ) e a produção de trigo de uma fazenda ( $y$ ). A seguir são apresentados dados representativos e o resultado do MINITAB:

x	1,2	1,3	1,5	1,8	2,1	2,3	2,5
y	790	950	740	1.230	1.000	1.465	1.370

x	2,9	3,1	3,2	3,3	3,9	4,0	4,3
y	1.420	1.625	1.600	1.720	1.500	1.550	1.560

A equação de regressão é  
 produção = -252 + 1.000 índice - 135 índicequad.

Previsor	Coef	DP	razão t	p
Constante	-251,6	285,1	-0,88	0,396
Índice	1.000,1	229,5	4,36	0,001
Índicequad	-135,44	41,97	-3,23	0,008

$s = 135,6$  R quadrado = 85,3%  
 R quad. (ajustado) = 82,6%

Análise de Variância

FONTE	GL	SQ	QM	F	p
Regressão	2	1170208	585104	31,83	0,000
Erro	11	202228	18384		
Total	13	1372435			

- a. Interprete o valor do coeficiente de determinação múltipla.
- b. Calcule o IC de 95% para o coeficiente do previsor quadrático.

- c. O desvio padrão estimado de  $\hat{\beta}_0 + \hat{\beta}_1x + \hat{\beta}_2x^2$  quando  $x = 2,5$  é 53,5. Teste  $H_0: \mu_{y,2,5} = 1.500$  versus  $H_a: \mu_{y,2,5} < 1.500$  usando  $\alpha = 0,01$ .
- d. Obtenha um IP de 95% para a produção de trigo quando o índice de abastecimento de água é 2,5 usando as informações fornecidas no item (c).

31. Os dados a seguir foram obtidos de um estudo sobre determinado método de preparar álcool puro proveniente da corrente das refinarias (“Direct Hydration of Olefins”, *Industrial and Eng. Chemistry*, 1961, p. 209–211). A variável independente  $x$  é a velocidade espacial horária do volume e a variável dependente  $y$  é a quantidade de conversão de isobutileno.

x	1	1	2	4	4	4	6
y	23,0	24,5	28,0	30,9	32,0	33,6	20,0

- a. Supondo que um modelo probabilístico quadrático seja apropriado, estime a função de regressão.
- b. Determine os valores previstos e os resíduos e construa um gráfico dos resíduos. O gráfico parece mais ou menos como o esperado quando o modelo quadrático é adequado? O gráfico indica que toda observação teve grande influência sobre a aderência? Um gráfico de dispersão identifica algum ponto de grande influência? Se sim, qual?
- c. Obtenha  $s^2$  e  $R^2$ . O modelo quadrático oferece uma boa aderência aos dados?
- d. No Exercício 11 observou-se que o valor previsto  $\hat{Y}_j$  e o resíduo  $Y_j - \hat{Y}_j$  são independentes entre si, de modo que  $\sigma^2 = V(Y_j) = V(\hat{Y}_j) + V(Y_j - \hat{Y}_j)$ . Uma listagem impressa de computador fornece os seguintes desvios padrão estimados dos valores previstos 0,955, 0,955, 0,712, 0,777, 0,777, 0,777 e 1,407. Use esses valores com  $s^2$  para calcular o desvio padrão estimado de cada resíduo. Em seguida, calcule os resíduos padronizados e represente-os graficamente em relação a  $x$ . O gráfico se parece mais com o gráfico do item (b)? Suponha que você tivesse padronizado os resíduos usando apenas  $s$  no denominador. A diferença entre os valores resultantes e os valores corretos seria maior?
- e. Usando as informações fornecidas no item (d), calcule um IP de 90% para a conversão de isobutileno quando a velocidade espacial horária do volume é 4.

32. Os dados a seguir são um subconjunto de dados obtido em um experimento realizado para investigar a relação entre o pH do solo  $x$  e  $y =$  concentração de alumínio/CE (“Root Responses of Three *Gramineae* Species to Soil Acidity in an Oxisol and an Ultisol”, *Soil Science*, 1973, p. 295–302):

x	4,01	4,07	4,08	4,10	4,18
y	1,20	0,78	0,83	0,98	0,65

x	4,20	4,23	4,27	4,30	4,41
y	0,76	0,40	0,45	0,39	0,30

x	4,45	4,50	4,58	4,68	4,70	4,77
y	0,20	0,24	0,10	0,13	0,07	0,04

Foi proposto um modelo cúbico no artigo, mas a versão do MINITAB usada pelo autor do presente texto recusou-se a incluir o termo  $x^3$  no modelo, afirmando que “ $x^3$  está altamente correlacionado com outras variáveis de previsão”. Para solucionar esse problema,  $\bar{x} = 4,3456$  foi subtraído de cada valor  $x$  para obter  $x' = x - \bar{x}$ . Portanto, foi solicitada uma regressão cúbica para ajustar o modelo cuja função de regressão era

$$y = \beta_0^* + \beta_1^*x' + \beta_2^*(x')^2 + \beta_3^*(x')^3$$

Foi obtido o seguinte resultado produzido por computador:

Parâmetro	Estimativa	DP Estimado
$\beta_0^*$	0,3463	0,0366
$\beta_1^*$	-1,2933	0,2535
$\beta_2^*$	2,3964	0,5699
$\beta_3^*$	-2,3968	2,4590

- Qual é a função de regressão estimada para o modelo “centralizado”?
  - Qual é o valor estimado do coeficiente  $\beta_3$  no modelo “não-centralizado” com função de regressão  $y = \beta_0 + \beta_1x + \beta_2x^2 + \beta_3x^3$ ? Qual é a estimativa de  $\beta_2$ ?
  - Usando o modelo cúbico, que valor de  $y$  seria previsto quando o nível de pH do solo for 4,5?
  - Execute um teste para determinar se o termo cúbico deve ser mantido no modelo.
33. Em vários problemas polinomiais, em vez de ajustar uma função de regressão “centralizada” usando  $x' = x - \bar{x}$ , é possível melhorar a precisão dos cálculos usando uma função da variável independente padronizada  $x' = (x - \bar{x})/s_x$ , onde  $s_x$  é o desvio padrão dos  $x_i$ s. Experimente ajustar a função de regressão cúbica  $y = \beta_0^* + \beta_1^*x' + \beta_2^*(x')^2 + \beta_3^*(x')^3$  aos dados a seguir, resultantes de um estudo sobre a relação entre a eficiência do empuxo  $y$  de foguetes de propulsão supersônica e o ângulo de semidivergência  $x$  da tubeira do foguete (“More on Correlating Data”, *CHEMTECH*, 1976, p. 266-270):

x	5	10	15	20	25	30	35
y	0,985	0,996	0,988	0,962	0,940	0,915	0,878

Parâmetro	Estimativa	DP Estimado
$\beta_0^*$	0,9671	0,0026
$\beta_1^*$	-0,0502	0,0051
$\beta_2^*$	-0,0176	0,0023
$\beta_3^*$	0,0062	0,0031

- Que valor de  $y$  seria previsto quando o ângulo de semidivergência fosse 20? Quando  $x = 25$ ?
- Qual é a função de regressão estimada de  $\hat{\beta}_0 + \hat{\beta}_1x + \hat{\beta}_2x^2 + \hat{\beta}_3x^3$  para o modelo “não-padronizado”?

- Use um teste de nível 0,05 para determinar se o termo cúbico deve ser excluído do modelo.
- O que você pode dizer sobre a relação entre as SQEs e os  $R^2$ s nos modelos padronizado e não-padronizado? Explique.
- A SQE do modelo cúbico é 0,00006300, ao passo que a SQE de um modelo quadrático é 0,00014367. Calcule  $R^2$  para cada modelo. A diferença entre ambos sugere que o termo cúbico pode ser excluído?

34. Os dados a seguir resultaram de um experimento para avaliar o potencial dos restos das minas de carvão não-queimado como meio para o desenvolvimento das plantas. As variáveis são  $x =$  cátions extraíveis do ácido e  $y =$  acidez trocável/capacidade de troca de cátions total (“Exchangeable Acidity in Unburnt Colliery Spoil”, *Nature*, 1969, p. 161):

x	-23	-5	16	26	30	38	52
y	1,50	1,46	1,32	1,17	0,96	0,78	0,77
x	58	67	81	96	100	113	
y	0,91	0,78	0,69	0,52	0,48	0,55	

Padronizando a variável independente  $x$  para obter  $x' = (x - \bar{x})/s_x$  e ajustando a função de regressão  $y = \beta_0^* + \beta_1^*x' + \beta_2^*(x')^2$ , obteve-se o seguinte resultado produzido por computador.

Parâmetro	Estimativa	DP Estimado
$\beta_0^*$	0,8733	0,0421
$\beta_1^*$	-0,3255	0,0316
$\beta_2^*$	0,0448	0,0319

- Estime  $\mu_{y,50}$ .
  - Calcule o valor do coeficiente de determinação múltipla. [Veja o Exercício 28(c).]
  - Qual é a função de regressão estimada  $\beta_0 + \hat{\beta}_1x + \hat{\beta}_2x^2$  usando a variável  $x$  não-padronizada?
  - Qual é o desvio padrão estimado de  $\hat{\beta}_2$  calculado no item (c)?
  - Execute um teste usando as estimativas padronizadas para determinar se o termo quadrático deve ser mantido no modelo. Repita usando as estimativas não-padronizadas. Suas conclusões são diferentes?
35. O artigo “The Respiration in Air and in Water of the Limpets *Patella caerulea* and *Patella lusitanica*” (*Comp. Biochemistry and Physiology*, 1975, p. 407-411) propôs um modelo de potência simples para a relação entre a taxa de respiração  $y$  e a temperatura  $x$  de *P. caerulea* no ar. Entretanto, um gráfico de  $\ln(y)$  versus  $x$  exibe um padrão curvo. Ajuste o modelo de potência quadrático  $Y = \alpha e^{\beta x + \gamma x^2} \cdot \epsilon$  para os dados a seguir.

x	10	15	20	25	30
y	37,1	70,1	109,7	177,2	222,6

## 13.4 | Análise de Regressão Múltipla

Na regressão múltipla, o objetivo é elaborar um modelo probabilístico que relacione uma variável dependente  $y$  a mais de uma variável independente ou de previsão. Seja  $k$  o número de variáveis de previsão ( $k \geq 2$ ) e representemos esses previsores por  $x_1, x_2, \dots, x_k$ . Por exemplo, ao tentar prever o preço de venda de uma casa, devemos ter  $k = 3$  com  $x_1 =$  tamanho (ft<sup>2</sup>),  $x_2 =$  idade (anos) e  $x_3 =$  número de ambientes.

### DEFINIÇÃO

A equação do modelo de regressão múltipla aditivo geral é

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \epsilon \quad (13.15)$$

onde  $E(\epsilon) = 0$  e  $V(\epsilon) = \sigma^2$ . Além disso, para testar hipóteses e calcular ICs ou IPs, supõe-se que  $\epsilon$  seja normalmente distribuído.

Sejam  $x_1^*, x_2^*, \dots, x_k^*$  os valores particulares de  $x_1, \dots, x_k$ . Portanto, (13.15) implica que

$$\mu_{Y|x_1^*, \dots, x_k^*} = \beta_0 + \beta_1 x_1^* + \cdots + \beta_k x_k^* \quad (13.16)$$

Logo, assim como  $\beta_0 + \beta_1 x$  descreve o valor  $Y$  médio como uma função de  $x$  na regressão linear simples, a **função de regressão real** (ou **populacional**)  $\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k$  fornece o valor esperado de  $Y$  como uma função de  $x_1, \dots, x_k$ . Os  $\beta_i$ s são os **coeficientes de regressão reais** (ou **populacionais**). O coeficiente de regressão  $\beta_1$  é interpretado como a mudança esperada em  $Y$  associada com um aumento de uma unidade em  $x_1$ , enquanto  $x_2, \dots, x_k$  são mantidos fixos. Interpretações análogas valem para  $\beta_2, \dots, \beta_k$ .

### Modelos com Interação e Previsores Quadráticos

Se um investigador tiver obtido observações sobre  $y, x_1$  e  $x_2$ , um modelo possível é  $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$ . Entretanto, outros modelos podem ser elaborados formando previsores que são funções matemáticas de  $x_1$  e/ou  $x_2$ . Por exemplo, com  $x_3 = x_1^2$  e  $x_4 = x_1 x_2$ , o modelo

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \epsilon$$

tem a forma geral de (13.15). Em geral, além de alguns previsores poderem ser funções matemáticas de outros, eles são também altamente desejáveis no sentido de que o modelo resultante seja mais adequado para explicar a variação em  $y$  do que qualquer modelo que não disponha desses previsores. Essa discussão mostra igualmente que a regressão polinomial é, na verdade, um caso especial de regressão múltipla. Por exemplo, o modelo quadrático  $Y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$  tem a forma de (13.15) com  $k = 2, x_1 = x$  e  $x_2 = x^2$ .

No caso de duas variáveis independentes,  $x_1$  e  $x_2$ , há quatro modelos úteis de regressão múltipla.

1. O modelo de primeira ordem:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

2. O modelo de segunda ordem sem interação:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1^2 + \beta_4 x_2^2 + \epsilon$$

3. O modelo com previsores de primeira ordem e interação:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \epsilon$$

4. O modelo completo de segunda ordem ou modelo quadrático completo:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1^2 + \beta_4 x_2^2 + \beta_5 x_1 x_2 + \epsilon$$

Compreender as diferenças entre esses modelos é um passo inicial fundamental na elaboração de modelos de regressão realistas, com base nas variáveis independentes em estudo.

O modelo de primeira ordem é a generalização mais direta da regressão linear simples. Ele expressa que, para um valor fixo de uma ou de outra variável, o valor esperado de  $Y$  é uma função linear da outra variável e que a mudança esperada em  $Y$  para um aumento de uma unidade em  $x_1$  ( $x_2$ ) é  $\beta_1$  ( $\beta_2$ ) independentemente do nível de  $x_2$  ( $x_1$ ). Logo, se representarmos graficamente a função de regressão como uma função de  $x_1$ , para vários valores diferentes de  $x_2$ , obteremos como contornos da função de regressão um conjunto de retas paralelas, como ilustrado na Figura 13.12(a). A função  $y = \beta_0 + \beta_1x_1 + \beta_2x_2$  especifica um plano em um espaço tridimensional; o modelo de primeira ordem expressa que cada valor observado da variável dependente corresponde a um ponto que se desvia verticalmente desse plano por um valor aleatório  $\epsilon$ .

De acordo com o modelo de segunda ordem sem interação, se  $x_2$  for fixo, a mudança esperada em  $Y$  para um aumento de 1 unidade em  $x_1$  é

$$\beta_0 + \beta_1(x_1 + 1) + \beta_2x_2 + \beta_3(x_1 + 1)^2 + \beta_4x_2^2 - (\beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1^2 + \beta_4x_2^2) = \beta_1 + \beta_3 + 2\beta_3x_1$$

Pelo fato de essa mudança esperada não depender de  $x_2$ , os contornos da função de regressão para diferentes valores de  $x_2$  permanecem paralelos entre si. Entretanto, a dependência da mudança esperada em relação ao valor de  $x_1$  significa que os contornos são agora curvas, e não linhas retas. Isso é ilustrado na Figura 13.12(b). Nesse caso, a superfície da regressão não é mais um plano em um espaço tridimensional, mas uma superfície curva.

Os contornos da função de regressão do modelo de primeira ordem com interação não são linhas retas paralelas. Isso ocorre porque a mudança esperada em  $Y$  quando  $x_1$  sofre um aumento de 1 é

$$\beta_0 + \beta_1(x_1 + 1) + \beta_2x_2 + \beta_3(x_1 + 1)x_2 - (\beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1x_2) = \beta_1 + \beta_3x_2$$

Essa mudança esperada depende do valor de  $x_2$ , de modo que cada linha de contorno deve ter um coeficiente angular diferente, como na Figura 13.12(c). A palavra *interação* reflete o fato de que uma mudança esperada em  $Y$ , quando uma variável aumenta em valor, depende do valor da outra variável.

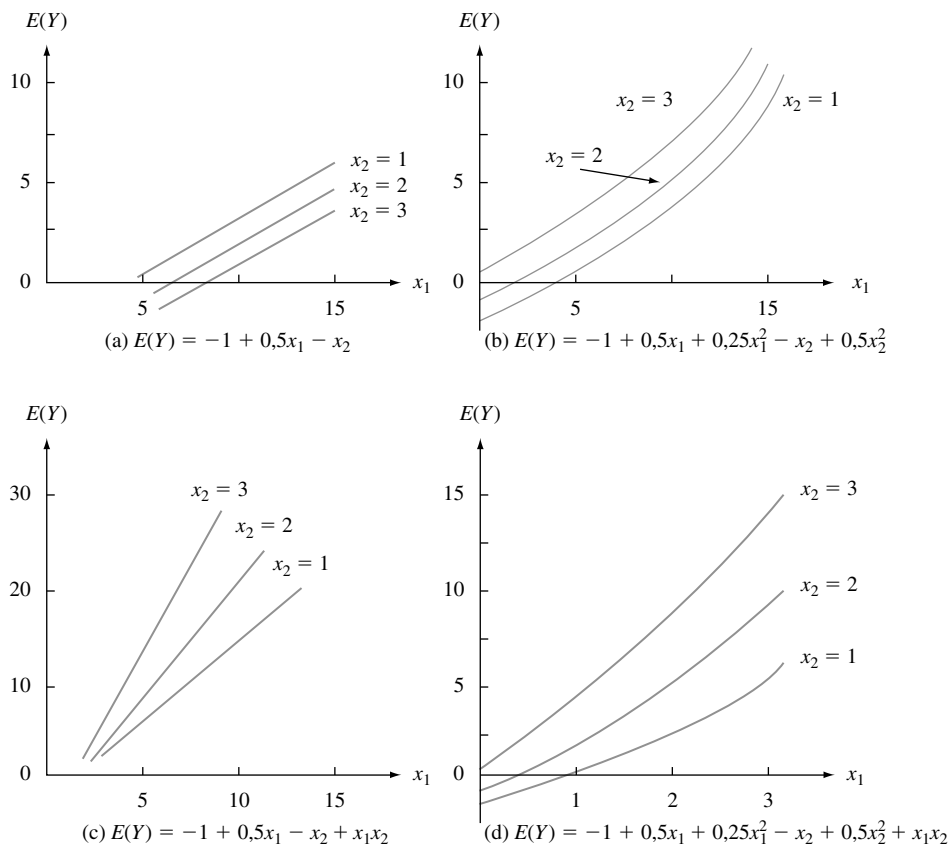


Figura 13.12 Contornos de quatro funções de regressão diferentes

Por fim, em relação ao modelo completo de segunda ordem, a mudança esperada em  $Y$  quando  $x_2$  é mantido fixo e  $x_1$  aumenta 1 unidade é  $\beta_1 + \beta_3 + 2\beta_3x_1 + \beta_5x_2$ , que é uma função tanto de  $x_1$  quanto de  $x_2$ . Isso implica que os contornos da função de regressão são curvos e não paralelos entre si, como ilustrado na Figura 13.12(d).

Considerações semelhantes aplicam-se a modelos formados com base em mais de duas variáveis independentes. Em geral, a existência de termos de interação no modelo implica que a mudança esperada em  $Y$  depende não apenas da variável que está sendo aumentada ou diminuída, mas também dos valores de algumas das variáveis fixas. Como na ANOVA, é possível ter termos de interação de ordem superior (por exemplo,  $x_1x_2x_3$ ), o que dificulta a interpretação do modelo.

Observe que, se o modelo contiver previsores de interação ou quadráticos, a interpretação genérica de  $\beta_i$  dada anteriormente não se aplicará normalmente. Isso ocorre porque, desse modo, não é possível aumentar  $x_i$  em 1 unidade e manter fixos os valores de todos os outros previsores.

## Modelos com Previsores para Variáveis Categóricas

Até aqui, avaliamos claramente a inclusão apenas de variáveis de previsão quantitativa (numérica) em um modelo de regressão múltipla. Usando uma codificação numérica simples, variáveis qualitativas (categóricas) — como o material usado nos mancais (alumínio ou cobre/chumbo) ou o tipo de madeira (pinho, carvalho, nogueira) — também podem ser incorporadas no modelo. Vamos nos voltar primeiro para o caso de uma variável dicotômica, aquela com apenas duas categorias possíveis — macho ou fêmea, fabricação americana ou estrangeira, e assim por diante. Tendo qualquer dessas variáveis, associamos uma **variável dummy** ou **indicadora**  $x$  cujos valores possíveis 0 e 1 indicam que categoria é relevante para qualquer observação em particular.

### Exemplo 13.11

O artigo “Estimating Urban Travel Times: A Comparative Study” (*Trans. Res.*, 1980, p. 173–175) descreve um estudo que relaciona a variável dependente  $y$  = tempo de percurso entre locais em uma determinada cidade e a variável independente  $x_2$  = distância entre os locais. Foram usados no estudo dois tipos de veículo, automóveis e caminhões. Seja

$$x_1 = \begin{cases} 1 & \text{se o veículo for um caminhão} \\ 0 & \text{se o veículo for um automóvel} \end{cases}$$

Um possível modelo de regressão múltipla é

$$Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \epsilon$$

O valor médio do tempo de percurso depende de o veículo ser um carro ou um caminhão:

$$\begin{aligned} \text{tempo médio} &= \beta_0 + \beta_2x_2 && \text{quando } x_1 = 0 \quad (\text{carros}) \\ \text{tempo médio} &= \beta_0 + \beta_1 + \beta_2x_2 && \text{quando } x_1 = 1 \quad (\text{caminhões}) \end{aligned}$$

O coeficiente  $\beta_1$  é a diferença em tempos médios entre caminhões e carros, mantendo-se a distância fixa; se  $\beta_1 > 0$ , em média os caminhões levarão um tempo maior para atravessar qualquer distância em particular do que os carros.

Uma segunda possibilidade é um modelo com interação entre os previsores:

$$Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1x_2 + \epsilon$$

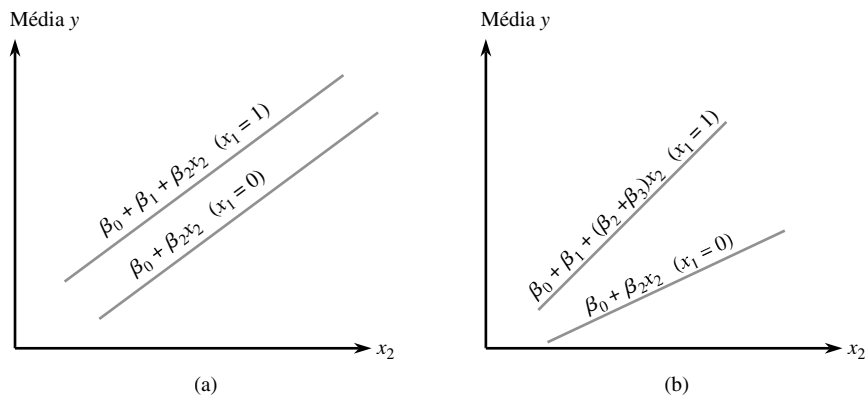
Agora, os tempos médios para os dois tipos de veículo são

$$\begin{aligned} \text{tempo médio} &= \beta_0 + \beta_2x_2 && \text{quando } x_1 = 0 \\ \text{tempo médio} &= \beta_0 + \beta_1 + (\beta_2 + \beta_3)x_2 && \text{quando } x_1 = 1 \end{aligned}$$

Para cada modelo, o gráfico do tempo médio versus distância é uma linha reta para ambos os tipos de veículo, como ilustrado na Figura 13.13. As duas retas são paralelas para o primeiro modelo (sem interação), mas em geral elas terão diferentes coeficientes angulares quando o segundo modelo estiver correto. Em relação



a esse último modelo, a mudança no tempo médio de percurso, com 1 milha de distância a mais, depende do tipo de veículo em questão — as duas variáveis “tipo de veículo” e “tempo de percurso” interagem. Na verdade, os dados coletados pelos autores do artigo citado sugerem que a interação existe. ■



**Figura 13.13** Funções de regressão para modelos com uma variável dummy ( $x_1$ ) e uma variável quantitativa  $x_2$ : (a) sem interação; (b) com interação

Você poderia pensar que a forma de lidar com uma situação de três categorias é definir uma única variável numérica com valores codificados tais como 0, 1 e 2 correspondendo às três categorias. Isso não está correto porque impõe uma ordenação nas categorias que não é necessariamente sugerida pelo contexto do problema. A abordagem correta para incorporar três categorias é definir duas diferentes variáveis dummy. Suponha, por exemplo, que  $y$  seja o tempo de vida de uma determinada ferramenta de corte,  $x_1$  seja a velocidade do corte e que três marcas de ferramenta estão sendo investigadas. Portanto, considere

$$x_2 = \begin{cases} 1 & \text{se for usada a ferramenta da marca A} \\ 0 & \text{caso contrário} \end{cases} \quad x_3 = \begin{cases} 1 & \text{se for usada a ferramenta da marca B} \\ 0 & \text{caso contrário} \end{cases}$$

Quando é feita uma observação sobre a ferramenta da marca A,  $x_2 = 1$  e  $x_3 = 0$ , ao passo que para uma ferramenta da marca B,  $x_2 = 0$  e  $x_3 = 1$ . Uma observação feita sobre a ferramenta da marca C tem  $x_2 = x_3 = 0$  e não é possível que  $x_2 = x_3 = 1$  porque uma ferramenta não pode ser simultaneamente da marca A e da marca B. O modelo sem interação teria apenas os previsores  $x_1$ ,  $x_2$  e  $x_3$ . O modelo de interação a seguir permite que a mudança média no tempo de vida, associada com 1 unidade de aumento na velocidade, dependa da marca da ferramenta:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_1 x_2 + \beta_5 x_1 x_3 + \epsilon$$

A construção de uma figura como a Figura 13.13 com um gráfico para cada um dos três pares possíveis ( $x_2$ ,  $x_3$ ) fornece três retas não paralelas (a não ser que  $\beta_4 = \beta_5 = 0$ ).

De maneira geral, incorporar uma variável categórica com  $c$  categorias possíveis em um modelo de regressão múltipla requer o uso de  $c - 1$  variáveis indicadoras (por exemplo, cinco marcas de ferramenta precisam usar quatro variáveis indicadoras). Logo, mesmo uma variável categórica pode adicionar vários previsores a um modelo.

## Estimando Parâmetros

Os dados em uma regressão linear simples compreendem  $n$  pares  $(x_1, y_1), \dots, (x_n, y_n)$ . Suponha que um modelo de regressão múltipla contenha duas variáveis de previsão,  $x_1$  e  $x_2$ . Então o conjunto de dados compreenderá  $n$  triplos  $(x_{11}, x_{21}, y_1), (x_{12}, x_{22}, y_2), \dots, (x_{1n}, x_{2n}, y_n)$ . Aqui, o primeiro subscripto em  $x$  refere-se ao predictor e o segundo ao número da observação. De maneira geral, com  $k$  previsores, os dados compreenderão  $n(k + 1)$ -tuplas  $(x_{11}, x_{21}, \dots, x_{k1}, y_1), (x_{12}, x_{22}, \dots, x_{k2}, y_2), \dots, (x_{1n}, x_{2n}, \dots, x_{kn}, y_n)$ , onde  $x_{ij}$  é o valor do  $i$ -ésimo predictor  $x_i$  associado com o valor observado  $y_j$ . Supõe-se que os  $y_j$ s foram observados independentemente, de acordo com

o modelo (13.15). Para estimar os parâmetros  $\beta_0, \beta_1, \dots, \beta_k$  usando o princípio dos mínimos quadrados, calcule a soma dos desvios quadrados dos  $y_j$ s observados com base em uma função teste  $y = b_0 + b_1x_1 + \dots + b_kx_k$ :

$$f(b_0, b_1, \dots, b_k) = \sum_j [y_j - (b_0 + b_1x_{1j} + b_2x_{2j} + \dots + b_kx_{kj})]^2 \quad (13.17)$$

As estimativas dos mínimos quadrados são aqueles valores dos *bis* que minimizam  $f(b_0, \dots, b_k)$ . Tomando a derivada parcial de  $f$  em relação a cada  $b_i$  ( $i = 0, 1, \dots, k$ ) e igualando todas as parciais a zero, obtemos o seguinte esquema de **equações normais**:

$$\begin{aligned} b_0n + b_1\sum x_{1j} + b_2\sum x_{2j} + \dots + b_k\sum x_{kj} &= \sum y_j \\ b_0\sum x_{1j} + b_1\sum x_{1j}^2 + b_2\sum x_{1j}x_{2j} + \dots + b_k\sum x_{1j}x_{kj} &= \sum x_{1j}y_j \\ \cdot & \cdot \cdot \\ \cdot & \cdot \cdot \\ \cdot & \cdot \cdot \\ b_0\sum x_{kj} + b_1\sum x_{1j}x_{kj} + \dots + b_{k-1}\sum x_{k-1,j}x_{kj} + b_k\sum x_{kj}^2 &= \sum x_{kj}y_j \end{aligned} \quad (13.18)$$

Essas equações são lineares nas incógnitas  $b_0, b_1, \dots, b_k$ . Resolvendo (13.18), obtemos as estimativas dos mínimos quadrados  $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$ . A solução do sistema de equações é mais adequada ao utilizar um software estatístico.

### Exemplo 13.12

O artigo “How to Optimize and Control the Wire Bonding Process: Part II” (*Solid State Technology*, janeiro de 1991, p. 67–72) descreve um experimento realizado para avaliar o impacto das variáveis  $x_1 =$  força (gm),  $x_2 =$  potência (mW),  $x_3 =$  temperatura ( $^{\circ}$ C) e  $x_4 =$  tempo (ms) sobre  $y =$  resistência de cisalhamento (corte) de ligação esférica (gm). Os dados<sup>3</sup> a seguir foram gerados para que sejam consistentes com as informações fornecidas no artigo:

Observação	Força	Potência	Temperatura	Tempo	Resistência
1	30	60	175	15	26,2
2	40	60	175	15	26,3
3	30	90	175	15	39,8
4	40	90	175	15	39,7
5	30	60	225	15	38,6
6	40	60	225	15	35,5
7	30	90	225	15	48,8
8	40	90	225	15	37,8
9	30	60	175	25	26,6
10	40	60	175	25	23,4
11	30	90	175	25	38,6
12	40	90	175	25	52,1
13	30	60	225	25	39,5
14	40	60	225	25	32,3
15	30	90	225	25	43,0
16	40	90	225	25	56,0
17	25	75	200	20	35,2
18	45	75	200	20	46,9
19	35	45	200	20	22,7

(continuação)

<sup>3</sup> Do livro *Statistics Engineering Problem Solving*, de Stephen Vardeman, uma excelente exposição da área coberta por nosso livro, embora em um nível mais elevado.

Observação	Força	Potência	Temperatura	Tempo	Resistência
20	35	105	200	20	58,7
21	35	75	150	20	34,5
22	35	75	250	20	44,0
23	35	75	200	10	35,7
24	35	75	200	30	41,8
25	35	75	200	20	36,5
26	35	75	200	20	37,6
27	35	75	200	20	40,3
28	35	75	200	20	46,0
29	35	75	200	20	27,8
30	35	75	200	20	40,3

Um software estatístico forneceu as seguintes estimativas de mínimos quadrados:

$$\hat{\beta}_0 = -37,48 \quad \hat{\beta}_1 = 0,2117 \quad \hat{\beta}_2 = 0,4983 \quad \hat{\beta}_3 = 0,1297 \quad \hat{\beta}_4 = 0,2583$$

Portanto, estimamos que 0,1297 gm seja a mudança média na resistência associada com um aumento de 1 grau na temperatura quando os outros três previsores são mantidos fixos; os outros coeficientes estimados são interpretados de maneira similar.

A equação da regressão estimada é

$$y = -37,48 + 0,2117x_1 + 0,4983x_2 + 0,1297x_3 + 0,2583x_4$$

Uma previsão pontual da resistência resultante de uma força de 35 gm, potência de 75 mW, temperatura de 200° e tempo de 20 ms é

$$\begin{aligned} \hat{y} &= -37,48 + (0,2117)(35) + (0,4983)(75) + (0,1297)(200) + (0,2583)(20) \\ &= 38,41 \text{ gm} \end{aligned}$$

Essa é também uma estimativa pontual do valor médio da resistência dos valores de força, potência, temperatura e tempo especificados. ■

## $\hat{\sigma}^2$ e $R^2$

Se substituirmos os valores dos previsores das observações sucessivas na equação da regressão estimada, obtaremos os **valores previstos** ou **ajustados**  $\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n$ . Por exemplo, visto que os valores dos quatro previsores da última observação no Exemplo 13.12 são 35, 75, 200 e 20, respectivamente, o valor previsto correspondente é  $\hat{y}_{30} = 38,41$ . Os **resíduos** são as diferenças  $y_1 - \hat{y}_1, \dots, y_n - \hat{y}_n$ . O último resíduo no Exemplo 13.12 é  $40,3 - 38,41 = 1,89$ . Quanto mais próximos os resíduos estiverem de 0, mais eficazmente a equação estimada preverá os valores de  $y$  que correspondem aos valores dos previsores em nossa amostra.

Da mesma maneira que na regressão linear simples e na regressão polinomial, a estimação de  $\sigma^2$  baseia-se na soma dos resíduos quadrados:

$$SQE = \sum (y_j - \hat{y}_j)^2 = \sum [y_j - (\hat{\beta}_0 + \hat{\beta}_1 x_{1j} + \dots + \hat{\beta}_k x_{kj})]^2$$

Uma fórmula eficiente para calcular SQE é empregada pela maioria dos softwares estatísticos. Pelo fato de  $k + 1$  parâmetros ( $\beta_0, \beta_1, \dots, \beta_k$ ) terem sido estimados,  $k + 1$  gl são perdidos, de modo que  $n - (k + 1)$  gl está associado a SQE, e

$$\hat{\sigma}^2 = s^2 = \frac{SQE}{n - (k + 1)} = QME$$

Sendo  $SQT = \sum (y_i - \bar{y})^2$ , a proporção da variação total explicada pelo modelo de regressão múltipla é  $R^2 = 1 - SSE/SQT$ , o coeficiente de determinação múltipla. Como na regressão polinomial,  $R^2$  é frequentemente ajustado ao número de parâmetros no modelo pela fórmula

$$R_a^2 = [(n - 1)R^2 - k]/[n - (k + 1)]$$

A raiz quadrada positiva do coeficiente de determinação múltipla denomina-se *coeficiente de correlação múltipla*  $R$ . É possível mostrar que  $R$  é o coeficiente de correlação amostral  $r$  entre os  $y_j$ s observados e os  $\hat{y}_j$ s previstos (isto é, usando  $x_j = \hat{y}_j$  na fórmula de  $r$  obtém-se  $r = R$ ).

### Exemplo 13.13

Determinados pesquisadores realizaram um estudo para verificar como várias características do concreto são influenciadas por  $x_1 = \%$  pó de calcário e  $x_2 =$  razão água-cimento, gerando os dados a seguir (“Durability of Concrete with Addition of Limestone Powder”, *Magazine of Concrete Research*, 1996, p. 131–137).

$x_1$	$x_2$	$x_1x_2$	Resist. à Comp. em 28 dias (MPa)	Capacidade de Adsorção (%)
21	0,65	13,65	33,55	8,42
21	0,55	11,55	47,55	6,26
7	0,65	4,55	35,00	6,74
7	0,55	3,85	35,90	6,59
28	0,60	16,80	40,90	7,28
0	0,60	0,00	39,10	6,90
14	0,70	9,80	31,55	10,80
14	0,50	7,00	48,00	5,63
14	0,60	8,40	42,30	7,43
			$\bar{y} = 39,317$ , $SQT = 278,52$	$\bar{y} = 7,339$ , $SQT = 18,356$

Considere primeiro a resistência à compressão como a variável dependente  $y$ . Ajustando o modelo de primeira ordem, obtemos

$$y = 84,82 + 0,1643x_1 - 79,67x_2 \quad SSE = 72,25 \text{ (gl} = 6) \quad R^2 = 0,741 \quad R_a^2 = 0,654$$

ao passo que incluindo um previsor de interação, obtemos

$$y = 6,22 + 5,779x_1 + 51,33x_2 - 9,357x_1x_2$$

$$SSE = 29,35 \text{ (gl} = 5) \quad R^2 = 0,895 \quad R_a^2 = 0,831$$

Com base nessa última calibração, a previsão da resistência à compressão, quando a % de calcário = 14 e a razão água-cimento = 0,60, é

$$\hat{y} = 6,22 + 5,779(14) + 51,33(0,60) - 9,357(8,4) = 39,32$$

A calibração da relação quadrática completa não produz praticamente nenhuma mudança no valor de  $R^2$ . Entretanto, quando a variável dependente é a capacidade de adsorção, são obtidos os resultados a seguir:  $R^2 = 0,747$  quando são usados apenas dois previsores, 0,802 quando se adiciona o previsor da interação e 0,889 quando são usados os cinco previsores para a relação quadrática completa. ■

Em geral,  $\hat{\beta}_1$  pode ser interpretado como uma estimativa da mudança média em  $y$  associada com um aumento de 1 unidade em  $x_1$ , ao passo que os valores de todos os outros previsores mantêm-se fixos. Entretanto, às vezes é difícil ou mesmo impossível aumentar o valor de um previsor e, ao mesmo tempo, manter todos os outros fixos. Em tais situações, há uma interpretação alternativa dos coeficientes da regressão estimada. Para dar um exemplo concreto, suponha que  $k = 2$  e que  $\hat{\beta}_1$  seja a estimativa de  $\beta_1$  na regressão de  $y$  aos dois previsores,  $x_1$  e  $x_2$ . Então

1. Regrida  $y$  apenas em relação a  $x_2$  (uma regressão linear simples) e represente os resíduos resultantes por  $g_1, g_2, \dots, g_n$ . Esses resíduos representam a variação em  $y$  após a remoção ou o ajuste dos efeitos de  $x_2$ .
2. Regrida  $x_1$  em relação a  $x_2$  (ou seja, considere  $x_1$  a variável dependente e  $x_2$  a variável independente nessa regressão linear simples) e represente os resíduos por  $f_1, \dots, f_n$ . Esses resíduos representam a variação em  $x_1$  após a remoção ou o ajuste dos efeitos de  $x_2$ .

Agora experimente representar graficamente os resíduos da primeira regressão em relação aos da segunda, ou seja, represente graficamente os pares  $(f_1, g_1), \dots, (f_n, g_n)$ . O resultado é denominado *gráfico dos resíduos parciais* ou *gráfico dos resíduos ajustados*. Se uma reta de regressão for ajustada aos pontos nesse gráfico, o coeficiente angular acabará sendo exatamente  $\hat{\beta}_1$  (além disso, os resíduos dessa reta são exatamente os resíduos  $e_1, \dots, e_n$  da regressão múltipla de  $y$  sobre  $x_1$  e  $x_2$ ). Portanto,  $\hat{\beta}_1$  pode ser interpretado como as mudanças estimadas em  $y$  associadas com um aumento de 1 unidade em  $x_1$  após a remoção ou ajuste dos efeitos de qualquer outro preditor do modelo. A mesma interpretação é válida para outros coeficientes estimados, independentemente do número de preditores no modelo (não há nada especial em  $k = 2$ ; o argumento precedente permanece válido se  $y$  for regredido em relação a todos os preditores diferentes de  $x_1$  na Etapa 1 e  $x_1$  for regredido em relação a outros  $k - 1$  preditores na Etapa 2).

Como exemplo, suponha que  $y$  seja o preço de venda de um prédio de apartamentos e que os preditores sejam o número de apartamentos, o tempo de vida, a área do terreno, o número de vagas na garagem e a área bruta construída (ft<sup>2</sup>). Talvez não seja sensato aumentar o número de apartamentos sem aumentar também a área bruta. Entretanto, se  $\hat{\beta}_5 = 16,00$ , podemos dizer que um aumento de \$16 no preço de venda está associado com cada pé quadrado extra da área bruta após o ajuste dos efeitos dos outros quatro preditores.

## Um Teste de Utilidade do Modelo

Quando os dados são multivariados, não existe nenhuma figura introdutória análoga a um gráfico de dispersão para indicar se um determinado modelo de regressão múltipla será considerado útil. O valor de  $R^2$  certamente transmite uma mensagem preliminar, mas esse valor é às vezes ilusório porque ele pode ser aumentado pelo uso de um número grande de preditores ( $k$  grande) em relação ao tamanho da amostra  $n$  (esse é o raciocínio lógico que justifica o ajuste de  $R^2$ ).

O teste de utilidade do modelo na regressão linear simples compreendia a hipótese nula  $H_0: \beta_1 = 0$ , segundo a qual não existe nenhuma relação útil entre  $y$  e o preditor único  $x$ . Aqui, consideramos a afirmação de que  $\beta_1 = 0, \beta_2 = 0, \dots, \beta_k = 0$ , o que expressa que não há nenhuma relação útil entre  $y$  e *qualquer* um dos  $k$  preditores. Se no mínimo um desses  $\beta$ s não for 0, o(s) preditor(es) correspondente(s) será(ão) útil(is). O teste baseia-se em uma estatística que tem uma determinada distribuição  $F$  quando  $H_0$  é verdadeira.

Hipótese nula:  $H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$

Hipótese alternativa:  $H_a$ : pelo menos um  $\beta_i \neq 0 \quad (i = 1, \dots, k)$

O valor da estatística de teste\*:  $f = \frac{R^2/k}{(1 - R^2)/[n - (k + 1)]}$  (13.19)

$$= \frac{SSR/k}{SSE/[n - (k + 1)]} = \frac{MSR}{MSE}$$

onde  $SQR =$  soma dos quadrados da regressão  $= SQT - SQE$

Região de rejeição para um teste de nível  $\alpha$ :  $f \geq F_{\alpha, k, n - (k + 1)}$

Exceto por um múltiplo constante, a estatística de teste, neste caso, é  $R^2/(1 - R^2)$ , a razão entre a variação explicada e não explicada. Se a proporção da variação explicada for alta em relação à inexplicada, tenderemos naturalmente a rejeitar  $H_0$  e a confirmar a utilidade do modelo. Contudo, se  $k$  for grande em relação a  $n$ , o fator  $[(n - (k + 1))/k]$  diminuirá  $f$  consideravelmente.

**Exemplo 13.14**

Retornando aos dados sobre a resistência de cisalhamento de ligação do Exemplo 13.12, um modelo com  $k = 4$  previsores foi calibrado de modo que as hipóteses relevantes são

$$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$$

$H_a$ : no mínimo um desses quatro  $\beta$ s não é 0

A Figura 13.14 mostra o resultado do software estatístico JMP. Os valores de  $s$  (Raiz do Erro Médio Quadrático),  $R^2$ , e  $R^2$  ajustado certamente sugerem um modelo útil. O valor da razão  $F$  de utilidade do modelo é

$$f = \frac{R^2/k}{(1 - R^2)/[n - (k + 1)]} = \frac{0,713959/4}{0,286041/(30 - 5)} = 15,60$$

Esse valor também aparece na coluna Razão F da tabela ANOVA da Figura 13.14. O maior valor crítico  $F$  com 4 gl no numerador e 25 gl no denominador na Tabela A.9 do Apêndice é 6,49, que inclui uma área de cauda superior de 0,001. Portanto, o valor  $P$  é  $< 0,001$ . A tabela ANOVA no resultado do JMP mostra que o valor  $P$  é  $< 0,0001$ . Esse resultado é altamente significativo. A hipótese nula deve ser rejeitada em qualquer nível de significância razoável. Concluimos que existe uma relação linear útil entre  $y$  e pelo menos um dos quatro previsores no modelo. Isso não significa que todos os quatro previsores são úteis; falaremos mais sobre isso subseqüentemente.

Resposta: resistência

Resumo da Calibração ou Aderência	
$R^2$	0,713959
$R^2$ Ajustado	0,668193
Raiz do Erro Médio Quadrático	5,157979
Resposta Média	38,40667
Observações (ou Soma dos Pesos)	30

Estimativas dos Parâmetros				
Termo	Estimativa	Erro-padrão	Razão t	Prob. >  t
constante	-37,47667	13,09964	-2,86	0,0084
força	0,2116667	0,210574	1,01	0,3244
potência	0,4983333	0,070191	7,10	<0,0001
temp.	0,1296667	0,042115	3,08	0,0050
tempo	0,2583333	0,210574	1,23	0,2313

Teste Integral do Modelo				
Análise de Variância				
Fonte	GL	Soma dos Quadrados	Quadrado Médio	Razão F
Modelo	4	1.660,1400	415,035	15,6000
Erro	25	665,1187	26,605	Prob.>F
Total C	29	2325,2587		<0,0001

**Figura 13.14** Resultado da regressão múltipla do JMP para os dados do Exemplo 13.14

## Inferência em Regressão Múltipla

Antes de testar as hipóteses, criar ICs e fazer previsões, deve-se primeiro examinar os gráficos de diagnóstico para verificar se o modelo precisa ser modificado ou se há *outliers* nos dados. Os gráficos sugeridos são os de resíduos (padronizados) *versus* cada variável independente, resíduos *versus*  $\hat{y}$ ,  $\hat{y}$  e de probabilidade normal dos resíduos padronizados. Os problemas possíveis são indicados pelos mesmos padrões discutidos na Seção 13.1. De particular importância é a identificação das observações cuja influência sobre a calibração é grande. Na seção seguinte, descrevemos várias ferramentas de diagnóstico adequadas para essa tarefa.

Pelo fato de cada  $\hat{\beta}_i$  ser uma função linear dos  $y_i$ s, o desvio padrão de cada  $\hat{\beta}_i$  é o produto de  $\sigma$  e uma função dos  $x_{ij}$ s, de modo que uma estimativa de  $s_{\hat{\beta}_i}$  é obtida substituindo-se  $s$  por  $\sigma$ . A função dos  $x_{ij}$ s é muito complexa, mas todos os softwares de regressão-padrão calculam e mostram os  $s_{\hat{\beta}_i}$ s. As inferências sobre um único  $\beta_i$  baseiam-se na variável padronizada

$$T = \frac{\hat{\beta}_i - \beta_i}{S_{\hat{\beta}_i}}$$

que tem uma distribuição  $t$  com  $n - (k + 1)$  gl.

A estimativa pontual de  $\mu_{Y, x_1^*, \dots, x_k^*}$ , o valor esperado de  $Y$  quando  $x_1 = x_1^*$ , ...,  $x_k = x_k^*$ , é  $\hat{\mu}_{Y, x_1^*, \dots, x_k^*} = \hat{\beta}_0 + \hat{\beta}_1 x_1^* + \dots + \hat{\beta}_k x_k^*$ . O desvio padrão estimado do estimador correspondente é, novamente, uma expressão complexa que compreende os  $x_{ij}$ s da amostra. Contudo, os melhores softwares estatísticos farão esse cálculo quando solicitados. As inferências sobre  $\mu_{Y, x_1^*, \dots, x_k^*}$  têm por base padronizar os respectivos estimadores para obter uma variável  $t$  com  $n - (k + 1)$  gl.

1. Um IC de  $100(1 - \alpha)\%$  para  $\beta_i$ , o coeficiente de  $x_i$  na função de regressão, é

$$\hat{\beta}_i \pm t_{\alpha/2, n-(k+1)} \cdot s_{\hat{\beta}_i}$$

2. Um teste de  $H_0: \beta_i = \beta_{i0}$  usa o valor  $t$  estatístico  $t = (\hat{\beta}_i - \beta_{i0})/s_{\hat{\beta}_i}$  com base em  $n - (k + 1)$  gl. O teste é de cauda superior, de cauda inferior ou bicaudal, dependendo da desigualdade de  $H_a$ , isto é, se  $>$ ,  $<$  ou  $\neq$ .
3. Um IC de  $100(1 - \alpha)\%$  para  $\mu_{Y, x_1^*, \dots, x_k^*}$  é

$$\hat{\mu}_{Y, x_1^*, \dots, x_k^*} \pm t_{\alpha/2, n-(k+1)} \cdot \{\text{DP estimado de } \hat{\mu}_{Y, x_1^*, \dots, x_k^*}\} = \hat{y} \pm t_{\alpha/2, n-(k+1)} \cdot s_{\hat{Y}}$$

onde  $\hat{Y}$  é a estatística  $\hat{\beta}_0 + \hat{\beta}_1 x_1^* + \dots + \hat{\beta}_k x_k^*$  e  $\hat{y}$  é o valor calculado de  $\hat{Y}$ .

4. Um IP de  $100(1 - \alpha)\%$  para um valor  $y$  futuro é

$$\hat{\mu}_{Y, x_1^*, \dots, x_k^*} \pm t_{\alpha/2, n-(k+1)} \cdot \{s^2 + (\text{DP estimado de } \hat{\mu}_{Y, x_1^*, \dots, x_k^*})^2\}^{1/2} = \hat{y} \pm t_{\alpha/2, n-(k+1)} \cdot \sqrt{s^2 + s_{\hat{Y}}^2}$$

Os intervalos simultâneos para os quais o nível de confiança ou previsão simultâneo é controlado podem ser obtidos aplicando-se a técnica de Bonferroni.

### Exemplo 13.15

A adsorção do solo e de sedimentos, quantidade que as substâncias químicas acumulam de forma condensada sobre a superfície, é uma importante característica que influencia na eficácia dos pesticidas e de várias substâncias agroquímicas. O artigo “Adsorption of Phosphate, Arsenate, Methanearsonate, and Cacodylate by Lake and Stream Sediments: Comparisons with Soils” (*J. of Environ. Qual.*, 1984, p. 499–504) apresenta os dados a seguir (Tabela 13.5) sobre  $y$  = índice de adsorção de fosfato,  $x_1$  = quantidade de ferro extraível e  $x_2$  = quantidade de alumínio extraível.

Tabela 13.5 Dados do Exemplo 13.5

Observação	$x_1 =$ Ferro Extraível	$x_2 =$ Alumínio Extraível	$y =$ Índice de Adsorção
1	61	13	4
2	175	21	18
3	111	24	14
4	124	23	18
5	130	64	26
6	173	38	26
7	169	33	21
8	169	61	30
9	160	39	28
10	244	71	36
11	257	112	65
12	333	88	62
13	199	54	40

O artigo propôs o modelo

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

Uma análise por computador gerou as seguintes informações:

Parâmetro $\beta_i$	Estimativa $\hat{\beta}_i$	DP Estimado $s_{\hat{\beta}_i}$
$\beta_0$	-7,351	3,485
$\beta_1$	0,11273	0,02969
$\beta_2$	0,34900	0,07131
$R^2 = 0,948$	$R^2$ ajustado = 0,938	$s = 4,379$
$\hat{\mu}_{Y-160,39} = \hat{y} = -7,351 + (0,11273)(160) + (0,34900)(39) = 24,30$		
DP estimado de $\hat{\mu}_{Y-160,39} = s_{\hat{y}} = 1,30$		

Um IC de 99% para  $\beta_1$ , a mudança na adsorção esperada associada a um aumento de 1 unidade no ferro extraível enquanto o alumínio mantém-se fixo, exige  $t_{0,005,13-(2+1)} = t_{0,005,10} = 3,169$ . O IC é

$$0,11273 \pm (3,169)(0,02969) = 0,11273 \pm 0,09409 \approx (0,019, 0,207)$$

De maneira semelhante, um intervalo de 99% para  $\beta_2$  é

$$0,34900 \pm (3,169)(0,07131) = 0,34900 \pm 0,22598 \approx (0,123, 0,575)$$

A técnica de Bonferroni implica que o nível de confiança simultâneo de ambos os intervalos é de 98%, no mínimo.

Um IC de 95% para  $\mu_{Y-160,39}$ , adsorção esperada quando o ferro extraível = 160 e o alumínio extraível = 39, é

$$24,30 \pm (2,228)(1,30) = 24,30 \pm 2,90 = (21,40; 27,20)$$

Um IP de 95% para um valor de adsorção futuro, a ser observado quando  $x_1 = 160$  e  $x_2 = 39$  é

$$24,30 \pm (2,228)\{(4,379)^2 + (1,30)^2\}^{1/2} = 24,30 \pm 10,18 = (14,12; 34,48)$$





Em geral, a hipótese de interesse tem a forma  $H_0: \beta_i = 0$  para um determinado  $i$ . Por exemplo, depois de ajustar o modelo de quatro previsores no Exemplo 13.12, o pesquisador pode desejar testar  $H_0: \beta_4 = 0$ . De acordo com  $H_0$ , na medida em que os previsores  $x_1, x_2$  e  $x_3$  permaneçam no modelo,  $x_4$  não contém nenhuma informação útil sobre  $y$ . O valor da estatística de teste é a **razão  $t$**   $\hat{\beta}_i/s_{\hat{\beta}_i}$ . Vários softwares estatísticos relatam a razão  $t$  e o valor  $P$  correspondente para cada predictor incluído no modelo. Por exemplo, a Figura 13.14 mostra que, na medida em que a potência, a temperatura e o tempo forem mantidos no modelo, o predictor  $x_1 =$  força pode ser excluído.

**Um Teste  $F$  para um Grupo de Previsores** O teste  $F$  de utilidade do modelo foi apropriado para testar se existem informações úteis sobre a variável dependente em qualquer um dos  $k$  previsores (isto é, se  $\beta_1 = \dots = \beta_k = 0$ ). Em várias situações, primeiramente se elabora um modelo contendo  $k$  previsores para, depois, querer saber se algum dos previsores em um determinado subconjunto fornece informações úteis sobre  $Y$ . Por exemplo, um modelo a ser usado para prever as notas de prova dos alunos deve incluir um grupo de variáveis como renda familiar e nível educacional e também outro grupo de variáveis com algumas características da escola, como tamanho da classe e gastos por aluno. Uma hipótese interessante é a de que os previsores das características da escola podem ser extraídos do modelo.

Denominemos os previsores de  $x_1, x_2, \dots, x_l, x_{l+1}, \dots, x_k$ , de modo que os previsores que estamos pensando em excluir do modelo sejam os últimos  $k - l$ . Então, queremos testar

$$H_0: \beta_{l+1} = \beta_{l+2} = \dots = \beta_k = 0$$

(de modo que o modelo “reduzido”  $Y = \beta_0 + \beta_1 x_1 + \dots + \beta_l x_l + \epsilon$  seja correto)

versus

$$H_a: \text{pelo menos um entre } \beta_{l+1}, \dots, \beta_k \text{ não é } 0$$

(portanto, no modelo “completo”  $Y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \epsilon$ , pelo menos um dos  $k - l$  previsores fornece informações úteis)

O teste é realizado calibrando-se tanto o modelo completo quanto o reduzido. Pelo fato de o modelo completo conter não apenas os previsores do modelo reduzido, mas também alguns outros previsores, é seguro que a sua calibração será pelo menos tão boa quanto a do modelo reduzido. Ou seja, se considerarmos  $SQE_k$  a soma dos resíduos quadrados do modelo completo e  $SQE_l$  a soma correspondente do modelo reduzido, então  $SQE_k \leq SQE_l$ .<sup>4</sup> Intuitivamente, se  $SQE_k$  for bem menor que  $SQE_l$ , o modelo completo fornecerá uma calibração bem melhor do que a do modelo reduzido; portanto, a estatística de teste apropriada deve então depender da redução  $SQE_l - SQE_k$  na variação inexplicada. O procedimento formal é

$SQE_k$  = variação inexplicada para o modelo completo

$SQE_l$  = variação inexplicada para o modelo reduzido

$$\text{Valor da estatística de teste: } f = \frac{(SQE_l - SQE_k)/(k - l)}{SQE_k/[n - (k + 1)]} \quad (13.20)$$

Região de rejeição:  $f \geq F_{\alpha, k-l, n-(k+1)}$

<sup>4</sup>As estimativas  $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$  em geral serão diferentes para os modelos completo e reduzido. Portanto, normalmente duas regressões múltiplas diferentes devem ser executadas para obter  $SQE_l$  e  $SQE_k$ . Contudo, se as variáveis forem relacionadas na ordem sugerida, a maioria dos softwares de computador fornece uma tabela ANOVA da “soma de quadrados sequencial” para o modelo completo que pode ser usada para evitar a necessidade de calibrar o modelo reduzido.

**Exemplo 13.16**

Os dados na Tabela 13.6 foram extraídos do artigo “Applying Stepwise Multiple Regression Analysis to the Reaction of Formaldehyde with Cotton Cellulose” (*Textile Research J.*, 1984, p. 157–165). A variável dependente  $y$  é a classificação de impressão firme, uma medida quantitativa da resistência ao enrugamento. As quatro variáveis independentes usadas no processo de elaboração do modelo são  $x_1$  = concentração de HCHO (formaldeído),  $x_2$  = razão do catalisador,  $x_3$  = temperatura de tratamento e  $x_4$  = tempo de tratamento.

**Tabela 13.6** Dados do Exemplo 13.16

Observação	$x_1$	$x_2$	$x_3$	$x_4$	$y$	Observação	$x_1$	$x_2$	$x_3$	$x_4$	$y$
1	8	4	100	1	1,4	16	4	10	160	5	4,6
2	2	4	180	7	2,2	17	4	13	100	7	4,3
3	7	4	180	1	4,6	18	10	10	120	7	4,9
4	10	7	120	5	4,9	19	5	4	100	1	1,7
5	7	4	180	5	4,6	20	8	13	140	1	4,6
6	7	7	180	1	4,7	21	10	1	180	1	2,6
7	7	13	140	1	4,6	22	2	13	140	1	3,1
8	5	4	160	7	4,5	23	6	13	180	7	4,7
9	4	7	140	3	4,8	24	7	1	120	7	2,5
10	5	1	100	7	1,4	25	5	13	140	1	4,5
11	8	10	140	3	4,7	26	8	1	160	7	2,1
12	2	4	100	3	1,6	27	4	1	180	7	1,8
13	4	10	180	3	4,5	28	6	1	160	1	1,5
14	6	7	120	7	4,7	29	4	1	100	1	1,3
15	10	13	180	3	4,8	30	7	10	100	7	4,6

Considere o modelo completo compreendendo  $k = 14$  preditores:  $x_1, x_2, x_3, x_4, x_5 = x_1^2, \dots, x_8 = x_4^2, x_9 = x_1x_2, \dots, x_{14} = x_3x_4$  (todos preditores de primeira e segunda ordens). A inclusão dos preditores de segunda ordem justifica-se? Ou seja, o modelo reduzido que compreende apenas os preditores  $x_1, x_2, x_3$  e  $x_4$  ( $l = 4$ ) deve ser usado? Segue o resultado da calibração dos dois modelos:

Parâmetro	Estimativa para o Modelo Reduzido	Estimativa para o Modelo Completo
$\beta_0$	-0,9122	-8,807
$\beta_1$	0,16073	0,1768
$\beta_2$	0,21978	0,7580
$\beta_3$	0,011226	0,10400
$\beta_4$	0,10197	0,5052
$\beta_5$	—	-0,04393
$\beta_6$	—	-0,035887
$\beta_7$	—	-0,00003271
$\beta_8$	—	-0,01646
$\beta_9$	—	0,00588
$\beta_{10}$	—	0,002702
$\beta_{11}$	—	0,01178
$\beta_{12}$	—	-0,0006547
$\beta_{13}$	—	0,00242
$\beta_{14}$	—	0,002526
$R^2$	0,692	0,921
SQE	17,4951	4,4782

As hipóteses a serem testadas são

$$H_0: \beta_5 = \beta_6 = \dots = \beta_{14} = 0$$

versus

$$H_a: \text{pelo menos um entre } \beta_5, \dots, \beta_{14} \text{ não é } 0$$

Com  $k = 14$  e  $l = 4$ , o valor crítico  $F$  para um teste com  $\alpha = 0,01$  é  $F_{0,01, 10, 15} = 3,80$ .

$$f = \frac{(17,4951 - 4,4782)/10}{4,4782/15} = \frac{1,3017}{0,2985} = 4,36$$

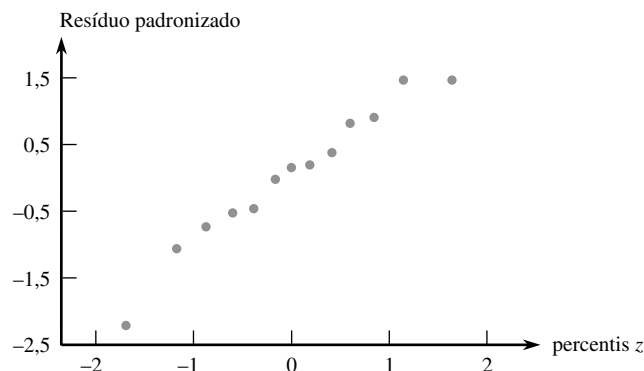
Visto que  $4,36 \geq 3,80$ ,  $H_0$  é rejeitada, concluímos que o modelo apropriado deve incluir no mínimo um dos pre-  
visores de segunda ordem. ■

### Avaliando a Adequação do Modelo

Os resíduos padronizados na regressão múltipla resultam da divisão de cada resíduo pelo seu desvio padrão estimado; a fórmula para esses desvios padrão é significativamente mais complexa do que no caso da regressão linear simples. Recomendamos um gráfico de probabilidade normal dos resíduos padronizados como base para validar a hipótese de normalidade. Os gráficos dos resíduos padronizados versus cada previsor e versus  $\hat{y}$  não devem mostrar qualquer padrão discernível. Os gráficos dos resíduos ajustados podem também ser úteis nessa tarefa. O livro de Neter et al. é uma referência extremamente útil.

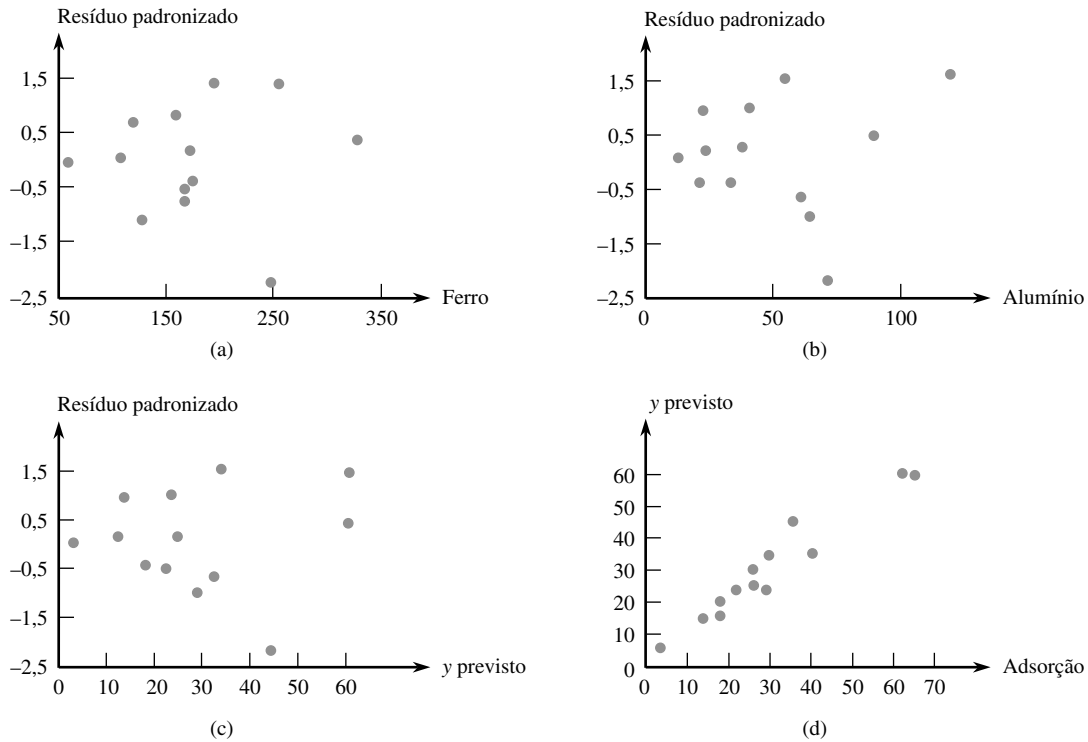
#### Exemplo 13.17

A Figura 13.15 mostra um gráfico de probabilidade normal dos resíduos padronizados dos dados de adsorção e do modelo ajustado do Exemplo 13.15. A retidão do gráfico não deixa praticamente nenhuma dúvida sobre a suposição de que o desvio aleatório  $\epsilon$  seja normalmente distribuído.



**Figura 13.15** Gráfico de probabilidade normal dos resíduos padronizados dos dados e do modelo do Exemplo 13.15

A Figura 13.16 mostra os outros gráficos sugeridos para os dados de adsorção. Visto que há apenas 13 observações no conjunto de dados, não existe muita evidência de padrão em nenhum dos três primeiros gráficos, a não ser a de aleatoriedade. O ponto na parte inferior de cada um desses três gráficos corresponde à observação cujo resíduo é grande. Falaremos mais sobre essas observações subseqüentemente. No momento, não há nenhuma razão convincente para uma ação corretiva.



**Figura 13.16** Gráficos de diagnóstico dos dados de adsorção: (a) resíduo padronizado versus  $x_1$ ; (b) resíduo padronizado versus  $x_2$ ; (c) resíduo padronizado versus  $\hat{y}$ ; (d)  $\hat{y}$  versus  $y$

**Exercícios** | Seção 13.4 (36–54)

**36.** A capacidade cardiorrespiratória é amplamente reconhecida como o principal componente do bem-estar físico global. A medição direta da entrada máxima de oxigênio ( $VO_2$  máx.) é a única medida adequada dessa capacidade, mas é trabalhosa e onerosa. Portanto, é desejável ter uma equação de previsão para  $VO_2$  máx. em função das quantidades facilmente obtidas. Considere as variáveis

- $y = VO_2$  máx. (L/min)
- $x_1 =$  peso (kg)
- $x_2 =$  idade (anos)
- $x_3 =$  tempo necessário para caminhar 1 milha (min.)
- $x_4 =$  taxa de batimentos cardíacos no fim da caminhada (batimentos/min.)

Eis aqui um modelo possível, para os alunos do sexo masculino, consistente com as informações fornecidas no artigo “Validation of the Rockport Fitness Walking Test in College Males and Females” (*Research Quarterly for Exercise and Sport*, 1994, p. 152–158):

$$Y = 5,0 + 0,01x_1 - 0,05x_2 - 0,13x_3 - 0,01x_4 + \epsilon$$

$\sigma = 0,4$

- a.** Interprete  $\beta_1$  e  $\beta_3$ .
- b.** Qual é o valor esperado de  $VO_2$  máx. com um peso de 76 kg, idade de 20 anos, tempo de caminhada de 12 min. e taxa de batimentos cardíacos de 140 b/m?

**c.** Qual é a probabilidade de  $VO_2$  máx. ficar entre 1,00 e 2,60 para uma única observação feita quando os valores dos previsores são tal como expresso na parte (b)?

**37.** Uma empresa de transporte de mercadorias por caminhão avaliou um modelo de regressão múltipla para relacionar a variável dependente  $y =$  tempo de percurso diário total de um de seus motoristas (horas) com os previsores  $x_1 =$  distância percorrida (milhas) e  $x_2 =$  número de entregas feitas. Suponha que a equação do modelo seja

$$Y = -0,800 + 0,060x_1 + 0,900x_2 + \epsilon$$

- a.** Qual é o valor médio do tempo de percurso quando uma distância de 50 milhas é percorrida e três entregas são feitas?
- b.** Como você interpretaria  $\beta_1 = 0,060$ , o coeficiente do previsor  $x_1$ ? Qual é a interpretação de  $\beta_2 = 0,900$ ?
- c.** Se  $\sigma = 0,5$  hora, qual é a probabilidade de o tempo de percurso ser de pelo menos 6 horas em relação a três entregas e a distância percorrida é de 50 milhas?

**38.** Sejam  $y =$  vida útil de um mancal,  $x_1 =$  viscosidade do óleo e  $x_2 =$  carga. Suponha que o modelo de regressão múltipla que relaciona a vida útil à viscosidade e à carga seja

$$Y = 125,0 + 0,775x_1 + 0,0950x_2 - 0,0090x_1x_2 + \epsilon$$

- a. Qual é o valor médio da vida útil em relação a uma viscosidade de 40 e carga de 1.100?
  - b. Quando a viscosidade é 30, qual é a mudança na vida média associada ao aumento de uma unidade na carga? Quando a viscosidade é 40, qual é a mudança na vida média associada com um aumento de uma unidade na carga?
39. Sejam  $y$  = vendas em uma lanchonete (milhares de \$),  $x_1$  = número de lanchonetes concorrentes no raio de 1 milha,  $x_2$  = população no raio de 1 milha (milhares de pessoas) e  $x_3$  = uma variável indicadora que é igual a 1 se a lanchonete tiver um balcão de atendimento para carros (*drive-thru*) e 0 se não tiver. Suponha que o modelo de regressão real seja:

$$Y = 10,00 - 1,2x_1 + 6,8x_2 + 15,3x_3 + \epsilon$$

- a. Qual é o valor médio das vendas em relação a 2 lojas concorrentes, 8.000 pessoas no raio de 1 milha, considerando que a lanchonete dispõe de balcão de atendimento para automóveis?
  - b. Qual é o valor médio das vendas em relação a uma lanchonete que não dispõe de balcão de atendimento para carros e tem três concorrentes e 5.000 pessoas no raio de 1 milha?
  - c. Interprete  $\beta_3$ .
40. O artigo "Readability of Liquid Crystal Displays: A Response Surface" (*Human Factors*, 1983, p. 185–190) usou um modelo de regressão múltipla com quatro variáveis independentes para investigar a precisão de leitura em monitores ou telas de cristal líquido (LCDs). As variáveis eram

$y$  = porcentagem de erros de leitura em um LCD de quatro dígitos

$x_1$  = nível da lâmpada traseira (no intervalo de 0 a 122 cd/m<sup>2</sup>)

$x_2$  = "subtense" do caractere (no intervalo de 0,025° a 1,34°)

$x_3$  = ângulo de visão (no intervalo de 0° a 60°)

$x_4$  = nível de luz ambiente (no intervalo de 20 a 1.500 lux)

A aderência do modelo aos dados foi  $Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + \epsilon$ . Os coeficientes estimados resultantes foram  $\hat{\beta}_0 = 1,52$ ,  $\hat{\beta}_1 = 0,02$ ,  $\hat{\beta}_2 = -1,40$ ,  $\hat{\beta}_3 = 0,02$  e  $\hat{\beta}_4 = -0,0006$ .

- a. Calcule uma estimativa da porcentagem de erros esperada quando  $x_1 = 10$ ,  $x_2 = 0,5$ ,  $x_3 = 50$  e  $x_4 = 100$ .
- b. Estime a porcentagem média de erros em relação a um nível 20 para a lâmpada traseira, 0,5 para a subtensa do caractere, 10 para o ângulo de visão e 30 para o nível de luz ambiente.
- c. Qual é a mudança esperada estimada na porcentagem de erros quando o nível de luz ambiente sofre um aumento de 1 unidade e todas as outras variáveis são fixadas nos valores fornecidos na parte (a)? Res-

ponda para um aumento de 100 unidades no nível de luz ambiente.

- d. Explique por que as respostas na parte (c) não dependem dos valores fixos de  $x_1$ ,  $x_2$  e  $x_3$ . Sob que circunstâncias haveria essa dependência?
  - e. O modelo estimado baseou-se em  $n = 30$  observações, com SQT = 39,2 e SQE = 20,0. Calcule e interprete o coeficiente de determinação múltipla e, em seguida, realize o teste de utilidade do modelo usando  $\alpha = 0,05$ .
41. A capacidade dos ecologistas de identificar regiões de maior riqueza de espécies poderia ter um impacto sobre a preservação da diversidade genética, um objetivo importante da Estratégia Mundial para a Conservação. O artigo "Prediction of Rarities from Habitat Variables: Coastal Plain Plants on Nova Scotian Lakeshores" (*Ecology*, 1992, p. 1.852–1.859) usou uma amostra de  $n = 37$  lagos para obter a equação de regressão estimada

$$y = 3,89 + 0,033x_1 + 0,024x_2 + 0,023x_3 - 0,0080x_4 - 0,13x_5 - 0,72x_6$$

onde  $y$  = riqueza de espécies,  $x_1$  = área da bacia hidrográfica,  $x_2$  = largura da costa,  $x_3$  = drenagem deficiente (%),  $x_4$  = cor da água (unidades de cores totais),  $x_5$  = areia (%) e  $x_6$  = alcalinidade. O coeficiente de determinação múltipla relatado foi  $R^2 = 0,83$ . Realize um teste de utilidade do modelo.

42. Uma investigação sobre um processo de fundição gerou os dados a seguir sobre  $x_1$  = temperatura da fornalha,  $x_2$  = tempo de moldagem da matriz e  $y$  = diferença de temperatura na superfície de moldagem da matriz ("A Multiple-Objective Decision-Making Approach for Assessing Simultaneous Improvement in Die Life and Casting Quality in a Die Casting Process", *Quality Engineering*, 1994, p. 371–383).

$x_1$	1.250	1.300	1.350	1.250	1.300
$x_2$	6	7	6	7	6
$y$	80	95	101	85	92

$x_1$	1.250	1.300	1.350	1.350
$x_2$	8	8	7	8
$y$	87	96	106	108

Aqui é apresentado o resultado do MINITAB do ajuste do modelo de regressão múltipla com os previsores  $x_1$  e  $x_2$ .

A equação de regressão é  
 $diftemp = -200 + 0,210 \text{ tempfornalha} + 3,00 \text{ tempofechamento}$

Previsor	Coef	DP	razão t	p
Constante	-199,56	11,64	-17,14	0,000
tempfornalha	0,210000	0,008642	24,30	0,000
tempofechamento	3,0000	0,4321	6,94	0,000

$s = 1,058$      $R^2 = 99,1\%$      $R^2$  (ajustado) = 98,8%

Análise de Variância

FONTE	GL	SQ	QM	F	p
Regressão	2	715,50	357,75	319,31	0,000
Erro	6	6,72	1,12		
Total	8	722,22			

- a. Execute o teste de utilidade do modelo.
- b. Calcule e interprete um intervalo de confiança de 95% para  $\beta_2$ , o coeficiente de regressão populacional de  $x_2$ .
- c. Quando  $x_1 = 1.300$  e  $x_2 = 7$ , o desvio padrão estimado de  $\hat{y}$  é  $s_{\hat{y}} = 0,353$ . Calcule um intervalo de confiança de 95% para a temperatura média real em relação a uma temperatura de fornalha de 1.300 e um tempo de moldagem de matriz de 7.
- d. Calcule um intervalo de previsão de 95% para a diferença de temperatura resultante de uma única operação experimental com uma temperatura de fornalha de 1.300 e um tempo de moldagem de matriz de 7.

43. Um experimento realizado para investigar o efeito do número de mols de cobalto ( $x_1$ ) e a temperatura de calcificação ( $x_2$ ) na área de superfície de um catalisador de hidróxido ferro-cobalto ( $y$ ) gerou os dados a seguir (“Structural Changes and Surface Properties of  $Co_xFe_{3-x}O_4$  Spinel”, *J. of Chemical Tech. and Biotech.*, 1994, p. 161–170). Quando solicitado para ajustar  $\beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3$ , onde  $x_3 = x_1x_2$  (um preditor de interação), o software SAS gerou o resultado a seguir.

$x_1$	0,6	0,6	0,6	0,6	0,6	1,0	1,0
$x_2$	200	250	400	500	600	200	250
$y$	90,6	82,7	58,7	43,2	25,0	127,1	112,3

Resultado do SAS para o Exercício 43

Variável Dependente: ÁREASUPERF

Análise de Variância

Fonte	GL	Soma dos Quadrados	Quadrado Médio	Valor F	Prob>F
Model	3	15223,52829	5074,50943	18,924	0,0001
Erro	16	4290,53971	268,15873		
Total C	19	19514,06800			

MQE Raiz	16,37555	R <sup>2</sup>	0,7801
Média Dep.	48,06000	R <sup>2</sup> Ajustado	0,7389
CV	34,07314		

Estimativas de Parâmetro

Variável	GL	Estimativa do Parâmetro	Erro-Padrão	T para H0: Parâmetro=0	Prob >  T
TERMO CONST.	1	185,485740	21,19747682	8,750	0,0001
VOLCOB	1	-45,969466	10,61201173	-4,332	0,0005
TEMP.	1	-0,301503	0,05074421	-5,942	0,0001
CONTEMP.	1	0,088801	0,02540388	3,496	0,0030

$x_1$	1,0	1,0	1,0	2,6	2,6	2,6	2,6
$x_2$	400	500	600	200	250	400	500
$y$	19,6	17,8	9,1	53,1	52,0	43,4	42,4

$x_1$	2,6	2,8	2,8	2,8	2,8	2,8
$x_2$	600	200	250	400	500	600
$y$	31,6	40,9	37,9	27,5	27,3	19,0

- a. Presuma o valor da área de superfície em relação a um volume de cobalto de 2,6 e a temperatura de 250, e calcule o valor do resíduo correspondente.
- b. Visto que  $\hat{\beta}_1 = -46,0$ , é apropriado concluir que, se o volume de cobalto sofrer um aumento de 1 unidade, mantendo-se fixos os valores dos outros preditores, é possível esperar que a área de superfície diminua em torno de 46 unidades? Explique seu raciocínio.
- c. Parece haver uma relação linear útil entre  $y$  e os preditores?
- d. Visto que o número de mols e a temperatura de calcificação permanecem no modelo, o preditor de interação  $x_3$  fornece informações úteis sobre  $y$ ? Expresse e teste as hipóteses apropriadas usando um nível de significância de 0,01.
- e. O desvio padrão estimado de  $\hat{Y}$  quando o número de mols é 2,0 e a temperatura de calcificação é 500 é  $s_{\hat{y}} = 4,69$ . Calcule um intervalo de confiança de 95% para o valor médio da área de superfície sob essas circunstâncias.

44. Os autores do artigo “An Ultracentrifuge Flour Absorption Method” (*Cereal Chemistry*, 1978, p. 96–101) investigaram a relação entre a absorção de água pela farinha de trigo e várias características da farinha. Mais especificamente, usaram um modelo de regressão linear múltipla de primeira ordem para relacionar a absorção  $y$  (%) à proteína da farinha  $x_1$  (%) e à perda de amido  $x_2$  (unidades Farrand). Os dados e o resultado do SPSS são apresentados a seguir:

$x_1$	$x_2$	$y$	$x_1$	$x_2$	$y$
8,5	2	30,9	12,9	24	47,0
8,9	3	32,7	12,0	25	46,8
10,6	3	36,7	12,9	28	45,9
10,2	20	41,9	13,1	28	48,8
9,8	22	40,9	11,4	32	46,2
10,8	20	42,9	13,2	28	47,8
11,6	31	46,3	11,6	35	49,2
12,0	32	47,6	12,1	34	48,3
12,5	31	47,2	11,3	35	48,6
10,9	28	44,0	11,1	40	50,2
12,2	36	47,7	11,5	45	49,6
11,9	28	43,9	11,6	50	53,2
11,3	30	46,8	11,7	55	54,3
13,0	27	46,2	11,7	57	55,8

- Interprete  $\hat{\beta}_1$  e  $\hat{\beta}_2$ .
- Que proporção de variação observada na absorção pode ser explicada pela relação do modelo?
- O modelo escolhido parece especificar uma relação linear útil entre a absorção e pelo menos um dos dois previsores?
- Se a proteína da farinha permanecer no modelo, a exclusão do preditor de perda de amido justifica-se?
- Quando  $x_1 = 10$  e  $x_2 = 25$ ,  $\hat{y} = 42,253$  e  $s_{\hat{y}} = 0,350$ . Calcule e interprete um intervalo de confiança e um intervalo de previsão.
- A inclusão de um preditor de interação  $x_3$  fornece  $\hat{\beta}_3 = -0,04304$  e  $s_{\hat{\beta}_3} = 0,01773$ . Esse preditor deveria ser mantido ao nível de significância 0,01?

Resultado do SPSS para o Exercício 44

R Múltiplo	0,98207
R <sup>2</sup>	0,96447
R <sup>2</sup> Ajustado	0,96163
Erro-Padrão	1,09412

Análise de Variância

	GL	Soma dos quadrados	Quadrado médio
Regressão	2	812,37959	406,18980
Resíduo	25	29,92755	1,19710

F = 339,31092      Signif F = ,0000

Variáveis na equação

Variável	B	EP B	95% Confiança	Intervalo B
AMIDO	0,33563	0,01814	0,29828	0,37298
FARINHA	1,44228	0,20764	1,01465	1,86991
(Constante)	19,43976	2,18829	14,93290	23,94662

45. O artigo “Analysis of the Modeling Methodologies for Predicting the Strength of Air-Jet Spun Yarns” (*Textile Res. J.*, 1997, p. 39–44) relatou um estudo realizado para relacionar a tenacidade da fibra ( $y$ , em g/tex.) com a contagem de fibras ( $x_1$ , em tex), porcentagem de poliéster ( $x_2$ ), pressão do primeiro bocal ( $x_3$ , em kg/cm<sup>2</sup>) e pressão do segundo bocal ( $x_4$ , em kg/cm<sup>2</sup>). A estimativa do termo constante na equação de regressão múltipla correspondente foi 6,121. Os coeficientes estimados dos quatro previsores foram  $-0,082$ ,  $0,113$ ,  $0,256$  e  $-0,219$ , respectivamente, e o coeficiente de determinação múltipla foi 0,946.

- Supondo que o tamanho da amostra tenha sido  $n = 25$ , expresse e teste as hipóteses apropriadas para determinar se o modelo ajustado especifica uma relação linear útil entre a variável dependente e no mínimo um entre os quatro previsores do modelo.
  - Novamente usando  $n = 25$ , calcule o valor de  $R^2$  ajustado.
  - Calcule um intervalo de confiança de 99% para a tenacidade média real da fibra quando a contagem de fibras for 16,5, a fibra contiver 50% de poliéster, a pressão do primeiro bocal for 3 e a pressão do segundo bocal for 5, e o desvio padrão estimado da tenacidade prevista sob essas circunstâncias for 0,350.
46. Uma análise de regressão foi realizada para relacionar  $y =$  tempo de reparo de um sistema de filtragem de água (h) e  $x_1 =$  tempo decorrido desde o serviço anterior (meses) e  $x_2 =$  tipo de reparo (1 se elétrico e 0 se mecânico) gerou o modelo a seguir baseado em  $n = 12$  observações:  $y = 0,950 + 0,400x_1 + 1,250x_2$ . Além disso,  $SQT = 12,72$ ,  $SQE = 2,09$  e  $s_{\hat{\beta}_2} = 0,312$ .
- Parece haver uma relação linear útil entre o tempo de reparo e os dois previsores do modelo? Realize um teste das hipóteses apropriadas usando um nível de significância de 0,05.
  - Visto que o tempo decorrido desde o último serviço permanece no modelo, o tipo de reparo fornece informações úteis sobre o tempo de reparo? Expresse e teste as hipóteses apropriadas usando um nível de significância de 0,01.

- c. Calcule e interprete um IC de 95% para  $\beta_2$ .
- d. O desvio padrão estimado de uma previsão do tempo de reparo, quando o tempo decorrido é igual a 6 meses e o tipo de reparo é elétrico, é 0,192. Presuma o tempo de reparo sob essas circunstâncias calculando um intervalo de previsão de 99%. O intervalo sugere que o modelo estimado fornecerá uma previsão precisa? Por que sim ou por que não?

47. O projeto eficaz de determinados tipos de incineradores de lixo municipal necessita de informações sobre o conteúdo de energia do lixo. Os autores do artigo “Modeling the Energy Content of Municipal Solid Waste Using Multiple Regression Analysis” (*J. of the Air and Waste Mgmt. Assoc.*, 1996, p. 650–656) gentilmente nos forneceram os dados a seguir sobre  $y$  = conteúdo de energia (kcal/kg), as três variáveis de composição física  $x_1$  = % plásticos por peso,  $x_2$  = % papel por peso e  $x_3$  = % de lixo por peso, e a variável de análise adjacente  $x_4$  = % umidade por peso para amostras de lixo obtidas de uma determinada região.

Obs	Plásticos	Papel	Lixo	Água	Conteúdo de Energia
1	18,69	15,65	45,01	58,21	947
2	19,43	23,51	39,69	46,31	1.407
3	19,24	24,23	43,16	46,63	1.452
4	22,64	22,20	35,76	45,85	1.553
5	16,54	23,56	41,20	55,14	989
6	21,44	23,65	35,56	54,24	1.162
7	19,53	24,45	40,18	47,20	1.466
8	23,97	19,39	44,11	43,82	1.656
9	21,45	23,84	35,41	51,01	1.254
10	20,34	26,50	34,21	49,06	1.336
11	17,03	23,46	32,45	53,23	1.097
12	21,03	26,99	38,19	51,78	1.266
13	20,49	19,87	41,35	46,69	1.401
14	20,45	23,03	43,59	53,57	1.223
15	18,81	22,62	42,20	52,98	1.216
16	18,28	21,87	41,50	47,44	1.334
17	21,41	20,47	41,20	54,68	1.155
18	25,11	22,59	37,02	48,74	1.453
19	21,04	26,27	38,66	53,22	1.278
20	17,99	28,22	44,18	53,37	1.153
21	18,73	29,39	34,77	51,06	1.225
22	18,49	26,58	37,55	50,66	1.237
23	22,08	24,88	37,07	50,72	1.327
24	14,28	26,27	35,80	48,24	1.229
25	17,74	23,61	37,36	49,92	1.205
26	20,54	26,58	35,40	53,58	1.221
27	18,25	13,77	51,32	51,38	1.138
28	19,09	25,62	39,54	50,13	1.295
29	21,25	20,63	40,72	48,67	1.391
30	21,62	22,71	36,22	48,19	1.372

Usando o MINITAB para ajustar um modelo de regressão múltipla com as quatro variáveis antes mencionadas como previsores do conteúdo de energia, foi obtido o seguinte resultado:

A equação de regressão é  

$$\text{contenerg.} = 2.245 + 28,9 \text{ plásticos} + 7,64 \text{ papel} + 4,30 \text{ lixo} - 37,4 \text{ água}$$

Previsor	Coef	DP	T	p
Constante	2244,9	177,9	12,62	0,000
plásticos	28,925	2,824	10,24	0,000
papel	7,644	2,314	3,30	0,003
lixo	4,297	1,916	2,24	0,034
água	-37,354	1,834	-20,36	0,000
s = 31,48	R <sup>2</sup> = 96,4	%R <sup>2</sup> (ajustado) = 95,8%		

Análise de Variância

Fonte	GL	SQ	QM	F	p
Regressão	4	664931	166233	167,71	0,000
Erro	25	24779	991		
Total	29	689710			

- a. Interprete os valores dos coeficientes de regressão estimados  $\hat{\beta}_1$  e  $\hat{\beta}_4$ .
  - b. Expresse e teste as hipóteses apropriadas para determinar se a aderência do modelo aos dados especifica uma relação linear útil entre o conteúdo de energia e pelo menos um entre os quatro previsores.
  - c. Visto que a % de plásticos, % de papel e % de água permanecem no modelo, a % de lixo fornece informações úteis sobre o conteúdo de energia? Expresse e teste as hipóteses apropriadas usando um nível de significância de 0,05.
  - d. Baseie-se no fato de que  $s_{\hat{y}} = 7,46$  quando  $x_1 = 20$ ,  $x_2 = 25$ ,  $x_3 = 40$  e  $x_4 = 45$  para calcular um intervalo de confiança de conteúdo médio real de energia sob essas circunstâncias. O intervalo resultante sugere que o conteúdo médio de energia foi estimado com precisão?
  - e. Use as informações fornecidas na parte (d) para prever o conteúdo de energia de uma amostra de lixo com as características especificadas de uma maneira que transmita informações sobre precisão e confiabilidade.
48. Um experimento realizado para investigar os efeitos de uma nova técnica para desengomar fibras da seda foi descrito no artigo “Some Studies in Degumming of Silk with Organic Acids” (*J. Society of Dyers and Colourists*, 1992, p. 79–86). Uma variável de resposta de interesse foi  $y$  = perda de peso (%). Os pesquisadores fizeram observações sobre a perda de peso para vários valores de três variáveis independentes:  $x_1$  = temperatura (°C) = 90, 100, 110;  $x_2$  = tempo de tratamento (min.) = 30, 75, 120;  $x_3$  = concentração de ácido tartárico (g/L) = 0, 8, 16. Nas análises de regressão, os três valores de cada variável foram codificados como -1, 0 e 1, respectivamente, gerando os dados a seguir (o valor  $y_8 = 19,3$  foi indicado, mas nosso valor  $y_8 = 20,3$  produziu um resultado de regressão idêntico ao reproduzido nesse artigo).



Obs	1	2	3	4	5	6	7	8
$x_1$	-1	-1	1	1	-1	-1	1	1
$x_2$	-1	1	-1	1	0	0	0	0
$x_3$	0	0	0	0	-1	1	-1	1
y	18,3	22,2	23,0	23,0	3,3	19,3	19,3	20,3
Obs	9	10	11	12	13	14	15	
$x_1$	0	0	0	0	0	0	0	
$x_2$	-1	-1	1	1	0	0	0	
$x_3$	-1	1	-1	1	0	0	0	
y	13,1	23,0	20,9	21,5	22,0	21,3	22,6	

Um modelo de regressão múltipla com  $k = 9$  previsores —  $x_1, x_2, x_3, x_4 = x_1^2, x_5 = x_2^2, x_6 = x_3^2, x_7 = x_1x_2, x_8 = x_1x_3, x_9 = x_2x_3$  — foi ajustado aos dados, gerando  $\hat{\beta}_0 = 21,967, \hat{\beta}_1 = 2,8125, \hat{\beta}_2 = 1,2750, \hat{\beta}_3 = 3,4375, \hat{\beta}_4 = -2,208, \hat{\beta}_5 = 1,867, \hat{\beta}_6 = -4,208, \hat{\beta}_7 = -0,975, \hat{\beta}_8 = -3,750, \hat{\beta}_9 = -2,325, SQE = 23,379$  e  $R^2 = 0,938$ .

- Esse modelo indica uma relação útil? Expresse e teste as hipóteses apropriadas usando um nível de significância de 0,01.
- O desvio padrão estimado de  $\hat{\mu}_Y$  quando  $x_1 = \dots = x_9 = 0$  (isto é, quando a temperatura = 100, tempo = 75 e concentração = 8) é 1,248. Calcule um IC de 95% para a perda de peso esperada em relação aos valores especificados da temperatura, do tempo e da concentração.
- Calcule um IP de 95% para um único valor de perda de peso a ser observado em relação aos valores 100, 75 e 8 de temperatura, tempo e concentração, respectivamente.
- Ajustando o modelo com apenas  $x_1, x_2$  e  $x_3$  como previsores, obtemos  $R^2 = 0,456$  e  $SQE = 203,82$ . Existe uma relação útil entre a perda de peso e pelo menos um dos previsores de segunda ordem  $x_4, x_5, \dots, x_9$ ? Expresse e teste as hipóteses apropriadas.

49. O artigo “The Influence of Temperature and Sunshine on the Alpha-Acid Contents of Hops” (*Agricultural Meteorology*, 1974, p. 375–382) relata os dados a seguir sobre a produção ( $y$ ), temperatura média no período entre a data de floração do lúpulo e a data de colheita ( $x_1$ ) e a porcentagem média de luz do sol durante o mesmo período ( $x_2$ ) para a variedade fuggles de lúpulo:

$x_1$	16,7	17,4	18,4	16,8	18,9	17,1
$x_2$	30	42	47	47	43	41
y	210	110	103	103	91	76
$x_1$	17,3	18,2	21,3	21,2	20,7	18,5
$x_2$	48	44	43	50	56	60
y	73	70	68	53	45	31

Apresentamos aqui um resultado parcial do MINITAB proveniente do ajuste do modelo de primeira ordem  $Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \epsilon$  usado nesse artigo:

Previsor	Coef	DP	razão t	p
Constante	415,11	82,52	5,03	0,000
Temp	-6,593	4,859	-1,36	0,208
Luz do Sol	-4,504	1,071	-4,20	0,002
s = 24,45		R <sup>2</sup> = 76,8%		R <sup>2</sup> (ajustado) = 71,6%

- Qual é o  $\hat{\mu}_{Y-18,9,43}$ , e qual é o resíduo correspondente?
  - Teste  $H_0: \beta_1 = \beta_2 = 0$  versus  $H_a$ : tanto  $\beta_1$  como  $\beta_2 \neq 0$  no nível 0,05.
  - O desvio padrão estimado de  $\hat{\beta}_0 + \hat{\beta}_1x_1 + \hat{\beta}_2x_2$  quando  $x_1 = 18,9$  e  $x_2 = 43$ , é 8,20. Use esses dados para obter um IC de 95% para  $\mu_{Y-18,9,43}$ .
  - Use as informações da parte (c) para obter um IP de 95% para a produção em um experimento futuro, em relação a  $x_1 = 18,9$  e  $x_2 = 43$ .
  - O MINITAB relatou que um IP de 95% para a produção, em relação a  $x_1 = 18$  e  $x_2 = 45$ , é (35,94, 151,63). Qual é o IP de 90% nessa situação?
  - Visto que  $x_2$  está no modelo, você manteria  $x_1$ ?
  - Ao ajustar o modelo  $Y = \beta_0 + \beta_2x_2 + \epsilon$  o valor resultante de  $R^2$  é 0,721. Verifique se a estatística  $F$  para testar  $H_0: Y = \beta_0 + \beta_2x_2 + \epsilon$  versus  $H_a: Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \epsilon$  satisfaz  $t^2 = f$ , onde  $t$  é o valor da estatística  $t$  da parte (f).
50. a. Quando o modelo  $Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1^2 + \beta_4x_2^2 + \beta_5x_1x_2 + \epsilon$  é ajustado aos dados sobre lúpulo do Exercício 49, a estimativa de  $\beta_5$  é  $\hat{\beta}_5 = 0,557$  com um desvio padrão estimado  $s_{\hat{\beta}_5} = 0,94$ . Teste  $H_0: \beta_5 = 0$  versus  $H_a: \beta_5 \neq 0$ .
- Cada razão  $t \hat{\beta}_i/s_{\hat{\beta}_i}$  ( $i = 1, 2, 3, 4, 5$ ) para o modelo da parte (a) é menor que 2 em valor absoluto, embora  $R^2 = 0,861$  para esse modelo. Seria correto extrair cada termo do modelo em virtude de sua pequena razão  $t$ ? Explique.
  - Usando  $R^2 = 0,861$  para o modelo da parte (a), teste  $H_0: \beta_3 = \beta_4 = \beta_5 = 0$  (o que expressa que todos os termos de segunda ordem podem ser excluídos).
51. O artigo “The Undrained Strength of Some Thawed Permafrost Soils” (*Canadian Geotechnical J.*, 1979, p. 420–427) contém os dados a seguir sobre a resistência ao corte sem drenagem de solos arenosos ( $y$ , em kPa), profundidade ( $x_1$ , em m) volume de água ( $x_2$ , em %).

	y	$x_1$	$x_2$	$\hat{y}$	$y - \hat{y}$	$e^*$
1	14,7	8,9	31,5	23,35	-8,65	-1,50
2	48,0	36,6	27,0	46,38	1,62	0,54
3	25,6	36,8	25,9	27,13	-1,53	-0,53
4	10,0	6,1	39,1	10,99	-0,99	-0,17
5	16,0	6,9	39,2	14,10	1,90	0,33
6	16,8	6,9	38,3	16,54	0,26	0,04
7	20,7	7,3	33,9	23,34	-2,64	-0,42
8	38,8	8,4	33,8	25,43	13,37	2,17
9	16,9	6,5	27,9	15,63	1,27	0,23
10	27,0	8,0	33,1	24,29	2,71	0,44
11	16,0	4,5	26,3	15,36	0,64	0,20
12	24,9	9,9	37,8	29,61	-4,71	-0,91
13	7,3	2,9	34,6	15,38	-8,08	-1,53
14	12,8	2,0	36,4	7,96	4,84	1,02

Para calcular os valores previstos e os resíduos, foi ajustado um modelo quadrático completo, que gerou a função de regressão estimada

$$y = -151,36 - 16,22x_1 + 13,48x_2 + 0,094x_1^2 - 0,253x_2^2 + 0,492x_1x_2$$

- a. Os gráficos de  $e^*$  versus  $x_1$ ,  $e^*$  versus  $x_2$  e  $e^*$  versus  $\hat{y}$  sugerem que o modelo quadrático completo deve ser modificado? Justifique sua resposta.
- b. O valor de  $R^2$  para o modelo quadrático completo é 0,759. Teste as hipóteses nulas no nível 0,05 estabelecendo que não há nenhuma relação linear entre a variável dependente e qualquer um dos cinco previsores.
- c. É possível mostrar que  $V(Y) = \sigma^2 = V(\hat{Y}) + V(Y - \hat{Y})$ . A estimativa de  $\sigma$  é  $\hat{\sigma} = s = 6,99$  (com base no modelo quadrático completo). Primeiro obtenha o desvio padrão estimado de  $Y - \hat{Y}$  e, em seguida, estime o desvio padrão de  $\hat{Y}$  (isto é,  $\hat{\beta}_0 + \hat{\beta}_1x_1 + \hat{\beta}_2x_2 + \hat{\beta}_3x_1^2 + \hat{\beta}_4x_2^2 + \hat{\beta}_5x_1x_2$ ) quando  $x_1 = 8,0$  e  $x_2 = 33,1$ . Por fim, calcule um IC de 95% para a resistência média. [Sugestão: Qual é  $(y - \hat{y})/e^*$ ?]
- d. Se calibrarmos o modelo de primeira ordem com a função de regressão  $\mu_{y,x_1,x_2} = \beta_0 + \beta_1x_1 + \beta_2x_2$ , obteremos SQE = 894,95. Usando o nível 0,05, teste a hipótese nula que expressa que todos os termos quadráticos podem ser excluídos do modelo.

52. A utilização da sacarose como fonte de carbono para a produção de substâncias químicas não é econômica. O melaço de beterraba é um substituto facilmente obtido e barato. O artigo “Optimization of the Production of  $\beta$ -Carotene from Molasses by *Blakeslea Trispora*” (*J. of Chemical Technology and Biotechnology*, 2002, p. 933–943) realizou uma análise de regressão múltipla para relacionar a variável dependente  $y$  = quantidade de  $\beta$ -caroteno ( $g/dm^3$ ) com três previsores: quantidade de ácido lineólico, quantidade de querosene e quantidade de antioxidante (todos em  $g/dm^3$ ).

Obs	Linoléico	Querosene	Antioxidante	Betacaroteno
1	30,00	30,00	10,00	0,7000
2	30,00	30,00	10,00	0,6300
3	30,00	30,00	18,41	0,0130
4	40,00	40,00	5,00	0,0490
5	30,00	30,00	10,00	0,7000
6	13,18	30,00	10,00	0,1000
7	20,00	40,00	5,00	0,0400
8	20,00	40,00	15,00	0,0065
9	40,00	20,00	5,00	0,2020
10	30,00	30,00	10,00	0,6300
11	30,00	30,00	1,59	0,0400
12	40,00	20,00	15,00	0,1320
13	40,00	40,00	15,00	0,1500
14	30,00	30,00	10,00	0,7000
15	30,00	46,82	10,00	0,3460
16	30,00	30,00	10,00	0,6300
17	30,00	13,18	10,00	0,3970
18	20,00	20,00	5,00	0,2690
19	20,00	20,00	15,00	0,0054
20	46,82	30,00	10,00	0,0640

- a. O ajuste do modelo completo de segunda ordem aos três previsores gerou  $R^2 = 0,987$  e  $R^2$  ajustado = 0,974, ao passo que o ajuste do modelo de primeira ordem gerou  $R^2 = 0,016$ . O que você concluiria sobre esses dois modelos?
- b. Para  $x_1 = x_2 = 30$ ,  $x_3 = 10$ , um software estatístico relatou que  $\hat{y} = 0,66573$ ,  $s_{\hat{y}} = 0,01785$  com base no modelo completo de segunda ordem. Presuma a quantidade de  $\beta$ -caroteno que resultaria de um único experimento realizado com os valores designados das variáveis independentes. Faça isso de uma maneira que lhe permita transmitir informações sobre precisão e confiabilidade.

53. Os *snowpacks* contêm um amplo espectro de poluentes que podem representar riscos ambientais. O artigo “Atmospheric PAH Deposition: Deposition Velocities and Washout Ratios” (*J. of Environmental Engineering*, 2002, p. 186–195) concentrou-se na deposição de hidrocarbonetos policíclicos aromáticos (HPAs). Os autores propuseram um modelo de regressão múltipla para relacionar a deposição ao longo de um período de tempo específico ( $y$ , em  $\mu g/m^2$ ) com dois previsores muito complexos  $x_1$  ( $\mu g\text{-sec}/m^3$ ) e  $x_2$  ( $\mu g/m^2$ ) definidos em função da concentração atmosférica de HPA para várias espécies, do tempo total e da quantidade total de precipitação. Aqui são apresentados os dados sobre a espécie fluoranteno e o resultado correspondente do MINITAB:

obs	x1	x2	flth
1	92017	0,0026900	278,78
2	51830	0,0030000	124,53
3	17236	0,0000196	22,65
4	15776	0,0000360	28,68
5	33462	0,0004960	32,66
6	243500	0,0038900	604,70
7	67793	0,0011200	27,69
8	23471	0,0006400	14,18
9	13948	0,0004850	20,64
10	8824	0,0003660	20,60
11	7699	0,0002290	16,61
12	15791	0,0014100	15,08
13	10239	0,0004100	18,05
14	43835	0,0000960	99,71
15	49793	0,0000896	58,97
16	40656	0,0026000	172,58
17	50774	0,0009530	44,25

A equação de regressão é

$$flth = -33,5 + 0,00205 x_1 + 29836 x_2$$

Previsor	Coef	Coef EP	T	P
Constante	-33,46	14,90	-2,25	0,041
x1	0,0020548	0,0002945	6,98	0,000
x2	29836	13654	2,19	0,046

$$S = 44,28 \quad R^2 = 92,3\% \quad R^2(\text{ajustado}) = 91,2\%$$

Análise de Variância

Fonte	GL	SQ	QM	F	P
Regressão	2	330989	165495	84,39	0,000
Erro Residual	14	27454	1961		
Total	16	358443			

Formule perguntas e execute análises apropriadas para tirar algumas conclusões.

54. Os dados a seguir resultaram de um estudo sobre a relação entre o brilho do papel manufaturado ( $y$ ) e as variáveis %  $H_2O_2$  por peso ( $x_1$ ), % NaOH por peso ( $x_2$ ), % silicato por peso ( $x_3$ ) e temperatura do processo ( $x_4$ )

“Advantages of CEHDP Bleaching for High Brightness Kraft Pulp Production”, *TAPPI*, 1964, p. 170A–173A). Permitiu-se que cada variável independente assumisse cinco diferentes valores, e esses valores foram codificados para a análise de regressão como -2, -1, 0, 1 e 2.

Teste No.	H <sub>2</sub> O <sub>2</sub> (x <sub>1</sub> )	NaOH Conc. (x <sub>2</sub> )	Silicato Conc. (x <sub>3</sub> )	Temp. (x <sub>4</sub> )	Brilho. (y)
1	-1	-1	-1	-1	83,9
2	+1	-1	-1	-1	84,9
3	-1	+1	-1	-1	83,4
4	+1	+1	-1	-1	84,2
5	-1	-1	+1	-1	83,8
6	+1	-1	+1	-1	84,7
7	-1	+1	+1	-1	84,0
8	+1	+1	+1	-1	84,8
9	-1	-1	-1	+1	84,5
10	+1	-1	-1	+1	86,0
11	-1	+1	-1	+1	82,6
12	+1	+1	-1	+1	85,1
13	-1	-1	+1	+1	84,5
14	+1	-1	+1	+1	86,0
15	-1	+1	+1	+1	84,0
16	+1	+1	+1	+1	85,4
17	-2	0	0	0	82,9
18	+2	0	0	0	85,5
19	0	-2	0	0	85,2
20	0	+2	0	0	84,5
21	0	0	-2	0	84,7
22	0	0	+2	0	85,0
23	0	0	0	-2	84,9
24	0	0	0	+2	84,0
25	0	0	0	0	84,5
26	0	0	0	0	84,7
27	0	0	0	0	84,6
28	0	0	0	0	84,9
29	0	0	0	0	84,9
30	0	0	0	0	84,5
31	0	0	0	0	84,6

Variáveis	-2	-1	0	+1	+2
x <sub>1</sub> Peróxido de hidrogênio (100%), %peso	0,1	0,2	0,3	0,4	0,5
x <sub>2</sub> NaOH, %peso	0,1	0,2	0,3	0,4	0,5
x <sub>3</sub> Silicato (41°Bé), %peso	5	1,5	2,5	3,5	4,5
x <sub>4</sub> Temperatura do processo, °F	130	145	160	175	190

a. Quando se ajustou um modelo (codificado) compreendendo todos os termos lineares, todos os termos quadráticos e todos os termos do produto cruzado, a função de regressão estimada era

$$y = 84,67 + 0,650x_1 - 0,258x_2 + 0,133x_3 + 0,108x_4 - 0,135x_1^2 + 0,028x_2^2 + 0,028x_3^2 - 0,072x_4^2 + 0,038x_1x_2 - 0,075x_1x_3 + 0,213x_1x_4 + 0,200x_2x_3 - 0,188x_2x_4 + 0,050x_3x_4$$

Use esse modelo estimado para prever o brilho em relação a 0,4% de H<sub>2</sub>O<sub>2</sub>, 0,4% de NaOH, 3,5% de silicato e uma temperatura de 175. Quais são os valores dos resíduos correspondentes a esses valores das variáveis?

b. Expresse a função de regressão estimada de maneira não codificada.

c. SQT = 17,2567 e R<sup>2</sup> para o modelo da parte (a) é 0,885. Ao ajustar um modelo que inclui apenas os quatro termos lineares, o valor resultante de R<sup>2</sup> é 0,721. Expresse e teste, com nível 0,05, a hipótese nula que estabelece que os coeficientes de todos os termos quadráticos e do produto cruzado na função de regressão são zero.

d. A função de regressão estimada (codificada) quando apenas termos lineares são incluídos é  $\hat{\mu}_{Y \cdot x_1 \cdot x_2 \cdot x_3 \cdot x_4} = 85,5548 + 0,6500x_1 - 0,2583x_2 + 0,1333x_3 + 0,1083x_4$ . Quando  $x_1 = x_2 = x_3 = x_4 = 0$ , o desvio padrão estimado de  $\hat{\mu}_{Y \cdot 0,0,0,0}$  é 0,0772. Suponha que se tenha acreditado que o brilho esperado para esses valores dos  $x_i$ s fosse de 85,0, no mínimo. As informações fornecidas contradizem essa convicção? Expresse e teste as hipóteses apropriadas.

### 13.5 Outras Questões em Regressão Múltipla

Nesta seção abordamos inúmeras questões que podem surgir ao realizar uma análise de regressão múltipla. Veja as referências do capítulo para obter um tratamento mais abrangente de qualquer tópico em particular.

#### Transformações em Regressão Múltipla

Em geral, considerações teóricas sugerem uma relação não-linear entre a variável dependente e duas ou mais variáveis independentes, ao passo que em outras circunstâncias os gráficos de diagnóstico propõem que seja usado algum tipo de função não-linear. Frequentemente, uma transformação linearizará o modelo.

**Exemplo 13.18**

Um artigo publicado no periódico *Lubrication Eng.* (“Accelerated Testing of Solid Film Lubricants”, 1972, p. 365–372) relatou uma pesquisa sobre o tempo de vida dos lubrificantes de películas sólidas. Três conjuntos de testes de mancais foram realizados em uma película Mil-L-8937 em cada uma das combinações de três cargas (3.000, 6.000 e 10.000 psi) e três velocidades (20, 60 e 100 rpm), e a vida útil (horas) foi registrada para cada operação, como mostrado na Tabela 13.7.

**Tabela 13.7** Dados sobre a vida útil do Exemplo 13.18

$s$	$l$ (1000's)	$w$	$s$	$l$ (1000's)	$w$
20	3	300,2	60	6	65,9
20	3	310,8	60	10	10,7
20	3	333,0	60	10	34,1
20	6	99,6	60	10	39,1
20	6	136,2	100	3	26,5
20	6	142,4	100	3	22,3
20	10	20,2	100	3	34,8
20	10	28,2	100	6	32,8
20	10	102,7	100	6	25,6
60	3	67,3	100	6	32,7
60	3	77,9	100	10	2,3
60	3	93,9	100	10	4,4
60	6	43,0	100	10	5,8
60	6	44,5			

Esse artigo contém uma observação que afirma que uma distribuição lognormal é apropriada para  $W$ , visto que  $\ln(W)$ , como se sabe, segue uma lei normal (lembre-se do Capítulo 4, em que se diz que isso é o que define uma distribuição lognormal). O modelo que aparece é  $W = (c/s^a l^b) \cdot \epsilon$ , onde  $\ln(W) = \ln(c) - a \ln(s) - b \ln(l) + \ln(\epsilon)$ ; portanto, com  $Y = \ln(W)$ ,  $x_1 = \ln(s)$ ,  $x_2 = \ln(l)$ ,  $\beta_0 = \ln(c)$ ,  $\beta_1 = -a$ , e  $\beta_2 = -b$ , temos um modelo de regressão linear múltipla. Depois de calcular  $\ln(w_i)$ ,  $\ln(s_i)$  e  $\ln(l_i)$  para os dados, um modelo de primeira ordem nas variáveis transformadas gerou os resultados mostrados na Tabela 13.8.

**Tabela 13.8** Coeficientes estimados e razões  $t$  do Exemplo 13.18

Parâmetro $\beta_i$	Estimativa $\hat{\beta}_i$	DP Estimativa $s_{\hat{\beta}_i}$	$t = \hat{\beta}_i/s_{\hat{\beta}_i}$
$\beta_0$	10,8719	0,7871	13,81
$\beta_1$	-1,2054	0,1710	-7,05
$\beta_2$	-1,3979	0,2327	-6,01

O valor do coeficiente de determinação múltipla (para a aderência transformada) é  $R^2 = 0,781$ . A função de regressão estimada para as variáveis transformadas é

$$\ln(w) = 10,87 - 1,21 \ln(s) - 1,40 \ln(l)$$

de modo que a função de regressão original é estimada como

$$w = e^{10,87} \cdot s^{-1,21} \cdot l^{-1,40}$$

A abordagem de Bonferroni pode ser usada para obter ICs simultâneos para  $\beta_1$  e  $\beta_2$ , e pelo fato de  $\beta_1 = -a$  e  $\beta_2 = -b$ , os intervalos para  $a$  e  $b$  são, portanto, obtidos imediatamente. ■

Na seção 13.2, o modelo de regressão logística foi introduzido para relacionar uma variável dicotômica  $y$  a um único previsor. Esse modelo pode ser facilmente estendido para incorporar mais de um previsor.

## Padronizando Variáveis

Na Seção 13.3, consideramos transformar  $x$  em  $x' = x - \bar{x}$  antes de ajustar um polinômio. Em relação à regressão múltipla, especialmente quando os valores das variáveis são de grande magnitude, é vantajoso aprofundar essa codificação. Sejam  $\bar{x}_j$  e  $s_j$  a média amostral e o desvio padrão amostral dos  $x_j$ s ( $j = 1, \dots, n$ ). Agora, vamos codificar cada variável  $x_i$  por  $x'_i = (x_i - \bar{x}_i)/s_i$ . A variável codificada  $x'_i$  simplesmente reexpressa qualquer valor  $x_i$  em unidades do desvio padrão acima ou abaixo da média. Portanto, se  $\bar{x}_i = 100$  e  $s_i = 20$ ,  $x_i = 130$  passa a ser  $x'_i = 1,5$  porque 130 está 1,5 desvios padrão acima da média dos valores de  $x_i$ . Por exemplo, o modelo completo de segunda ordem codificado com duas variáveis independentes tem a função de regressão

$$\begin{aligned} E(Y) &= \beta_0 + \beta_1 \left( \frac{x_1 - \bar{x}_1}{s_1} \right) + \beta_2 \left( \frac{x_2 - \bar{x}_2}{s_2} \right) + \beta_3 \left( \frac{x_1 - \bar{x}_1}{s_1} \right)^2 \\ &\quad + \beta_4 \left( \frac{x_2 - \bar{x}_2}{s_2} \right)^2 + \beta_5 \left( \frac{x_1 - \bar{x}_1}{s_1} \right) \left( \frac{x_2 - \bar{x}_2}{s_2} \right) \\ &= \beta_0 + \beta_1 x'_1 + \beta_2 x'_2 + \beta_3 x'^2_1 + \beta_4 x'^2_2 + \beta_5 x'_1 x'_2 \end{aligned}$$

Os benefícios da codificação são (1) maior precisão numérica em todos os cálculos (em decorrência do menor número de erros de arredondamento por parte do computador) e (2) estimativa mais precisa, em comparação com os parâmetros do modelo não-codificado, pelo fato de os parâmetros individuais do modelo codificado evidenciarem o comportamento da função de regressão próximo ao centro dos dados, e não próximo da origem.

### Exemplo 13.19

O artigo “The Value and the Limitations of High-Speed Turbo-Exhausters for the Removal of Tar-Fog from Carburetted Water-Gas” (*Soc. Chemical Industry J.*, 1946, p. (166-168) apresenta os dados (na Tabela 13.9) sobre  $y =$  volume de alcatrão (grãos/100 ft<sup>3</sup>) de uma corrente de gás como uma função de  $x_1 =$  velocidade do rotor (rpm) e  $x_2 =$  temperatura de entrada do gás (°F). Além disso, esses dados são também investigados no artigo “Some Aspects of Nonorthogonal Data Analysis” (*J. Quality Technology*, 1973, p. 67-79), que recomenda o modelo codificado descrito anteriormente.

**Tabela 13.9** Dados do Exemplo 13.19

Operação	$y$	$x_1$	$x_2$	$x'_1$	$x'_2$
1	60,0	2400	54,5	-1,52428	-0,57145
2	61,0	2450	56,0	-1,39535	-0,35543
3	65,0	2450	58,5	-1,39535	0,00461
4	30,5	2500	43,0	-1,26642	-2,22763
5	63,5	2500	58,0	-1,26642	-0,06740
6	65,0	2500	59,0	-1,26642	0,07662
7	44,0	2700	52,5	-0,75070	-0,85948
8	52,0	2700	65,5	-0,75070	1,01272
9	54,5	2700	68,0	-0,75070	1,37276
10	30,0	2750	45,0	-0,62177	-1,93960
11	26,0	2775	45,5	-0,55731	-1,86759
12	23,0	2800	48,0	-0,49284	-1,50755
13	54,0	2800	63,0	-0,49284	0,65268
14	36,0	2900	58,5	-0,23499	0,00461
15	53,5	2900	64,5	-0,23499	0,86870
16	57,0	3000	66,0	0,02287	1,08472
17	33,5	3075	57,0	0,21627	-0,21141
18	34,0	3100	57,5	0,28073	-0,13941
19	44,0	3150	64,0	0,40966	0,79669

(continuação)

Operação	y	x <sub>1</sub>	x <sub>2</sub>	x' <sub>1</sub>	x' <sub>2</sub>
20	33,0	3200	57,0	0,53859	-0,21141
21	39,0	3200	64,0	0,53859	0,79669
22	53,0	3200	69,0	0,53859	1,51677
23	38,5	3225	68,0	0,60305	1,37276
24	39,5	3250	62,0	0,66752	0,50866
25	36,0	3250	64,5	0,66752	0,86870
26	8,5	3250	48,0	0,66752	-1,50755
27	30,0	3500	60,0	1,31216	0,22063
28	29,0	3500	59,0	1,31216	0,07662
29	26,5	3500	58,0	1,31216	-0,06740
30	24,5	3600	58,0	1,57002	-0,06740
31	26,5	3900	61,0	2,34360	0,36465

As médias e os desvios padrão são  $\bar{x}_1 = 2.991,13$ ,  $s_1 = 387,81$ ,  $\bar{x}_2 = 58,468$  e  $s_2 = 6,944$ , de modo que  $x'_1 = (x_1 - 2.991,13)/387,81$  e  $x'_2 = (x_2 - 58,468)/6,944$ . Com  $x'_3 = (x'_1)^2$ ,  $x'_4 = (x'_2)^2$ ,  $x'_5 = x'_1 \cdot x'_2$ , a calibração do modelo completo de segunda ordem requer a solução de um sistema de seis equações normais em seis incógnitas. Uma análise por computador gerou  $\hat{\beta}_0 = 40,2660$ ,  $\hat{\beta}_1 = -13,4041$ ,  $\hat{\beta}_2 = 10,2553$ ,  $\hat{\beta}_3 = 2,3313$ ,  $\hat{\beta}_4 = -2,3405$  e  $\hat{\beta}_5 = 2,5978$ . A equação da regressão estimada é, portanto,

$$\hat{y} = 40,27 - 13,40x'_1 + 10,26x'_2 + 2,33x'_3 - 2,34x'_4 + 2,60x'_5$$

Logo se  $x_1 = 3.200$  e  $x_2 = 57,0$ ,  $x'_1 = 0,539$ ,  $x'_2 = -0,211$ ,  $x'_3 = (0,539)^2 = 0,2901$ ,  $x'_4 = (-0,211)^2 = 0,0447$  e  $x'_5 = (0,539)(-0,211) = -0,1139$ , de modo que

$$\begin{aligned} \hat{y} &= 40,27 - (13,40)(0,539) + (10,26)(-0,211) + (2,33)(0,2901) \\ &\quad - (2,34)(0,0447) + (2,60)(-0,1139) = 31,16 \end{aligned}$$

■

## Seleção de Variáveis

Em geral, o pesquisador terá dados sobre um grande número de previsores e, portanto, desejará construir um modelo de regressão que compreenda um subconjunto desses previsores. O uso desse subconjunto tornará o modelo resultante mais fácil de gerenciar, em especial se existir a intenção de coletar mais dados posteriormente, e também mais fácil de interpretar e compreender do que um modelo com um número bem maior de previsores. Existem duas perguntas fundamentais em relação à seleção de variáveis:

1. Se pudermos examinar as regressões que compreendem todos os possíveis subconjuntos de previsores para os quais dispomos de dados, que critérios devem ser usados para selecionar um modelo?
2. Se o número de previsores for muito grande para que todas as regressões sejam examinadas, existe alguma maneira de examinar um número menor de subconjuntos dentre os quais seja encontrado um bom modelo (ou modelos)?

Respondendo à Pergunta (1) em primeiro lugar, se o número de previsores for pequeno (digamos,  $\leq 5$ ), não seria tão tedioso examinar todas as possíveis regressões usando qualquer um dos softwares estatísticos que podem ser facilmente obtidos. Se for possível obter dados sobre seis previsores, no mínimo, todas as possíveis regressões compreenderão pelo menos 64 ( $= 2^6$ ) modelos diferentes. Para qualquer  $m$  especificado entre 1 e 10, vários softwares fornecerão os  $m$  melhores modelos com um único predictor, os  $m$  melhores modelos com dois previsores, e assim por diante ("melhor", aqui, significa a menor SQE ou, de maneira equivalente, o maior  $R^2$ ). O MINITAB fará isso para 20 previsores, no máximo, ao passo que o BMDP manipulará no máximo 27. Portanto, as SQEs correspondentes (ou as respectivas funções) podem ser comparadas de acordo com qualquer critério descrito a seguir. O motivo para especificar um  $m$  maior que 1 é verificar se nos melhores modelos os valores de SQE e  $R^2$  são similares.

**Critérios de Seleção de Variáveis** Uma vez mais usamos um subscrito  $k$  para representar uma quantidade (digamos,  $SQE_k$ ) calculada com base em um modelo com  $k$  previsores (e, portanto,  $k + 1$   $\beta$ 's, porque  $\beta_0$  sempre será incluído). Para um valor fixo de  $k$ , é sensato identificar o melhor modelo como aquele que tem a  $SQE_k$  mínima. O maior problema está em comparar as  $SQE_k$ s de diferentes valores de  $k$ . Três critérios distintos, cada um deles uma função de  $SQE_k$ , são amplamente usados.

1.  $R_k^2$ , o coeficiente de determinação múltipla de um modelo de  $k$  previsores. Pelo fato de  $R_k^2$  sempre aumentar à medida que  $k$  aumenta (e nunca poder diminuir), não estamos interessados no  $k$  que maximiza  $R_k^2$ . Na verdade, desejamos identificar um  $k$  pequeno para o qual  $R_k^2$  seja praticamente tão grande quanto  $R^2$  para todos os previsores no modelo.
2.  $QME_k = SQE/(n - k - 1)$ , o erro médio quadrático para um modelo de  $k$  previsores. Esse critério é freqüentemente usado no lugar de  $R_k^2$ , porque, embora  $R_k^2$  nunca diminua à medida que  $k$  aumenta, uma pequena diminuição em  $SQE_k$  obtida com um predictor adicional pode ser mais do que compensada por um decréscimo de 1 no denominador de  $QME_k$ . Portanto, o objetivo é encontrar o modelo que tenha o  $QME_k$  mínimo. Visto que o  $R_k^2$  ajustado  $= 1 - QME_k/QMT$ , onde  $QMT = SQT/(n - 1)$  é constante em  $k$ , a avaliação de  $R_k^2$  ajustado é equivalente à avaliação de  $QME_k$ .
3. O raciocínio lógico do terceiro critério,  $C_k$ , é mais difícil de compreender, mas esse critério é amplamente usado pelos analistas de dados. Suponha que o modelo de regressão real seja especificado por  $m$  previsores — ou seja,

$$Y = \beta_0 + \beta_1 x_1 + \cdots + \beta_m x_m + \epsilon \quad V(\epsilon) = \sigma^2$$

de modo que

$$E(Y) = \beta_0 + \beta_1 x_1 + \cdots + \beta_m x_m$$

Experimente ajustar um modelo usando um subconjunto de  $k$  desses  $m$  previsores; para simplificar, suponha que usemos  $x_1, x_2, \dots, x_k$ . Em seguida, solucionando o esquema de equações normais, obtemos as estimativas  $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$  (mas não, obviamente, as estimativas de qualquer  $\beta$  que corresponda aos previsores que não se encontram no modelo ajustado). O valor esperado real  $E(y)$  pode ser, portanto estimado por  $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \cdots + \hat{\beta}_k x_k$ . Considere agora o **erro total esperado normalizado da estimativa**

$$\Gamma_k = \frac{E\left(\sum_{i=1}^n [\hat{Y}_i - E(Y_i)]^2\right)}{\sigma^2} = \frac{E(SQE_k)}{\sigma^2} + 2(k + 1) - n \quad (13.21)$$

A segunda igualdade em (13.21) deve ser levada em conta porque exige um argumento de valor esperado enganoso. Portanto, um determinado subconjunto será interessante se seu valor  $\Gamma_k$  for pequeno. Infelizmente, contudo,  $E(SQE_k)$  e  $\sigma^2$  não são conhecidos. Para resolver isso, consideremos  $s^2$  a estimativa de  $\sigma^2$  com base no modelo que inclui todos os previsores para os quais os dados estão disponíveis e definem

$$C_k = \frac{SQE_k}{s^2} + 2(k + 1) - n$$

Portanto, um modelo interessante é especificado por um subconjunto de previsores para os quais  $C_k$  é pequeno.

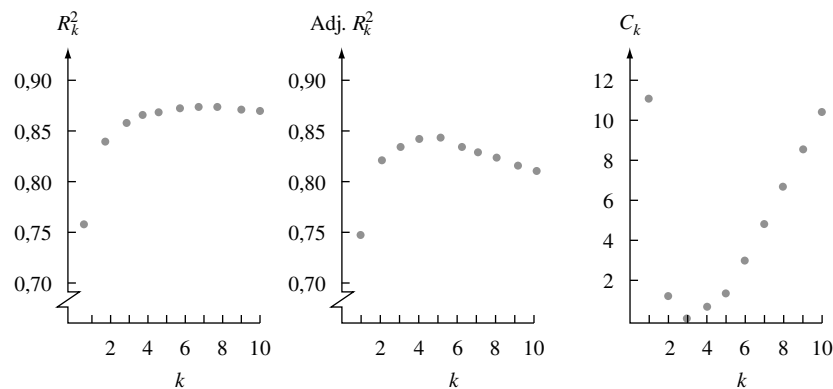
### Exemplo 13.20

O artigo de avaliação crítica de Ron Hocking relacionado na bibliografia deste capítulo apresenta uma análise sobre os dados extraídos das edições de 1974 da revista *Motor Trend*. A variável dependente era  $y =$  consumo

de gasolina por milha, havia  $n = 32$  observações e os previsores para os quais foram obtidos dados eram  $x_1 =$  formato do motor (1 = reto e 0 = V),  $x_2 =$  número de cilindros,  $x_3 =$  tipo de transmissão (1 = manual e 0 = automática),  $x_4 =$  número de velocidades de transmissão,  $x_5 =$  tamanho do motor,  $x_6 =$  cavalo-vapor,  $x_7 =$  número de carcaças do carburador,  $x_8 =$  coeficiente de transmissão final,  $x_9 =$  peso e  $x_{10} =$  tempo em um quarto de milha. Na Tabela 13.10 apresentamos um resumo das informações dessa análise. Para cada  $k$ , a tabela descreve o subconjunto com  $SQE_k$  mínima; em ordem decrescente, a coluna de variáveis indica que a variável é adicionada ao passar de  $k$  para  $k + 1$  (ao passar de  $k = 2$  para  $k = 3$ , tanto  $x_3$  quanto  $x_{10}$  são adicionadas, ao passo que  $x_2$  é excluída). A Figura 13.17 logo a seguir apresenta gráficos de  $R_k^2$  e  $R_k^2$  ajustado e  $C_k$  em relação a  $k$ ; esses gráficos são fundamentais como suporte visual para a seleção de um subconjunto. A estimativa de  $\sigma^2$  é  $s^2 = 6,24$ , que é  $QME_{10}$ . Um modelo simples, mas altamente favorável a todos os critérios é aquele que contém os previsores  $x_3, x_9$  e  $x_{10}$ .

**Tabela 13.10** Melhores subconjuntos dos dados para consumo de gasolina/milha do Exemplo 13.20

$k =$ Número de Previsores	Variáveis	$SQE_k$	$R_k^2$	$R_k^2$ ajustado	$C_k$
1	9	247,2	0,756	0,748	11,6
2	2	169,7	0,833	0,821	1,2
3	3, 10, -2	150,4	0,852	0,836	0,1
4	6	142,3	0,860	0,839	0,8
5	5	136,2	0,866	0,840	1,8
6	8	133,3	0,869	0,837	3,4
7	4	132,0	0,870	0,832	5,2
8	7	131,3	0,871	0,826	7,1
9	1	131,1	0,871	0,818	9,0
10	2	131,0	0,871	0,809	11,0



**Figura 13.17**  $R_k^2$  e  $C_k$  para os dados sobre consumo de gasolina/milha

Genericamente falando, quando se usa um subconjunto de  $k$  previsores ( $k < m$ ) para ajustar um modelo, os estimadores  $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$  serão influenciados por  $\beta_0, \beta_1, \dots, \beta_k$  e  $\hat{Y}$  também será um estimador tendencioso para a  $E(Y)$  real (tudo isso porque estão faltando no modelo ajustado  $m - k$  previsores). Contudo, como medido pelo erro normal total esperado  $\Gamma_k$ , a precisão das estimativas baseadas em um subconjunto é maior do que a que seria obtida se fossem usados todos os previsores possíveis; fundamentalmente, essa maior precisão implica um custo, a introdução de um viés nos estimadores. Um valor de  $k$  para o qual  $C_k \approx k + 1$  indica que o desvio associado com este modelo de  $k$  previsores seria pequeno.



**Exemplo 13.21**

Os dados sobre a resistência de cisalhamento de ligação do Exemplo 13.12 contêm o valor de quatro diferentes variáveis independentes,  $x_1 - x_4$ . Descobrimos que o modelo que compreende apenas essas quatro variáveis como previsores era útil, e não há nenhuma razão convincente para pensarmos em incluir previsores de segunda ordem. A Figura 13.18 apresenta a resposta à solicitação feita ao MINITAB para que identificasse os dois melhores modelos de cada tamanho especificado.

O melhor modelo de dois previsores, em que constam os previsores potência e temperatura, parece ser uma ótima escolha em todas as contagens:  $R^2$  é significativamente maior do que para os modelos com poucos previsores, ainda que quase tão grande quanto para qualquer modelo maior,  $R^2$  ajustado alcançou praticamente seu valor máximo em relação a esses dados e  $C_k$  é pequeno e próximo de  $2 + 1 = 3$ .

A resposta é resistência					t e m p e r r a t o r e s			
Variáveis	$R^2$	$R^2$ Ajustado	C-p	s				
1	57,7	56,2	11,0	5,9289	X			
1	10,8	7,7	51,9	8,6045		X		
2	68,5	66,2	3,5	5,2070	X	X		
2	59,4	56,4	11,5	5,9136	X		X	
3	70,2	66,8	4,0	5,1590	X	X	X	
3	69,7	66,2	4,5	5,2078	X	X	X	
4	71,4	66,8	5,0	5,1580	X	X	X	X

**Figura 13.18** Resultado da opção Melhores Subconjuntos do MINITAB

**Regressão Passo a Passo** Quando o número de previsores é muito grande para permitir a avaliação explícita ou implícita de todos os possíveis subconjuntos, vários procedimentos alternativos de seleção geralmente identificarão bons modelos. O procedimento mais simples é o método de **eliminação retroativa (ER)**. Esse método começa com o modelo no qual são usados todos os previsores em avaliação. Seja  $x_1, \dots, x_m$  o conjunto de todos esses previsores. Portanto, cada razão  $\hat{\beta}_i/s_{\hat{\beta}_i}$  ( $i = 1, \dots, m$ ) apropriada para testar  $H_0: \beta_i = 0$  versus  $H_a: \beta_i \neq 0$  é examinada. Se a razão  $t$  com o menor valor absoluto for menor que uma constante  $t_{out}$  pré-especificada, ou seja, se

$$\min_{i=1, \dots, m} \left| \frac{\hat{\beta}_i}{s_{\hat{\beta}_i}} \right| < t_{out}$$

então, o predictor correspondente a essa razão menor é eliminado do modelo. O modelo reduzido é então ajustado, as  $m - 1$  razões  $t$  são novamente examinadas, e outro predictor é eliminado se corresponder à menor razão  $t$  absoluta inferior a  $t_{out}$ . Nesse sentido, o algoritmo continua até determinada etapa; todas as razões  $t$  absolutas são no mínimo  $t_{out}$ . O modelo usado é aquele que contém todos os previsores que não foram eliminados. O valor  $t_{out} = 2$  em geral é recomendado, visto que a maioria dos valores  $t_{0,05}$  são aproximadamente 2. Alguns softwares concentram-se nos valores  $P$ , em vez de nas razões  $t$ .

**Exemplo 13.22** (Continuação do Exemplo 13.19)

Em relação ao modelo quadrático completo codificado, no qual  $y =$  volume de alcatrão, os cinco previsores possíveis são  $x'_1, x'_2, x'_3 = x_1'^2, x'_4 = x_2'^2$  e  $x'_5 = x_1'x_2'$  (de modo que  $m = 5$ ). Sem especificar  $t_{out}$ , o predictor com a menor razão  $t$  absoluta (com asterisco) foi eliminado em cada uma das etapas, gerando a seqüência de modelos mostrada na Tabela 13.11.

**Tabela 13.11** Resultado da eliminação para os dados do Exemplo 13.19

Etapa	Previsores	razão $t$				
		1	2	3	4	5
1	1, 2, 3, 4, 5	16,0	10,8	2,9	2,8	1,8*
2	1, 2, 3, 4	15,4	10,2	3,7	2,0*	—
3	1, 2, 3	14,5	12,2	4,3*	—	—
4	1, 2	10,9	9,1*	—	—	—
5	1	4,4*	—	—	—	—

Se usássemos  $t_{\text{out}} = 2$ , o modelo resultante teria por base  $x'_1, x'_2$  e  $x'_3$ , visto que na Etapa 3 nenhum predictor poderia ser eliminado. É possível confirmar se cada subconjunto é de fato o melhor subconjunto do respectivo tamanho, embora isto não aconteça sempre. ■

Uma alternativa ao procedimento ER é a **seleção progressiva (SP)**. A SP inicia-se sem parâmetro algum no modelo e experimenta ajustar o modelo sucessivamente com apenas  $x_1$ , apenas  $x_2, \dots$ , e, por fim, apenas  $x_m$ . A variável que, quando ajustada, gera a maior razão  $t$  absoluta é introduzida no modelo contanto que essa razão ultrapasse a constante especificada  $t_{\text{in}}$ . Suponha que  $x_1$  seja introduzido no modelo. Logo, os modelos com  $(x_1, x_2), (x_1, x_3), \dots, (x_1, x_m)$  são por sua vez considerados. O maior  $|\hat{\beta}_j/s_{\hat{\beta}_j}|$  ( $j = 2, \dots, m$ ) portanto especifica o predictor introduzido contanto que esse valor máximo também exceda  $t_{\text{in}}$ . O processo continua até a etapa em que nenhuma razão  $t$  absoluta exceda  $t_{\text{in}}$ . Os predictores inseridos, portanto, especificam o modelo. O valor  $t_{\text{in}} = 2$ , em geral, é usado pelo mesmo motivo que  $t_{\text{out}} = 2$  é usado na ER. Em relação aos dados sobre o volume de alcatrão, a SP gerou a seqüência de modelos apresentada nas etapas 5, 4, ..., 1 da Tabela 13.11 e, desse modo, está de acordo com a ER. Isso nem sempre ocorre.

O procedimento retroativo mais amplamente usado é uma combinação de SP e ER, chamada de PR. Esse procedimento inicia-se com a seleção progressiva, adicionando variáveis ao modelo, mas depois de adicionar cada variável ele examina as variáveis anteriormente introduzidas para verificar se alguma deve ser eliminada. Por exemplo, se houvesse oito predictores em avaliação e o conjunto atual compreendesse  $x_2, x_3, x_5$  e  $x_6$ , com  $x_5$  tendo acabado de ser adicionado, as razões  $t$   $\hat{\beta}_2/s_{\hat{\beta}_2}, \hat{\beta}_3/s_{\hat{\beta}_3}$  e  $\hat{\beta}_6/s_{\hat{\beta}_6}$  seriam examinadas. Se a menor razão absoluta for menor que  $t_{\text{out}}$ , a variável correspondente é eliminada do modelo. A idéia por trás do procedimento PR é que, com a seleção progressiva, uma única variável pode estar relacionada com  $y$  de uma maneira mais significativa que duas ou mais variáveis individualmente, mas, mais tarde, a combinação dessas variáveis pode tornar a variável única redundante. Isso de fato ocorreu com os dados sobre o consumo de gasolina/milha do Exemplo 13.20, em que  $x_2$  foi introduzido e subseqüentemente extraído do modelo.

O procedimento PR faz parte de todos os softwares de computador-padrão. O software BMDP especifica  $t_{\text{in}} = 2$  e  $t_{\text{out}} = \sqrt{3,9}$  (a maioria dos softwares na verdade usa  $f = t^2$  em vez de  $t$  em si).

Embora, na maioria das situações esses procedimentos de seleção automáticos identifiquem um bom modelo, não há nenhuma garantia de que serão gerados o melhor ou o mais próximo do melhor modelo. É necessário examinar minuciosamente os conjuntos de dados para os quais parecem haver relações significativas entre alguns dos possíveis predictores; falaremos mais sobre isso em breve.

## Identificação de Observações Influentes

Na regressão linear simples, é fácil reconhecer uma observação cujo valor  $x$  é bem maior ou bem menor que outros valores  $x$  na amostra. As observações desse tipo podem exercer grande influência sobre a equação da regressão estimada (se ela de fato depender do grau de consistência do valor  $y$  correspondente em relação aos dados restantes). Na regressão múltipla, é também desejável saber se os valores dos predictores de uma determinada observação são tais que ela tem o potencial de exercer grande influência sobre a equação estimada. Um método para identificar observações potencialmente influentes vale-se da constatação de que, pelo fato de cada  $\hat{\beta}_i$  ser uma função linear de  $y_1, y_2, \dots, y_n$ , cada valor  $y$  previsto na forma  $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_k x_k$

é também uma função linear dos  $y_j$ s. Mais especificamente, os valores previstos correspondentes às observações da amostra podem ser escritos da seguinte forma:

$$\begin{aligned} \hat{y}_1 &= h_{11}y_1 + h_{12}y_2 + \cdots + h_{1n}y_n \\ \hat{y}_2 &= h_{21}y_1 + h_{22}y_2 + \cdots + h_{2n}y_n \\ &\vdots \\ \hat{y}_n &= h_{n1}y_1 + h_{n2}y_2 + \cdots + h_{nn}y_n \end{aligned}$$

Cada coeficiente  $h_{ij}$  é uma função apenas dos  $x_j$ s na amostra, e não dos  $y_j$ s. É possível mostrar que  $h_{ij} = h_{ji}$  e que  $0 \leq h_{ij} \leq 1$ .

Concentremo-nos nos coeficientes da “diagonal”  $h_{11}, h_{22}, \dots, h_{nn}$ . O coeficiente  $h_{jj}$  é o peso dado a  $y_j$  no cálculo do valor  $\hat{y}_j$  previsto correspondente. Essa quantidade pode ainda ser expressa como uma medida da distância entre o ponto  $(x_{1j}, \dots, x_{kj})$  no espaço  $k$ -dimensional e o centro dos dados  $(\bar{x}_1, \dots, \bar{x}_k)$ . Portanto, é natural caracterizar uma observação cujo  $h_{jj}$  é relativamente tão grande quanto um potencialmente influente. A menos que haja uma relação linear perfeita entre os  $k$  preditores,  $\sum_{j=1}^n h_{jj} = k + 1$ , de modo que a média dos  $h_{jj}$ s é  $(k + 1)/n$ . Alguns estatísticos sugerem que, se  $h_{jj} > 2(k + 1)/n$ , a  $j$ -ésima observação será citada como potencialmente influente; outros usam  $3(k + 1)/n$  como a linha divisora.

**Exemplo 13.23**

Os dados a seguir foram apresentados no artigo “Testing for the Inclusion of Variables in Linear Regression by a Randomization Technique” (*Technometrics*, 1966, p. 695–699) e foram novamente analisados por Hoaglin e Welsch, em “The Hat Matrix in Regression and ANOVA” (*Amer. Statistician*, 1978, p. 17–23). Os  $h_{ij}$ s (os elementos abaixo da diagonal foram omitidos por simetria) resultam dos dados.

Número de Vigas	Gravidade Específica ( $x_1$ )		Quantidade de Umidade ( $x_2$ )		Resistência ( $y$ )	
1	0,499		11,1		11,14	
2	0,558		8,9		12,74	
3	0,604		8,8		13,13	
4	0,441		8,9		11,51	
5	0,550		8,8		12,38	
6	0,528		9,9		12,60	
7	0,418		10,7		11,13	
8	0,480		10,5		11,70	
9	0,406		10,5		11,02	
10	0,467		10,7		11,41	

	1	2	3	4	5	6	7	8	9	10
1	0,418	-0,002	0,079	-0,274	-0,046	0,181	0,128	0,222	0,050	0,242
2		0,242	0,292	0,136	0,243	0,128	-0,041	0,033	-0,035	0,004
3			0,417	-0,019	0,273	0,187	-0,126	0,044	-0,153	0,004
4				0,604	0,197	-0,038	0,168	-0,022	0,275	-0,028
5					0,252	0,111	-0,030	0,019	-0,010	-0,010
6						0,148	0,042	0,117	0,012	0,111
7							0,262	0,145	0,277	0,174
8								0,154	0,120	0,168
9									0,315	0,148
10										0,187

Aqui,  $k = 2$ , de modo que  $(k + 1)/n = 3/10 = 0,3$ ; visto que  $h_{44} = 0,604 > 2(0,3)$ , o quarto ponto de dados é identificado como potencialmente influente. ■

Outra técnica para avaliar a influência da  $j$ -ésima observação que leva em conta  $y_j$  e igualmente os valores do previsor envolve a exclusão da  $j$ -ésima observação do conjunto de dados e que se realize uma regressão com base nas observações remanescentes. Se os coeficientes estimados da regressão da “observação excluída” diferirem significativamente das estimativas baseadas nos dados completos, a  $j$ -ésima observação sem dúvida terá exercido um impacto substancial sobre a calibração. Uma maneira de avaliar se os coeficientes estimados mudam significativamente é expressar cada mudança em relação ao desvio padrão estimado do coeficiente:

$$\frac{(\hat{\beta}_i \text{ antes da exclusão}) - (\hat{\beta}_i \text{ após a exclusão})}{s_{\hat{\beta}_i}} = \frac{\text{mudança em } \hat{\beta}_i}{s_{\hat{\beta}_i}}$$

Existem fórmulas computacionais eficientes que permitem que todas essas informações sejam obtidas da regressão em que “não há exclusão”, de modo que as  $n$  regressões adicionais não são necessárias.

**Exemplo 13.24** (Continuação do Exemplo 13.23)

Experimente excluir separadamente as observações 1 e 6, cujos resíduos são os maiores, e a observação 4, onde  $h_{jj}$  é grande. A Tabela 13.12 contém as informações pertinentes.

**Tabela 13.12** Mudanças nos coeficientes estimados do Exemplo 13.24

Parâmetro	Estimativas Sem Exclusão	DP Estimado	Mudança quando o ponto $j$ é excluído		
			$j = 1$	$j = 4$	$j = 6$
$\beta_0$	10,302	1,896	2,710	-2,109	-0,642
$\beta_1$	8,495	1,784	-1,772	1,695	0,748
$\beta_2$	0,2663	0,1273	-0,1932	0,1242	0,0329
		$e_j$ :	-3,25	-0,96	2,20
		$h_{jj}$ :	0,418	0,604	0,148

Para excluir tanto o ponto 1 quanto o ponto 4, a mudança em cada estimativa está no intervalo de 1–1,5 desvios padrão, que é razoavelmente significativa (isso não nos diz o que aconteceria se ambos os pontos fossem omitidos simultaneamente). Entretanto, em relação ao ponto 6, a mudança é de aproximadamente 0,25 desvio padrão. Portanto, os pontos 1 e 4, mas não o 6, poderiam muito bem ser omitidos no cálculo da equação da regressão. ■

### Multicolinearidade

Em vários conjuntos de dados de regressão múltipla, os previsores  $x_1, x_2, \dots, x_k$  são altamente interdependentes. Suponha que consideremos o modelo usual

$$Y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \epsilon$$

com os dados  $(x_{1j}, \dots, x_{kj}, y_j)$  ( $j = 1, \dots, n$ ) disponíveis para calibração. Se usarmos o princípio dos mínimos quadrados para retroceder  $x_i$  sobre os outros previsores  $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_k$ , obtendo

$$\hat{x}_i = a_0 + a_1 x_1 + \dots + a_{i-1} x_{i-1} + a_{i+1} x_{i+1} + \dots + a_k x_k$$

será possível mostrar que

$$V(\hat{\beta}_i) = \frac{\sigma^2}{\sum_{j=1}^n (x_{ij} - \hat{x}_{ij})^2} \tag{13.22}$$

Quando os valores  $x_i$  da amostra puderem ser previstos muito bem com base nos outros valores de previsão, o denominador de (13.22) será pequeno, de modo que  $V(\hat{\beta}_i)$  será bem grande. Se isso ocorrer pelo menos com um previsor, diz-se que os dados exibem **multicolinearidade**. A multicolinearidade em geral é sugerida por um resultado de regressão por computador no qual  $R^2$  é grande, mas algumas das razões  $t \hat{\beta}_i / s_{\hat{\beta}_i}$  são pequenas para

previsores que, com base em informações anteriores e na intuição, parecem importantes. Outro indício de que existe multicolinearidade está relacionado com o valor em  $\hat{\beta}_i$  cujo sinal é oposto ao que a intuição sugeriria, o que indica que outro predictor ou conjunto de preditores está funcionando como “substituto” de  $x_i$ .

Uma avaliação da extensão da multicolinearidade pode ser obtida regressando cada predictor seqüencialmente sobre os  $k - 1$  preditores remanescentes. Seja  $R_i^2$  o valor do  $R^2$  na regressão com a variável dependente  $x_i$  e os preditores  $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_k$ . Diz-se que a multicolinearidade severa está presente quando  $R_i^2 > 0,9$  para qualquer  $i$ . O MINITAB recusa-se a incluir um predictor no modelo quando seu valor  $R_i^2$  é muito próximo de 1.

Infelizmente, não existe consenso entre os estatísticos sobre quais soluções são apropriadas quando a multicolinearidade está presente. Uma possibilidade é continuar a usar um modelo que inclua todos os preditores, mas estimar os parâmetros usando um princípio diferente dos mínimos quadrados. Veja a referência deste capítulo para obter mais detalhes.

## Exercícios | Seção 13.5 (55–64)

**55.** O artigo “Bank Full Discharge of Rivers” (*Water Resources J.*, 1978, p. 1.141–1.154) apresenta os dados a seguir sobre a quantidade de descarga ( $q$ , em  $m^3/s$ ), área de vazão ( $a$ , em  $m^2$ ) e coeficiente angular (declive) da superfície da água ( $b$ , em  $m/m$ ), obtidos em inúmeras estações de planícies de inundação. Segue um subconjunto dos dados. O artigo propôs um modelo potencial multiplicativo  $Q = \alpha a^\beta b^\gamma \epsilon$ .

$q$	17,6	23,8	5,7	3,0	7,5
$a$	8,4	31,6	5,7	1,0	3,3
$b$	0,0048	0,0073	0,0037	0,0412	0,0416
$q$	89,2	60,9	27,5	13,2	12,2
$a$	41,1	26,2	16,4	6,7	9,7
$b$	0,0063	0,0061	0,0036	0,0039	0,0025

- Use uma transformação apropriada para tornar o modelo linear e, em seguida, estime os parâmetros de regressão para o modelo transformado. Por fim, estime  $\alpha$ ,  $\beta$  e  $\gamma$  (os parâmetros do modelo original). Qual seria sua previsão de quantidade de descarga em relação a uma área de vazão de 10 e coeficiente angular de 0,01?
  - Sem de fato realizar nenhuma análise, como você ajustaria um modelo exponencial multiplicativo  $Q = \alpha e^{\beta a} e^{\gamma b} \epsilon$ ?
  - Após a transformação para obter linearidade na parte (a), o IC de 95% obtido de um resultado de computador para o valor da função de regressão transformada em relação a  $a = 3,3$  e  $b = 0,046$  foi (0,217, 1,755). Obtenha um IC de 95% para  $\alpha a^\beta b^\gamma$  quando  $a = 3,3$  e  $b = 0,0046$ .
- 56.** Em um experimento para estudar os fatores que influenciam a gravidade específica da madeira (“Anatomical Factors Influencing Wood Specific Gravity of Slash Pines and the Implications for the Development of a High-Quality Pulpwood”, TAPPI, 1964, p. 401–404), foi obtida uma amostra de 20 espécies de madeira madura. Além disso, foi medido o número de fibras/ $mm^2$  na madeira primaveril ( $x_1$ ), o número de fibras/ $mm^2$  na madeira de verão ( $x_2$ ), % madeira primaveril ( $x_3$ ), absorção de luz pela madeira primaveril ( $x_4$ ) e absorção de luz pela madeira de verão ( $x_5$ ).
- O ajuste da função de regressão  $\mu_{Y \cdot x_1, x_2, x_3, x_4, x_5} = \beta_0 + \beta_1 x_1 + \dots + \beta_5 x_5$  produziu um  $R^2 = 0,769$ . Os dados indicam que existe uma relação linear entre a gravidade específica e pelo menos um dos preditores? Realize um teste, usando  $\alpha = 0,01$ .
  - Quando  $x_2$  é extraído do modelo, o valor de  $R^2$  permanece em 0,769. Calcule o  $R^2$  ajustado para o modelo completo e o modelo em que  $x_2$  foi excluído.
  - Quando  $x_1$ ,  $x_2$  e  $x_4$  são todos excluídos, o valor resultante de  $R^2$  é 0,654. A soma total dos quadrados é  $SQT = 0,0196610$ . Os dados sugerem que todos os  $x_1$ ,  $x_2$  e  $x_4$  têm coeficientes zero no modelo de regressão real? Teste as hipóteses apropriadas no nível 0,05.
  - A média e o desvio padrão de  $x_3$  foram 52,540 e 5,4447, respectivamente, ao passo que os de  $x_5$  eram 89,195 e 3,6660, respectivamente. Quando o modelo que compreendia essas duas variáveis padronizadas foi ajustado, a equação da regressão estimada foi  $y = 0,5255 - 0,0236x'_3 + 0,0097x'_5$ . Que valor da gravidade específica você presumiria em relação a uma amostra com % madeira primaveril = 50 e % absorção de luz na madeira de verão = 90?
  - O desvio padrão estimado do coeficiente estimado  $\hat{\beta}_3$  de  $x'_3$  (isto é, para  $\hat{\beta}_3$  do modelo padronizado) era 0,0046. Obtenha um IC de 95% para  $\beta_3$ .
  - Usando as informações das partes (d) e (e), qual é o coeficiente estimado de  $x_3$  no modelo não padronizado (usando apenas os preditores  $x_3$  e  $x_5$ ) e qual é o desvio padrão estimado do estimador de coeficientes (isto é,  $s_{\hat{\beta}_3}$  para  $\hat{\beta}_3$  no modelo não padronizado)?
  - A estimativa de  $\sigma$  para o modelo de dois preditores é  $s = 0,02001$ , ao passo que o desvio padrão estimado de  $\hat{\beta}_0 + \hat{\beta}_3 x'_3 + \hat{\beta}_5 x'_5$  quando  $x'_3 = -0,3747$  e  $x'_5 = -0,2769$  (isto é, quando  $x_3 = 50,5$  e  $x_5 = 88,9$ ) é 0,00482. Calcule um IP de 95% para a gravidade específica em relação a % madeira primaveril = 50,5 e % absorção de luz na madeira de verão = 88,9.

57. Na tabela a seguir, apresentamos a menor SQE para cada um dos números de previsores  $k$  ( $k = 1, 2, 3, 4$ ) de um problema de regressão no qual  $y =$  aquecimento cumulativo para o endurecimento do cimento,  $x_1 =$  % aluminato tricálcio,  $x_2 =$  % silicato tricálcio,  $x_3 =$  % ferrato de alumínio e  $x_4 =$  % silicato dicálcio.

Número de Previsores $k$	Previsor(es)	SQE
1	$x_4$	880,85
2	$x_1, x_2$	58,01
3	$x_1, x_2, x_3$	49,20
4	$x_1, x_2, x_3, x_4$	47,86

Além disso,  $n = 13$  e  $SQT = 2.715,76$ .

- Use os critérios discutidos no texto para indicar um determinado modelo de regressão.
- A seleção progressiva resultaria no melhor modelo de dois previsores? Explique.

58. O artigo “Creep and Fatigue Characteristics of Ferrocement Slabs” (J. *Ferrocement*, 1984, p. 309–322) relatou os dados sobre  $y =$  resistência à tração (MPa),  $x_1 =$  espessura da laje (cm),  $x_2 =$  carga (kg),  $x_3 =$  período sob carga (dias) e  $x_4 =$  duração do teste (dias) resultantes de testes de tensão de  $n = 9$  lajes de concreto reforçado. Os resultados da aplicação do método de eliminação ER de seleção de variáveis estão resumidos no formato tabular a seguir. Explique o que ocorreu em cada etapa do procedimento.

Etapa Constante	1 8,496	2 12,670	3 12,989
$x_1$	-0,29	-0,42	-0,49
RAZÃO T	-1,33	-2,89	-3,14
$x_2$	0,0104	0,0110	0,0116
RAZÃO T	6,30	7,40	7,33
$x_3$	0,0059		
RAZÃO T	0,83		
$x_4$	-0,023	-0,023	
RAZÃO T	-1,48	-1,53	
S	0,533	0,516	0,570
R <sup>2</sup>	95,81	95,10	92,82

59. A opção Melhor Regressão do MINITAB foi usada nos dados sobre a gravidade específica da madeira apresentados no Exercício 56, gerando o resultado a seguir. Na sua opinião, que modelo(s) deveriam ser investigados mais detalhadamente?

A resposta é gravidade da madeira primaveril

Variáveis	R <sup>2</sup>	R <sup>2</sup> (ajustado)	C-p	s	fibramadeiraprimaveril	fibramadeiradeverão	%madeiraprimaveril	orelhasmadprimav.	orelhasmad.verão
1	56,4	53,9	10,6	0,021832			X		
1	10,6	5,7	38,5	0,031245				X	
1	5,3	0,1	41,7	0,032155	X				
2	65,5	61,4	7,0	0,019975		X		X	
2	62,1	57,6	9,1	0,020950	X	X			

(Continua na próxima coluna)

2	60,3	55,6	10,2	0,021439	X	X			
3	72,3	67,1	4,9	0,018461	X	X	X		
3	71,2	65,8	5,6	0,018807			X	X	X
3	71,1	65,7	5,6	0,018846			X	X	X
4	77,0	70,9	4,0	0,017353	X	X	X	X	
4	74,8	68,1	5,4	0,018179		X	X	X	X
4	72,7	65,4	6,7	0,018919	X	X	X	X	X
5	77,0	68,9	6,0	0,017953	X	X	X	X	X

60. Os dados a seguir processados no MINITAB resultaram da aplicação dos métodos de eliminação retroativa e de seleção progressiva aos dados sobre a gravidade específica da madeira apresentados no Exercício 56. Para cada método, explique o que ocorreu a cada iteração do algoritmo.

A resposta é gravidade da madeira primaveril sob 5 previsores, com N = 20

Etapa Constante	1 0,4421	2 0,4384	3 0,4381	4 0,5179
fibramadeiraprimaveril	0,00011	0,00011	0,00012	
Valor total	1,17	1,95	1,98	
fibramadeiradeverão	0,00001			
Valor total	0,12			
%madeiraprimaveril	-0,00531	-0,00526	-0,00498	-0,00438
Valor total	-5,70	-6,56	-5,96	-5,20
orelhasmadprimav	-0,0018	-0,0019		
Valor total	-1,63	-1,76		
orelhasmadverão	0,0044	0,0044	0,0031	0,0027
Valor total	3,01	3,31	2,63	2,12
S	0,0180	0,0174	0,0185	0,0200
R <sup>2</sup>	77,05	77,03	72,27	65,50
Etapa Constante	1 0,7585	2 0,5179		
%madeiraprimaveril	-0,00444	-0,00438		
Valor total	-4,82	-5,20		
orelhasmadverão		0,0027		
Valor total		2,12		
S	0,0218	0,0200		
R <sup>2</sup>	56,36	65,50		

61. Reveja os dados sobre a gravidade específica da madeira apresentados no Exercício 56. Os valores de  $R^2$  a seguir resultaram da regressão de cada previsor sobre os outros quatro previsores (na primeira regressão, a variável dependente foi  $x_1$  e os previsores foram  $x_2-x_5$  etc.): 0,628, 0,711, 0,341, 0,403 e 0,403. A multicolinearidade parece ser um problema significativo? Explique.

62. Um estudo realizado para investigar a relação entre a variável de resposta relativa a quedas de pressão em uma coluna de bolhas de uma chapa térmica e os previsores  $x_1 =$  velocidade do fluido superficial,  $x_2 =$  viscosidade do líquido e  $x_3 =$  tamanho da abertura da malha gerou os dados a seguir (“A Correlation of TwoPhase Pressure Drops in Screen-Plate Bubble Column”, *Canad. J. of Chem. Engr.*, 1993, p. 460–463). Os resíduos padronizados e os valores  $h_{ii}$  resultaram do modelo com apenas  $x_1, x_2$  e  $x_3$  como previsores. Existe alguma observação incomum?

## Dados do Exercício 62

Observação	Velocidade	Viscosidade	Tamanho da malha	Resposta	Resíduo Padronizado	$h_{ii}$
1	2,14	10,00	0,34	28,9	2,01721	0,202242
2	4,14	10,00	0,34	26,1	1,34706	0,066929
3	8,15	10,00	0,34	22,8	0,96537	0,274393
4	2,14	2,63	0,34	24,2	1,29177	0,224518
5	4,14	2,63	0,34	15,7	-0,68311	0,079651
6	8,15	2,63	0,34	18,3	0,23785	0,267959
7	5,60	1,25	0,34	18,1	0,06456	0,076001
8	4,30	2,63	0,34	19,1	0,13131	0,074927
9	4,30	2,63	0,34	15,4	-0,74091	0,074927
10	5,60	10,10	0,25	12,0	-1,38857	0,152317
11	5,60	10,10	0,34	19,8	-0,03585	0,068468
12	4,30	10,10	0,34	18,6	-0,40699	0,062849
13	2,40	10,10	0,34	13,2	-1,92274	0,175421
14	5,60	10,10	0,55	22,8	-1,07990	0,712933
15	2,14	112,00	0,34	41,8	-1,19311	0,516298
16	4,14	112,00	0,34	48,6	1,21302	0,513214
17	5,60	10,10	0,25	19,2	0,38451	0,152317
18	5,60	10,10	0,25	18,4	0,18750	0,152317
19	5,60	10,10	0,25	15,0	-0,64979	0,152317

63. O resultado da regressão múltipla do MINITAB para os dados sobre HPA do Exercício 53 na seção anterior incluiu as seguintes informações:

Observações incomuns

Obs	x1	flth	Aderência	Aderência EP	Resíduo	Resíduo-padrão
6	243500	604,7	582,9	40,7	21,8	1,25X
7	67793	27,7	139,3	12,3	-111,6	-2,62R

R representa uma observação com um grande resíduo padronizado

X representa uma observação cujo valor X exerce sobre ela grande influência

O que isso sugere sobre a aplicabilidade de se usar a equação ajustada apresentada anteriormente como base para inferências? Os pesquisadores na verdade excluíram a observação 7 e realizaram novamente a regressão. Isso faz sentido?

64. Veja os dados sobre a descarga de água apresentados no Exercício 55 e considere  $y = \ln(q)$ ,  $x_1 = \ln(a)$  e  $x_2 =$

$\ln(b)$ . Experimente ajustar o modelo  $Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \epsilon$ .

a. Os  $h_{ii}$ s resultantes são 0,138, 0,302, 0,266, 0,604, 0,464, 0,360, 0,215, 0,153, 0,214 e 0,284. Alguma observação parece influente?

b. Os coeficientes estimados são  $\hat{\beta}_0 = 1,5652$ ,  $\hat{\beta}_1 = 0,9450$ ,  $\hat{\beta}_2 = 0,1815$ , e os desvios padrão estimados correspondentes são  $s_{\hat{\beta}_0} = 0,7328$ ,  $s_{\hat{\beta}_1} = 0,1528$  e  $s_{\hat{\beta}_2} = 0,1752$ . O segundo resíduo padronizado é  $e_2^* = 2,19$ . Quando a segunda observação é omitida do conjunto de dados, os coeficientes estimados resultantes são  $\hat{\beta}_0 = 1,8982$ ,  $\hat{\beta}_1 = 1,025$  e  $\hat{\beta}_2 = 0,308$ . Alguma dessas mudanças indica que a segunda observação é influente?

c. Ao excluir a quarta observação (por quê?), obtemos  $\hat{\beta}_0 = 1,4592$ ,  $\hat{\beta}_1 = 0,9850$ , e  $\hat{\beta}_2 = 0,1515$ . Essa observação é influente?

## Exercícios Suplementares (65–82)

65. Sabe-se que o tratamento do concreto é vulnerável a vibrações por choque, que podem provocar rachaduras ou danos imperceptíveis no material. Como parte de um estudo sobre os fenômenos da vibração, o artigo “Shock Vibration Test of Concrete” (*ACI Materials J.*, 2002, p. 361–370) relatou os dados a seguir sobre a velocidade de pico da partícula (mm/s) e razão da velocidade do pulso ultra-sônico antes e após um impacto em prismas de concreto.

As rachaduras transversais apareceram nos últimos 12 prismas, mas não se observou nenhuma rachadura nos primeiros 18 prismas.

a. Elabore um boxplot comparativo da vpp para os prismas que sofreram e não sofreram rachaduras e comente. Em seguida, avalie a diferença entre a vpp média real para os prismas com e sem rachaduras de maneira que transmita informações sobre precisão e confiabilidade.

b. Os pesquisadores ajustaram o modelo de regressão linear simples ao conjunto de dados de 30 observações, sendo a vpp a variável independente e a razão a variável dependente. Use um software estatístico para ajustar vários modelos de regressão diferentes e faça inferências apropriadas.

Obs.	vpp	Razão	Obs.	vpp	Razão
1	160	0,996	16	708	0,990
2	164	0,996	17	806	0,984
3	178	0,999	18	884	0,986
4	252	0,997	19	526	0,991
5	293	0,993	20	490	0,993
6	289	0,997	21	598	0,993
7	415	0,999	22	505	0,993
8	478	0,997	23	525	0,990
9	391	0,992	24	675	0,991
10	486	0,985	25	1.211	0,981
11	604	0,995	26	1.036	0,986
12	528	0,995	27	1.000	0,984
13	749	0,994	28	1.151	0,982
14	772	0,994	29	1.144	0,962
15	532	0,987	30	1.068	0,986

66. Os autores do artigo “Long-Term Effects of Cathodic Protection on Prestressed Concrete Structures” (*Corrosion*, 1997, p. 891–908) apresentou um gráfico de dispersão de  $y =$  fluxo de permeação em estado constante ( $\mu\text{A}/\text{cm}^2$ ) versus  $x =$  espessura da lâmina invertida ( $\text{cm}^{-1}$ ); foi usado como base um padrão linear significativo para que se pudesse obter uma conclusão considerável sobre o comportamento do material. A seguir é apresentado o resultado do MINITAB da calibração do modelo de regressão linear simples.

A equação de regressão é  
fluxo = -0,398 + 0,260 lâmina invertida

Previsor	Coef	DP	razão t	p
Constante	-0,3982	0,5051	-0,79	0,460
lâmina invertida	0,26042	0,01502	17,34	0,000

$s = 0,4506$   $R^2 = 98,0\%$   $R^2(\text{ajustado}) = 97,7\%$

Análise de Variância

Fonte	DF	SQ	QM	F	p
Regressão	1	61,050	61,050	300,64	0,000
Erro	6	1,218	0,203		
Total	7	62,269			

Obs.	Lâmina invertida		Aderência		Resíduo	
	fluxo	Aderência	DP	Resíduo	Padrão	
1	19,8	4,3	4,758	0,242	-0,458	-1,20
2	20,6	5,6	4,966	0,233	0,634	1,64
3	23,5	6,1	5,722	0,203	0,378	0,94
4	26,1	6,2	6,399	0,182	-0,199	-0,48
5	30,3	6,9	7,493	0,161	-0,593	-1,41
6	43,5	11,2	10,930	0,236	0,270	0,70
7	45,0	11,3	11,321	0,253	-0,021	-0,06
8	46,5	11,7	11,711	0,271	-0,011	-0,03

- a. Interprete o coeficiente angular estimado e o coeficiente de determinação.
- b. Calcule uma estimativa pontual do fluxo médio real em relação a uma lâmina invertida com espessura de 23,5.
- c. O modelo parece útil?
- d. Presuma o fluxo em relação a uma espessura invertida de 45 de maneira que transmita informações sobre precisão e confiabilidade.
- e. Investigue a adequação do modelo.

67. O artigo “Validation of the Rockport Fitness Walking Test in College Males and Females” (*Research Quar-*

*terly for Exercise and Sport*, 1994, p. 152–158) sugeriu a equação de regressão estimada a seguir para relacionar  $y = \text{VO}_2 \text{ máx. (L/min., medida da capacidade cardiorrespiratória)}$  para os previsores  $x_1 =$  gênero (feminino = 0, masculino = 1),  $x_2 =$  peso (lb),  $x_3 =$  tempo de caminhada de 1 milha (min.) e  $x_4 =$  taxa de batimentos cardíacos no fim da caminhada (batimentos/min.):

$$y = 3,5959 + 0,6566x_1 + 0,0096x_2 - 0,0996x_3 - 0,0080x_4$$

- a. Como você interpretaria o coeficiente estimado  $\hat{\beta}_3 = -0,0996$ ?
- b. Como você interpretaria o coeficiente estimado  $\hat{\beta}_1 = 0,6566$ ?
- c. Suponha que uma observação sobre um indivíduo do sexo masculino com peso de 170 lb, tempo de caminhada de 11 min. e taxa de batimentos cardíacos de 140 batimentos/min. resultou em  $\text{VO}_2 \text{ máx.} = 3,15$ . O que você presumiria de  $\text{VO}_2 \text{ máx.}$  nessa situação e qual é o valor do resíduo correspondente?
- d. Usando  $\text{SQE} = 30,1033$  e  $\text{SQT} = 102,3922$ , que proporção da variação observada em  $\text{VO}_2 \text{ máx.}$  pode ser atribuída à relação do modelo?
- e. Adotando um tamanho de amostra de  $n = 20$ , realize um teste de hipóteses para determinar se o modelo escolhido especifica uma relação útil entre  $\text{VO}_2 \text{ máx.}$  e pelo menos um dos previsores.

68. O reconhecimento de características dos modelos de superfície de peças complexas está se tornando cada vez mais importante no desenvolvimento de sistemas de projeto auxiliado por computador (computer-aided design—CAD) eficientes. O artigo “A Computationally Efficient Approach to Feature Abstraction in Design-Manufacturing Integration” (*J. of Engr. for Industry*, 1995, p. 16–27) contém um gráfico de  $\log_{10}$  (tempo total de reconhecimento), em que o tempo é apresentado em s, versus  $\log_{10}$  (número de bordas de uma peça), do qual foram reproduzidos os seguintes dados representativos:

<b>Log(bordas)</b>	1,1	1,5	1,7	1,9	2,0	2,1
<b>Log(tempo)</b>	0,30	0,50	0,55	0,52	0,85	0,98
<b>Log(bordas)</b>	2,2	2,3	2,7	2,8	3,0	3,3
<b>Log(tempo)</b>	1,10	1,00	1,18	1,45	1,65	1,84
<b>Log(bordas)</b>	3,5	3,8	4,2	4,3		
<b>Log(tempo)</b>	2,05	2,46	2,50	2,76		

- a. Um gráfico de dispersão de  $\log(\text{tempo})$  versus  $\log(\text{bordas})$  indica uma relação linear aproximada entre essas duas variáveis?
- b. Que modelo probabilístico para relacionar  $y =$  tempo de reconhecimento com  $x =$  número de bordas é subentendido pela relação de regressão linear simples entre as variáveis transformadas?
- c. As quantidades calculadas com base nos dados, são

$$n = 16 \quad \sum x'_i = 42,4 \quad \sum y'_i = 21,69$$

$$\sum (x'_i)^2 = 126,34 \quad \sum (y'_i)^2 = 38,5305$$

$$\sum x'_i y'_i = 68,640$$



Calcule as estimativas de parâmetro do modelo na parte (b) e, em seguida, obtenha uma previsão pontual do tempo em relação a um número de bordas de 300.

69. Foram medidas a pressão atmosférica (psi) e a temperatura (°F) de um processo de compressão em um determinado dispositivo de cilindro com êmbolo, gerando os dados a seguir (de *Introduction to Engineering Experimentation*, Prentice-Hall, Inc., 1996, p. 153):

<b>Pressão</b>	20,0	40,4	60,8	80,2	100,4
<b>Temperatura</b>	44,9	102,4	142,3	164,8	192,2
<b>Pressão</b>	120,3	141,1	161,4	181,9	201,4
<b>Temperatura</b>	221,4	228,4	249,5	269,4	270,8
<b>Pressão</b>	220,8	241,8	261,1	280,4	300,1
<b>Temperatura</b>	291,5	287,3	313,3	322,3	325,8
<b>Pressão</b>	320,6	341,1	360,8		
<b>Temperatura</b>	337,0	332,6	342,9		

- a. Você ajustaria o modelo de regressão linear simples aos dados e o usaria como base para prever a temperatura de acordo com a pressão? Por quê ou por quê não?
- b. Encontre um modelo probabilístico adequado e use-o para prever o valor da temperatura resultante de uma pressão de 200 da maneira mais informativa possível.
70. Um aluno de engenharia aeronáutica realizou um experimento para investigar como  $y$  = coeficiente de sustentação/arrasto está relacionado com as variáveis  $x_1$  = posição de uma determinada superfície de sustentação dianteira em relação à asa principal e  $x_2$  = disposição da cauda em relação à asa principal, obtendo os dados a seguir (*Statistics for Engineering Problem Solving*, p. 133):

$x_1$ (pol.)	$x_2$ (pol.)	$y$
-1,2	-1,2	0,858
-1,2	0	3,156
-1,2	1,2	3,644
0	-1,2	4,281
0	0	3,481
0	1,2	3,918
1,2	-1,2	4,136
1,2	0	3,364
1,2	1,2	4,018

$\bar{y} = 3,428, SQT = 8,55$

- a. Ajustando-se o modelo de primeira ordem, obtém-se  $SQE = 5,18$ , ao passo que incluindo  $x_3 = x_1x_2$  como predictor obtém-se  $SQE = 3,07$ . Calcule e interprete o coeficiente de determinação múltipla de cada modelo.
- b. Realize um teste de utilidade do modelo usando  $\alpha = 0,05$  para cada um dos modelos descritos na parte (a). Por acaso algum dos resultados surpreendeu você?
71. Um banho de amônia é um dos mais amplamente usados para sedimentar revestimentos de liga Pd-Ni. O artigo “Modelling of Palladium and Nickel in an Am-

monia Bath in a Rotary Device” (*Plating and Surface Finishing*, 1997, p. 102–104) apresentou um estudo sobre como as características de composição do banho influenciam nas propriedades do revestimento. Considere os dados a seguir sobre  $x_1$  = concentração de Pd (g/dm<sup>3</sup>),  $x_2$  = concentração de Ni (g/dm<sup>3</sup>),  $x_3$  = pH,  $x_4$  = temperatura (°C),  $x_5$  = densidade de corrente catódica (A/dm<sup>2</sup>) e  $y$  = conteúdo de paládio (%) do revestimento.

Obs	concpd	concni	pH	temp	denscor	contpdd
1	16	24	9,0	35	5	61,5
2	8	24	9,0	35	3	51,0
3	16	16	9,0	35	3	81,0
4	8	16	9,0	35	5	50,9
5	16	24	8,0	35	3	66,7
6	8	24	8,0	35	5	48,8
7	16	16	8,0	35	5	71,3
8	8	16	8,0	35	3	62,8
9	16	24	9,0	25	3	64,0
10	8	24	9,0	25	5	37,7
11	16	16	9,0	25	5	68,7
12	8	16	9,0	25	3	54,1
13	16	24	8,0	25	5	61,6
14	8	24	8,0	25	3	48,0
15	16	16	8,0	25	3	73,2
16	8	16	8,0	25	5	43,3
17	4	20	8,5	30	4	35,0
18	20	20	8,5	30	4	69,6
19	12	12	8,5	30	4	70,0
20	12	28	8,5	30	4	48,2
21	12	20	7,5	30	4	56,0
22	12	20	9,5	30	4	77,6
23	12	20	8,5	20	4	55,0
24	12	20	8,5	40	4	60,6
25	12	20	8,5	30	2	54,9
26	12	20	8,5	30	6	49,8
27	12	20	8,5	30	4	54,1
28	12	20	8,5	30	4	61,2
29	12	20	8,5	30	4	52,5
30	12	20	8,5	30	4	57,1
31	12	20	8,5	30	4	52,5
32	12	20	8,5	30	4	56,6

- a. Ajuste o modelo de primeira ordem aos cinco previsores e avalie sua utilidade. Todos os previsores parecem importantes?
- b. Ajuste o modelo completo de segunda ordem e avalie sua utilidade.
- c. O grupo de previsores de segunda ordem (de interação e quadráticos) parecem fornecer informações mais úteis sobre  $y$  do que os previsores de primeira ordem? Realize um teste apropriado das hipóteses.
- d. Os autores do artigo citado sugeriram que fossem usados todos os cinco previsores de primeira ordem mais o predictor extra  $x_6 = (\text{pH})^2$ . Ajuste esse modelo. Todos os seis previsores parecem importantes?
72. O artigo “An Experimental Study of Resistance Spot Welding in 1 mm Thick Sheet of Low Carbon Steel” (*J. of Engr. Manufacture*, 1996, p. 341–348) abordou uma análise estatística cujo principal objetivo era estabelecer uma relação que pudesse explicar a variação na resistência à solda ( $y$ ) relacionando a resistência às seguintes características do processo: corrente de solda (cs), tempo de solda (ts) e força do eletrodo (fe).

- a.  $SQT = 16,18555$ . Ajustando o modelo completo de segunda ordem, obtém-se  $SQE = 0,80017$ . Calcule e interprete o coeficiente de determinação múltipla.
- b. Supondo que  $n = 37$ , realize um teste de utilidade do modelo (a tabela ANOVA do artigo expressa que  $n - (k + 1) = 1$ , mas outras informações fornecidas contradizem essa afirmação e estão consistentes com o tamanho da amostra que sugerimos).
- c. A razão  $F$  fornecida para a interação corrente e tempo foi 2,32. Se todos os previsores forem mantidos no modelo, esse predictor de interação poderá ser eliminado? *Sugestão:* Assim como na regressão linear simples, uma razão  $F$  para um coeficiente é o quadrado de suas razões  $t$ .
- d. Os autores propuseram eliminar dois previsores de interação e um predictor quadrático, e sugeriram a equação estimada  $y = 3,352 + 0,098cs + 0,222ts + 0,297fe - 0,0102(ts)^2 - 0,037(fe)^2 + 0,0128(cs)(ts)$ . Considere uma corrente de solda de 10 kA, um tempo de solda de 12 ciclos ca e uma força de eletrodo de 6 kN. Supondo que o desvio padrão estimado da resistência prevista nessa situação seja 0,0750, calcule um IP de 95% para a resistência. O intervalo sugere que o valor da resistência pode ser previsto com precisão?

73. Os dados a seguir sobre  $x =$  frequência (MHz) e  $y =$  energia de saída (W) de uma determinada configuração de laser foram reproduzidos de um gráfico apresentado no artigo “Frequency Dependence in RF Discharge Excited Waveguide CO<sub>2</sub> Lasers” (*IEEE J. Quantum Electronics*, 1984, p. 509–514).

$x$	60	63	77	100	125	157	186	222
$y$	16	17	19	21	22	20	15	5

Uma análise realizada por computador forneceu as informações a seguir para um modelo de regressão quadrática:  $\hat{\beta}_0 = -1,5127, \hat{\beta}_1 = 0,391901, \hat{\beta}_2 = -0,00163141, s_{\hat{\beta}_2} = 0,00003391, SQE = 0,29, SQT = 202,88$  e  $s_{\hat{y}} = 0,1141$  quando  $x = 100$ .

- a. O modelo quadrático parece adequado para explicar a variação observada na saída de energia relacionando-a à frequência?
  - b. O modelo de regressão linear simples seria quase tão satisfatório quanto o modelo quadrático?
  - c. Na sua opinião, valeria a pena avaliar um modelo cúbico?
  - d. Calcule um IC de 95% para a saída de energia esperada em relação a uma frequência de 100.
  - e. Use um IP de 95% para prever a energia de uma única operação experimental em relação a uma frequência de 100.
74. A condutividade é uma importante característica do vidro. O artigo “Structure and Properties of Rapidly Quenched Li<sub>2</sub>O–Al<sub>2</sub>O<sub>3</sub>–Nb<sub>2</sub>O<sub>5</sub> Glasses” (*J. Amer Ceramic Soc.*, 1983, p. 890–892) descreve os dados a seguir sobre  $x =$  conteúdo de Li<sub>2</sub>O de um determinado tipo de vidro e  $y =$  condutividade a 500 K.

$x$	19	20	24	27	29	30
$y$	$10^{-8,0}$	$10^{-7,1}$	$10^{-7,2}$	$10^{-6,7}$	$10^{-6,2}$	$10^{-6,8}$
$x$	31	39	40	43	45	50
$y$	$10^{-5,8}$	$10^{-5,3}$	$10^{-6,0}$	$10^{-4,7}$	$10^{-5,4}$	$10^{-5,1}$

(Este é um subconjunto dos dados apresentados no artigo.) Proponha um modelo adequado para relacionar  $y$  com  $x$ , estime os parâmetros do modelo e presuma a condutividade em relação a um conteúdo de Li<sub>2</sub>O igual a 35.

75. O efeito do manganês (Mn) sobre o crescimento do trigo é examinado no artigo “Manganese Deficiency and Toxicity Effects on Growth, Development and Nutrient Composition in Wheat” (*Agronomy J.*, 1984, p. 213–217). Um modelo de regressão quadrática foi usado para relacionar  $y =$  altura da muda (cm) com  $x = \log_{10}$  (Mn adicionado), sendo  $\mu M$  as unidades de Mn adicionado. Os dados a seguir foram reproduzidos de um gráfico de dispersão apresentado nesse artigo:

$x$	-1,0	-0,4	0	0,2	1,0
$y$	32	37	44	45	46
$x$	2,0	2,8	3,2	3,4	4,0
$y$	42	42	40	37	30

Além disso,  $\hat{\beta}_0 = 41,7422, \hat{\beta}_1 = 6,581, \hat{\beta}_2 = -2,3621, s_{\hat{\beta}_0} = 0,8522, s_{\hat{\beta}_1} = 1,002, s_{\hat{\beta}_2} = 0,3073$ , e  $SQE = 26,98$ .

- a. O modelo quadrático é útil para descrever a relação entre  $x$  e  $y$ ? [*Sugestão:* A regressão quadrática é um caso especial de regressão múltipla com  $k = 2, x_1 = x$  e  $x_2 = x^2$ .] Aplique um procedimento apropriado.
  - b. O predictor quadrático deve ser eliminado?
  - c. Estime a altura esperada do trigo tratado com 10  $\mu M$  de Mn usando um IC de 90%. [*Sugestão:* O desvio padrão estimado de  $\beta_0 + \beta_1 + \beta_2$  é 1,031.]
76. O artigo “Chemithermomechanical Pulp from Mixed High Density Hardwoods” (TAPPI, julho de 1988, p. 145–146) relata um estudo no qual foram obtidos os dados a seguir para relacionar  $y =$  área de superfície específica (cm<sup>3</sup>/g) com  $x_1 =$  % de NaOH usada como tratamento químico prévio e  $x_2 =$  tempo de tratamento (min.) para um lote de polpa.

$x_1$	$x_2$	$y$
3	30	5,95
3	60	5,60
3	90	5,44
9	30	6,22
9	60	5,85
9	90	5,61
15	30	8,36
15	60	7,30
15	90	6,43

Os resultados do MINITAB a seguir foram obtidos de uma solicitação de ajuste do modelo  $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$ .

A equação de regressão é  
 ÁREA = 6,05 + 0,142 NAOH - 0,0169 TEMPO

Previsor	Coef	DP	razão t	p
Constante	6,0483	0,5208	11,61	0,000
NAOH	0,14167	0,03301	4,29	0,005
TEMPO	-0,016944	0,006601	-2,57	0,043

s = 0,4851 R<sup>2</sup> = 80,7% R<sup>2</sup> (ajustado) 74,2%

Análise de Variância

FONTE	GL	SQ	QM	F	p
Regressão	2	5,8854	2,9427	12,51	0,007
Erro	6	1,4118	0,2353		
Total	8	7,2972			

- a. Que proporção de variação observada na área de superfície específica pode ser explicada pela relação do modelo?
- b. O modelo escolhido parece especificar uma relação útil entre a variável dependente e os previsores?
- c. Se % de NaOH permanecesse no modelo, você sugeriria que o previsor tempo de tratamento fosse eliminado?
- d. Calcule um IC de 95% para a mudança esperada na área de superfície específica associada a um aumento de 1% de NaOH, mantendo o tempo de tratamento fixo.
- e. O MINITAB informou que o desvio padrão estimado de  $\hat{\beta}_0 + \hat{\beta}_1(9) + \hat{\beta}_2$  é 0,162. Calcule um intervalo de previsão para o valor da área de superfície específica a ser observado quando % NaOH = 9 e tempo de tratamento = 60.

77. Uma análise de regressão múltipla foi realizada para relacionar y = resistência à tração de uma amostra de fibra sintética com as variáveis x<sub>1</sub> = porcentagem de algodão e x<sub>2</sub> = tempo de secagem. O conjunto de dados compreendia n = 12 observações.

- a. Os coeficientes estimados eram  $\hat{\beta}_0 = 180,00$ ,  $\hat{\beta}_1 = 1,000$  e  $\hat{\beta}_2 = 10,500$ . Calcule uma estimativa pontual da resistência à tensão esperada quando a porcentagem de algodão = 15 e o tempo de secagem = 3,5.
- b. As somas dos quadrados foram SQT = 1.210,30 e SQE = 117,40. Que proporção da variação observada na resistência à tensão pode ser atribuída à relação do modelo?
- c. Use as informações na parte (b) para determinar se o modelo com as variáveis x<sub>1</sub> e x<sub>2</sub> especifica uma relação útil.
- d. O desvio padrão estimado de  $\hat{\beta}_0 + \hat{\beta}_1x_1 + \hat{\beta}_2x_2$  quando x<sub>1</sub> = 18 e x<sub>2</sub> = 3,0 foi 1,20. Calcule um IP de 95% para a resistência à tensão de uma amostra de tecido para a qual x<sub>1</sub> = 18 e x<sub>2</sub> = 3,0.

78. Um estudo foi realizado para relacionar o tempo até a falha (y) de um determinado componente de uma máquina com as variáveis tensão operacional (x<sub>1</sub>), velocidade do motor (x<sub>2</sub>) e temperatura operacional (x<sub>3</sub>). O conjunto de dados resultante compreendia n = 20 observações. Quando o modelo com as três variáveis x<sub>1</sub>, x<sub>2</sub> e x<sub>3</sub> foi ajustado aos dados, o valor da soma dos quadrados do erro foi 8.212,5. Ajustando o modelo de

interação de segunda ordem (com previsores quadráticos e todos os produtos dos pares de variáveis), obteve-se uma soma de quadrados do erro de 5.027,1. Pelo menos um dos previsores quadráticos ou de interação deveria ser mantido no modelo? Expresse e teste as hipóteses relevantes.

79. O artigo “A Statistical Analysis of the Notch Toughness of 9% Nickel Steels Obtained from Production Heats” (*J. of Testing and Eval.*, 1987, p. 355–363) relatou os resultados de uma análise de regressão múltipla relacionando y = rigidez Charpy de um entalhe em v (joules) com as seguintes variáveis: x<sub>1</sub> = espessura da placa (mm), x<sub>2</sub> = conteúdo de carbono (%), x<sub>3</sub> = conteúdo de manganês (%), x<sub>4</sub> = conteúdo de fósforo (%), x<sub>5</sub> = conteúdo de enxofre (%), x<sub>6</sub> = conteúdo de silício (%), x<sub>7</sub> = conteúdo de níquel (%), x<sub>8</sub> = intensidade de resistência (Pa) e x<sub>9</sub> = resistência à tensão (Pa).

a. Os melhores subconjuntos possíveis exigiram que as variáveis fossem adicionadas na seguinte ordem: x<sub>5</sub>, x<sub>8</sub>, x<sub>6</sub>, x<sub>3</sub>, x<sub>2</sub>, x<sub>7</sub>, x<sub>9</sub>, x<sub>1</sub> e x<sub>4</sub>. Os valores de R<sub>k</sub><sup>2</sup>, QME<sub>k</sub> e C<sub>k</sub> são os seguintes:

Nº de Previsores	1	2	3	4
R <sub>k</sub> <sup>2</sup>	0,354	0,453	0,511	0,550
QME <sub>k</sub>	2.295	1.948	1.742	1.607
C <sub>k</sub>	314	173	89,6	35,7

Nº de Previsores	5	6	7	8	9
R <sub>k</sub> <sup>2</sup>	0,562	0,570	0,572	0,575	0,575
QME <sub>k</sub>	1.566	1.541	1.535	1.530	1.532
C <sub>k</sub>	19,9	11,0	9,4	8,2	10,0

Que modelo você recomendaria? Explique o raciocínio lógico por trás de sua escolha.

b. Os autores também avaliaram modelos de segunda ordem compreendendo os previsores x<sub>i</sub><sup>2</sup> e x<sub>i</sub>x<sub>j</sub>. As informações sobre o melhor dentre esses modelos, começando com as variáveis x<sub>2</sub>, x<sub>3</sub>, x<sub>5</sub>, x<sub>6</sub>, x<sub>7</sub> e x<sub>8</sub>, são as seguintes (ao passar do melhor modelo de quatro previsores para o melhor modelo de cinco previsores, x<sub>8</sub> foi excluída e tanto x<sub>2</sub>x<sub>6</sub> quanto x<sub>7</sub>x<sub>8</sub> foram introduzidas; x<sub>8</sub> foi novamente introduzida em um estágio posterior):

Nº de Previsores	1	2	3	4	5
R <sub>k</sub> <sup>2</sup>	0,415	0,541	0,600	0,629	0,650
QME <sub>k</sub>	2.079	1.636	1.427	1.324	1.251
C <sub>k</sub>	433	109	104	52,4	16,5

Nº de Previsores	6	7	8	9	10
R <sub>k</sub> <sup>2</sup>	0,652	0,655	0,658	0,659	0,659
QME <sub>k</sub>	1.246	1.237	1.229	1.229	1.230
C <sub>k</sub>	14,9	11,2	8,5	9,2	11,0

Qual desses modelos você recomendaria e por quê? (*Observação:* Os modelos baseados em oito das va-

riáveis originais não apresentaram uma melhoria notável em relação aos que foram avaliados aqui.)

**80.** Uma amostra de  $n = 20$  empresas foi selecionada e foram determinados os valores de  $y =$  preço da ação e  $k = 15$  variáveis de previsão (como dividendo trimestral, lucros do ano anterior, quociente da dívida). Quando o modelo de regressão múltipla com esses 15 previsores foi ajustado aos dados, obteve-se um  $R^2 = 0,90$ .

- O modelo parece especificar uma relação útil entre  $y$  e as variáveis de previsão? Execute um teste usando o nível de significância 0,05. [Sugestão: O valor crítico  $F$  para o numerador 15 e o denominador 4 gl é 5,86.]
- Com base no resultado da parte (a), um valor  $R^2$  alto por si só sugere que o modelo seja útil? Sob que circunstâncias você suspeitaria de um modelo com um valor  $R^2$  alto?
- Com  $n$  e  $k$  especificados previamente, de que tamanho  $R^2$ ? deveria ser para o modelo ser considerado útil ao nível de significância 0,05?

**81.** A exposição à poluição atmosférica diminui a expectativa de vida? Essa pergunta foi examinada no artigo “Does Air Pollution Shorten Lives?” (*Statistics and Public Policy*, Reading, MA, Addison Wesley, 1977). Dados sobre

$y =$  taxa de mortalidade total (mortes por 10.000)

$x_1 =$  leitura média de partículas suspensas ( $\mu\text{g}/\text{m}^3$ )

$x_2 =$  menor leitura de sulfato ( $[\mu\text{g}/\text{m}^3] \times 10$ )

$x_3 =$  densidade populacional (pessoas/ $\text{mi}^2$ )

$x_4 =$  (porcentagem de não-brancos)  $\times 10$

$x_5 =$  (porcentagem acima de 65)  $\times 10$

para o ano de 1960 foram registrados para  $n = 117$  áreas estatísticas metropolitanas selecionadas aleatoriamente. A equação da regressão estimada é

$$y = 19,607 + 0,041x_1 + 0,071x_2 + 0,001x_3 + 0,041x_4 + 0,687x_5$$

- Para esse modelo,  $R^2 = 0,827$ . Usando um nível de significância  $\alpha = 0,05$ , execute um teste de utilidade do modelo.
  - O desvio padrão estimado de  $\hat{\beta}_1$  foi 0,016. Calcule e interprete um IC de 90% para  $\beta_1$ .
  - Visto que o desvio padrão estimado de  $\hat{\beta}_4$  é 0,007, determine se a porcentagem de não-brancos é uma variável importante no modelo. Use um nível de significância de 0,01.
  - Em 1960, os valores de  $x_1, x_2, x_3, x_4$  e  $x_5$  registrados para a cidade de Pittsburgh (Estados Unidos) foram 166, 60, 788, 68 e 95, respectivamente. Use a equação de regressão fornecida para prever a taxa de mortalidade de Pittsburgh. Como sua previsão se compara com a quantidade real de 1960 de 103 mortes por 10.000?
- 82.** Visto que  $R^2 = 0,723$  para o modelo contendo os previsores  $x_1, x_4, x_5$  e  $x_8$  e  $R^2 = 0,689$  no modelo com os previsores  $x_1, x_3, x_5$  e  $x_6$ , o que você pode dizer sobre  $R^2$  no modelo contendo os previsores
- $x_1, x_3, x_4, x_5, x_6$  e  $x_8$ ? Explique.
  - $x_1$  e  $x_4$ ? Explique.

---

## Bibliografia

- CHATTERJEE, S. e PRICE, Bertram. *Regression Analysis by Example*. Nova York: Wiley, 1990, 2ª ed. Uma discussão breve, mas informativa, sobre tópicos selecionados, especialmente multicolinearidade e o uso de métodos de estimativa tendenciosos.
- DANIEL, Cuthbert e WOOD, Fred. *Fitting Equations to Data*. Nova York, Wiley, 1980, 2ª ed. Contém várias idéias e métodos desenvolvidos pela ampla experiência dos autores em consultoria.
- DRAPER, Norman e SMITH, Harry. *Applied Regression Analysis*. Nova York, Wiley, 1999, 3ª ed. Veja a bibliografia do Capítulo 12.
- HOAGLIN, David e WELSCH, Roy. “The Hat Matrix in Regression and ANOVA”. *American Statistician*, 1978, p. 17–23. Descreve métodos para detectar observações influentes em um conjunto de dados de regressão.
- HOCKING, Ron. “The Analysis and Selection of Variables in Linear Regression”. *Biometrics*, 1976, p. 1–49. Uma excelente pesquisa sobre esse tema.
- NETER, John, KUTNER, Michael, NACHTSHEIM, Christopher e WASSERMAN, William. *Applied Linear Statistical Models*. Irwin, Homewood, IL, 1996 (4. ed). Veja a bibliografia do Capítulo 12.

# Testes de Aderência e Análise de Dados Categorizados

## Introdução

No tipo mais simples de situação considerada neste capítulo, cada observação em um exemplo é classificada como pertencendo a um número finito de categorias (por exemplo, o tipo sanguíneo pode pertencer a uma entre quatro categorias O, A, B ou AB). Com  $p_i$  indicando a probabilidade de que qualquer observação particular pertença a uma categoria  $i$  (ou a proporção da população que pertence à categoria  $i$ ), desejamos testar a hipótese nula que especifique completamente os valores de todos os  $P_i$ 's (tal como  $H_0: p_1 = 0,45, p_2 = 0,35, p_3 = 0,15, p_4 = 0,05$ , quando houver quatro categorias). O teste estatístico será uma medida da discrepância entre os números observados nas categorias e os números esperados quando  $H_0$  é verdadeiro. Como a decisão será tomada pela comparação do valor calculado do teste estatístico com o valor crítico da distribuição qui-quadrado, o procedimento é chamado teste de aderência qui-quadrado.

Algumas vezes, a hipótese nula especifica que os  $p_i$ s dependem de um número menor de parâmetros sem especificar os valores destes parâmetros. Por exemplo, com três categorias a hipótese nula pode dizer que  $p_1 = \theta^2, p_2 = 2\theta(1 - \theta)$  e  $p_3 = (1 - \theta)^2$ . Para um teste qui-quadrado ser realizado, os valores de quaisquer parâmetros não especificados devem ser estimados a partir dos dados de amostra. Estes problemas são discutidos na Seção 14.2. Os métodos são então aplicados para testar a hipótese nula que defina que a amostra tem origem em uma família específica de distribuições, tais como a família de Poisson (com  $\lambda$  estimado a partir da amostra) ou uma família normal (com  $\mu$  e  $\sigma$  estimados).

Os testes qui-quadrado para duas situações diferentes são apresentados na Seção 14.3. Na primeira, a hipótese nula define que os  $p_i$ s são os mesmos para várias populações diferentes. O segundo tipo de situação envolve obter uma amostra a partir de uma única população e classificar cada indivíduo em relação a dois fatores de categorias diferentes (tais como preferência religiosa e filiação a partido político). A hipótese nula nesta situação é a de que os dois fatores são independentes dentro da população.

## 14.1 Testes de Aderência Quando as Probabilidades das Categorias são Completamente Especificadas

Um experimento binomial consiste de uma seqüência de testes independentes nos quais cada um pode produzir um de dois possíveis resultados,  $S$  (para sucesso) e  $F$  (para falha). A probabilidade de sucesso, indicada por  $p$ , é tida como constante de teste para teste e o número  $n$  de testes é fixado pela definição do experimento. No Capítulo 8, apresentamos um teste  $z$  com uma amostra grande para testar  $H_0: p = p_0$ . Observe que esta hipótese nula especifica ambos  $P(S)$  e  $P(F)$ , uma vez que se  $P(S) = p_0$ , então  $P(F) = 1 - p_0$ . Indicando  $P(F)$  por  $q$  e  $1 - p_0$  por  $q_0$ , a hipótese nula pode ser escrita de forma alternativa como  $H_0: p = p_0, q = q_0$ . O teste  $z$  é bicaudal quando a alternativa de interesse é  $p \neq p_0$ .

Um **experimento multinomial** generaliza um experimento binomial permitindo que cada teste resulte em um de  $k$  possíveis resultados, onde  $k > 2$ . Por exemplo, suponha que uma loja aceite três tipos diferentes de cartões de crédito. Um experimento multinomial resultará da observação do tipo de cartão de crédito utilizado — tipo 1, tipo 2 ou tipo 3 — para cada um dos próximos  $n$  clientes que pagarem com cartão de crédito. Em geral, faremos referência aos  $k$  possíveis resultados em qualquer realização de um experimento como categorias, e  $p_i$  indica a probabilidade de um experimento resultar na categoria  $i$ . Se o experimento consiste da seleção de  $n$  indivíduos ou objetos de uma população e da categorização de cada um desses indivíduos, então  $p$  é a proporção da população que pertence a  $i$ -ésima categoria (um experimento será aproximadamente multinomial desde que  $n$  seja muito menor do que o tamanho da população).

A hipótese nula de interesse especificará o valor de cada  $p_i$ . Por exemplo, no caso  $k = 3$ , podemos ter  $H_0: p_1 = 0,5, p_2 = 0,3, p_3 = 0,2$ . A hipótese alternativa registrará que  $H_0$  não é verdadeira — ou seja, pelo menos um dos  $p_i$ s tem um valor diferente daquele definido por  $H_0$  (neste caso, pelo menos dois devem ser diferentes, uma vez que a soma deles é igual a 1). O símbolo  $p_{i0}$  representará o valor de  $p_i$  especificado na hipótese nula. No exemplo acima,  $p_{10} = 0,5, p_{20} = 0,3$  e  $p_{30} = 0,2$ .

Antes de o experimento multinomial ser realizado, o número de testes que resultarão na categoria  $i$  ( $i = 1, 2, \dots$ , ou  $k$ ) é uma variável aleatória — tal como o número de sucessos e o número de falhas em um experimento binomial são variáveis aleatórias. Essa variável aleatória será indicada por  $N_i$  e o seu valor observado por  $n_i$ . Como cada teste resulta em exatamente uma das categorias de  $k$ ,  $\sum N_i = n$ , e o mesmo é verdadeiro para os  $n_i$ s. Como um exemplo, um experimento com  $n = 100$  e  $k = 3$  pode produzir  $N_1 = 46, N_2 = 35$  e  $N_3 = 19$ .

Os números esperados de sucessos e de falhas em um experimento binomial são  $np$  e  $nq$ , respectivamente. Quando  $H_0: p = p_0$  e  $q = q_0$  for verdadeira, os números esperados de sucessos e falhas serão  $np_0$  e  $nq_0$ , respectivamente. Da mesma maneira, em um experimento multinomial, o número esperado de testes que resultam na categoria  $i$  é  $E(N_i) = np_i$  ( $i = 1, \dots, k$ ). Quando  $H_0: p_1 = p_{10}, \dots, p_k = p_{k0}$  for verdadeira, estes valores esperados tornam-se  $E(N_1) = np_{10}, E(N_2) = np_{20}, \dots, E(N_k) = np_{k0}$ . Para o caso  $k = 3, H_0: p_1 = 0,5, p_2 = 0,3, p_3 = 0,2$  e  $n = 100, E(N_1) = 100(0,5) = 50, E(N_2) = 30$  e  $E(N_3) = 20$  quando  $H_0$  for verdadeira. Os  $n_i$ s são freqüentemente mostrados em um formato de tabela que consiste de uma linha de  $k$  células, uma para cada categoria, como ilustrado na Tabela 14.1. Os valores esperados quando  $H_0$  for verdadeira são mostrados abaixo dos valores observados. Os  $N_i$ s e  $n_i$ s são normalmente chamados de contagem das células observadas (ou freqüências das células observadas), e  $np_{10}, np_{20}, \dots, np_{k0}$  são as contagens de células esperadas correspondentes sob  $H_0$ .

**Tabela 14.1** A contagem das células observadas e esperadas

Categoria	$i = 1$	$i = 2$	$\dots$	$i = k$	Total da linha
Observada	$n_1$	$n_2$	$\dots$	$n_k$	$n$
Esperada	$np_{10}$	$np_{20}$	$\dots$	$np_{k0}$	$n$

Por um lado, os  $n_i$ s devem estar todos relativamente próximos dos  $np_{i0}$ s correspondentes quando  $H_0$  é verdadeira. Por outro lado, várias das contagens observadas podem diferir substancialmente das contagens esperadas quando os valores reais dos  $p_i$ s diferirem significativamente do que a hipótese nula afirma. O procedimento de teste envolve o cálculo da discrepância entre os  $n_i$ s e os  $np_{i0}$ s, sendo  $H_0$  rejeitada quando a discrepância for suficientemente grande. É natural tomar como base uma medida de discrepância entre os desvios quadrados  $(n_1 - np_{10})^2, (n_2 - np_{20})^2, \dots, (n_k - np_{k0})^2$ . Uma maneira óbvia de combinar isto em uma média geral é adicioná-los para obter  $\sum (n_i - np_{i0})^2$ . Entretanto, suponha que  $np_{10} = 100$  e  $np_{20} = 10$ . Então se  $n_1 = 95$  e  $n_2 = 5$ , as duas categorias contribuem com os mesmos desvios quadrados para a medida proposta. Contudo,  $n_1$  é apenas 5% menor do que seria esperado quando  $H_0$  é verdadeira, enquanto  $n_2$  é 50% menor. Para levar em consideração as magnitudes relativas dos desvios, dividiremos cada desvio quadrado pelo valor correspondente esperado e depois vamos combiná-los.

Antes de fornecer uma descrição mais detalhada, devemos discutir um tipo de distribuição de probabilidade chamada *distribuição qui-quadrado*. Essa distribuição foi apresentada pela primeira vez na Seção 4.4 e utilizada no Capítulo 7 para obter um intervalo de confiança para a variância  $\sigma^2$  de uma população normal. A distribuição qui-quadrado tem um único parâmetro  $\nu$ , chamado de número de graus de liberdade (gl) da distribuição, com valores possíveis 1, 2, 3, . . . . Analogamente ao valor crítico  $t_{\alpha, \nu}$  para a distribuição  $t$ ,  $\chi^2_{\alpha, \nu}$  é o valor tal qual  $\alpha$  da área sob a curva  $\chi^2$  com  $\nu$  gl está localizado à direita de  $\chi^2_{\alpha, \nu}$  (veja a Figura 14.1). Os valores selecionados de  $\chi^2_{\alpha, \nu}$  são dados na Tabela A.7 do Apêndice.

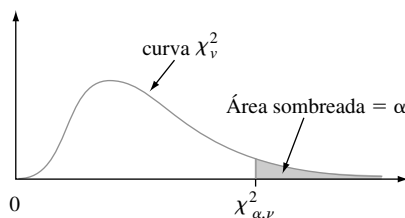


Figura 14.1 Um valor crítico para uma distribuição qui-quadrado

TEOREMA

Se  $np_i \geq 5$  para todo  $i$  ( $i = 1, 2, \dots, k$ ), a variável

$$\chi^2 = \sum_{i=1}^k \frac{(N_i - np_i)^2}{np_i} = \sum_{\text{todas as células}} \frac{(\text{observado} - \text{esperado})^2}{\text{esperado}}$$

tem aproximadamente um distribuição qui-quadrado com  $k - 1$  gl.

O fato de que  $gl = k - 1$  é uma conseqüência da restrição  $\sum N_i = n$ . Embora haja  $k$  contagens de células observadas, uma vez que  $k - 1$  são conhecidas, a remanescente é determinada unicamente. Ou seja, há somente  $k - 1$  contagens das células “livremente determinadas” e, assim,  $k - 1$  gl.

Se  $np_{i0}$  for substituído por  $np_i$  em  $\chi^2$ , o teste estatístico resultante possui uma distribuição qui-quadrado quando  $H_0$  é verdadeira. A rejeição de  $H_0$  é apropriada quando  $\chi^2 \geq c$  (pois grandes discrepâncias entre as contagens observadas e esperadas levam a um valor grande de  $\chi^2$ ), e a escolha  $c = \chi^2_{\alpha, k-1}$  produz um teste com nível de significância  $\alpha$ .

Hipótese nula:  $H_0: p_1 = p_{10}, p_2 = p_{20}, \dots, p_k = p_{k0}$

Hipótese alternativa:  $H_a$ : pelo menos um  $p_i$  é diferente de  $p_{i0}$

Valor do teste estatístico:  $\chi^2 = \sum_{\text{todas as células}} \frac{(\text{observado} - \text{esperado})^2}{\text{esperado}} = \sum_{i=1}^k \frac{(n_i - np_{i0})^2}{np_{i0}}$

Região de Rejeição:  $\chi^2 \geq \chi^2_{\alpha, k-1}$

**Exemplo 14.1**

Se concentramos nossa atenção em duas características diferentes de um organismo, cada um controlado por um único gene, e cruzamos uma família pura que tem genótipo AABB com uma família pura com genótipo aabb (letras maiúsculas indicam alelos dominantes e letras minúsculas, alelos recessivos), o resultado será AaBb. Se estes organismos da primeira geração são então cruzados entre si (um cruzamento di-híbrido), haverá quatro fenótipos dependentes da presença de um alelo dominante. As leis de Mendel sobre hereditariedade implicam que estes quatro fenótipos devem ter probabilidades  $\frac{9}{16}, \frac{3}{16}, \frac{3}{16}$  e  $\frac{1}{16}$  de surgirem em qualquer cruzamento di-híbrido dado.

O artigo “Linkage Studies of the Tomato” (*Trans. Royal Canadian Institute*, 1931, p. 1-19) informa os seguintes dados sobre os fenótipos de um cruzamento di-híbrido de tomates longos de folhas curtas com tomates anões de folhas com formato de batata. Há  $k = 4$  categorias que correspondem aos quatro fenótipos possíveis, com a hipótese nula sendo

$$H_0: p_1 = \frac{9}{16}, p_2 = \frac{3}{16}, p_3 = \frac{3}{16}, p_4 = \frac{1}{16}$$

As contagens de células esperadas são  $9n/16, 3n/16, 3n/16$  e  $n/16$  e o teste se baseia em  $k - 1 = 3$  gl. O tamanho total da amostra foi de  $n=1.611$ . As contagens observadas e esperadas são dadas na Tabela 14.2

**Tabela 14.2** Contagem de células observadas e esperadas para o Exemplo 14.1

	<i>i</i> = 1 Longo, folha curta	<i>i</i> = 2 Longo, folha de batata	<i>i</i> = 3 Anão, folha curta	<i>i</i> = 4 Anão, folha de batata
$n_i$	926	288	293	104
$np_{i0}$	906,2	302,1	302,1	100,7

A contribuição para  $\chi^2$  da primeira célula é

$$\frac{(n_1 - np_{10})^2}{np_{10}} = \frac{(926 - 906,2)^2}{906,2} = 0,433$$

As contribuições das células 2, 3 e 4 na fórmula são iguais a 0,658, 0,274 e 0,108, respectivamente, de forma que  $\chi^2 = 0,433 + 0,658 + 0,274 + 0,108 = 1,473$ . Um teste com nível de significância 0,10 requer  $\chi^2_{0,10,3}$ , o número na Tabela A.7 do Apêndice com 3 gl na linha e 0,10 na coluna. Este valor crítico é 6,251. Como 1,473 é menor do que 6,251,  $H_0$  não pode ser rejeitada mesmo neste nível elevado de significância. Os dados são bem consistentes com as leis de Mendel. ■

Embora tenhamos desenvolvido o teste qui-quadrado para situações nas quais  $k > 2$ , ele também pode ser utilizado quando  $k = 2$ . A hipótese nula neste caso pode ser estabelecida quando  $H_0: p_1 = p_{10}$ , desde que as relações  $p_2 = 1 - p_1$  e  $p_{20} = 1 - p_{10}$  tornem a inclusão  $p_2 = p_{20}$  em  $H_0$  redundante. A hipótese alternativa é  $H_a: p_1 \neq p_{10}$ . Estas hipóteses podem também ser testadas com a utilização de um teste estatístico  $z$  bicaudal

$$Z = \frac{(N_1/n) - p_{10}}{\sqrt{\frac{p_{10}(1 - p_{10})}{n}}} = \frac{\hat{p}_1 - p_{10}}{\sqrt{\frac{p_{10}p_{20}}{n}}}$$

Surpreendentemente, os dois procedimentos de teste são completamente equivalentes. Isto acontece porque pode ser mostrado que  $Z^2 = \chi^2$  e  $(z_{\alpha/2})^2 = \chi^2_{1,\alpha}$ , tal que  $\chi^2 \geq \chi^2_{1,\alpha}$  se, e somente se,  $|Z| \geq z_{\alpha/2}$ .<sup>1</sup>

<sup>1</sup> O fato de que  $(z_{\alpha/2})^2 = \chi^2_{1,\alpha}$  é uma consequência da relação entre a distribuição normal padronizada e a distribuição qui-quadrado com 1 gl, se  $Z \sim N(0, 1)$ , então  $Z^2$  tem distribuição qui-quadrado com  $\nu = 1$ .



Se a hipótese alternativa for tanto  $H_a: p_1 > p_{10}$  como  $H_a: p_1 < p_{10}$ , o teste qui-quadrado não poderá ser utilizado. Deve-se então usar um teste  $z$  com cauda única superior ou inferior.

Como acontece com todos os procedimentos de teste, deve-se tomar cuidado para não confundir significância estatística com significância prática. Um  $\chi^2$  calculado que exceda  $\chi_{\alpha, k-1}^2$  pode ser o resultado de uma amostra de tamanho muito grande, em detrimento de quaisquer diferenças práticas entre os  $p_{i0}$ s supostos e os  $p_i$ s verdadeiros. Assim, se  $p_{10} = p_{20} = p_{30} = \frac{1}{3}$ , mas os  $p_i$ s verdadeiros têm valores 0,330, 0,340 e 0,330, um valor grande de  $\chi^2$  com certeza surgirá com um  $n$  suficientemente grande. Antes de rejeitar  $H_0$ , os  $\hat{p}_i$ s devem ser examinados para verificar se eles sugerem um modelo diferente daquele de  $H_0$  a partir de um ponto de vista prático.

## Valores $P$ para Teste Qui-Quadrado

Os testes qui-quadrado neste capítulo são todos com cauda superior, então concentraremos nossa atenção neste caso. Como o valor  $P$  para um teste  $t$  com cauda superior é a área sob a curva  $t_v$  à direita do  $t$  calculado, o valor  $P$  para um teste qui-quadrado com cauda superior é a área sob a curva à direita de  $\chi_v^2$  calculado. A Tabela A.7 do Apêndice fornece informações limitadas sobre os valores  $P$ , pois apenas cinco valores críticos com cauda superior são tabelados para cada  $\nu$  distinto. Temos, portanto, outra tabela no apêndice, análoga à Tabela A.8, que facilita a produção de valores  $P$  mais precisos.

O fato de que as curvas  $t$  foram todas centradas em zero nos permite tabular as áreas de cauda da curva  $t$  de maneira relativamente compacta, com a margem esquerda fornecendo os valores que variam de 0,0 a 4,0 na escala  $t$  horizontal e várias colunas mostrando áreas de cauda superior correspondentes para várias gls. O movimento à direita das curvas qui-quadrado quando gl aumenta requer um tipo diferente de tabulação. A margem esquerda da Tabela A.11 do Apêndice mostra várias áreas de cauda superior: 0,100, 0,095, 0,090, ..., 0,005 e 0,001. Cada coluna da tabela se destina a um valor diferente de gl e as entradas da tabela mostram os valores no eixo qui-quadrado horizontal que captura estas áreas de cauda correspondentes. Por exemplo, movendo-se para baixo até a área de cauda 0,085 e atravessando para a coluna 2 gl, vemos que a área à direita de 4,93 sob a curva qui-quadrado de 2 gl é 0,085 (veja a Figura 14.2).

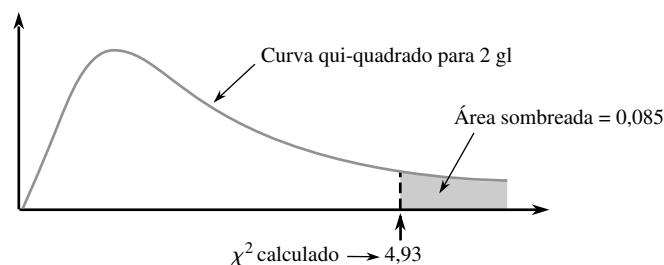


Figura 14.2 Valor  $P$  para um teste qui-quadrado com cauda superior

Para capturar esta mesma área com cauda superior sob a curva 10 gl, devemos ir até 16,54. Na coluna 2 gl, a linha superior mostra que se o valor calculado da variável qui-quadrado for menor do que 4,60, a área da cauda capturada (o valor  $P$ ) excede 0,10. Da mesma forma, a última linha desta coluna indica que se o valor calculado excede 13,81, a área da cauda é menor do que 0,001 (valor  $P < 0,001$ ).

## $\chi^2$ Quando o $p_i$ s São Funções de Outros Parâmetros

Freqüentemente os  $p_i$ s são supostos dependentes de um número menor de parâmetros  $\theta_1, \dots, \theta_m$  ( $m < k$ ). Então uma hipótese específica envolvendo os  $\theta_i$ s produz  $p_{i0}$ s específicos, que são então usados no teste  $\chi^2$ .

### Exemplo 14.2

Em um artigo sobre genética bastante conhecido (“The Progeny in Generations  $F_{12}$  e  $F_{17}$  of a Cross Between a Yellow-Wrinkled and a Green-Round Seeded Pea”, *J. Genetics* 1923, p. 255-331), o estatístico G. U. Yule analisou os dados resultantes do cruzamento de ervilhas de jardim. Os alelos dominantes no experimento foram

Y = cor amarela e R = formato redondo, resultando no dominante duplo YR. Yule examinou 269 vagens com quatro sementes resultantes de um cruzamento di-híbrido e contou o número de sementes YR em cada vagem. Definindo  $X$  como o número de YRs em uma vagem aleatoriamente selecionada, os valores de  $X$  possíveis são 0, 1, 2, 3, 4 que identificamos com as células 1, 2, 3, 4 e 5 de uma tabela retangular (ou seja, uma vagem com  $X = 4$  produz um número observado na célula 5).

A hipótese de que as leis de Mendel funcionam e de que genótipos de sementes individuais dentro de uma vagem são independentes uns dos outros implica que  $X$  tem uma distribuição binomial com  $n = 4$  e  $\theta = \frac{9}{16}$ . Assim, desejamos testar  $H_0: p_1 = p_{10}, \dots, p_5 = p_{50}$ , onde

$$p_{i0} = P(i - 1 \text{ YRs entre 4 sementes quando } H_0 \text{ é verdadeira}) \\ = \binom{4}{i - 1} \theta^{i-1} (1 - \theta)^{4-(i-1)} \quad i = 1, 2, 3, 4, 5; \theta = \frac{9}{16}$$

Os dados de Yule e os cálculos estão na Tabela 14.3 com as contagens de células esperadas  $np_{i0} = 269p_{i0}$ .

**Tabela 14.3** Contagens de células observadas e esperadas para o Exemplo 14.2

Célula $i$ ervilhas/vagens YR	1 0	2 1	3 2	4 3	5 4
Observada	16	45	100	82	26
Esperadas	9,86	50,68	97,75	83,78	26,93
$\frac{(\text{observada} - \text{esperada})^2}{\text{esperada}}$	3,823	0,637	0,052	0,038	0,032

Assim  $\chi^2 = 3,823 + \dots + 0,032 = 4,582$ . Como  $\chi_{0,01,k-1}^2 = \chi_{0,01,4}^2 = 13,277$ ,  $H_0$  não é rejeitada ao nível 0,01. A Tabela A.11 do Apêndice mostra que como  $4,582 < 7,77$ , o valor  $P$  para o teste excede 0,10.  $H_0$  não deve ser rejeitada em qualquer nível de significância razoável. ■

### $\chi^2$ Quando a Distribuição Subjacente É Contínua

Temos há muito trabalhado com a hipótese de que as  $k$  categorias são naturalmente definidas no contexto do experimento em consideração. O teste  $\chi^2$  também pode ser utilizado para checar se uma amostra provém de uma distribuição contínua subjacente específica. Vamos indicar por  $X$  a variável sendo amostrada e vamos supor que a fdp suposta para  $X$  é  $f_0(x)$ . Como na construção de uma distribuição de frequência no Capítulo 1, subdividimos a escala de medida de  $X$  em  $k$  intervalos  $[a_0, a_1), [a_1, a_2), \dots, [a_{k-1}, a_k)$ , onde o intervalo  $[a_{i-1}, a_i)$  inclui o valor  $a_{i-1}$  mas não  $a_i$ . As probabilidades das células especificadas por  $H_0$  são então

$$p_{i0} = P(a_{i-1} \leq X < a_i) = \int_{a_{i-1}}^{a_i} f_0(x) dx$$

As células devem ser escolhidas tal que  $np_{i0} \geq 5$  para  $i = 1, \dots, k$ . Frequentemente elas são selecionadas tal que os  $np_{i0}$ s sejam iguais.

#### Exemplo 14.3

Para verificar se o tempo de acesso ao trabalho de parto entre mães grávidas está uniformemente distribuído ao longo de um dia de 24 horas, podemos dividir um dia em  $k$  períodos, cada um deles de comprimento  $24/k$ . A hipótese nula define que  $f(x)$  é uma fdp uniforme no intervalo  $[0,24]$ , tal que  $p_{i0} = 1/k$ . O artigo “The Hour of Birth” (*British J. Preventive and Social Medicine*, 1953, p. 43-59) informa 1.186 tempos de parto que foram categorizados em  $k = 24$  intervalos de 1 hora iniciando à meia-noite, resultando em contagens de células de 52, 73,

89, 88, 68, 47, 58, 47, 48, 53, 47, 34, 21, 31, 40, 24, 37, 31, 47, 34, 36, 44, 78 e 59. Cada contagem de célula esperada é  $1.186 \cdot \frac{1}{24} = 49,42$ , e o valor resultante de  $\chi^2$  é 162,77. Como  $\chi^2_{0,01,23} = 41,637$ , o valor calculado é bastante significativo e a hipótese nula é claramente rejeitada. Genericamente falando, fica a impressão de que é muito mais provável que o trabalho de parto inicie tarde da noite do que durante as horas normais diurnas. ■

Para testar se uma amostra provém de uma distribuição normal específica, os parâmetros fundamentais  $\theta_1 = \mu$  e  $\theta_2 = \sigma$ , e cada  $p_{i0}$  será uma função desses parâmetros.

**Exemplo 14.4**

Em uma determinada universidade, os exames finais devem durar 2 horas. O departamento de psicologia estruturou um departamento para um curso elementar que visava satisfazer o seguinte critério: (1) o tempo real que se leva para completar o exame é normalmente distribuído, (2)  $\mu = 100$  min. e (3) exatamente 90% de todos os estudantes terminarão dentro do período de 2 horas. Para verificar se este é realmente o caso, 120 estudantes foram selecionados aleatoriamente, e seus tempos de conclusão registrados. Decidiu-se que  $k = 8$  intervalos seriam utilizados. O critério implica que o nonagésimo percentil da distribuição do tempo de conclusão é  $\mu + 1,28\sigma = 120$ . Como  $\mu = 100$ , isto implica que  $\sigma = 15,63$ .

Os oito intervalos que dividem a escala normal padrão em oito segmentos apropriadamente iguais são  $[0; 0,32)$ ,  $[0,32; 0,675)$ ,  $[0,675; 1,15)$ ,  $[1,15; \infty)$ , e suas quatro partes simétricas do outro lado de 0. Para  $\mu = 100$  e  $\sigma = 15,63$ , estes intervalos se tornam  $[100, 105)$ ,  $[105; 110,55)$ ,  $[110,55; 117,97)$  e  $[117,97; \infty)$ . Assim  $p_{i0} = \frac{1}{8} = 0,125$  ( $i = 1, \dots, 8$ ), de forma que cada contagem de célula esperada é  $np_{i0} = 120(0,125) = 15$ . As contagens de célula observadas são 21, 17, 12, 16, 10, 15, 19 e 10, resultando em um  $\chi^2$  de 7,73. Como  $\chi^2_{0,10,7} = 12,017$  e 7,73 não é  $\geq 12,017$ , não há evidência para concluir que o critério não foi atingido. ■

**Exercícios** | Seção 14.1 (1–11)

1. Qual conclusão seria apropriada para um teste qui-quadrado de cauda superior em cada uma das seguintes situações?
  - a.  $\alpha = 0,05$ ,  $df = 4$ ,  $\chi^2 = 12,25$
  - b.  $\alpha = 0,01$ ,  $df = 3$ ,  $\chi^2 = 8,54$
  - c.  $\alpha = 0,10$ ,  $df = 2$ ,  $\chi^2 = 4,36$
  - d.  $\alpha = 0,01$ ,  $k = 6$ ,  $\chi^2 = 10,20$
2. Fale o que for possível sobre o valor  $P$  para um teste qui-quadrado com cauda superior em cada uma das seguintes situações:
  - a.  $\chi^2 = 7,5$ ,  $df = 2$
  - b.  $\chi^2 = 13,0$ ,  $df = 6$
  - c.  $\chi^2 = 18,0$ ,  $df = 9$
  - d.  $\chi^2 = 21,3$ ,  $k = 5$
  - e.  $\chi^2 = 5,0$ ,  $k = 4$
3. Um departamento de estatística de uma grande universidade mantém um serviço de acompanhamento para estudantes em seus cursos iniciais. O serviço foi organizado com a expectativa de que 40% de seus clientes seria do curso de estatística de negócios, 30% de estatística de engenharia, 20% do curso de estatística para estudantes de ciências sociais e outros 10% para o curso de estudantes de agricultura. Uma amostra aleatória de  $n = 120$  clientes revelou 52, 38, 21 e 9 dos quatro cursos. Estes dados sugerem que as porcentagens nas quais o curso foi baseado não estão corretas? Defina e teste a hipótese relevante utilizando  $\alpha = 0,05$ .
4. Suponha que quando pombos-correio são desorientados de uma determinada maneira, eles não mostram

preferência por uma direção específica de vôo após a partida (tal que a direção  $X$  deva ser uniformemente distribuída no intervalo de  $0^\circ$  a  $360^\circ$ ). Para realizar este teste, 120 pombos foram desorientados, soltos e a direção do vôo de cada um foi registrada; os dados resultantes estão a seguir. Utilize o teste qui-quadrado no nível 0,10 para verificar se os dados suportam a hipótese feita

<b>Direção</b>	$0 - < 45^\circ$	$45 - < 90^\circ$	$90 - < 135^\circ$
<b>Frequência</b>	12	16	17
<b>Direção</b>	$135 - < 180^\circ$	$180 - < 225^\circ$	$225 - < 270^\circ$
<b>Frequência</b>	15	13	20
<b>Direção</b>	$270 - < 315^\circ$	$315 - < 360^\circ$	
<b>Frequência</b>	17	10	

5. Um sistema de recuperação de informações tem dez localizações de armazenagem. A informação é armazenada com a expectativa de que, no longo prazo, a proporção de solicitações por localização  $i$  seja dada por  $p_i = (5,5 - |i - 5,5|)/30$ . Uma amostra de 200 solicitações de recuperação de informações forneceu as seguintes frequências para as localizações de 1-10, respectivamente: 4, 15, 23, 25, 38, 31, 32, 14, 10 e 8. Utilize um teste qui-quadrado ao nível de significância 0,10 para decidir se os dados são consistentes com as proposições iniciais (utilize o valor  $P$  aproximado).

6. O sorgo é um importante cereal cujas qualidade e aparência podem ser afetadas pela presença de pigmentos no pericarpo (as paredes do ovário da planta). O artigo “A Genetic and Biochemical Study on Pericarp Pigments in a Cross Between Two Cultivars of Grain Sorghum,”. *Sorghum Bicolor*”(*Heredity*, 1976, p 413-416 ) informa sobre um experimento que envolveu um cruzamento inicial entre o sorgo CK60 (uma variedade americana com sementes brancas) e Abu Taima (uma variedade etíope com sementes amarelas) para produzir plantas com sementes vermelhas e então um autocruzamento de plantas com sementes vermelhas. De acordo com a teoria genética, este cruzamento F2 deveria produzir plantas com sementes vermelhas, amarelas e brancas na razão 9 : 3 : 4. Os dados do experimento seguem adiante, eles confirmam ou contradizem a teoria genética? Teste ao nível 0,05 utilizando o método do valor  $P$ .

Cor da Semente	Vermelha	Amarela	Branca
Frequência Observada	195	73	100

7. Criminologistas têm debatido com frequência se há uma relação entre as condições climáticas e a incidência de crime violento. O autor do artigo “Is There a Season for Homicide?” (*Criminology*, 1988, p. 287-296) classificou 1.361 homicídios de acordo com a estação, resultando nos dados a seguir. Teste a hipótese nula de proporções iguais usando  $\alpha = 0,01$  utilizando a tabela qui-quadrado para dizer o que for possível sobre o valor  $P$ .

Inverno	Primavera	Verão	Outono
328	334	372	327

8. O artigo “Psychiatric and Alcoholic Admissions Do Not Occur Disproportionately Close to Patients Birthdays” (*Psychological Reports*, 1992, p. 944-946) concentra-se na existência de alguma relação entre a data de admissão do paciente para tratamento de alcoolismo e o aniversário do paciente. Assumindo um ano de 365 dias (ou seja, excluindo anos bissextos), na ausência de qualquer relação, a data de admissão de paciente é igualmente provável em qualquer um dos 365 possíveis dias. Os pesquisadores estabeleceram quatro diferentes categorias de admissão: (1) até 7 dias do aniversário; (2) entre 8 e 30 dias, inclusive, a partir do aniversário; (3) entre 31 e 90 dias, inclusive, a partir do aniversário; (4) mais de 90 dias do aniversário. Uma amostra de 200 pacientes teve frequências observadas de 11, 24, 69 e 96 para as categorias 1, 2, 3 e 4, respectivamente. Defina e teste a hipótese relevante utilizando um nível de significância de 0,01.

9. O tempo de resposta de um sistema de computador a uma requisição para um determinado tipo de informação é concebido para ter distribuição exponencial

com parâmetro  $\lambda = 1$  s (tal que se  $X =$  tempo de resposta, a fdp de  $X$  sob  $H_0$  é  $f_0(x) = e^{-x}$  para  $x \geq 0$ ).

a. Se você observou  $X_1, X_2, \dots, X_n$  e quis utilizar o teste qui-quadrado com cinco intervalos de classes tendo iguais probabilidades sob  $H_0$ , quais seriam os intervalos de classe resultantes?

b. Realize o teste teste qui-quadrado utilizando os seguintes dados resultantes de uma amostra aleatória de 40 tempos de resposta:

0,10	0,99	1,14	1,26	3,24	0,12	0,26	0,80
0,79	1,16	1,76	0,41	0,59	0,27	2,22	0,66
0,71	2,21	0,68	0,43	0,11	0,46	0,69	0,38
0,91	0,55	0,81	2,51	2,77	0,16	1,11	0,02
2,13	0,19	1,21	1,13	2,93	2,14	0,34	0,44

10. a. Demonstre que outra expressão para a estatística qui-quadrado é

$$\chi^2 = \sum_{i=1}^k \frac{N_i^2}{np_{i0}} - n$$

Por que é mais eficiente calcular  $\chi^2$  utilizando esta fórmula?

b. Quando a hipótese nula é  $H_0: p_1 = p_2 = \dots = p_k = 1/k$  (ou seja,  $p_{i0} = 1/k$  para todo  $i$ ), como a fórmula da parte (a) pode ser simplificada? Utilize a expressão simplificada para calcular  $\chi^2$  para os dados de direção dos pombos do Exercício 4.

11. a. Tendo obtido uma amostra aleatória de uma população, você deseja utilizar um teste qui-quadrado para decidir se a distribuição populacional é normal padronizada. Se você baseia o teste em seis intervalos de classe tendo probabilidades iguais sob  $H_0$ , quais devem ser os intervalos de classe?

b. Se você deseja utilizar um teste qui-quadrado para testar  $H_0$ : a distribuição populacional é normal com  $\mu = 0,5$ ,  $\sigma = 0,002$  e o teste deve ser baseado em seis intervalos de classe equiprováveis (sob  $H_0$ ), quais devem ser estes intervalos?

c. Utilize o teste qui-quadrado com os intervalos da parte (b) para decidir, com base nos seguintes diâmetros dos 45 parafusos, se o diâmetro do parafuso é uma variável distribuída normalmente com  $\mu = 0,5$  in.,  $\sigma = 0.002$  in.

0,4974	0,4976	0,4991	0,5014	0,5008	0,4993
0,4994	0,5010	0,4997	0,4993	0,5013	0,5000
0,5017	0,4984	0,4967	0,5028	0,4975	0,5013
0,4972	0,5047	0,5069	0,4977	0,4961	0,4987
0,4990	0,4974	0,5008	0,5000	0,4967	0,4977
0,4992	0,5007	0,4975	0,4998	0,5000	0,5008
0,5021	0,4959	0,5015	0,5012	0,5056	0,4991
0,5006	0,4987	0,4968			

## 14.2 Testes de Aderência para Hipóteses Compostas

Na seção anterior, apresentamos um teste de aderência baseado em uma estatística  $\chi^2$  para decidir entre  $H_0: p_1 = p_{10}, \dots, p_k = p_{k0}$  e a alternativa  $H_a$  definindo que  $H_0$  não é verdadeiro. A hipótese nula foi uma hipótese simples no sentido de que cada  $p_{i0}$  foi um número especificado, tal que as contagens de células esperadas quando  $H_0$  era verdadeiro foram unicamente números determinados.

Em muitas situações, há  $k$  categorias ocorrendo naturalmente, mas  $H_0$  define apenas que  $p_i$ s são funções de outros parâmetros  $\theta_1, \dots, \theta_m$ , sem especificar os valores destes  $\theta$ s. Por exemplo, uma população pode estar em equilíbrio em relação às proporções dos três genótipos AA, Aa e aa. Com  $p_1, p_2$  e  $p_3$  indicando estas proporções (probabilidades), pode-se desejar testar

$$H_0: p_1 = \theta^2, p_2 = 2\theta(1 - \theta), p_3 = (1 - \theta)^2 \quad (14.1)$$

onde  $\theta$  representa a proporção do gene A na população. Esta hipótese é composta, pois saber que  $H_0$  é verdadeiro não determina as probabilidades das células e contagens de células esperadas, mas apenas sua forma geral. Para realizar um teste  $\chi^2$ , os  $\theta$ s desconhecidos devem primeiramente ser estimados.

Da mesma maneira, podemos estar interessados em testar a fim de verificar se uma amostra provém de uma família particular de distribuições sem especificar qualquer membro específico da família. Para utilizar o teste  $\chi^2$  a fim de verificar se a distribuição é de Poisson, por exemplo, o parâmetro  $\lambda$  deve ser estimado. Além disso, como há na realidade um número infinito de valores possíveis de uma variável de Poisson, estes valores devem ser agrupados tal que exista um número finito de células. Se  $H_0$  define que a distribuição subjacente é normal, o uso de um teste  $\chi^2$  deve ser precedido por uma escolha de células e a estimativa de  $\mu$  e  $\sigma$ .

### $\chi^2$ Quando os Parâmetros São Estimados

Como dito antes,  $k$  indicará o número de categorias ou células e  $p_i$  indicará a probabilidade de uma observação cair na  $i$ -ésima célula. A hipótese nula agora define que cada  $p_i$  é uma função de um número pequeno de parâmetros  $\theta_1, \dots, \theta_m$  com os  $\theta$ s por outro lado não especificados:

$$H_0: p_1 = \pi_1(\boldsymbol{\theta}), \dots, p_k = \pi_k(\boldsymbol{\theta}) \quad \text{onde } \boldsymbol{\theta} = (\theta_1, \dots, \theta_m) \quad (14.2)$$

$H_a$ : a hipótese  $H_0$  não é verdadeira

Por exemplo, para  $H_0$  de (14.1)  $m = 1$  (há apenas um  $\theta$ ),  $\pi_1(\theta) = \theta^2$ ,  $\pi_2(\theta) = 2\theta(1 - \theta)$ , e  $\pi_3(\theta) = (1 - \theta)^2$ .

No caso  $k = 2$ , há na verdade apenas uma única variável aleatória  $N_1$  (desde que  $N_1 + N_2 = n$ ), que tem uma distribuição binomial. A probabilidade conjunta em que  $N_1 = n_1$  e  $N_2 = n_2$  é então

$$P(N_1 = n_1, N_2 = n_2) = \binom{n}{n_1} p_1^{n_1} \cdot p_2^{n_2} \propto p_1^{n_1} \cdot p_2^{n_2}$$

onde  $p_1 + p_2 = 1$  e  $n_1 + n_2 = n$ . Para um  $k$  qualquer, a distribuição conjunta de  $N_1, \dots, N_k$  é a distribuição multinomial (Seção 5.1) com

$$P(N_1 = n_1, \dots, N_k = n_k) \propto p_1^{n_1} \cdot p_2^{n_2} \cdot \dots \cdot p_k^{n_k} \quad (14.3)$$

Quando  $H_0$  é verdadeiro, (14.3) torna-se

$$P(N_1 = n_1, \dots, N_k = n_k) \propto [\pi_1(\boldsymbol{\theta})]^{n_1} \cdot \dots \cdot [\pi_k(\boldsymbol{\theta})]^{n_k} \quad (14.4)$$

Para aplicar o teste qui-quadrado,  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_m)$  deve ser estimado.

**MÉTODO DE ESTIMAÇÃO**

Sejam  $n_1, n_2, \dots, n_k$  indicar os valores observados de  $N_1, \dots, N_k$ . Então  $\hat{\theta}_1, \dots, \hat{\theta}_m$  são aqueles valores de  $\theta_i$ s que maximizam (14.4).

Os estimadores resultantes  $\hat{\theta}_1, \dots, \hat{\theta}_m$  são os estimadores de máxima verossimilhança de  $\theta_1, \dots, \theta_m$ ; este princípio de estimativa foi discutido na Seção 6.2.

**Exemplo 14.5**

Em humanos, há um grupo sanguíneo, o grupo MN, que é composto de indivíduos que têm um dos três tipos sanguíneos M, MN e N. O tipo é determinado por dois alelos, e não há dominância, então os três genótipos possíveis dão origem a três fenótipos. Uma população que consiste de indivíduos no grupo MN está em equilíbrio se

$$\begin{aligned} P(\text{M}) &= p_1 = \theta^2 \\ P(\text{MN}) &= p_2 = 2\theta(1 - \theta) \\ P(\text{N}) &= p_3 = (1 - \theta)^2 \end{aligned}$$

para algum  $\theta$ . Suponha uma amostra a partir da qual uma população produza os resultados mostrados na Tabela 14.4.

**Tabela 14.4** Contagens observadas para o Exemplo 14.5

Tipo	M	MN	N	
Observada	125	225	150	$n = 500$

Então

$$\begin{aligned} [\pi_1(\theta)]^{n_1} [\pi_2(\theta)]^{n_2} [\pi_3(\theta)]^{n_3} &= [(\theta^2)]^{n_1} [2\theta(1 - \theta)]^{n_2} [(1 - \theta)^2]^{n_3} \\ &= 2^{n_2} \cdot \theta^{2n_1 + n_2} \cdot (1 - \theta)^{n_2 + 2n_3} \end{aligned}$$

Maximizando essa expressão em relação a  $\theta$  (ou, de maneira equivalente, maximizando o logaritmo natural desta quantidade, o que é mais fácil do que diferenciar) produz

$$\hat{\theta} = \frac{2n_1 + n_2}{[(2n_1 + n_2) + (n_2 + 2n_3)]} = \frac{2n_1 + n_2}{2n}$$

Com  $n_1 = 125$  e  $n_2 = 225$ ,  $\hat{\theta} = 475/1000 = 0,475$ . ■

Uma vez que  $\theta = (\theta_1, \dots, \theta_m)$  tenha sido estimado por  $\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_m)$ , as contagens de células esperadas estimadas são  $n\pi_i(\hat{\theta})$ s. Estas são agora utilizadas no lugar dos  $np_{i0}$ s da Seção 14.1 para especificar uma estatística  $\chi^2$ .

**TEOREMA**

Sob condições gerais de “regularidade” em  $\theta_1, \dots, \theta_m$  e o  $\pi_i(\theta)$ s, se  $\theta_1, \dots, \theta_m$  são estimados pelo método de máxima verossimilhança como descrito previamente e  $n$  é grande,

$$\chi^2 = \sum_{\text{todas as células}} \frac{(\text{observado} - \text{estimado esperado})^2}{\text{estimado esperado}} = \sum_{i=1}^k \frac{[N_i - n\pi_i(\hat{\theta})]^2}{n\pi_i(\hat{\theta})}$$

tem aproximadamente uma distribuição qui-quadrado com  $k - 1 - m$  gl quando  $H_0$  de (14.2) é verdadeiro. Um nível aproximado de um teste  $\alpha$  de  $H_0$  versus  $H_a$  é então rejeitar  $H_0$  se  $\chi^2 \geq \chi_{\alpha, k-1-m}^2$ . Na prática, o teste pode ser utilizado se  $n\pi_i(\hat{\theta}) \geq 5$  para todo  $i$ .

Observe que o número de graus de liberdade é reduzido pelo número de  $\theta$ s estimado.

**Exemplo 14.6** (Continuação do Exemplo 14.5)

Com  $\hat{\theta} = 0,475$  e  $n = 500$ , as contagens de célula esperadas estimadas são  $n\pi_1(\hat{\theta}) = 500(\hat{\theta})^2 = 112,81$ ,  $n\pi_2(\hat{\theta}) = (500)(2)(0,475)(1 - 0,475) = 249,38$ , e  $n\pi_3(\hat{\theta}) = 500 - 112,81 - 249,38 = 137,81$ . Então

$$\chi^2 = \frac{(125 - 112,81)^2}{112,81} + \frac{(225 - 249,38)^2}{249,38} + \frac{(150 - 137,81)^2}{137,81} = 4,78$$

Como  $\chi^2_{0,05,k-1-m} = \chi^2_{0,05,3-1-1} = \chi^2_{0,05,1} = 3,843$  e  $4,78 \geq 3,843$ ,  $H_0$  é rejeitado. A Tabela A.11 do Apêndice mostra que o valor  $P \approx 0,029$ . ■

**Exemplo 14.7**

Considere uma série de jogos entre dois times, I e II, que termina assim que um dos times tenha ganhado quatro jogos (sem possibilidade de empate). Um modelo de probabilidade simples para uma série assume que os resultados de jogos sucessivos são independentes e que a probabilidade do time I ganhar qualquer jogo em particular é constante e igual a  $\theta$ . Designamos arbitrariamente I como o melhor time, tal que  $\theta \geq 0,5$ . Qualquer série em particular pode então terminar depois de 4, 5, 6 ou 7 jogos. Sejam  $\pi_1(\theta)$ ,  $\pi_2(\theta)$ ,  $\pi_3(\theta)$ ,  $\pi_4(\theta)$  as probabilidades de terminar em 4, 5, 6 e 7 jogos, respectivamente. Então

$$\begin{aligned} \pi_1(\theta) &= P(\text{I ganhar em 4 jogos}) + P(\text{II ganhar em 4 jogos}) \\ &= \theta^4 + (1 - \theta)^4 \\ \pi_2(\theta) &= P(\text{I ganhar 3 dos primeiros 4 e o quinto}) \\ &\quad + P(\text{I perder 3 dos primeiros 4 e o quinto}) \\ &= \binom{4}{3} \theta^3(1 - \theta) \cdot \theta + \binom{4}{1} \theta(1 - \theta)^3 \cdot (1 - \theta) \\ &= 4\theta(1 - \theta)[\theta^3 + (1 - \theta)^3] \\ \pi_3(\theta) &= 10\theta^2(1 - \theta)^2[\theta^2 + (1 - \theta)^2] \\ \pi_4(\theta) &= 20\theta^3(1 - \theta)^3 \end{aligned}$$

O artigo “Seven-Game Series in Sports” escrito por Groeneveld e Meeden (*Mathematics Magazine*, 1975, p. 187-192) testou o ajuste deste modelo aos resultados das finais da Liga de Hóquei Nacional durante o período de 1943-1967 (quando os membros da liga eram estáveis). Os dados aparecem na Tabela 14.5.

**Tabela 14.5** Contagens observadas e esperadas para o modelo simples

Célula	1	2	3	4	
Número de jogos realizados	4	5	6	7	
Frequência observada	15	26	24	18	$n = 83$
Frequência esperada estimada	16,351	24,153	23,240	19,256	

As contagens de células esperadas estimadas são  $83\pi_i(\hat{\theta})$ , onde  $\hat{\theta}$  é o valor de  $\theta$  que maximiza

$$\begin{aligned} &\{\theta^4 + (1 - \theta)^4\}^{15} \cdot \{4\theta(1 - \theta)[\theta^3 + (1 - \theta)^3]\}^{26} \\ &\quad \cdot \{10\theta^2(1 - \theta)^2[\theta^2 + (1 - \theta)^2]\}^{24} \cdot \{20\theta^3(1 - \theta)^3\}^{18} \end{aligned} \tag{14.5}$$

Os métodos de cálculo-padrão falham em produzir uma boa fórmula para a maximização do valor  $\hat{\theta}$ , então deve ser calculado com a utilização dos métodos numéricos. O resultado é  $\hat{\theta} = 0,654$ , a partir do qual  $\pi_i(\hat{\theta})$  e as

contagens de células esperadas estimadas são calculadas. O valor calculado de  $\chi^2$  é 0,360 e (como  $k - 1 - m = 4 - 1 - 1 = 2$ )  $\chi^2_{0,10,2} = 4,605$ . Não há, dessa maneira, razão para rejeitar o modelo simples como aplicado às séries finais da LHN.

O artigo citado também considera os dados das Séries Mundiais para o período de 1903-1973. Para o modelo simples,  $\chi^2 = 5,97$ , então o modelo não parece apropriado. A razão sugerida para isto é que para o modelo simples

$$P(\text{séries com duração de seis jogos} \mid \text{séries com duração de pelo menos seis jogos}) \geq 0,5 \tag{14.6}$$

onde das 38 séries que realmente tiveram pelo menos seis jogos de duração, apenas 13 tiveram a duração de exatamente seis. O seguinte modelo alternativo é então apresentado:

$$\begin{aligned} \pi_1(\theta_1, \theta_2) &= \theta_1^4 + (1 - \theta_1)^4 \\ \pi_2(\theta_1, \theta_2) &= 4\theta_1(1 - \theta_1)[\theta_1^3 + (1 - \theta_1)^3] \\ \pi_3(\theta_1, \theta_2) &= 10\theta_1^2(1 - \theta_1)^2\theta_2 \\ \pi_4(\theta_1, \theta_2) &= 10\theta_1^2(1 - \theta_1)^2(1 - \theta_2) \end{aligned}$$

Os dois primeiros  $\pi_i$ s são idênticos ao modelo simples, enquanto  $\theta_2$  é a probabilidade condicional de (14.6) (que pode agora ser qualquer número entre 0 e 1). Os valores de  $\hat{\theta}_1$  e  $\hat{\theta}_2$  que maximizam a expressão análoga à expressão (14.5) são determinados numericamente como  $\hat{\theta}_1 = 0,614$ ,  $\hat{\theta}_2 = 0,342$ . Um resumo aparece na Tabela 14.6, e  $\chi^2 = 0,384$ . Como dois parâmetros são estimados,  $gl = k - 1 - m = 1$  com  $\chi^2_{0,10,1} = 2,706$ , indicando uma boa aderência dos dados a este novo modelo.

**Tabela 14.6** Contagens observadas e esperadas para o modelo mais complexo

Número de jogos realizados	4	5	6	7
Frequência observada	12	16	13	25
Frequência esperada estimada	10,85	18,08	12,68	24,39

Uma das condições de regularidade sobre os  $\theta_i$ s no teorema é que eles são funcionalmente independentes um do outro. Ou seja, um único  $\theta_i$  não pode ser determinado a partir do valor de outro  $\theta_j$ , tal que  $m$  é o número de parâmetros de funcionalidade independentes estimado. Uma regra prática geral para graus de liberdade em testes qui-quadrado é a seguinte.

$$\chi^2 \text{ df} = \left( \begin{array}{c} \text{contagens de células} \\ \text{determinadas livremente} \end{array} \right) - \left( \begin{array}{c} \text{número parâmetros} \\ \text{independentes estimados} \end{array} \right)$$

Esta regra será utilizada em conexão com vários testes qui-quadrado diferentes na próxima seção.

### Aderência para Distribuições Discretas

Muitos experimentos envolvem a observação de uma amostra aleatória  $X_1, X_2, \dots, X_n$ , a partir de alguma distribuição discreta. Pode-se então desejar investigar se a distribuição em questão é um membro de uma família particular, tal como a família de Poisson ou a binomial negativa. No caso de ambas as distribuições de Poisson e binomial negativa, o conjunto de valores possíveis é infinito, então os valores devem ser agrupados em  $k$  subconjuntos antes de um teste qui-quadrado poder ser utilizado. Os agrupamentos devem ser feitos para que a frequência esperada em cada célula (grupo) seja pelo menos 5. A última célula corresponderá então aos valores de  $X$  de  $c, c + 1, c + 2, \dots$  para algum valor  $c$ .



Este agrupamento pode complicar consideravelmente o cálculo dos  $\hat{\theta}_i$ s e das contagens de células esperadas estimadas. Isto porque o teorema requer que os  $\hat{\theta}_i$ s sejam obtidos das contagens de células  $N_1, \dots, N_k$ , em vez de valores de amostra  $X_1, \dots, X_n$ .

**Exemplo 14.8**

A Tabela 14.7 apresenta dados de contagem sobre o número de plantas *Larrea divaricata* encontradas em cada um dos 48 quadrantes de amostra, como informado no artigo “Some Sampling Characteristics of Plants and Arthropods of the Arizona Desert” (*Ecology*, 1962, p. 567-571).

**Tabela 14.7** Contagens observadas para o Exemplo 14.8

Célula	1	2	3	4	5
Número de plantas	0	1	2	3	$\geq 4$
Frequência	9	9	10	14	6

O autor ajusta uma distribuição de Poisson aos dados. Seja  $\lambda$  o parâmetro de Poisson e suponha por enquanto que as seis contagens na célula 5 foram realmente 4, 4, 5, 5, 6, 6. Então, indicando os valores da amostra por  $x_1, \dots, x_{48}$ , nove dos  $x_i$ s foram 0, nove foram 1 e assim por diante. A função de verossimilhança da amostra observada é

$$\frac{e^{-\lambda}\lambda^{x_1}}{x_1!} \dots \frac{e^{-\lambda}\lambda^{x_{48}}}{x_{48}!} = \frac{e^{-48\lambda}\lambda^{\sum x_i}}{x_1! \dots x_{48}!} = \frac{e^{-48\lambda}\lambda^{101}}{x_1! \dots x_{48}!}$$

O valor de  $\lambda$  para o qual ele é maximizado é  $\hat{\lambda} = \sum x_i/n = 101/48 = 2,10$  (o valor informado no artigo).

Entretanto, o  $\hat{\lambda}$  exigido para  $\chi^2$  é obtido pela maximização da Expressão (14.4) em vez da função de verossimilhança da amostra toda. As probabilidades de célula são

$$\pi_i(\lambda) = \frac{e^{-\lambda}\lambda^{i-1}}{(i-1)!} \quad i = 1, 2, 3, 4$$

$$\pi_5(\lambda) = 1 - \sum_{i=0}^3 \frac{e^{-\lambda}\lambda^i}{i!}$$

então o lado direito de (14.4) torna-se

$$\left[ \frac{e^{-\lambda}\lambda^0}{0!} \right]^9 \left[ \frac{e^{-\lambda}\lambda^1}{1!} \right]^9 \left[ \frac{e^{-\lambda}\lambda^2}{2!} \right]^{10} \left[ \frac{e^{-\lambda}\lambda^3}{3!} \right]^{14} \left[ 1 - \sum_{i=0}^3 \frac{e^{-\lambda}\lambda^i}{i!} \right]^6$$

Não há uma boa fórmula para  $\hat{\lambda}$ . O valor de maximização de  $\lambda$ , nesta última expressão, portanto, deve ser obtido numericamente. ■

*Como as estimativas dos parâmetros normalmente são muito mais difíceis de serem calculadas a partir de dados agrupados do que de amostras inteiras, elas são geralmente calculadas utilizando-se o método anterior. Quando estes estimadores “completos” são utilizados na estatística qui-quadrado, a distribuição da estatística é alterada e o teste nível  $\alpha$  não é mais especificado pelo valor crítico  $\chi^2_{\alpha, k-1-m}$ .*

**TEOREMA**

Sejam  $\hat{\theta}_1, \dots, \hat{\theta}_m$  os estimadores de máxima verossimilhança de  $\theta_1, \dots, \theta_m$ , baseado na amostra completa  $X_1, \dots, X_n$ , e seja  $\chi^2$  os dados estatísticos baseados nestes estimadores. Então o valor crítico  $c_\alpha$  especifica um teste de cauda superior nível  $\alpha$  que satisfaz

$$\chi^2_{\alpha, k-1-m} \leq c_\alpha \leq \chi^2_{\alpha, k-1} \tag{14.7}$$

O procedimento de teste implicado por este teorema é o seguinte:

Se  $\chi^2 \geq \chi^2_{\alpha, k-1}$ , rejeita  $H_0$ .  
 Se  $\chi^2 \leq \chi^2_{\alpha, k-1-m}$ , não rejeita  $H_0$ .  
 Se  $\chi^2_{\alpha, k-1-m} < \chi^2 < \chi^2_{\alpha, k-1}$ , retarda o julgamento. (14.8)

**Exemplo 14.9** (Continuação do Exemplo 14.8)

Utilizando  $\hat{\lambda} = 2,10$ , as contagens de célula esperadas estimadas são calculadas a partir de  $n\pi_i(\hat{\lambda})$ , onde  $n = 48$ . Por exemplo,

$$n\pi_1(\hat{\lambda}) = 48 \cdot \frac{e^{-2,1}(2,1)^0}{0!} = (48)(e^{-2,1}) = 5,88$$

Da mesma maneira,  $n\pi_2(\hat{\lambda}) = 12,34$ ,  $n\pi_3(\hat{\lambda}) = 12,96$ ,  $n\pi_4(\hat{\lambda}) = 9,07$  e  $n\pi_5(\hat{\lambda}) = 48 - 5,88 - \dots - 9,07 = 7,75$ . Então

$$\chi^2 = \frac{(9 - 5,88)^2}{5,88} + \dots + \frac{(6 - 7,75)^2}{7,75} = 6,31$$

Como  $m = 1$  e  $k = 5$ , no nível 0,05 precisamos de  $\chi^2_{0,05,3} = 7,815$  e  $\chi^2_{0,05,4} = 9,488$ . Como  $6,31 \leq 7,815$ , não rejeitamos  $H_0$ ; ao nível de 5%, a distribuição de Poisson fornece um ajuste razoável aos dados. Observe que  $\chi^2_{0,10,3} = 6,251$  e  $\chi^2_{0,10,4} = 7,779$ , portanto no nível 0,10 teremos que retardar a avaliação se a distribuição de Poisson foi apropriada. ■

Algumas vezes, mesmo com as estimativas de máxima verossimilhança baseadas na amostra inteira são difíceis de serem calculadas. Este é o caso, por exemplo, dos dois parâmetros para a distribuição binomial negativa (generalizada). Em tais situações, as estimativas do método dos momentos são frequentemente utilizadas e o  $\chi^2$  resultante comparado ao  $\chi^2_{\alpha, k-1-m}$ , embora não se saiba como o uso dos estimadores de momentos afeta o valor crítico verdadeiro.

**Aderência para Distribuições Discretas**

O teste qui-quadrado pode também ser utilizado para testar se uma amostra provém de uma família específica de distribuições contínuas, tais como a família exponencial ou a família normal. A escolha de células (intervalos de classe) é ainda mais arbitrária no caso contínuo do que no caso discreto. Para garantir que o teste qui-quadrado seja válido, as células devem ser escolhidas de maneira independente de observações de amostra. Uma vez que escolhidas, é quase sempre difícil estimar parâmetros não especificados (tais como  $\mu$  e  $\sigma$  no caso normal) a partir de contagens de células observadas, então, em vez disso, os estimadores de máxima verossimilhança baseados em amostras inteiras são calculados. O valor crítico  $c_\alpha$  novamente satisfaz (14.7) e o procedimento de teste é dado por (14.8).

**Exemplo 14.10**

O The Institute of Nutrition of Central America and Panama (INCAP) realizou estudos extensivos sobre dietas e projetos de pesquisa na América Central. Em um estudo informado em novembro de 1964, publicado no *American Journal of Clinical Nutrition* (“The Blood Viscosity of Various Socioeconomic Groups in Guatemala”), medidas de colesterol total de soro para uma amostra de 49 índios de baixa renda da zona rural foram informadas como segue (em mg/L):

204	108	140	152	158	129	175	146	157	174	192	194	144	152	135	223	145
231	115	131	129	142	114	173	226	155	166	220	180	172	143	148	171	143
124	158	144	108	189	136	136	197	131	95	139	181	165	142	162		

É plausível que o nível de colesterol no soro seja normalmente distribuído por esta população? Suponha que previamente para a amostragem acreditava-se que os valores plausíveis para  $\mu$  e  $\sigma$  eram de 150 e 30, respectivamente. Os sete intervalos de classe equiprováveis para a distribuição normal padrão são:  $(-\infty; -1,07)$ ,  $(-1,07; -0,57)$ ,  $(-0,57; -0,18)$ ,  $(-0,18; 0,18)$ ,  $(0,18; 0,57)$ ,  $(0,57; 1,07)$  e  $(1,07; \infty)$  com cada extremidade fornecendo também a distância em desvios padrão em relação à média para qualquer outra distribuição normal. Para  $\mu = 150$  e  $\sigma = 30$ , estes intervalos são  $(-\infty; 117,9)$ ,  $(117,9; 132,9)$ ,  $(132,9; 144,6)$ ,  $(144,6; 155,4)$ ,  $(155,4; 167,1)$ ,  $(167,1; 182,1)$  e  $(182,1; \infty)$ .

Para obter as probabilidades de células estimadas  $\pi_1(\hat{\mu}, \hat{\sigma}), \dots, \pi_7(\hat{\mu}, \hat{\sigma})$ , primeiro precisamos dos estimadores de máxima verossimilhança  $\hat{\mu}$  e  $\hat{\sigma}$ . No Capítulo 6, o estimador de máxima verossimilhança de  $\sigma$  foi mostrado como  $[\sum(x_i - \bar{x})^2/n]^{1/2}$  (ao invés de  $s$ ), tal que  $s = 31,75$ ,

$$\hat{\mu} = \bar{x} = 157,02 \quad \hat{\sigma} = \left[ \frac{\sum(x_i - \bar{x})^2}{n} \right]^{1/2} = \left[ \frac{(n - 1)s^2}{n} \right]^{1/2} = 31,42$$

Cada  $\pi_i(\hat{\mu}, \hat{\sigma})$  é então a probabilidade de uma  $va X$  normal com média 157,02 e desvio padrão 31,42 cair no  $i$ -ésimo intervalo de classe. Por exemplo,

$$\pi_2(\hat{\mu}, \hat{\sigma}) = P(117,9 \leq X \leq 132,9) = P(-1,25 \leq Z \leq -0,77) = 0,1150$$

então  $n\pi_2(\hat{\mu}, \hat{\sigma}) = 49(0,1150) = 5,64$ . As contagens de células observadas e estimadas são mostradas na Tabela 14.8.

O  $\chi^2$  calculado é 4,60. Com  $k = 7$  células e  $m = 2$  parâmetros estimados,  $\chi^2_{0,05,k-1} = \chi^2_{0,05,6} = 12,592$  e  $\chi^2_{0,05,k-1-m} = \chi^2_{0,05,4} = 9,488$ . Como  $4,60 \leq 9,488$ , uma distribuição normal fornece um ajuste razoável para os dados.

**Tabela 14.8** Contagens de células observadas e esperadas para o Exemplo 14.10

Célula	$(-\infty; 117,9)$	$(117,9; 132,9)$	$(132,9; 144,6)$	$(144,6; 155,4)$
Observada	5	5	11	6
Estimado esperado	5,17	5,64	6,08	6,64

Célula	$(155,4; 167,1)$	$(167,1; 182,1)$	$(182,1; \infty)$
Observada	6	7	9
Estimado esperado	7,12	7,97	10,38

**Exemplo 14.11**

O artigo “Some Studies on Tuft Weight Distribution in the Opening Room” (*Textile Research J.*, 1976, p. 567-573) informa os dados de acompanhamento da distribuição do peso dos ramos de saída  $X$  (mg) das fibras de algodão para o peso de entrada  $x_0 = 70$ .

Intervalo	0-8	8-16	16-24	24-32	32-40	40-48	48-56	56-64	64-70
Frequência observada	20	8	7	1	2	1	0	1	0
Frequência esperada	18,0	9,9	5,5	3,0	1,8	0,9	0,5	0,3	0,1

Os autores postularam uma distribuição exponencial truncada:

$$H_0: f(x) = \frac{\lambda e^{-\lambda x}}{1 - e^{-\lambda x_0}} \quad 0 \leq x \leq x_0$$

A média desta distribuição é

$$\mu = \int_0^{x_0} x f(x) dx = \frac{1}{\lambda} - \frac{x_0 e^{-\lambda x_0}}{1 - e^{-\lambda x_0}}$$

O parâmetro  $\lambda$  foi estimado pela substituição de  $\mu$  por  $\bar{x} = 13,086$  e a equação resultante foi resolvida para obter  $\hat{\lambda} = 0,0742$  (então  $\hat{\lambda}$  é uma estimativa pelo método dos momentos e não um estimador de máxima verossimilhança). Então, com  $\hat{\lambda}$  substituindo  $\lambda$  em  $f(x)$ , as freqüências de células esperadas estimadas como mostradas anteriormente são calculadas como

$$40\pi_i(\hat{\lambda}) = 40P(a_{i-1} \leq X < a_i) = 40 \int_{a_{i-1}}^{a_i} f(x) dx = \frac{40(e^{-\hat{\lambda}a_{i-1}} - e^{-\hat{\lambda}a_i})}{1 - e^{-\hat{\lambda}x_0}}$$

onde  $[a_{i-1}, a_i)$  é o  $i$ -ésimo intervalo de classe. Para obter as contagens de célula esperadas de pelo menos 5, as últimas seis células são combinadas para produzir as contagens observadas de 20, 8, 7, 5 e as contagens esperadas de 18,0; 9,9; 5,5; 6,6. O valor calculado de qui-quadrado é então  $\chi^2 = 1,34$ . Como  $\chi_{0,05,2}^2 = 5,992$ ,  $H_0$  não é rejeitada, então o modelo exponencial truncado fornece um bom ajuste. ■

### Um Teste Especial para Normalidade

Gráficos de probabilidade foram introduzidos na Seção 4.6 como um método informal para avaliar a plausibilidade de qualquer distribuição de população especificada como aquela a partir da qual a amostra dada foi selecionada. Quanto maior a linearidade do gráfico de probabilidade, mais plausível é a distribuição na qual o gráfico está baseado. Um gráfico de probabilidade normal é utilizado para verificar se qualquer membro de uma família de distribuição normal é plausível. Representemos a amostra  $x_i$  quando ordenada a partir do menor para o maior por  $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ . Então o gráfico sugerido para a verificação da normalidade foi um gráfico de pontos  $(x_{(i)}, y_i)$ , onde  $y_i = \Phi^{-1}((i - 0,5)/n)$ .

Uma medida quantitativa da extensão nos quais agrupamentos de pontos em torno de uma reta é o coeficiente de correlação da amostra  $r$  introduzido no Capítulo 12. Considere o cálculo de  $r$  para  $n$  pares  $(x_{(1)}, y_1), \dots, (x_{(n)}, y_n)$ . Os  $y_i$ s aqui não são valores observados em uma amostra aleatória de uma população  $y$ , então as propriedades deste  $r$  são muito diferentes daquelas descritas na Seção 12.5. Entretanto, é verdade que quanto maior os desvios  $r$  de 1, menor a probabilidade de o gráfico parecer-se com uma reta (lembre-se de que um gráfico de probabilidade deve ter coeficiente angular positivo). Esta idéia pode ser estendida para produzir um procedimento de teste formal: Rejeitar a hipótese da normalidade da população se  $r \leq c_\alpha$ , onde  $c_\alpha$  é um valor crítico escolhido para produzir o nível  $\alpha$  de significância desejado. Ou seja, o valor crítico é escolhido de modo que quando a distribuição de população é realmente normal, a probabilidade de se obter um valor, que é no máximo  $c_\alpha$  (e assim rejeitando incorretamente  $H_0$ ), é o  $\alpha$  desejado. Os desenvolvedores do software estatístico de computador MINITAB forneceram valores críticos para  $\alpha = 0,10; 0,05; e 0,01$  em combinação com os tamanhos de amostras diferentes. Estes valores críticos estão baseados em uma definição ligeiramente diferente dos  $y_i$ s do que a previamente dada.

O MINITAB construirá também um gráfico de probabilidade normal baseado nestes  $y_i$ s. O gráfico será quase idêntico na aparência àquele baseado nos  $y_i$ s anteriores. Quando houver vários  $x_{(i)}$ s juntos, o MINITAB calcula  $r$  utilizando a média dos  $y_i$ s correspondentes como o segundo número em cada par.

Seja  $y_i = \Phi^{-1}[(i - 0,375)/(n + 0,25)]$  e calcule o coeficiente de correlação da amostra de  $r$  para  $n$  pares  $(x_{(1)}, y_1), \dots, (x_{(n)}, y_n)$ . O teste Ryan-Joiner de

$H_0$ : a distribuição de população é normal

versus

$H_a$ : a distribuição de população não é normal

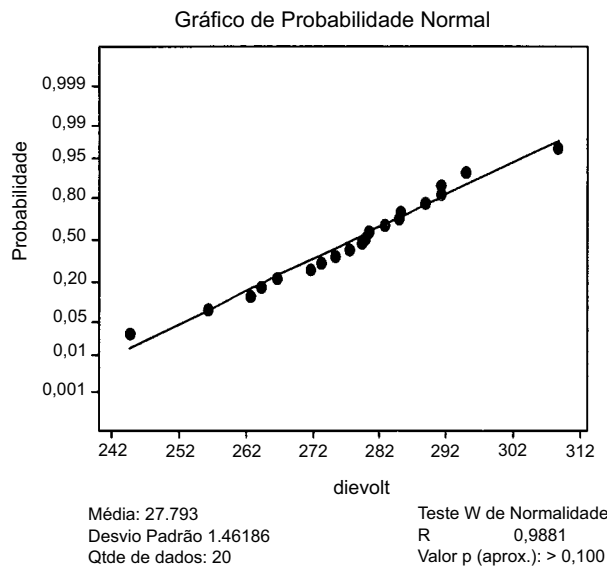
consiste da rejeição de  $H_0$  quando  $r \leq c_\alpha$ . Valores críticos de  $c_\alpha$  são dados na Tabela A.12 do Apêndice para vários níveis de significância  $\alpha$  e tamanhos  $n$  de amostras.

**Exemplo 14.12**

A seguinte amostra de  $n = 20$  observações de voltagem dielétrica de quebra de peças de resina de epóxi apareceu anteriormente no Exemplo 4.29

$y_i$	-1,871	-1,404	-1,127	-0,917	-0,742	-0,587	-0,446	-0,313	-0,186	-0,062
$x_{(i)}$	24,46	25,61	26,25	26,42	26,66	27,15	27,31	27,54	27,74	27,94
$y_i$	0,062	0,186	0,313	0,446	0,587	0,742	0,917	1,127	1,404	1,871
$x_{(i)}$	27,98	28,04	28,28	28,49	28,50	28,87	29,11	29,13	29,50	30,88

Realizou-se no MINITAB o teste de Ryan-Joiner, e o resultado aparece na Figura 14.3. O valor da estatística do teste é  $r = 0,9881$ , e a Tabela A.12 do Apêndice fornece 0,9600 como o valor crítico 0,10 corresponde à área sob cauda inferior da curva de distribuição de amostragem  $r$  quando  $n = 20$  e a distribuição em questão é realmente normal. Como  $0,9881 > 0,9600$ , a hipótese nula de normalidade não pode ser rejeitada mesmo para um nível de significância tão grande quanto 0,10.



**Figura 14.3** A saída do MINITAB a partir do teste de Ryan-Joiner para os dados do Exemplo 14.12

**Exercícios** | Seção 14.2 (12–23)

12. Considere uma grande população de famílias nas quais cada uma tem exatamente três crianças. Se os sexos das três crianças em qualquer família são independentes uns dos outros, o número de meninos em uma família selecionada aleatoriamente terá uma distribuição binomial baseada em três tentativas.

a. Suponha que uma amostra aleatória de 160 famílias produza os seguintes resultados. Teste a hipótese relevante procedendo como no Exemplo 14.5.

<b>Número de Meninos</b>	0	1	2	3
<b>Frequência</b>	14	66	64	16

b. Suponha que uma amostra aleatória de família em uma população não-humana tenha resultado em frequências observadas de 15, 20, 12 e 3, respectivamente. O teste qui-quadrado poderia ser baseado no mesmo número de graus de liberdade do teste do item (a)? Explique.

13. Um estudo de esterilidade na mosca de fruta (“Hybrid Dysgenesis em *Drosophila melanogaster*: The Biology of Female and Male Sterility”, *Genetics*, 1979, p. 161-174) informa os seguintes dados sobre o número de ovários desenvolvidos por cada mosca fêmea em uma amostra de tamanho 1.388. Um modelo para esterilidade unilateral afirma que cada ovário desenvolve-se com alguma probabilidade  $p$  independentemente do outro ovário. Testar a aderência deste modelo utilizando  $\chi^2$ .

<b>x = Número de Ovários Desenvolvidos</b>	0	1	2
<b>Contagem Observada</b>	1.212	118	58

14. O artigo “Feeding Ecology of the Red-Eyed Vireo and Associated Foliage-Gleaning Birds”) (*Ecological Monographs*, 1971, p. 129-152) apresenta os dados associados à variável  $X$  = o número de saltos antes do primeiro vôo e precedido de outro vôo. O autor então propôs e ajustou uma distribuição de probabilidade geométrica [ $p(x) = P(X = x) = p^{x-1} \cdot q$  para  $x = 1, 2, \dots$  em que  $q = 1 - p$ ], aos dados. O tamanho total da amostra foi  $n = 130$ .

<b>x</b>	1	2	3	4	5	6	7	8	9	10	11	12
<b>Número de Vezes que x foi Observada</b>	48	31	20	9	6	5	4	2	1	1	2	1

a. A função de verossimilhança é  $(p^{x-1} \cdot q) \cdot \dots \cdot (p^{x-1} \cdot q) = p^{\sum x_i - n} \cdot q^n$ . Mostre que o estimador de máxima verossimilhança de  $p$  é dado por  $\hat{p} = (\sum x_i - n) / \sum x_i$ , e calcule  $\hat{p}$  para os dados fornecidos.

b. Estime as contagens de células esperadas utilizando  $\hat{p}$  do item (a) [contagens de células espe-

radas =  $n \cdot (\hat{p})^{x-1} \cdot \hat{q}$  para  $x = 1, 2, \dots$ ], e teste o ajuste do modelo utilizando um teste  $\chi^2$  combinando as contagens para  $x = 7, 8, \dots$ , e 12 dentro de uma célula ( $x \geq 7$ ).

15. Um determinado tipo de lanterna é vendido com quatro baterias incluídas. Uma amostra aleatória de 150 lanternas é obtida, e o número de baterias com defeito em cada uma é determinado, resultando nos seguintes dados:

<b>Número com Defeito</b>	0	1	2	3	4
<b>Frequência</b>	26	51	47	16	10

Seja  $X$  o número de baterias com defeito em uma amostra selecionada aleatoriamente de lanternas. Teste a hipótese nula de que a distribuição de  $X$  é  $\text{Bin}(4, \theta)$ . Ou seja, com  $p_i = P(i \text{ defeitos})$ , teste

$$H_0: p_i = \binom{4}{i} \theta^i (1 - \theta)^{4-i} \quad i = 0, 1, 2, 3, 4$$

[*Sugestão*: Para obter o estimador de máxima verossimilhança de  $\theta$ , escreva a função de verossimilhança (a função a ser maximizada) como  $\theta^u (1 - \theta)^v$ , onde os expoentes  $u$  e  $v$  são funções lineares das contagens de células. Então tome o log natural, diferencie em relação a  $\theta$ , iguale o resultado a 0 e resolva para  $\hat{\theta}$ ].

16. Em um experimento genético, os pesquisadores observaram 300 cromossomos de um tipo particular e contaram o número de cromátides irmãs trocadas em cada um (“On the Nature of Sister-Chromatid Exchanges in 5-Bromodeoxyuridine Substituted Chromosomes”, *Genetics*, 1979, p. 1251-1264). Um modelo de Poisson foi suposto para a distribuição do número de trocas. Testar o ajuste de uma distribuição de Poisson aos dados estimando primeiramente  $\lambda$  e então combinando as contagens para  $x = 8$  e  $x = 9$  em uma célula.

<b>X = Número de Trocas</b>	0	1	2	3	4	5	6	7	8	9
<b>Contagens Observadas</b>	6	24	42	59	62	44	41	14	6	2

17. Um artigo em *Annals of Mathematical Statistics* informa os seguintes dados sobre o número de moluscos em cada 120 grupos de moluscos. A fmp de Poisson fornece um modelo plausível para a distribuição do número de moluscos verdadeiros em um grupo? (*Sugestão*: Some as frequências para 7, 8, ..., 12, para estabelecer uma única categoria “ $\geq 7$ ”).

<b>Número de Moluscos</b>	0	1	2	3	4	5	6	7	8	9	10	11	12
<b>Frequência</b>	24	16	16	18	15	9	6	5	3	4	3	0	1

18. O artigo “A Probabilistic Analysis of Dissolved Oxygen-Biochemical Oxygen Demand Relationship in Streams” (*J. Water Resources Control Fed.*, 1969, p.73-90) informa os dados sobre a razão de oxigenação em fluxos a 20°C em uma determinada região. A média e o desvio padrão da amostra foram calculados como  $\bar{x} = 0,173$  e  $s = 0,066$ , respectivamente. Baseado na distribuição de frequência associada, pode-se concluir que a razão de oxigenação é uma variável distribuída normalmente? Utilize o teste qui-quadrado no nível  $\alpha = 0,05$ .

Razão (por dia)	Frequência
Abaixo de 0,100	12
0,100 – abaixo de 0,150	20
0,150 – abaixo de 0,200	23
0,200 – abaixo de 0,250	15
0,250 ou mais	13

19. Cada farol de automóvel passa por uma inspeção anual que pode ser classificada em muito alta ( $H$ ), muito baixa ( $L$ ) ou apropriada ( $N$ ). Verificando os dois faróis simultaneamente (e não fazendo distinção entre esquerda e direita) os resultados são dados em seis possibilidades  $HH, LL, NN, HL, HN$  e  $LN$ . Se as probabilidades (proporções de população) para a única direção do foco do farol são  $P(H) = \theta_1, P(L) = \theta_2$  e  $P(N) = 1 - \theta_1 - \theta_2$  e a classificação dos dois faróis independem uma da outra, as probabilidades dos seis resultados para uma seleção aleatória de um carro são as seguintes:

$$\begin{aligned}
 p_1 &= \theta_1^2 & p_2 &= \theta_2^2 & p_3 &= (1 - \theta_1 - \theta_2)^2 \\
 p_4 &= 2\theta_1\theta_2 & p_5 &= 2\theta_1(1 - \theta_1 - \theta_2) \\
 p_6 &= 2\theta_2(1 - \theta_1 - \theta_2)
 \end{aligned}$$

Utilize os dados correspondentes para testar a hipótese nula

$$H_0: p_1 = \pi_1(\theta_1, \theta_2), \dots, p_6 = \pi_6(\theta_1, \theta_2)$$

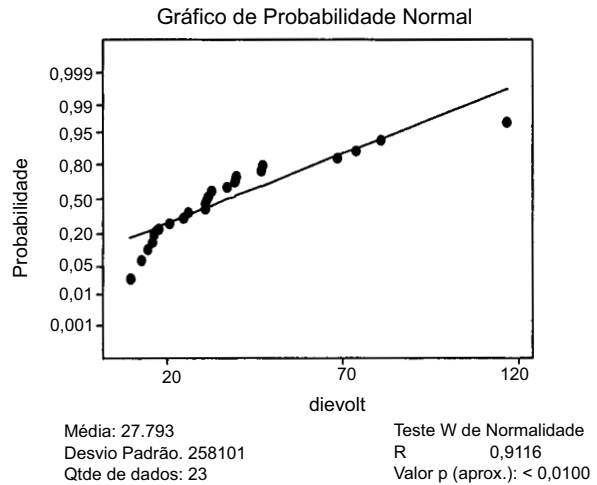
onde o  $\pi_i(\theta_1, \theta_2)$ s são dados previamente

Saída	HH	LL	NN	HL	HN	LN
Frequência	49	26	14	20	53	38

(Sugestão: Escreva a função de verossimilhança como uma função de  $\theta_1$  e  $\theta_2$ , calcule o log natural e então calcule  $\partial/\partial\theta_1$  e  $\partial/\partial\theta_2$ , iguale-os a 0 e resolva para  $\hat{\theta}_1, \hat{\theta}_2$ .)

20. O artigo “Compatibility of Outer and Fusible Interlining Fabrics in Tailored Garments” (*Textile Res. J.*, 1997, p. 137-142) fornece as seguintes observações de rigidez de dobramento ( $\mu N \cdot m$ ) para tipos de tecidos de média qualidade, a partir dos quais os resultados do MINITAB correspondentes foram obtidos:

24,6	12,7	14,4	30,6	16,1	9,5	31,5	17,2
46,9	68,3	30,8	116,7	39,5	73,8	80,6	20,3
25,8	30,9	39,2	36,8	46,6	15,6	32,3	



Você usaria um intervalo de confiança  $t$  de uma amostra para estimar a média real da rigidez de dobramento? Explique seu raciocínio.

21. O artigo a partir do qual os dados no Exercício 20 foram obtidos também forneceu os dados correspondentes à razão massa de composição/massa de saída de tipos do tecido de alta qualidade.

1,15	1,40	1,34	1,29	1,36	1,26	1,22	1,40
1,29	1,41	1,32	1,34	1,26	1,36	1,36	1,30
1,28	1,45	1,29	1,28	1,38	1,55	1,46	1,32

O MINITAB forneceu  $r = 0,9852$  como o valor do teste estatístico de Ryan-Joiner e informou que o valor  $P > 0,10$ . Você usaria o teste  $t$  de uma amostra para checar a hipótese a cerca do valor da razão média real? Por quê?

22. Os conjuntos de dados a seguir consistem de 25 observações sobre resistência a fratura do prato de base do aço temperado com níquel a 18% (de “Fracture Testing of Weldments”, *ASTM Special Publ. No. 381*, 1965, p. 328-356). Utilize o gráfico de probabilidade normal do teste de coeficiente de correlação para decidir se uma distribuição normal fornece um modelo plausível para resistência à fratura.

69,5	71,9	72,6	73,1	73,3	73,5	74,1	74,2	75,3
75,5	75,7	75,8	76,1	76,2	76,9	77,0	77,9	78,1
79,6	79,7	79,9	80,1	82,2	83,7	93,7		

23. O artigo “Nonbloated Burned Clay Aggregate Concrete” (*J. Materials*, 1972, p. 555-563) informa os seguintes dados sobre a resistência à flexão em amostras do concreto não-inchado agregado com argila (psi) com 7 dias:

257	327	317	300	340	340	343	374	377	386
383	393	407	407	434	427	440	407	450	440
456	460	456	476	480	490	497	526	546	700

Testar no nível 0,10 para decidir se a resistência à flexão é uma variável normalmente distribuída.

### 14.3 Tabelas de Contingência de Entrada Dupla

Nas duas seções anteriores, discutimos problemas de inferência nos quais os dados contados eram apresentados em uma tabela de células retangular. Cada tabela consistia de uma linha e um número especificado de colunas, onde as colunas corresponderiam às categorias nas quais a população tinha sido dividida. Dessa vez, estudamos os problemas nos quais os dados também consistem de contagens ou frequências, mas a tabela de dados terá agora  $I$  linhas ( $I \geq 2$ ) e  $J$  colunas, então  $IJ$  células. Há duas situações comumente encontradas nas quais tais dados aparecem:

1. Há  $I$  populações de interesse, cada uma correspondendo a uma linha diferente da tabela, e cada população é dividida nas mesmas  $J$  categorias. Uma amostra é tirada da  $i$ -ésima população ( $i = 1, \dots, I$ ), e as contagens são colocadas em células na  $i$ -ésima linha da tabela. Por exemplo, os clientes de cada uma das  $I = 3$  cadeias de lojas de departamentos podem ter disponíveis as mesmas  $J = 5$  categorias de pagamento: dinheiro, cheque, cartão de crédito da loja, Visa e MasterCard.
2. Há uma população única de interesse, com cada indivíduo em uma população categorizada em relação a dois fatores diferentes. Há  $I$  categorias associadas ao primeiro fator e  $J$  categorias associadas do segundo fator. Uma única amostra é tomada e o número de indivíduos que pertence a ambas as categorias  $i$  do fator 1 e categoria  $j$  do fator 2 é colocado na célula na linha  $i$ , coluna  $j$  ( $i = 1, \dots, I$ ;  $j = 1, \dots, J$ ). Como um exemplo, clientes que fazem uma compra são classificados de acordo com ambos os departamentos nos quais a compra foi feita, com  $I = 6$  departamentos, e de acordo com o método de pagamento, com  $J = 5$  como em (1) acima.

Seja  $n_{ij}$  o número de indivíduos na(s) amostra(s) que caem na célula  $(i, j)$ -ésima (linha  $i$ , coluna  $j$ ) da tabela, ou seja, a contagem de célula  $(i, j)$ -ésima. A tabela mostra os  $n_{ij}$ s sendo chamados de uma **tabela de contingência de entrada dupla**; um protótipo é mostrado na Tabela 14.9.

**Tabela 14.9** Tabela de contingência de dupla entrada

	1	2	...	$j$	...	$J$
1	$n_{11}$	$n_{12}$	...	$n_{1j}$	...	$n_{1J}$
2	$n_{21}$					$\vdots$
$\vdots$	$\vdots$					
$i$	$n_{i1}$	...		$n_{ij}$	...	
$\vdots$	$\vdots$					
$I$	$n_{I1}$	...				$n_{IJ}$

Nas situações do tipo 1, queremos investigar se as proporções nas diferentes categorias são as mesmas para todas as populações. A hipótese nula afirma que as populações são homogêneas em relação a essas categorias. Nas situações do tipo 2, investigamos se as categorias dos dois fatores ocorrem independentemente uma da outra na população.

#### Teste de Homogeneidade

Assumimos que cada indivíduo em cada uma das  $I$  populações pertence a exatamente uma de  $J$  categorias. Uma amostra de  $n_i$  indivíduos é tomada da  $i$ -ésima população; sejam  $n = \sum n_i$  e



$$n_{ij} = \text{número de indivíduos na } i\text{-ésima amostra que cai na categoria } j$$

$$n_{\cdot j} = \sum_{i=1}^I n_{ij} = \text{número total de indivíduos entre o } n\text{-ésimo amostrado que cai na categoria } j$$

Os  $n_{ij}$ s são registrados em uma tabela de contingência de entrada dupla com  $I$  linhas e  $J$  colunas. A contagem dos  $n_{ij}$ s na  $i$ -ésima linha é  $n_i$ , portanto a contagem das entradas na  $j$ -ésima coluna é  $n_{\cdot j}$ .

Seja

$$p_{ij} = \text{a proporção de indivíduos na população } i \text{ que cai na categoria } j$$

Assim, para a população 1, as proporções  $J$  são  $p_{11}, p_{12}, \dots, p_{1J}$  (cuja soma é 1) e da mesma maneira para as outras populações. A hipótese nula de homogeneidade afirma que a proporção de indivíduos na categoria  $j$  é a mesma para cada população e que isto é verdade para toda categoria, ou seja, para todo  $j$ ,  $p_{1j} = p_{2j} = \dots = p_{Ij}$ .

Quando  $H_0$  é verdadeiro, podemos utilizar  $p_1, p_2, \dots, p_J$  para indicar as proporções da população nas diferentes  $J$  categorias, estas proporções são comuns para todas as populações  $I$ . O número esperado de indivíduos na  $i$ -ésima amostra que caem na  $j$ -ésima categoria quando  $H_0$  é verdadeiro é então  $E(N_{ij}) = n_i \cdot p_j$ . Para estimar  $E(N_{ij})$ , devemos primeiro estimar  $p_j$ , a proporção na categoria  $j$ . Entre o total de amostras de  $n$  indivíduos,  $N_{\cdot j}$  cai na categoria  $j$ , portanto usamos  $\hat{p}_j = N_{\cdot j}/n$  como o estimador (isto pode ser mostrado como o estimador de máxima verossimilhança de  $p_j$ ). A substituição da estimativa  $\hat{p}_j$  por  $p_j$  em  $n_i p_j$  produz uma fórmula simples para as contagens esperadas estimadas sob  $H_0$ :

$$\begin{aligned} \hat{e}_{ij} &= \text{contagem esperada estimada na célula } (i, j) = n_i \cdot \frac{n_{\cdot j}}{n} \\ &= \frac{(\text{total da } i\text{-ésima linha})(\text{total da } j\text{-ésima coluna})}{n} \end{aligned} \quad (14.9)$$

O teste estatístico também tem a mesma forma das situações dos problemas anteriores. O número de graus de liberdade vem da regra prática geral. Em cada linha da Tabela 14.9 há  $J - 1$  contagens de células determinadas livremente (cada tamanho  $n_i$  de amostra é fixo), portanto há um total de  $I(J - 1)$  células determinadas livremente. Parâmetros  $p_1, \dots, p_J$  são estimados, mas porque  $\sum p_i = 1$ , somente  $J - 1$  destes são independentes. Assim  $gl = I(J - 1) - (J - 1) = (J - 1)(I - 1)$ .

Hipótese nula:  $H_0: p_{1j} = p_{2j} = \dots = p_{Ij} \quad j = 1, 2, \dots, J$

Hipótese alternativa:  $H_a: H_0$  não é verdadeiro

Valor do teste estatístico:

$$\chi^2 = \sum_{\text{todas as células}} \frac{(\text{observado} - \text{estimado esperado})^2}{\text{estimado esperado}} = \sum_{i=1}^I \sum_{j=1}^J \frac{(n_{ij} - \hat{e}_{ij})^2}{\hat{e}_{ij}}$$

Região de rejeição:  $\chi^2 \geq \chi_{\alpha, (I-1)(J-1)}^2$

A informação do valor  $P$  pode ser obtida como descrito na Seção 14.1. O teste pode ser aplicado com segurança desde que  $\hat{e}_{ij} \geq 5$  para todas as células.

### Exemplo 14.13

Uma empresa embala um certo produto em latas de três tamanhos diferentes, cada uma utilizando uma linha de produção distinta. A maioria das latas atende às especificações, mas um engenheiro de controle de qualidade identificou as seguintes razões para não-conformidade:

#### 1. Mancha na lata

2. Rachadura na lata
3. Localização imprópria da aba de abrir
4. Falta da aba de abrir
5. Outros

Uma amostra de unidades em não-conformidade é selecionada de cada uma das três linhas, e cada unidade é classificada de acordo com a razão para a não-conformidade, resultando na seguinte tabela de dados de contingência:

	Razões para a Não-Conformidade					Tamanho da Amostra	
	Mancha	Rachadura	Localização	Faltando	Outros		
Linha de Produção	1	34	65	17	21	13	150
	2	23	52	25	19	6	125
	3	32	28	16	14	10	100
	<b>Total</b>	89	145	58	54	29	375

Os dados sugerem que as proporções que caem em várias categorias de não-conformidade não são as mesmas para as três linhas? Os parâmetros de interesse são as várias proporções, e as hipóteses relevantes são

$H_0$ : as linhas de produção são homogêneas em relação às cinco categorias de não conformidade, ou seja,  $p_{1j} = p_{2j} = p_{3j}$  para  $j = 1, \dots, 5$

$H_a$ : as linhas de produção não são homogêneas em relação às cinco categorias

As frequências esperadas estimadas (assumindo homogeneidade) devem agora ser calculadas. Considere a primeira categoria de não-conformidade para a primeira linha de produção. Quando as linhas são homogêneas,

o número esperado estimado entre as 150 unidades selecionadas que estão manchadas

$$= \frac{(\text{total da 1ª linha})(\text{total da 1ª coluna})}{\text{total dos tamanhos da amostra}} = \frac{(150)(89)}{375} = 35,60$$

A contribuição da célula no canto superior esquerdo para  $\chi^2$  é, então,

$$\frac{(\text{observado} - \text{estimado esperado})^2}{\text{estimado esperado}} = \frac{(34 - 35,60)^2}{35,60} = 0,072$$

As outras contribuições são calculadas de maneira semelhante. A Figura 14.4 mostra a saída do MINITAB para o teste qui-quadrado. A contagem observada está no número superior em cada célula e diretamente abaixo está a contagem esperada estimada. A contribuição de cada célula para  $\chi^2$  aparece abaixo das contagens e o valor do teste estatístico é  $\chi^2 = 14,159$ . Todas as contagens esperadas estimadas são pelo menos 5, portanto combinar categorias é desnecessário. O teste está baseado em  $(3 - 1)(5 - 1) = 8$  gl. A Tabela A.11 do Apêndice mostra que os valores que capturam as áreas de cauda superior de 0,08 e 0,075 sob a curva 8 gl são 14,06 e 14,26, respectivamente. Assim o valor  $P$  está entre 0,075 e 0,08; o MINITAB fornece o valor  $P = 0,079$ . A hipótese nula de homogeneidade não deve ser rejeitada nos níveis usuais de significância de 0,05 ou 0,01, mas deve ser rejeitada para níveis  $\alpha$  maiores que 0,10.

Contagens esperadas são impressas abaixo das contagens observadas

	mancha	rachadura	localização	falta	outros	Total
1	34 35,60	65 58,00	17 23,20	21 21,60	13 11,60	150
2	23 29,67	52 48,33	25 19,33	19 18,00	6 9,67	125
3	32 23,73	28 38,67	16 15,47	14 14,40	10 7,73	100
Total	89	145	58	54	29	375

Qui-quadrado = 0,072 + 0,845 + 1,657 + 0,017 + 0,169 + 1,498 + 0,278 + 1,661 + 0,056 + 1,391 + 2,879 + 2,943 + 0,018 + 0,011 + 0,664 = 14,159

gl = 8, p = 0,079

**Figura 14.4** Saída do MINITAB para o teste qui-quadrado do Exemplo 14.13

## Teste de Independência

Abordaremos agora a relação entre dois fatores diferentes em uma única população. O número de categorias do primeiro fator será indicado por  $I$  e o número de categorias do segundo fator por  $J$ . Cada indivíduo na população pertence a exatamente uma das categorias  $I$  associadas ao primeiro fator e exatamente uma das categorias  $J$  associadas ao segundo fator. Por exemplo, a população de interesse pode consistir de todos os indivíduos que assistem regularmente ao noticiário nacional na televisão, com o primeiro fator sendo a preferência da rede (ABC, CBS, NBC ou PBS, portanto  $I = 4$ ) e o segundo fator a filosofia política (liberal, moderada ou conservadora, resultando em  $J = 3$ ).

Para uma amostra de  $n$  indivíduos tomados da primeira população, seja  $n_{ij}$  o número entre os  $n$  que caem na categoria  $i$  do primeiro fator e na categoria  $j$  do segundo fator. Os  $n_{ij}$ s são mostrados em uma tabela de contingência de entrada dupla com  $I$  linhas e  $J$  colunas. No caso de homogeneidade para  $I$  populações, os totais de linha foram fixados de maneira avançada, e apenas os totais das colunas  $J$  são aleatórios. Agora, apenas o tamanho total da amostra é fixado, e ambos os  $n_{i\cdot}$ s e os  $n_{\cdot j}$ s são valores observados de variáveis aleatórias. Para afirmar a hipótese de interesse, seja

$$\begin{aligned} p_{ij} &= \text{a proporção de indivíduos na população que pertencem à categoria } i \text{ do fator 1 e a} \\ &\quad \text{categoria } j \text{ do fator 2} \\ &= P(\text{um indivíduo selecionado aleatoriamente que cai em ambas categoria } i \text{ do fator} \\ &\quad \text{1 e categoria } j \text{ do fator 2}) \end{aligned}$$

Então

$$\begin{aligned} p_{i\cdot} &= \sum_j p_{ij} = P(\text{um indivíduo selecionado aleatoriamente que cai na categoria } i \text{ do fator 1}) \\ p_{\cdot j} &= \sum_i p_{ij} = (\text{um indivíduo selecionado aleatoriamente que cai na categoria } j \text{ do fator 2}) \end{aligned}$$

Lembre-se de que dois eventos  $A$  e  $B$  são independentes se  $P(A \cap B) = P(A) \cdot P(B)$ . A hipótese nula, diz que categoria de indivíduo em relação a um fator 1 é independente da categoria em relação ao fator 2. Em símbolos, isto torna-se  $p_{ij} = p_{i\cdot} \cdot p_{\cdot j}$  para todo par  $(i, j)$ .

A contagem esperada na célula  $(i, j)$  é  $n \cdot p_{ij}$ , então quando a hipótese nula é verdadeira,  $E(N_{ij}) = n \cdot p_{i\cdot} \cdot p_{\cdot j}$ . Para obter uma estatística qui-quadrado devemos, portanto, estimar os  $p_{i\cdot}$ s ( $i = 1, \dots, I$ ) e  $p_{\cdot j}$ s ( $j = 1, \dots, J$ ). As estimativas (máxima probabilidade) são

$$\hat{p}_{i\cdot} = \frac{n_{i\cdot}}{n} = \text{proporção da amostra para categoria } i \text{ do fator 1}$$

e

$$\hat{p}_{\cdot j} = \frac{n_{\cdot j}}{n} = \text{proporção da amostra para a categoria } j \text{ do fator 2}$$

Isto fornece as contagens de células esperadas estimadas de maneira idêntica àquelas no caso homogêneo.

$$\begin{aligned} \hat{e}_{ij} &= n \cdot \hat{p}_{i\cdot} \cdot \hat{p}_{\cdot j} = n \cdot \frac{n_{i\cdot}}{n} \cdot \frac{n_{\cdot j}}{n} = \frac{n_{i\cdot} \cdot n_{\cdot j}}{n} \\ &= \frac{(\text{total da } i\text{-ésima linha})(\text{total da } j\text{-ésima coluna})}{n} \end{aligned}$$

O teste estatístico é também idêntico àquele utilizado no teste para homogeneidade, assim como o número de graus de liberdade. Isto acontece porque o número das contagens de células determinadas livremente é  $IJ - 1$ , visto que apenas o total  $n$  é fixado de maneira avançada. Existem  $I$   $p_{i\cdot}$ s estimados, mas apenas  $I - 1$  são independentemente estimados desde que  $\sum p_{i\cdot} = 1$ , e de maneira semelhante  $J - 1$   $p_{\cdot j}$ s são independentemente estimados, portanto  $I + J - 2$  parâmetros são estimados independentemente. A regra prática agora produz  $gl = IJ - 1 - (I + J - 2) = IJ - I - J + 1 = (I - 1) \cdot (J - 1)$ .

Hipótese nula:  $H_0: p_{ij} = p_{i.} \cdot p_{.j} \quad i = 1, \dots, I; j = 1, \dots, J$

Hipótese alternativa:  $H_a: H_0$  não é verdadeiro

Valor do teste estatístico:

$$\chi^2 = \sum_{\text{todas as células}} \frac{(\text{observada} - \text{estimada esperada})^2}{\text{estimado esperado}} = \sum_{i=1}^I \sum_{j=1}^J \frac{(n_{ij} - \hat{e}_{ij})^2}{\hat{e}_{ij}}$$

Região de rejeição:  $\chi^2 \geq \chi^2_{\alpha, (I-1)(J-1)}$

Novamente, a informação do valor  $P$  pode ser obtida como descrito na Seção 14.1. O teste pode ser aplicado com segurança enquanto  $\hat{e}_{ij} \geq 5$ , para todas as células.

**Exemplo 14.14**

Um estudo da relação entre as condições das instalações nos postos de gasolina e a agressividade na composição do preço da gasolina (“An Analysis of Price Aggressiveness in Gasoline Marketing”, *J. Marketing Research*, 1970, p. 36-42) informa os dados correspondentes baseados em uma amostra de  $n = 441$  postos. No nível 0,01, os dados sugerem que as condições das instalações e a política de preços são independente uma da outra? Contagens de células observadas e estimadas são mostradas na Tabela 14.10.

**Tabela 14.10** Contagens de células observadas e esperadas para o Exemplo 14.14

	Política de Preço Observada			$n_i$	Política de Preço Esperada			
	Agressiva	Neutra	Não-agressiva		$n_i$	Agressiva	Neutra	
<b>Subpadrão</b>	24	15	17	56	17,02	22,10	16,89	56
<b>Condição Padrão</b>	52	73	80	205	62,29	80,88	61,83	205
<b>Moderno</b>	58	86	36	180	54,69	71,02	54,29	180
$n_j$	134	174	133	441	134	174	133	441

Desta forma

$$\chi^2 = \frac{(24 - 17,02)^2}{17,02} + \dots + \frac{(36 - 54,29)^2}{54,29} = 22,47$$

e porque  $\chi^2_{0,01,4} = 13,277$ , a hipótese de independência é rejeitada.

Concluimos que o conhecimento da política de preço de um posto fornece informação sobre as facilidades nele. Em particular, postos com uma política agressiva de preços parecem ser mais propensos a ter facilidades subpadrão do que os postos com uma política não-agressiva ou neutra de preços. ■

Os modelos e métodos para a análise de dados nos quais cada indivíduo é categorizado em relação a três ou mais fatores (tabelas de contingência multidimensional) são discutidos em vários dos capítulos de referência.

**Exercícios** | Seção 14.3 (24–36)

24. A tabela correspondente de entrada dupla foi construída utilizando-se dados do artigo “Television Viewing and Physical Fitness in Adults” (*Research Quarterly for Exercise and Sport*, 1990, p. 315-320). O autor esperava determinar se o tempo gasto assistindo à televisão está associado com o estado cardiovascular. Os entrevistados foram questionados sobre seus hábitos de assistir à televisão e foram classificados como em boa condição física se sua pontuação estava na categoria excelente ou muito boa em um teste de caminhada. Incluamos os resultados do MINITAB a partir de uma análise qui-quadrado. Os quatro grupos de TV corresponderam às diferentes quantidades de tempo por dia gastas assistindo à TV (0, 1-2, 3-4 ou 5 ou mais horas). Os 168 indivíduos representados na primeira coluna foram aqueles julgados em boas condições físicas. As contagens esperadas aparecem abaixo das contagens observadas e o MINITAB mostra a contribuição para  $\chi^2$  a partir de cada célula. Defina e teste a hipótese relevante utilizando  $\alpha = 0,05$ .

	1	2	Total
1	35 25,48	147 156,52	182
2	101 102,20	629 627,80	730
3	28 35,00	222 215,00	250
4	4 5,32	34 32,68	38
Total	168	1.032	1.200

$$\text{Qui-quadrado} = 3,557 + 0,579 + 0,014 + 0,002 + 1,400 + 0,228 + 0,328 + 0,053 = 6,161$$

gl = 3

25. Os dados correspondentes referem-se a marcas de folhas encontradas em amostras selecionadas a partir de áreas de pastagens e jardins (“The Biology of the Leaf Mark Polymorphism in *Trifolium repens* L.”, *Heredity*, 1976, p. 306-325). Utilize um teste  $\chi^2$  para decidir se as proposições verdadeiras de marcas diferentes são idênticas para os dois tipos de regiões.

	Tipo de Marca				Outros	Tamanho da Amostra
	LG	LL	Y + YL	O		
Áreas de Pastagens	409	11	22	7	277	726
Áreas de Jardim	512	4	14	11	220	761

26. Os dados a seguir resultaram de um experimento para o estudo dos efeitos da remoção das folhas na habilidade da fruta de um determinado tipo de amadurecer (“Fruit

Set, Herbivory, Fruit Reproduction, and the Fruiting Strategy of *Catalpa speciosa*” *Ecology*, 1980, p. 57-64):

Tratamento	Número de Frutas Maduras	Número de Frutas Abortadas
Controle	141	206
Duas folhas removidas	28	69
Quatro folhas removidas	25	73
Seis folhas removidas	24	78
Oito folhas removidas	20	82

Os dados sugerem que a chance do amadurecimento da fruta é afetada pelo número de folhas removidas? Enuncie e teste a hipótese apropriada no nível 0,01.

27. O artigo “Human Lateralization from Head to Foot: Sex-Related Factors” (*Science*, 1978, p. 1291-1292 ) apresenta uma amostra de homens destros e outra de mulheres destros; o texto informa o número de indivíduos cujos pés eram do mesmo tamanho, os que tinham o pé esquerdo maior do que o direito (uma diferença de meio número de calçado ou mais), ou os que tinham o pé direito maior do que o esquerdo.

	E > D	E = D	E < D	Tamanho da Amostra
	Homens	2	10	28
Mulheres	55	18	14	87

Os dados indicam que o sexo tem um efeito importante no desenvolvimento da assimetria do pé? Defina as hipóteses nula e alternativa apropriadas, calcule o valor de  $\chi^2$  e obtenha a informação sobre o valor *P*.

28. O artigo “Susceptibility of Mice to Audiogenic Seizure Is Increased by Handling Their Dams During Gestation” (*Science*, 1976, p. 427-428 ) informa sobre a pesquisa do efeito de diferentes tratamentos aplicados sobre as frequências de captação de sons.

Tratamento	Sem Resposta	Corrida Selvagem	Captação Clônica	Captação Tônica
Thienylalanine	21	7	24	44
Solvente	15	14	20	54
Simulação	23	10	23	48
Não manipulado	47	13	28	32

Os dados sugerem que as porcentagens verdadeiras em diferentes categorias de resposta dependem da natureza

do tratamento aplicado? Defina e teste a hipótese relevante utilizando  $\alpha = 0,005$ .

29. Os dados correspondentes às combinações de sexo de duas recombinações resultantes de seis diferentes genótipos masculinos aparecem no artigo “A New Method for Distinguishing Between Meiotic and Premeiotic Recombinational Events in *Drosophila inelanogaster*” (*Genetics*, 1979, p. 543-554). Tais dados suportam a hipótese de que a distribuição de frequência entre as três combinações de sexo é homogênea em relação aos diferentes genótipos? Defina os parâmetros de interesse, enuncie o  $H_0$  e  $H_a$  apropriados e realize as análises.

	Combinação de Sexo			
	M/M	M/F	F/F	
<b>Genótipo</b> <b>Macho</b>	1	35	80	39
	2	41	84	45
	3	33	87	31
	4	8	26	8
	5	5	11	6
	6	30	65	20

30. Três configurações de modelos diferentes estão sendo consideradas para um componente particular. Há quatro possibilidades de modos de falha para o componente. Um engenheiro obteve os seguintes dados sobre o número de falhas em cada modo para cada uma das três configurações. A configuração tem algum efeito sobre o tipo de falha?

	Modo de Falha				
	1	2	3	4	
<b>Configuração</b>	1	20	44	17	9
	2	4	17	7	12
	3	10	31	14	5

31. Uma amostra aleatória de indivíduos que dirigem sozinho para o trabalho em uma área metropolitana grande foi obtida, e cada indivíduo foi categorizado em relação ao tamanho do carro e à distância percorrida. Os dados correspondentes sugerem que a distância percorrida e o tamanho do carro são relatados na amostra de população? Defina a hipótese apropriada e utilize um teste qui-quadrado de nível 0,05.

		Distância Percorrida		
		0-<10	10-<20	≥20
<b>Tamanho de Carro</b>	<b>Subcompacto</b>	6	27	19
	<b>Compacto</b>	8	36	17
	<b>Tamanho médio</b>	21	45	33
	<b>Tamanho grande</b>	14	18	6

32. Cada indivíduo em uma amostra aleatória de estudantes da escola secundária e da faculdade foi classificado de

maneira cruzada em relação a ponto de vista político e o uso de maconha, resultando nos dados mostrados na tabela de dupla entrada correspondente (“Attitudes About Marijuana and Political Views”, *Psychological Reports*, 1973, p. 1051-1054). Os dados suportam a hipótese de que o ponto de vista político e o nível de utilização de maconha são independentes dentro da população? Teste a hipótese apropriada utilizando o nível de significância de 0,01.

		Nível de Utilização		
		Nunca	Raramente	Freqüentemente
<b>Ponto de Vista Político</b>	<b>Liberal</b>	479	173	119
	<b>Conservador</b>	214	47	15
	<b>Outros</b>	172	45	85

33. Mostre que a estatística qui-quadrado para o teste de independência pode ser escrita na forma

$$\chi^2 = \sum_{i=1}^I \sum_{j=1}^J \left( \frac{N_{ij}^2}{\hat{E}_{ij}} \right) - n$$

Por que esta fórmula é mais eficiente nos cálculos do que a fórmula definida para  $\chi^2$ ?

34. Suponha que no Exercício 32 cada estudante tenha sido categorizado em relação a ponto de vista político, uso de maconha e preferência religiosa, com as categorias deste último fator sendo Protestante, Católico e Outros. Os dados podem ser mostrados em três diferentes tabelas de dupla entrada, cada uma correspondendo a uma categoria do terceiro fator. Com  $p_{ijk} = P(\text{categoria política } i, \text{ categoria maconha } j \text{ e categoria religiosa } k)$ , a hipótese nula de independência de todos os três fatores afirma que  $p_{ijk} = p_{i\cdot} \cdot p_{\cdot j} \cdot p_{\cdot k}$ . Represente por  $n_{ijk}$  a frequência observada na célula  $(i, j, k)$ . Mostre como estimar as contagens de células esperadas assumindo que  $H_0$  é verdadeiro ( $\hat{e}_{ijk} = n \hat{p}_{ijk}$ , portanto o  $\hat{p}_{ijk}$ s deve ser determinado). Então utilize a regra prática geral para determinar o número de graus de liberdade para a estatística qui-quadrada.

35. Suponha que em um estado com quatro regiões distintas, uma amostra aleatória de  $n_k$  eleitores é obtida a partir da  $k$ -ésima região para  $k = 1, 2, 3$  e 4. Cada eleitor é então classificado de acordo com qual candidato (1, 2 ou 3) ele ou ela preferem e de acordo com o registro do eleitor (1 = Dem., 2 = Rep., 3 = Indep.). Represente por  $p_{ijk}$  a proporção de eleitores na região  $k$  que pertence à categoria do candidato  $i$  e categoria de registro  $j$ . A hipótese nula das regiões homogêneas é  $H_0: p_{ij1} = p_{ij2} = p_{ij3} = p_{ij4}$  para todo  $i, j$  (ou seja, a proporção dentro da combinação candidato/registo é a mesma para todas as quatro regiões). Assumindo que  $H_0$  é verdadeiro, determine  $\hat{p}_{ijk}$  e  $\hat{e}_{ijk}$  como funções dos  $n_{ijk}$ s observados e utilize a regra prática geral para obter o número de graus de liberdade para o teste qui-quadrado.

36. Considere a tabela  $2 \times 3$  correspondente mostrando as proporções da amostra que caem em várias combinações

de categorias (por exemplo, 13% daqueles na amostra que estavam na primeira categoria de ambos os fatores).

	1	2	3
1	0,13	0,19	0,28
2	0,07	0,11	0,22

- Suponha que a amostra considerada seja de  $n = 100$  pessoas. Utilize o teste qui-quadrado para independência com nível de significância 0,10.
- Repita a parte (a) assumindo que o tamanho da amostra foi de  $n = 1.000$ .
- Qual é o menor tamanho de amostra  $n$  para a qual estas proporções observadas resultarão na rejeição da hipótese de independência?

### Exercícios Suplementares (37–48)

37. O artigo “Birth Order and Political Success” (*Psych. Reports*, 1971, p. 1239-1242) informa que entre 31 candidatos à política selecionados aleatoriamente que descendem de famílias com quatro filhos, 12 eram primogênitos, 11 eram filhos do meio e 8 eram caçulas. Utilize esses dados para testar a hipótese nula que um candidato à política de tal família é igualmente provável a estar em qualquer uma das quatro posições ordinais.

38. Os resultados de um experimento para avaliar o efeito do petróleo sobre parasitas de peixes foram descritos no artigo “Effects of Crude Oils on the Gastrointestinal Parasites of Two Species of Marine Fish” (*J. Wildlife Diseases*, 1983, p. 253-258). Três tratamentos (correspondendo a populações no procedimento descrito) foram comparados: (1) sem contaminação, (2) contaminação por óleo derramado há 1 ano e (3) contaminação por óleo novo. Para cada condição de tratamento, uma amostra de peixe foi tomada e, então, cada peixe foi classificado como tendo o parasita ou não. Dados compatíveis com estes do artigo são fornecidos. Os dados indicam que os três tratamentos diferem em relação à proporção real de peixe com parasitas ou sem? Teste utilizando  $\alpha = 0,01$ .

Tratamento	Parasitado	Não-parasitado
Controle	30	3
Óleo antigo	16	8
Óleo novo	16	16

39. Qualificações de treinadores atléticos colegiais titulares e assistentes de ambos os sexos foram comparadas no artigo “Sex Bias and the Validity of Believed Differences Between Male and Female Interscholastic Athletic Coaches” (*Research Quarterly for Exercise and Sport*, 1990, p. 259-267). Cada pessoa em uma amostra aleatória de 2.225 treinadores e 1.141 treinadoras foi classificada de acordo com o número de anos de experiência como treinador para obter a tabela de dupla entrada correspondente. Há evidência suficiente para concluir que as proporções que caem nas

categorias de experiência são diferentes para homens e mulheres? Use  $\alpha = 0,01$ .

#### Anos de Experiência

Sexo	1–3	4–6	7–9	10–12	13+
Masculino	202	369	482	361	811
Feminino	230	251	238	164	258

40. Os autores do artigo “Predicting Professional Sports Game Outcomes from Intermediate Game Scores” (*Chance*, 1992, p. 18-22) utilizaram um teste qui-quadrado para determinar se havia qualquer mérito na idéia de que os jogos de basquete não são decididos antes do último tempo, enquanto os de baseball terminam até o sétimo turno. Eles também consideraram o futebol e o hóquei. Os dados foram coletados de 189 jogos de basquete, 92 baseball, 80 de hóquei e 93 de futebol. Os jogos analisados foram amostrados aleatoriamente a partir de todos os jogos realizados durante a temporada de 1990 de baseball e futebol, e para a temporada de 1990-1991 de basquete e hóquei. Para cada jogo, o líder foi determinado e então foi observado se ele realmente acabou como vencedor. Estas informações estão resumidas na tabela a seguir.

Esporte	Líder do jogo Líder Vence	Líder do jogo Líder Perde
Basquete	150	39
Baseball	86	6
Hóquei	65	15
Futebol	72	21

Os autores afirmam que “O líder do jogo é definido como o time que é líder após três quartos em basquete e futebol; dois períodos em hóquei; e sete turnos em baseball. O valor qui-quadrado em três graus de liberdade é 10,52 ( $P < 0,015$ )”.

- Defina a hipótese relevante e encontre uma conclusão utilizando  $\alpha = 0,05$ .

b. Você acredita que a sua conclusão na parte (a) pode ser atribuída a um único esporte como sendo uma anomalia?

41. A tabela de frequência de dupla entrada correspondente no artigo “Marijuana Use in College”(*Youth and Society*; 1979, p. 323-334). Cada um dos 445 estudantes universitários foram classificados de acordo com a frequência no uso da maconha e o uso, pelos pais, de álcool e drogas psicotrópicas. Os dados sugerem que o uso pelos pais e estudantes são independentes na população a partir da qual a amostra foi tomada? Utilize o método do valor *P* para chegar a uma conclusão.

		Nível-padrão de uso de maconha		
		Nunca	Ocasional	Regular
Uso de álcool e drogas pelos pais	Nenhum	141	54	40
	Um	68	44	51
	Ambos	17	11	19

42. Em um estudo de 2.989 mortes por câncer, o local da morte (casa, hospital ou UTI) e a idade na data da

morte foram registrados, resultando na tabela de frequência de dupla entrada (“Where Cancer Patients Die”, *Public Health Reports*, 1983, p. 173). Utilizando o nível de significância  $\alpha = 0,01$ , teste a hipótese nula de que a idade na data da morte e o local da morte são independentes.

	Local			
	Casa	Hospital	UTI	
Idade	15–54	94	418	23
	55–64	116	524	34
	65–74	156	581	109
	Mais de 74	138	558	238

43. Em um estudo para investigar a extensão na qual os indivíduos estão prevenidos contra odores industriais em uma determinada região (“Annoyance and Health Reactions to Odor from Refineries and Other Industries in Carson, California”, *Environment Research*, 1978, p. 119-132), uma amostra de indivíduos foi obtida a partir de cada uma das três diferentes áreas próximas às instalações industriais. Cada um foi perguntado se ele ou ela percebia odores (1) todo dia, (2) pelo menos uma vez por semana, (3) pelo menos uma vez por mês, (4) menos freqüente que uma vez por mês ou (5) não sentiam odores, resultando nos dados e resultados de SPSS a seguir. Defina e teste uma hipótese

Tabulação cruzada: ÁREA POR CATEGORIA

CATEGORIA DA ÁREA	Linha Pct	Coluna Pct	1,00	2,00	3,00	4,00	5,00	Total da Linha
			1,00	20 12,7 20,6% 52,6%	28 24,7 28,9% 37,8%	23 18,0 23,7% 42,6%	14 16,0 14,4% 29,2%	
2,00	14 12,4 14,7% 36,8%	34 24,2 35,8% 45,9%	21 17,6 22,1% 38,9%	14 15,7 14,7% 29,2%	12 25,1 12,6% 15,6%	95 32,6%		
3,00	4 12,9 4,0% 10,5%	12 25,2 12,1% 16,2%	10 18,4 10,1% 18,5%	20 16,3 20,2% 41,7%	53 26,2 53,5% 68,8%	99 34,0%		
Total da Coluna		38 13,1%	74 25,4%	54 18,6%	48 16,5%	77 26,5%	291 100,0%	
Qui-quadrado	G.L.	Significância	E.F.Mín.	Células com E.F. < 5				
70,64156	8	0,0000	12,405	Nenhum				



44. Muitos compradores expressaram desagrado porque as lojas deixaram de colocar preços nos itens individuais. O artigo “The Impact of Item Price Removal on Grocery Shopping Behavior” (*J. Marketing*, 1980, p. 73-93) informa sobre um estudo no qual cada comprador em uma amostra foi classificado por idade e se ele ou ela sentiam a necessidade do preço no item. Baseado nos dados correspondentes, a necessidade do preço do item aparecer é independente da idade?

	Idade				
	<30	30-39	40-49	50-59	≥60
Número na Amostra	150	141	82	63	49
Número que deseja preço no item	127	118	77	61	41

45. Represente por  $p_1$  a proporção de sucessos em uma população particular. O valor do teste estatístico no Capítulo 8 para o teste  $H_0: p_1 = p_{10}$  foi  $z = (\hat{p}_1 - p_{10})/\sqrt{p_{10}p_{20}/n}$ , onde  $p_{20} = 1 - p_{10}$ . Mostre que para o caso  $k = 2$ , o valor do teste estatístico qui-quadrado da Seção 14.1 satisfaz  $\chi^2 = z^2$ . [Sugestão: Mostre primeiro que  $(n_1 - np_{10})^2 = (n_2 - np_{20})^2$ .]

46. O torneio de basquete da NCAA inicia com 64 times que são divididos em quatro torneios regionais, cada um envolvendo 16 times. Estes, em cada região, são então classificados (espalhados) de 1 a 16. Durante o período de 12 anos de 1991 até 2002, o time no topo da classificação venceu seu torneio regional 22 vezes, o time classificado em segundo lugar venceu 10 vezes, o time classificado no terceiro lugar venceu 5 vezes e os 11 torneios regionais remanescentes foram vencidos pelos times classificados abaixo de 3. Represente por  $P_{ij}$  a probabilidade que o time classificado como  $i$  em sua região seja vitorioso em seu jogo contra o time classificado como  $j$ . Uma vez que os  $P_{ij}$ s estejam disponíveis, é possível calcular a probabilidade de que cada classificado particular vença seu torneio regional (um cálculo complicado porque o número de resultados no espaço amostral é um tanto grande). O artigo “Probability Models for the NCAA Regional Basketball Tournaments” (*The American Statistician*, 1991, p. 35-38) propõe vários modelos diferentes para os  $P_{ij}$ s.

- a. Um modelo postulou  $P_{ij} = 0,5 + \lambda(i - j)$  com  $\lambda = 1/32$  (a partir do qual  $P_{16,1} = \lambda$ ,  $P_{16,2} = 2\lambda$  etc.). Baseado nisto,  $P(\text{classificado em primeiro lugar vence}) = 0,27477$ ,  $P(\text{classificado em segundo lugar vence}) = 0,20834$  e  $P(\text{classificado em terceiro lugar vence}) = 0,15429$ . Este modelo parece fornecer um bom ajuste aos dados?
- b. Um modelo mais sofisticado tem  $P_{ij} = 0,5 + 0,2813625(z_i - z_j)$ , onde os  $z$ s são medidas de resistência relativa relacionadas com percentuais normais padronizados (percentis para times sucessivos

altamente classificados estão mais próximos do que o caso dos times com a classificação mais baixa, e 0,2813625 garante que a variação de probabilidades seja a mesma como para o modelo na parte (a)). As probabilidades resultantes das classificações 1, 2 ou 3 de vencer seus torneios regionais são 0,45883, 0,18813 e 0,11032, respectivamente. Calcule o ajuste deste modelo.

47. Você já imaginou se os jogadores de futebol sofressem efeitos adversos ao cabecear a bola? Os autores do artigo “No Evidence of Impaired Neurocognitive Performance in Collegiate Soccer Players” (*The Amer. J. of Sports Medicine*, 2002, p. 157-162) investigaram esta questão a partir de várias perspectivas.

- a. O artigo informou que 45 dos 91 jogadores de futebol em sua amostra sofreram pelo menos uma concussão, 28 de 96 atletas não-jogadores de futebol sofreram pelo menos uma concussão, e apenas 8 de 53 estudantes revelaram o mesmo quadro. Analise este dado e tire as conclusões apropriadas.
- b. Para os jogadores de futebol, o coeficiente de correlação da amostra calculado a partir dos valores de  $x =$  exposição do jogador (número total de temporadas jogadas antes de inscrever-se no estudo) e  $y =$  classificação em um teste de memória imediata foi  $r = -0,220$ . Interprete este resultado.
- c. Aqui está um resumo das informações sobre classificações em um teste de associação de palavras oral controlado para jogadores de futebol e atletas não-jogadores de futebol:

$$n_1 = 26, \bar{x}_1 = 37,50, s_1 = 9,13, n_2 = 56, \bar{x}_2 = 39,63, s_2 = 10,19$$

Analise este dado e tire as conclusões apropriadas.

- d. Considerando o número de concussões anteriores em não-jogadores de futebol, os valores da média  $\pm$  desvio padrão para os três grupos foram  $0,30 \pm 0,67$ ;  $0,49 \pm 0,87$  e  $0,19 \pm 0,48$ . Analise estes dados e tire as conclusões apropriadas.
48. Os dígitos sucessivos em uma expansão decimal de  $\pi$  se comportam como se fossem selecionados de uma tabela de números aleatórios (ou provenientes de um computador gerador de números aleatórios)?
- a. Represente por  $p_0$  a grande quantidade de dígitos na expansão como igual a 0 e defina  $p_1, \dots, p_9$  de maneira semelhante. Qual hipótese a respeito destas proporções deve ser testada, e qual é o gl para o teste qui-quadrado?
  - b.  $H_0$  da parte (a) não deve ser rejeitado para a seqüência não-aleatória 012...901...901... Considere grupos não-sobrepostos de dois dígitos e represente por  $p_{ij}$  a grande proporção de grupos para os quais o primeiro dígito é  $i$  e o segundo dígito é  $j$ . Qual hipótese sobre estas proporções deve ser testada, e qual é o gl para o teste qui-quadrado?

- c. Considere grupos não-sobrepostos de 5 dígitos. Um teste qui-quadrado de hipótese apropriada sobre os  $P_{ijklm}$ s pode ser baseado nos primeiros 100.000 dígitos? Explique.
- d. O artigo “Are the Digits of  $\pi$  an Independent and Identically Distributed Sequence?” (*The American*

*Statistician*, 2000, p. 12-16) considerou os primeiros 1.254.540 dígitos de  $\pi$  e informou os seguintes valores  $P$  para tamanhos de grupo de 1, ..., 5: 0,572; 0,078; 0,529; 0,691; 0,298. O que você poderia concluir?

---

## Bibliografia

AGRESTI, Alan. *An Introduction to Categorical Data Analysis*, Nova York: Wiley 1996. Um tratamento excelente de vários aspectos da análise de dados categorizada por um ou mais pesquisadores importantes nesta área.

EVERITT, B. S. *The Analysis of Contingency Tables*, Nova York: Halsted Press, 1992 (2. ed.). Uma compacta, mas

informativa pesquisa, dos métodos para análise categórica de dados, exposta com um mínimo de matemática.

MOSTELLER, Frederick e ROURKE, Richard, *Sturdy Statistics*, Addison-Wesley, Reading, MA, 1973. Contém vários capítulos legíveis sobre os vários usos do qui-quadrado.

## Procedimentos Livres de Distribuição

### Introdução

Quando a população ou populações subjacentes não são normais, os testes  $t$  e  $F$  e os intervalos de confiança  $t$  dos Capítulos 7-13 terão em geral níveis reais de significância ou níveis de confiança diferentes dos níveis nominais (aqueles obtidos pelo modelador por meio da escolha de, digamos,  $t_{0,025}$ ,  $F_{0,01}$  etc.) e  $100(1 - \alpha)\%$ , embora a diferença entre os níveis real e nominal possa não ser grande quando a distância da normalidade não for tão grande. Como os procedimentos  $t$  e  $F$  exigem a hipótese de distribuição normal, eles não são procedimentos de livre distribuição, pois estão baseados em uma família paramétrica particular de distribuições (normal). Eles não são, portanto, procedimentos “não-paramétricos”.

Neste capítulo descrevemos os procedimentos que são válidos [nível real  $\alpha$  ou nível de confiança  $100(1 - \alpha)\%$ ] simultaneamente para muitos tipos diferentes de **distribuições subjacentes**. Tais procedimentos são chamados de livres de distribuição ou não-paramétricos. Na Seção 15.1 discutimos um procedimento de teste para analisar uma amostra única de dados, a Seção 15.2 apresenta um procedimento de teste para ser utilizado em problemas de dupla amostra. Na Seção 15.3 desenvolvemos intervalos de confiança livres de distribuição para  $\mu$  e  $\mu_1 - \mu_2$ . A Seção 15.4 descreve os procedimentos ANOVA livres de distribuição. Esses procedimentos são todos concorrentes dos procedimentos paramétricos ( $t$  e  $F$ ) descritos em capítulos anteriores, portanto é importante comparar o desempenho dos dois tipos sob ambos os modelos de população normal e não-normal. Falando genericamente, os procedimentos livres de distribuição têm desempenho quase tão bom quanto seus concorrentes  $t$  e  $F$  quando a distribuição for normal, e freqüentemente oferecem uma melhora considerável sob condições não-normais.

## 15.1 | O Teste de Postos com Sinais de Wilcoxon

Um pesquisador realizou certo experimento químico um total de dez vezes sob condições idênticas, obtendo os seguintes valores ordenados da temperatura da reação:

$$-0,57 \quad -0,19 \quad -0,05 \quad 0,76 \quad 1,30 \quad 2,02 \quad 2,17 \quad 2,46 \quad 2,68 \quad 3,02$$

A distribuição de temperatura da reação é certamente contínua. Suponha que o pesquisador disposto a assumir que a distribuição de temperatura da reação seja simétrica, ou seja, há um ponto de simetria tal que a curva de densidade à esquerda daquele ponto seja a imagem espelhada da curva de densidade à sua direita. Este ponto de simetria é a média da distribuição, como ilustrado na Figura 15.1 (e é também o valor médio  $\mu$  desde que a média seja finita). A hipótese de simetria pode, em um primeiro momento, parecer muito forte, mas lembre-se de que a distribuição normal é simétrica, portanto a simetria é realmente uma hipótese mais fraca do que a normalidade.

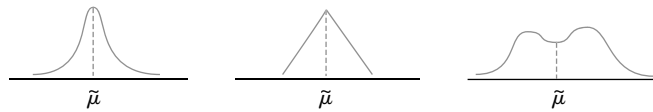


Figura 15.1 Várias distribuições simétricas

Vamos agora considerar o teste de hipótese de que a média da distribuição da temperatura de reação seja zero, ou seja,  $H_0: \tilde{\mu} = 0$ . Este resultado diz que uma temperatura de qualquer magnitude, por exemplo, 1,50, não é mais provável de ser positiva (+1,50) do que de ser negativa (-1,50). Um olhar geral sobre os dados sugere que esta hipótese não é muito convincente, por exemplo, a média amostral é 1,66, o que é muito maior do que a magnitude de qualquer uma das três observações negativas.

A Figura 15.2 mostra duas fdps simétricas, uma para cada  $H_0$  verdadeira e uma para cada  $H_a$  verdadeira. Quando  $H_0$  é verdadeira, esperamos que as magnitudes das observações negativas no exemplo sejam comparáveis às magnitudes das observações positivas. Se, entretanto,  $H_0$  é “grosseiramente” falsa como na Figura 15.2(b), então as observações de magnitude absoluta grande tenderão a ser positivas em vez de negativas.

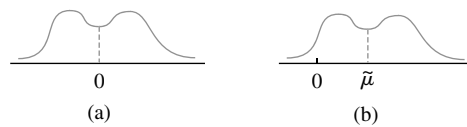


Figura 15.2 (a)  $\tilde{\mu} = 0$ ; (b)  $\tilde{\mu} > 0$

Para as dez observações em nosso exemplo, suponha que procedamos como a seguir:

1. Desconsiderando os sinais das observações, classificamos a coleção de dez números em ordem crescente de magnitude.
2. Utilizamos como teste estatístico  $S_+$  = soma dos postos associados às observações positivas.

<b>Magnitude Absoluta</b>	0,05	0,19	0,57	0,76	1,30	2,02	2,17	2,46	2,68	3,02
<b>Classificação</b>	1	2	3	4	5	6	7	8	9	10
<b>Sinal Classificação</b>	-1	-2	-3	4	5	6	7	8	9	10

O valor observado de  $S_+$  é  $s_+ = 4 + 5 + 6 + 7 + 8 + 9 + 10 = 49$ .  $H_0$  é agora rejeitada em favor de  $H_a$ :  $\tilde{\mu} > 0$  quando  $s_+$  for muito grande porque um valor grande de  $s_+$  indica que a maioria das observações com magnitude absoluta grande são positivas, o que, por outro lado, indica uma média muito maior do que 0. Se, em contrapartida, as observações com sinais negativos são alternadas com observações positivas quando os sinais são desprezados, então o  $s_+$  observado não será muito grande.

### A Distribuição de $S_+$

Para decidir entre  $H_0$  e  $H_a$  devemos saber como 49 é um valor “surpreendentemente grande” de  $S_+$  quando a hipótese nula é verdadeira. Isto requer a obtenção da distribuição de probabilidade de  $S_+$  quando  $H_0$  é verdadeira e a seleção da região de rejeição como um conjunto de grandes valores de  $s_+$  com pequena probabilidade sob  $H_0$  ( $\alpha$  pequeno). Se o nosso  $s_+$  observado estiver na região de rejeição, podemos então rejeitar  $H_0$  ao nível  $\alpha$  desejado e, caso contrário, não rejeitar  $H_0$ .

**Tabela 15.1** Possíveis seqüências de classificação com sinais para  $n = 5$

Seqüência	$s_+$	Seqüência	$s_+$
-1 -2 -3 -4 -5	0	-1 -2 -3 +4 -5	4
+1 -2 -3 -4 -5	1	+1 -2 -3 +4 -5	5
-1 +2 -3 -4 -5	2	-1 +2 -3 +4 -5	6
-1 -2 +3 -4 -5	3	-1 -2 +3 +4 -5	7
+1 +2 -3 -4 -5	3	+1 +2 -3 +4 -5	7
+1 -2 +3 -4 -5	4	+1 -2 +3 +4 -5	8
-1 +2 +3 -4 -5	5	-1 +2 +3 +4 -5	9
+1 +2 +3 -4 -5	6	+1 +2 +3 +4 -5	10
-1 -2 -3 -4 +5	5	-1 -2 -3 +4 +5	9
+1 -2 -3 -4 +5	6	+1 -2 -3 +4 +5	10
-1 +2 -3 -4 +5	7	-1 +2 -3 +4 +5	11
-1 -2 +3 -4 +5	8	-1 -2 +3 +4 +5	12
+1 +2 -3 -4 +5	8	+1 +2 -3 +4 +5	12
+1 -2 +3 -4 +5	9	+1 -2 +3 +4 +5	13
-1 +2 +3 -4 +5	10	-1 +2 +3 +4 +5	14
+1 +2 +3 -4 +5	11	+1 +2 +3 +4 +5	15

Considere o caso  $n = 5$ . A observação-chave é que quando  $H_0$  é verdadeira, qualquer conjunto de postos com sinais é tão provável como qualquer outro conjunto. Em outras palavras, a menor observação na magnitude absoluta é igualmente provável de ser tanto positiva como negativa, a segunda menor observação em grandeza absoluta e igualmente provável de ser tanto positiva como negativa, aplicando-se o mesmo raciocínio para as outras observações. Assim, para  $n = 5$ , a seqüência  $-1, +2, +3, -4, +5$  de postos com sinais é tão provável de ser observada (quando  $H_0$  é verdadeira) quanto a seqüência  $-1, -2, -3, -4, +5$  ou qualquer outra seqüência. A Tabela 15.1 lista as possíveis seqüências de postos com sinais para  $n = 5$  e o valor de  $s_+$  associado a cada seqüência. Como há 32 dessas seqüências (dois sinais possíveis para a menor observação, dois para a segunda menor e assim por diante ou  $2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 = 32$  seqüências), cada seqüência tem probabilidade  $1/32$  de ser observada quando  $H_0$  é verdadeira. A partir disto o fmp de  $S_+$  sob  $H_0$  pode ser imediatamente obtida, como na Tabela 15.2. Observe que para  $n = 10$  observações, há  $2^{10} = 1.024$  possíveis seqüências classificadas com sinais e listar estas seqüências poderia ser muito tedioso. Cada seqüência, entretanto, receberia a probabilidade  $1/1.024$  sob  $H_0$ , a partir da qual a distribuição nula (distribuição quando  $H_0$  é verdadeira) de  $S_+$  pode ser facilmente obtida.

**Tabela 15.2** Distribuição nula de  $S_+$  quando  $n = 5$

$s_+$	0	1	2	3	4	5	6	7
$p(s_+)$	1/32	1/32	1/32	2/32	2/32	3/32	3/32	3/32
$s_+$	8	9	10	11	12	13	14	15
$p(s_+)$	3/32	3/32	3/32	2/32	2/32	1/32	1/32	1/32

Para dar um exemplo de como as probabilidades na Tabela 15.2 foram obtidas, centralizaremos nossa atenção no valor 7 de  $S_+$ ; há três seqüências classificadas com sinais para as quais  $S_+ = 7$  e cada uma recebe a probabilidade  $1/32$  sob  $H_0$ , de forma que  $P(S_+ = 7 \text{ quando } H_0 \text{ é verdadeira}) = 3/32$ .

A Tabela 15.2 pode agora ser utilizada para especificar a região de rejeição para  $H_0: \tilde{\mu} = 0$  versus  $H_a: \tilde{\mu} > 0$  para a qual  $\alpha$  pode ser calculado. Considere a região de rejeição  $R = \{13, 14, 15\}$ . Então

$$\begin{aligned} \alpha &= P(\text{rejeitar } H_0 \text{ quando } H_0 \text{ é verdadeira}) \\ &= P(S_+ = 13, 14, \text{ ou } 15 \text{ quando } H_0 \text{ é verdadeira}) \\ &= 1/32 + 1/32 + 1/32 = 3/32 \\ &= 0,094 \end{aligned}$$

tal que  $R = \{13, 14, 15\}$  especifica um teste com um nível aproximadamente de 0,1. Para a região de rejeição  $\{14, 15\}$ ,  $\alpha = 2/32 = 0,063$ . Para a amostra  $x_1 = 0,58, x_2 = 2,50, x_3 = -0,21, x_4 = 1,23$  e  $x_5 = 0,97$ , a seqüência de classificação com sinais é  $-1, +2, +3, +4, +5$ , portanto  $s_+ = 14$  e no nível 0,063  $H_0$  poderia ser rejeitada.

### Uma Descrição Geral do Teste de Postos com Sinais de Wilcoxon

Como a distribuição subjacente é assumida simétrica,  $\mu = \tilde{\mu}$ , afirmaremos a hipótese de interesse em termos de  $\mu$  em vez de  $\tilde{\mu}$ .<sup>1</sup>

HIPÓTESE

$X_1, X_2, \dots, X_n$  é uma amostra aleatória de uma distribuição de probabilidade contínua e simétrica com média (e mediana)  $\mu$ .

Quando o valor hipotético de  $\mu$  é  $\mu_0$ , as diferenças absolutas  $|x_1 - \mu_0|, \dots, |x_n - \mu_0|$  devem ser classificadas da menor para a maior.

Hipótese nula:  $H_0: \mu = \mu_0$

Valor do teste estatístico:  $s_+ =$  a soma dos postos associados com os valores positivos de  $(x_i - \mu_0)$ s

Hipótese Alternativa

Região de Rejeição para o Teste Nível  $\alpha$

$H_a: \mu > \mu_0$

$s_+ \geq c_1$

$H_a: \mu < \mu_0$

$s_+ \leq c_2$  [onde  $c_2 = n(n + 1)/2 - c_1$ ]

$H_a: \mu \neq \mu_0$

tanto  $s_+ \geq c$  como  $s_+ \leq n(n + 1)/2 - c$

onde os valores críticos  $c_1$  e  $c$  obtidos da Tabela A.13 do Apêndice satisfazem  $P(S_+ \geq c_1) \approx \alpha$  e  $P(S_+ \geq c) \approx \alpha/2$  quando  $H_0$  é verdadeira.

#### Exemplo 15.1

Um fabricante de ferros elétricos, desejando testar a precisão do controle de termostato na marca de 500°F, instruiu o engenheiro de teste para obter temperaturas reais naquela marca para 15 ferros utilizando um termômetro acoplado. As medidas resultantes são as seguintes:

494,6	510,8	487,5	493,2	502,6	485,0	495,9	498,2
501,6	497,3	492,0	504,3	499,2	493,5	505,8	

<sup>1</sup>Se as caudas da distribuição forem “muito pesadas”, como foi o caso com a distribuição de Cauchy, mencionada no Capítulo 6, então  $\mu$  não existirá. Em tais casos, o teste de Wilcoxon ainda será válido para testes concernentes a  $\tilde{\mu}$ .

O engenheiro acredita que é razoável assumir que um desvio de temperatura a partir de 500°F de qualquer magnitude específica possa ser tanto positivo quanto negativo (a hipótese de simetria), mas deseja proteger-se contra possíveis não-normalidades da distribuição de temperatura real, portanto decide utilizar o teste de classificação com sinal de Wilcoxon para verificar se os dados sugerem fortemente a calibração incorreta do ferro.

As hipóteses são  $H_0: \mu = 500$  versus  $H_a: \mu \neq 500$ , onde  $\mu =$  a média verdadeira da temperatura real a 500°F. Subtraindo 500 de cada  $x_i$  resulta em

-5,6   10,8   -12,5   -6,8   2,6   -15,0   -4,1   -1,8   1,6   -2,7  
-8,0   4,3   -0,8   -6,5   5,8

Os postos são obtidos pela ordenação do menor para o maior desprezando o sinal.

<b>Magnitude Absoluta</b>	0,8	1,6	1,8	2,6	2,7	4,1	4,3	5,6	5,8	6,5	6,8	8,0	10,8	12,5	15,0
<b>Classificação</b>	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
<b>Sinal</b>	-	+	-	+	-	-	+	-	+	-	-	-	+	-	-

Assim,  $s_+ = 2 + 4 + 7 + 9 + 13 = 35$ . Da Tabela A.13 do Apêndice,  $P(S_+ \geq 95) = P(S_+ \leq 25) = 0,024$  quando  $H_0$  é verdadeira, portanto o teste de duas caudas com o nível aproximado de 0,05 rejeita  $H_0$  quando  $s_+ \geq 95$  ou  $\leq 25$  [o  $\alpha$  exato é  $2(0,024) = 0,048$ ]. Como  $s_+ = 35$  não está na região de rejeição, não se pode concluir que ao nível 0,05  $\mu$  seja diferente de 500. Ainda ao nível 0,094 (aproximadamente 0,1),  $H_0$  não é rejeitada, uma vez que  $P(S_+ \leq 30) = 0,047$  implica que os valores de  $s_+$  entre 30 e 90 não são significativos a este nível. O valor  $P$  dos dados é, assim, maior do que 0,1. ■

## Observações Pareadas

Quando os dados consistem de pares  $(X_1, Y_1), \dots, (X_n, Y_n)$  e as diferenças  $D_1 = X_1 - Y_1, \dots, D_n = X_n - Y_n$  eram normalmente distribuídas, no Capítulo 9 utilizamos um teste  $t$  pareado para testar hipóteses a respeito da diferença esperada  $\mu_D$ . Se não fizermos a hipótese de normalidade, as hipóteses sobre  $\mu_D$  podem ser testadas utilizando-se o teste de classificação com sinal de Wilcoxon sobre os  $D_i$ s desde que a distribuição das diferenças seja contínua e simétrica. Se ambos  $X_i$  e  $Y_i$  têm distribuições contínuas que diferem apenas em relação às suas médias (portanto, a distribuição  $Y$  é a distribuição  $X$  deslocada por  $\mu_1 - \mu_2 = \mu_D$ ), então  $D_i$  terá uma distribuição simétrica contínua (não é necessário que as distribuições  $X$  e  $Y$  sejam simétricas individualmente). A hipótese nula é  $H_0: \mu_D = \Delta_0$ , e o teste estatístico  $S_+$  é a soma dos postos associados aos  $(D_i - \Delta_0)$ s positivos.

### Exemplo 15.2

Um experimento para comparar a capacidade de dois solventes diferentes para extrair creosoto impregnado em testes de registros envolvidos no uso de oito diferentes tipos de registros. Depois de dividir cada registro em dois segmentos, um deles foi selecionado aleatoriamente para aplicação do primeiro solvente, com o segundo solvente utilizado no outro segmento, produzindo os seguintes dados:

Registro	1	2	3	4	5	6	7	8
Solvente 1	3,92	3,79	3,70	4,08	3,87	3,95	3,55	3,76
Solvente 2	4,25	4,20	4,41	3,89	4,39	3,75	4,20	3,90
Diferença	-0,33	-0,41	-0,71	0,19	-0,52	0,20	-0,65	-0,14
Classificação com sinais	-4	-5	-8	2	-6	3	-7	-1

O primeiro solvente é aquele utilizado atualmente, e o segundo é uma nova formulação projetada para resultar em uma melhor capacidade de extração. Estes dados sugerem que o real valor da média da quantidade extraída

pelo segundo solvente excede aquele do primeiro solvente? As hipóteses relevantes são  $H_0: \mu_D = 0$  versus  $H_a: \mu_D < 0$ . A Tabela A.13 do Apêndice mostra que para um teste com nível de significância de aproximadamente 0,05, a hipótese nula deveria ser rejeitada se  $s_+ \leq (8)(9)/2 - 30 = 6$ . O valor do teste estatístico é  $2 + 3 = 5$ , caindo na região de rejeição. Portanto, rejeitamos  $H_0$  ao nível de significância 0,05 em favor da conclusão que o novo solvente é mais eficiente do que o utilizado atualmente. A saída do MINITAB correspondente resulta no valor do teste estatístico e também no valor  $P$  correspondente, que é  $P(S_+ \leq 5$  quando  $H_0$  é verdadeira).

Teste da mediana = 0,000000 versus mediana < 0,000000

	N	N para Teste	Wilcoxon Estatístico	P	Mediana Estimada
diferença	8	8	5,0	0,040	-0,3025

### Empates no Teste de Postos com Sinais de Wilcoxon

Embora uma implicação teórica da continuidade da distribuição subjacente seja a de que os empates não ocorrerão, na prática eles freqüentemente acontecem por causa da medida discreta dos instrumentos. Se houvesse vários valores de dados com a mesma magnitude absoluta, então poderia ser atribuída a média dos postos que eles receberiam se diferissem muito pouco uns dos outros. Por exemplo, se no Exemplo 15.1,  $x_8 = 498,2$  for alterado para 498,4, então dois diferentes valores de  $(x_i - 500)$  teriam magnitude absoluta 1,6. Os postos cuja média deverá ser calculada seriam 2 e 3, portanto a cada uma poderia ser atribuída a classificação 2,5.

### Aproximação de Amostras Grandes

A Tabela A.13 do Apêndice fornece os valores críticos para testes  $\alpha$  de nível somente quando  $n \leq 20$ . Para  $n > 20$ , pode ser demonstrado que  $S_+$  tem aproximadamente uma distribuição normal com

$$\mu_{s_+} = \frac{n(n + 1)}{4} \quad \sigma_{s_+}^2 = \frac{n(n + 1)(2n + 1)}{24}$$

quando  $H_0$  é verdadeira.

A média e a variância resultam da percepção de que quando  $H_0$  é verdadeira (a distribuição simétrica é centrada em  $\mu_0$ ), então a classificação  $i$  pode receber um sinal de + da mesma maneira como pode receber um sinal de -. Desta forma

$$S_+ = W_1 + W_2 + W_3 + \dots + W_n$$

onde

$$W_1 = \begin{cases} 1 & \text{com probabilidade } 0,5 \\ 0 & \text{com probabilidade } 0,5 \end{cases} \quad \dots \quad W_n = \begin{cases} n & \text{com probabilidade } 0,5 \\ 0 & \text{com probabilidade } 0,5 \end{cases}$$

( $W_i = 0$  é equivalente a classificar  $i$  associado a um -, portanto  $i$  não contribui para  $S_+$ .)

$S_+$  é então uma soma de variáveis aleatórias e quando  $H_0$  é verdadeira, estes  $W_i$ s podem ser mostrados como sendo independentes. As aplicações das regras do valor esperado e da variância fornecem a média e a variância de  $S_+$ . Como os  $W_i$ s não são identicamente distribuídos, nossa versão do Teorema do Limite Central não pode ser aplicada, mas há uma versão mais geral de tal teorema que pode ser utilizada para justificar a conclusão de normalidade.

O teste estatístico de amostra grande é agora dado por

$$Z = \frac{S_+ - n(n + 1)/4}{\sqrt{n(n + 1)(2n + 1)/24}} \tag{15.1}$$



Para as três alternativas-padrão, os valores críticos para o nível do teste  $\alpha$  são os valores normais padrão  $z_{\alpha}$ ,  $-z_{\alpha}$ , e  $\pm z_{\alpha/2}$ .

### Exemplo 15.3

Um tipo particular de suporte de aço é projetado para suportar uma força de compressão (lb/in<sup>2</sup>) de pelo menos 50.000. Para cada suporte em uma amostra de 25 suportes, a força de compressão foi determinada e é dada na Tabela 15.3. Supondo que a força de compressão real é distribuída simetricamente em relação ao valor médio verdadeiro, utilize o teste de Wilcoxon para decidir se a média verdadeira da força de compressão é menor do que o valor especificado. Ou seja, teste  $H_0: \mu = 50.000$  versus  $H_a: \mu < 50.000$  (favorecendo a prerrogativa de que a força média de compressão é de pelo menos 50.000).

**Tabela 15.3** Dados do Exemplo 15.3

$x_i - 50.000$	Classificação com sinais	$x_i - 50.000$	Classificação com sinais	$x_i - 50.000$	Classificação com sinais
-10	-1	-99	-10	165	+18
-27	-2	113	+11	-178	-19
36	+3	-127	-12	-183	-20
-55	-4	-129	-13	-192	-21
73	+5	136	+14	-199	-22
-77	-6	-150	-15	-212	-23
-81	-7	-155	-16	-217	-24
90	+8	-159	-17	-229	-25
-95	-9				

A soma da classificação com sinais tomando valores positivos é  $3 + 5 + 8 + 11 + 14 + 18 = 59$ ,  $n(n + 1)/4 = 162,5$  e  $n(n + 1)(2n + 1)/24 = 1.381,25$ , então

$$z = \frac{59 - 162,5}{\sqrt{1.381,25}} = -2,78$$

O teste de nível 0,01 com cauda inferior rejeita  $H_0$  se  $z \leq -2,33$ . Como  $-2,78 \leq -2,33$ ,  $H_0$  é rejeitada em favor da conclusão de que a média verdadeira da força de compressão é menor do que 50.000. ■

Quando há empates nas magnitudes absolutas, tal que a média dos postos deve ser utilizada, é ainda correto padronizar  $S_+$  subtraindo  $n(n + 1)/4$ , mas a seguinte fórmula corrigida para variância deve ser aplicada:

$$\sigma_{S_+}^2 = \frac{1}{24}n(n + 1)(2n + 1) - \frac{1}{48}\sum(\tau_i - 1)(\tau_i)(\tau_i + 1) \quad (15.2)$$

onde  $\tau_i$  é o número de empates no  $i$ -ésimo conjunto de valores empatados e a soma é sobre todos os conjuntos de valores empatados. Se, por exemplo,  $n = 10$  e a classificação com sinais é 1, 2, -4, -4, 4, 6, 7, 8,5, 8,5 e 10, então há dois conjuntos empatados com  $\tau_1 = 3$  e  $\tau_2 = 2$ , de tal forma que a somatória é  $(2)(3)(4) + (1)(2)(3) = 30$  e  $\sigma_{S_+}^2 = 96,25 - 30/48 = 95,62$ . O denominador em (15.1) deve ser substituído pela raiz quadrada de (15.2), embora, como este exemplo mostra, a correção seja normalmente insignificante.

### Eficiência do Teste de Postos com Sinais de Wilcoxon

Quando a distribuição subjacente sendo amostrada é normal, tanto o teste  $t$  como o teste de classificação com sinais pode ser utilizado para testar uma hipótese sobre  $\mu$ . O teste  $t$  é o melhor teste em tal situação porque entre todos os testes de nível  $\alpha$ , este é o que tem um  $\beta$  mínimo. Como normalmente se acredita que há muitas situações

experimentais nas quais a hipótese de normalidade pode ser razoavelmente aceita, assim como algumas nas quais não deveria ser, há duas questões que devem ser levadas em consideração para comparar os dois testes:

1. Quando a distribuição subjacente é normal (a “base” do teste  $t$ ), quanto é perdido com a utilização do teste de classificação com sinais?
2. Quando a distribuição subjacente não é normal, uma melhora significativa pode ser atingida pela utilização do teste de classificação com sinais?

Se o teste de Wilcoxon não sofre muito em relação ao teste  $t$  na “base” anterior, e tem desempenho significativamente melhor do que o teste  $t$  para uma grande quantidade de outras distribuições, então haverá um forte argumento para a utilização do teste de Wilcoxon.

Infelizmente, não há uma resposta simples para estas duas questões. Refletindo, não é de se surpreender que o teste  $t$  possa ter um desempenho pobre quando a distribuição subjacente tem “caudas pesadas” (ou seja, quando os valores observados que caem longe de  $\mu$  são relativamente mais prováveis do que quando sua distribuição é normal). Isto porque o comportamento do teste  $t$  depende da média amostral, que pode ser muito instável na presença de caudas pesadas. A dificuldade em fornecer respostas às duas questões é que  $\beta$  do teste de Wilcoxon é muito difícil de obter e estudar para qualquer distribuição subjacente, e o mesmo pode ser dito do teste  $t$  quando a distribuição não é normal. Ainda que  $\beta$  fosse facilmente obtido, qualquer medida de eficiência dependeria claramente de qual distribuição subjacente fosse postulada. Várias medidas eficientes diferentes têm sido propostas pelos estatísticos; uma que muitos consideram confiável é chamada de **eficiência relativa assintótica** (ERA). A ERA de um teste em relação a outro é essencialmente o quociente dos tamanhos das amostras necessários para obter probabilidades de erro idênticas para os dois testes. Assim se o ERA de um teste em relação ao segundo é igual a 0,5, então quando o tamanho da amostra for grande, o tamanho da amostra do primeiro teste deverá ser o dobro do tamanho da amostra do segundo para que ele tenha um desempenho tão bom quanto este. Embora o ERA não caracterize o desempenho do teste para tamanhos da amostra pequenos, os resultados a seguir podem ser demonstrados como verdadeiros:

1. Quando a distribuição subjacente é normal, o ERA do teste de Wilcoxon em relação ao teste  $t$  é de aproximadamente 0,95.
2. Para qualquer distribuição, o ERA será de pelo menos 0,86 e para várias distribuições será muito maior do que 1.

Podemos resumir estes resultados dizendo que, em problemas com amostras grandes, o teste de Wilcoxon nunca é muito menos eficiente do que o teste  $t$  e poderia ser muito mais eficiente se a distribuição subjacente fosse distante da normal. Ainda que a questão esteja longe de ser resolvida no caso de tamanhos da amostra obtidos em muitos problemas práticos, estudos têm mostrado que o teste de Wilcoxon funciona razoavelmente e é dessa maneira uma alternativa viável para o teste  $t$ .

---

## Exercícios | Seção 15.1 (1–9)

1. Reconsidere a situação descrita no Exercício 32 da Seção 8.2, e utilize o teste de Wilcoxon com  $\alpha = 0,05$  para testar a hipótese relevante.
2. Utilize o teste de Wilcoxon para analisar os dados fornecidos no Exemplo 8.9.
3. Os dados correspondentes referem-se a um subconjunto das informações reportadas no artigo “Synovial Fluid pH, Lactate, Oxygen and Carbon Dioxide Partial Pressure in Various Joint Diseases” (*Arthritis and Rheumatism*, 1971, p. 476-477). As observações são valores de pH do fluido sinovial (que lubrifica juntas e tendões)

tomados a partir de joelhos de indivíduos que sofrem de artrite. Assumindo que a média verdadeira de pH para indivíduos que não têm artrite seja 7,39, o nível 0,05 do teste verifica se os dados indicam uma diferença entre os valores médio de pH para indivíduos com artrite e sem artrite.

7,02	7,35	7,34	7,17	7,28	7,77	7,09
7,22	7,45	6,95	7,40	7,10	7,32	7,14

4. Uma amostra aleatória de 15 mecânicos automotivos certificados para trabalharem em um tipo determinado

de carro foi selecionada, e o tempo (em minutos) necessário para diagnosticar cada problema em particular foi determinado, resultando nos dados a seguir:

30,6 30,1 15,6 26,7 27,1 25,4 35,0 30,8  
31,9 53,2 12,5 23,2 8,8 24,9 30,2

Utilize o teste de Wilcoxon ao nível 0,10 de significância para decidir se os dados sugerem que a média verdadeira do tempo de diagnóstico é menor do que 30 minutos.

5. Ambos os métodos gravimétrico e espectrofotométrico estão sendo considerados para determinar o conteúdo de fosfato contido em um material particular. Doze amostras do material são obtidas, cada uma é dividida ao meio e uma determinação é feita em cada metade utilizando um dos dois métodos, resultando nos dados a seguir:

Amostra	1	2	3	4
Gravimétrica	54,7	58,5	66,8	46,1
Espectrofotométrica	55,0	55,7	62,9	45,5
Amostra	5	6	7	8
Gravimétrica	52,3	74,3	92,5	40,2
Espectrofotométrica	51,1	75,4	89,6	38,4
Amostra	9	10	11	12
Gravimétrica	87,3	74,8	63,2	68,5
Espectrofotométrica	86,8	72,5	62,3	66,0

Utilize o teste de Wilcoxon para decidir se uma técnica fornece um valor diferente para a média do que a outra para este tipo de material.

6. Reconsidere a situação descrita no Exercício 41 da Seção 9.3 e utilize o teste de Wilcoxon para testar as hipóteses apropriadas.
7. Utilize a versão de amostras grandes do teste de Wilcoxon ao nível 0,05 de significância sobre os dados do Exercício 37 na Seção 9.3 para decidir se a diferença da média verdadeira entre as concentrações externas e internas excedem 0,20.

8. As 25 observações correspondentes sobre resistência à fratura do prato de base do aço temperado com níquel a 18% foram reportadas no artigo “Fracture Testing of Weldments”, (*ASTM Special Publ. No. 381, 1965, p. 328-356*). Suponha que uma empresa concordará em comprar este aço para uma aplicação especial somente se puder ser demonstrado com veemência por uma evidência experimental que a resistência média verdadeira excede 75. Assumindo que a distribuição de resistência à fratura seja simétrica, defina e teste a hipótese apropriada ao nível 0,05 e calcule o valor  $P$ .

69,5 71,9 72,6 73,1 73,3 73,5 74,1 74,2 75,3  
75,5 75,7 75,8 76,1 76,2 76,2 76,9 77,0 77,9  
78,1 79,6 79,7 80,1 82,2 83,7 93,7

9. Suponha que as observações  $X_1, X_2, \dots, X_n$  são feitas em um processo nos instantes 1, 2, ...,  $n$ . Com base nesses dados, desejamos testar

$H_0$ : os  $X_i$ s constituem uma seqüência distribuída independente e identicamente

versus

$H_a$ :  $X_{i+1}$  tende a ser maior do  $X_i$  para  $i = 1, \dots, n$  (uma tendência a crescimento)

Suponha que os  $X_i$ s são classificados de 1 até  $n$ . Então quando  $H_a$  é verdadeira, postos mais altos tendem a ocorrer no fim da seqüência, ao passo que se  $H_0$  é verdadeira, postos grandes e pequenos tendem a ocorrer ao mesmo tempo. Seja  $R_i$  a classificação de  $X_i$  e considere o teste estatístico  $D = \sum_{i=1}^n (R_i - i)^2$ . Então os valores pequenos de  $D$  dão suporte a  $H_a$  (por exemplo, o menor valor é 0 para  $R_1 = 1, R_2 = 2, \dots, R_n = n$ ), portanto  $H_0$  deveria ser rejeitada em favor de  $H_a$  se  $d \leq c$ . Quando  $H_0$  é verdadeira, qualquer seqüência de postos tem probabilidade  $1/n!$ . Utilize isto a fim de encontrar  $c$  para o qual o teste tem um nível o mais próximo de 0,10 quanto possível no caso de  $n = 4$ . [Sugestão: Liste as seqüências de classificação de 4!, calcule  $d$  para cada uma e então obtenha a distribuição nula de  $D$ . Veja o livro de Lehmann (na Bibliografia) para obter mais informações.]

## 15.2 | O Teste de Soma de Classificação de Wilcoxon

Quando pelo menos um dos tamanhos da amostra em um problema de duas amostras é pequeno, o teste  $t$  exige a hipótese de normalidade (pelo menos aproximadamente). Há situações nas quais um pesquisador deseja utilizar um teste que seja válido mesmo se as distribuições subjacentes sejam não-normais. Descrevemos agora este teste, chamado de **teste de soma de classificação de Wilcoxon**. Um nome alternativo para o procedimento é o teste de Mann-Whitney, embora o teste estatístico de Mann-Whitney seja algumas vezes expresso de maneira ligeiramente diferente daquela do teste de Wilcoxon. O procedimento do teste de Wilcoxon é livre de distribuição porque terá o nível desejado de significância para uma classe muito grande de distribuições subjacentes.

**HIPÓTESES**

$X_1, \dots, X_m$  e  $Y_1, \dots, Y_n$  são duas amostras aleatórias independentes a partir de distribuições contínuas com médias  $\mu_1$  e  $\mu_2$ , respectivamente. As distribuições  $X$  e  $Y$  têm a mesma forma e extensão, a única diferença possível entre as duas está no valor de  $\mu_1$  e  $\mu_2$ .

Quando  $H_0: \mu_1 - \mu_2 = \Delta_0$  é verdadeira, a distribuição  $X$  é deslocada pela quantidade  $\Delta_0$  para a direita da distribuição  $Y$ , embora quando  $H_0$  é falsa, o deslocamento ocorrerá por uma quantidade diferente de  $\Delta_0$ .

**Desenvolvimento do Teste Quando  $m = 3, n = 4$**

Considere primeiro o teste de  $H_0: \mu_1 - \mu_2 = 0$ . Se  $\mu_1$  é realmente muito maior do que  $\mu_2$ , então a maioria dos  $x$ s observados cairá à direita dos  $y$ s observados. Entretanto, se  $H_0$  é verdadeira, então os valores observados a partir das duas amostras deverão ser alternados. O teste estatístico fornecerá uma quantificação da alternância nas duas amostras.

Considere o caso  $m = 3$  e  $n = 4$ . Então, se todos os três  $x$ s observados estiverem à direita de todos os quatro  $y$ s observados, isto fornecerá uma evidência forte de rejeição de  $H_0$ , em favor de  $H_a: \mu_1 - \mu_2 \neq 0$ , com uma conclusão semelhante sendo apropriada se todos os três  $x$ s caem abaixo de todos os quatro  $y$ s. Suponha um conjunto de  $x$ s e  $y$ s em uma amostra combinada de tamanho  $m + n = 7$  e classifique estas observações da menor para a maior, com a menor recebendo a classificação 1 e a maior, a classificação 7. Se tanto a maioria dos maiores postos como a maioria dos menores postos fossem associadas com observações  $X$ , iríamos iniciar com o  $H_0$  duvidoso. Isto sugere o teste estatístico

$$W = \text{a soma dos postos na amostra combinada associada às } X \text{ observações} \tag{15.3}$$

Para os valores de  $m$  e  $n$  sendo considerados, o menor valor possível de  $W$  é  $w = 1 + 2 + 3 = 6$  (se todos os três  $x$ s são menores do que todos os quatro  $y$ s) e o maior valor possível é  $w = 5 + 6 + 7 = 18$  (se todos os três  $x$ s são maiores do que todos os quatro  $y$ s).

Como exemplo, suponha  $x_1 = -3,10, x_2 = 1,67, x_3 = 2,01, y_1 = 5,27, y_2 = 1,89, y_3 = 3,86$  e  $y_4 = 0,19$ . Então o exemplo ordenado é  $-3,10; 0,19; 1,67; 1,89; 2,01; 3,86$  e  $5,27$ . Os postos de  $X$  para este exemplo são 1 (para  $-3,10$ ), 3 (para  $1,67$ ) e 5 (para  $2,01$ ) portanto o valor calculado de  $W$  é  $w = 1 + 3 + 5 = 9$ .

O procedimento de teste baseado em estatística (15.3) é rejeitar  $H_0$  se o valor de  $w$  calculado for “muito extremo”, ou seja,  $\geq c$  para o teste de cauda superior e  $\leq c$  para o teste de cauda inferior e ambos  $\geq c_1$  e  $\leq c_2$  para um teste de dupla cauda. A(s) constante(s) crítica(s)  $c$  ( $c_1, c_2$ ) deve(m) ser escolhida(s) tal que o teste tenha o nível de significância desejado  $\alpha$ . Para verificar como isto deve ser feito, lembre-se de que quando  $H_0$  é verdadeira, todas as sete observações vêm da mesma população. Isto significa que sob  $H_0$ , qualquer tripla possível de postos associados com os três  $x$ s — tais como (1, 4, 5), (3, 5, 6) ou (5, 6, 7) — têm a mesma probabilidade que qualquer outra tripla de classificação possível. Como há  $\binom{7}{3} = 35$  possíveis triplas de classificação sob  $H_0$ , cada uma tem probabilidade  $\frac{1}{35}$ . A partir de uma lista de todas as 35 triplas de classificação e o valor de  $w$  associado com cada uma, a distribuição de probabilidade de  $W$  pode ser imediatamente determinada. Por exemplo, há quatro triplas de classificação que têm o valor de  $w$  igual a 11 — (1, 3, 7), (1, 4, 6), (2, 3, 6) e (2, 4, 5) — portanto  $P(W = 11) = \frac{4}{35}$ . O resumo da listagem e dos cálculos aparece na Tabela 15.4.

**Tabela 15.4** Distribuição de probabilidade de  $W$  ( $m = 3, n = 4$ ) quando  $H_0$  é verdadeira

$w$	6	7	8	9	10	11	12	13	14	15	16	17	18
$P(W = w)$	$\frac{1}{35}$	$\frac{1}{35}$	$\frac{2}{35}$	$\frac{3}{35}$	$\frac{4}{35}$	$\frac{4}{35}$	$\frac{5}{35}$	$\frac{4}{35}$	$\frac{4}{35}$	$\frac{3}{35}$	$\frac{2}{35}$	$\frac{1}{35}$	$\frac{1}{35}$

A distribuição da Tabela 15.4 é simétrica em relação ao valor  $w = (6 + 18)/2 = 12$ , que é o valor médio na lista ordenada de possíveis valores de  $W$ . Isto acontece porque as duas triplas classificadas ( $r, s, t$ ) (com

$r < s < t$ ) e  $(8 - t, 8 - s, 8 - r)$  têm valores de  $w$  simétricos em relação a 12, portanto para cada tripla com o valor de  $w$  abaixo de 12, há uma tripla com o valor de  $w$  acima de 12 na mesma quantidade.

A hipótese alternativa é  $H_a: \mu_1 - \mu_2 > 0$ , então  $H_0$  deve ser rejeitada em favor de  $H_a$  para valores de  $W$  grandes. Escolhendo como a região de rejeição, o conjunto de valores de  $W \{17, 18\}$ ,  $\alpha = P(\text{tipo de erro I}) = P(\text{rejeita } H_0 \text{ quando } H_0 \text{ é verdadeira}) = P(W = 17 \text{ ou } 18 \text{ quando } H_0 \text{ é verdadeira}) = \frac{1}{35} + \frac{1}{35} = \frac{2}{35} = 0,057$ ; a região  $\{17, 18\}$  específica, portanto, um teste com nível de significância de aproximadamente 0,05. Da mesma maneira, a região  $\{6, 7\}$  que é apropriada para  $H_a: \mu_1 - \mu_2 < 0$  tem  $\alpha = 0,057 \approx 0,05$ . A região  $\{6, 7, 17, 18\}$  que é apropriada para a alternativa de dois lados, tem  $\alpha = \frac{4}{35} = 0,114$ . O valor de  $W$  para os dados fornecidos vários parágrafos atrás foi  $w = 9$ , que é muito próximo do valor médio 12, portanto  $H_0$  não seria rejeitada em um nível  $\alpha$  razoável para qualquer uma das três  $H_a$ s.

### Descrição Geral do Teste de Classificação com Sinal de Wilcoxon

A hipótese nula  $H_0: \mu_1 - \mu_2 = \Delta_0$  é manipulada subtraindo  $\Delta_0$  de cada  $X_i$  e usando o  $(X_i - \Delta_0)$ s como os  $X_i$ s usados anteriormente. Lembrando que para qualquer  $K$  positivo inteiro, a soma dos primeiros  $K$  inteiros é  $K(K + 1)/2$ , o menor valor possível  $W$  da estatística é  $m(m + 1)/2$ , que ocorre quando os  $(X_i - \Delta_0)$ s estão todos à esquerda da amostra  $Y$ . O maior valor possível de  $W$  ocorre quando os  $(X_i - \Delta_0)$ s localizam-se inteiramente à direita do  $Y$ s; neste caso,  $W = (n + 1) + \dots + (m + n) = (\text{soma dos primeiros } m + n \text{ inteiros}) - (\text{soma dos primeiros } n \text{ inteiros})$ , o que resulta em  $m(m + 2n + 1)/2$ . Como no caso especial  $m = 3$  e  $n = 4$ , a distribuição de  $W$  é simétrica em relação ao número que está entre o maior e o menor valor; este valor médio é  $m(m + n + 1)/2$ . Por causa desta simetria, as probabilidades envolvendo os valores críticos de cauda inferior podem ser obtidas a partir dos valores de cauda superior correspondentes.

Hipótese nula:  $H_0: \mu_1 - \mu_2 = \Delta_0$

Valor do teste estatístico:  $w = \sum_{i=1}^m r_i$  onde  $r_i =$  classificação de  $(x_i - \Delta_0)$  na amostra combinada de  $m + n$   $(x - \Delta_0)$ s e  $y$ s

#### Hipótese Alternativa

$$H_a: \mu_1 - \mu_2 > \Delta_0$$

$$H_a: \mu_1 - \mu_2 < \Delta_0$$

$$H_a: \mu_1 - \mu_2 \neq \Delta_0$$

#### Região de Rejeição

$$w \geq c_1$$

$$w \leq m(m + n + 1) - c_1$$

$$\text{tanto } w \geq c \text{ como } w \leq m(m + n + 1) - c$$

onde  $P(W \geq c_1 \text{ quando } H_0 \text{ é verdadeira}) \approx \alpha$ ,  $P(W \geq c \text{ quando } H_0 \text{ é verdadeira}) \approx \alpha/2$ .

Como  $W$  tem uma distribuição de probabilidade discreta, não existirá sempre um valor crítico que corresponda exatamente a um dos níveis usuais de significância. A Tabela A.14 do Apêndice fornece os valores críticos de cauda superior para as probabilidades o mais próximas de 0,05; 0,025; 0,01 e 0,005, a partir dos quais os testes de dupla cauda de nível 0,05 ou 0,01 podem ser obtidos. A tabela fornece informação somente para  $m = 3, 4, \dots, 8$  e  $n = m, m + 1, \dots, 8$  (ou seja,  $3 \leq m \leq n \leq 8$ ). Para valores de  $m$  e  $n$  que excedam 8, uma aproximação normal pode ser utilizada. Para utilizar a tabela para  $m$  e  $n$  pequenos, entretanto, *as amostras de  $X$  e  $Y$  devem ser rotuladas tal que  $m \leq n$ .*

#### Exemplo 15.4

A concentração de fluoreto urinário (partes por milhão) foi medida ambos para uma amostra de pastagem de criação de gado em uma área previamente exposta à poluição de fluoreto e para uma amostra semelhante de pastagem em uma região não-poluída:

<b>Poluída</b>	21,3	18,7	23,0	17,1	16,8	20,9	19,7
<b>Não-poluída</b>	14,2	18,3	17,2	18,4	20,0		

Os dados indicam fortemente que a média real de concentração de fluoreto para o pasto de criação de gado na região poluída é maior do que na região não-poluída? Utilize o teste de classificação com sinal de Wilcoxon ao nível  $\alpha = 0,01$ .

Os tamanhos da amostra neste caso são 7 e 5. Para obter  $m \leq n$ , identifique as observações não-poluídas como os  $x$ s ( $x_1 = 14,2, \dots, x_5 = 20,0$ ) e as observações poluídas como os  $y$ s. Assim  $\mu_1$  é a média real da concentração de fluoreto sem poluição, e  $\mu_2$  é a média real da concentração com poluição. A hipótese alternativa é  $H_a: \mu_1 - \mu_2 < 0$  (a poluição causa aumento na concentração), portanto um teste de cauda inferior é apropriado. Da Tabela A.14 do Apêndice com  $m = 5$  e  $n = 7$ ,  $P(W \geq 47$  quando  $H_0$  é verdadeira)  $\approx 0,01$ . O valor crítico para o teste de cauda inferior é, portanto,  $m(m + n + 1) - 47 = 5(13) - 47 = 18$ ;  $H_0$  será agora rejeitada se  $w \leq 18$ . A amostra ordenada é mostrada a seguir; o  $W$  calculado é  $w = r_1 + r_2 + \dots + r_5$  (onde  $r_i$  é a classificação de  $x_i$ )  $= 1 + 5 + 4 + 6 + 9 = 25$ . Como 25 não é  $\leq 18$ ,  $H_0$  não é rejeitado no nível 0,01 (aproximadamente).

$x$	$y$	$y$	$x$	$x$	$x$	$y$	$y$	$x$	$y$	$y$	$y$
14,2	16,8	17,1	17,2	18,3	18,4	18,7	19,7	20,0	20,9	21,3	23,0
1	2	3	4	5	6	7	8	9	10	11	12

### Uma Aproximação Normal para $W$

Quando ambos  $m$  e  $n$  excedem 8, a distribuição de  $W$  pode ser aproximada por uma curva normal apropriada, e esta aproximação pode ser utilizada no lugar da Tabela A.14 do Apêndice. Para obter a aproximação, precisamos de  $\mu_W$  e  $\sigma_W^2$  quando  $H_0$  é verdadeira. Neste caso, a classificação  $R_i$  de  $X_i - \Delta_0$  é igualmente provável para ser qualquer um dos valores possíveis 1, 2, 3, . . . ,  $m + n$  ( $R_i$  tem uma distribuição uniforme discreta nos primeiros  $m + n$  inteiros positivos), portanto  $\mu_{R_i} = (m + n + 1)/2$ . Isto resulta em, visto que  $W = \sum R_i$ ,

$$\mu_W = \mu_{R_1} + \mu_{R_2} + \dots + \mu_{R_m} = \frac{m(m + n + 1)}{2} \tag{15.4}$$

A variância de  $R_i$  é também facilmente calculada por ser  $(m + n + 1)(m + n - 1)/12$ . Entretanto, como os  $R_i$ s não são variáveis independentes,  $V(W) \neq mV(R_i)$ . Utilizando o fato de que, para quaisquer dois inteiros distintos  $a$  e  $b$  entre 1 e  $m + n$  inclusive,  $P(R_i = a, R_j = b) = 1/[(m + n)(m + n - 1)]$  (dois inteiros estão sendo amostrados sem substituição),  $Cov(R_i, R_j) = -(m + n + 1)/12$ , o que produz

$$\sigma_W^2 = \sum_{i=1}^m V(R_i) + \sum_{i \neq j} Cov(R_i, R_j) = \frac{mn(m + n + 1)}{12} \tag{15.5}$$

O Teorema de Limite Central pode ser utilizado para concluir que quando  $H_0$  é verdadeira, o teste estatístico

$$Z = \frac{W - m(m + n + 1)/2}{\sqrt{mn(m + n + 1)/12}}$$

tem aproximadamente uma distribuição normal padrão. Esta estatística é utilizada em conjunto com os valores críticos  $z_\alpha$ ,  $-z_\alpha$  e  $\pm z_{\alpha/2}$  para testes com cauda superior, inferior ou com duas caudas, respectivamente.

## Exemplo 15.5

Um artigo no *Journal of Applied Physiology* (“Histamine Content in Sputum from Allergic and Non-Allergic Individuals”, 1969, p. 535-539) informa os seguintes dados sobre níveis de histamina na saliva (peso seco de saliva em  $\mu\text{g/g}$ ) para uma amostra de 9 indivíduos classificados como alérgicos e outra amostra de 13 indivíduos classificados como não-alérgicos:

Alérgicos	67,6	39,6	1.651,0	100,0	65,9	1.112,0	31,0	102,4	64,7					
Não-alérgicos	34,3	27,3	35,4	48,1	5,2	29,1	4,7	41,7	48,0	6,6	18,9	32,4	45,5	

Os dados indicam que há uma diferença na média real do nível de histamina na saliva entre alérgicos e não-alérgicos?

Como ambos os tamanhos das amostras excedem 8, utilizamos a aproximação normal. A hipótese nula é  $H_0: \mu_1 - \mu_2 = 0$  e os postos observados dos  $x$ s são  $r_1 = 18, r_2 = 11, r_3 = 22, r_4 = 19, r_5 = 17, r_6 = 21, r_7 = 7, r_8 = 20$  e  $r_9 = 16$ , portanto  $w = \sum r_i = 151$ . A média e a variância de  $W$  são dadas por  $\mu_w = 9(23)/2 = 103,5$  e  $\sigma_w^2 = 9(13)(23)/12 = 224,25$ . Desta forma

$$z = \frac{151 - 103,5}{\sqrt{224,25}} = 3,17$$

A hipótese alternativa é  $H_a: \mu_1 - \mu_2 \neq 0$ , portanto ao nível 0,01  $H_0$  é rejeitada se tanto  $z \geq 2,58$  como  $z \leq -2,58$ . Como  $3,17 \geq 2,58$ ,  $H_0$  é rejeitado e concluímos que há uma diferença na média real dos níveis de histamina na saliva (o artigo também utilizou o teste de Wilcoxon). ■

## Empates no Teste de Soma de Postos de Wilcoxon

Teoricamente, a hipótese de continuidade de duas distribuições garante que todos os  $m + n$ ,  $x$ s e  $y$ s observados terão valores diferentes. Na prática, entretanto, haverá frequentemente empates nos valores observados. Assim como no teste classificado com sinais de Wilcoxon, a prática comum em lidar com empates é atribuir cada uma das observações em um conjunto particular de empates a média dos postos que eles receberiam se diferissem muito ligeiramente um do outro.

Se ambos os tamanhos das amostras excederem 8, o numerador de  $Z$  é ainda apropriado, mas o denominador deveria ser substituído pela raiz quadrada da variância ajustada

$$\sigma_w^2 = \frac{mn(m+n+1)}{12} - \frac{mn}{12(m+n)(m+n-1)} \sum (\tau_i - 1)(\tau_i)(\tau_i + 1) \quad (15.6)$$

onde  $\tau_i$  é o número de empates no  $i$ -ésimo conjunto de valores de empates e a soma de todos os conjuntos de valores de empates. A menos que haja um grande número de empates, há pequenas diferenças entre as Equações (15.6) e (15.5).

## Eficiência do Teste de Soma de Postos de Wilcoxon

Quando as distribuições que estão sendo amostradas são ambas normais com  $\sigma_1 = \sigma_2$ , e, portanto tendo as mesmas formas e distribuições, tanto o teste  $t$  como o teste de Wilcoxon podem ser utilizados (o teste  $t$  de duas amostras assume a normalidade, mas não as variâncias iguais, portanto as hipóteses subjacentes de seu uso são mais restritivas em um sentido e menos em outro do que aquele para o teste de Wilcoxon). Nesta situação, o teste  $t$  é o melhor entre todos os testes possíveis no sentido de minimizar  $\beta$  para qualquer  $\alpha$  fixado. Entretanto, um pesquisador nunca pode estar absolutamente certo de que as hipóteses subjacentes são satisfeitas. É, portanto,

relevante perguntar (1) o quanto é perdido com a utilização do teste de Wilcoxon em vez do teste  $t$  quando as distribuições são normais com variâncias iguais e (2) como  $W$  se compara com  $T$  em situações não-normais.

A noção da eficiência do teste foi discutida na seção anterior em conexão com o teste  $t$  de uma amostra e o outro de classificação com sinais de Wilcoxon. Os resultados para os testes de duas amostras são os mesmos que os de uma amostra. Quando tanto a normalidade quanto as variâncias iguais satisfazem, o teste de soma de classificação é aproximadamente 95% tão eficiente quanto o teste  $t$  em amostras grandes. Ou seja, o teste  $t$  fornecerá as mesmas probabilidades de erro que o teste de Wilcoxon utilizando tamanhos da amostra ligeiramente menores. Por outro lado, o teste de Wilcoxon será sempre pelo menos 86% tão eficiente quanto o teste  $t$ , podendo ser ainda mais eficiente se as distribuições subjacentes forem bastante não-normais, especialmente com caudas pesadas. A comparação do teste de Wilcoxon com o teste  $t$  de duas amostras é menos definida. O teste  $t$  não é reconhecido como o melhor teste em qualquer situação, portanto parece seguro concluir que na medida em que as distribuições de população têm formas e distribuições semelhantes, o comportamento do teste de Wilcoxon deve ser comparado preferencialmente ao teste  $t$  de duas amostras.

Finalmente, observamos que os cálculos de  $\beta$  para o teste de Wilcoxon são um tanto difíceis, porque a distribuição de  $W$  quando  $H_0$  é falsa depende não apenas de  $\mu_1 - \mu_2$ , mas também das formas das duas distribuições. Para a maioria das distribuições subjacentes, a distribuição não-nula de  $W$  é virtualmente intratável. Esta é a razão por que os estatísticos desenvolveram a eficiência de amostras grandes (assintótica relativa) como um meio de comparação dos testes. Com os recursos de um software de computador atual, outra abordagem para o cálculo de  $\beta$  é a realização de um experimento de simulação.

**Exercícios** | Seção 15.2 (10–16)

**10.** Em um experimento para comparar o poder de colagem de dois adesivos diferentes, cada um foi utilizado em cinco amostras, para grudar duas superfícies. A força necessária para separar tais superfícies foi determinada para cada amostra. Para o adesivo 1, os valores resultantes foram 229, 286, 245, 299 e 250, enquanto para o adesivo 2, foram obtidas as observações 213, 179, 163, 247 e 225. Seja  $\mu_i$  a média real da força de adesão do adesivo do tipo  $i$ . Utilize o teste de soma de postos de Wilcoxon ao nível 0,05 para testar  $H_0: \mu_1 = \mu_2$  versus  $H_a: \mu_1 > \mu_2$ .

**11.** O artigo “A Study of Wood Stove Particulate Emissions” (*J. Air Pollution Control Assn.*, 1979, p. 724-728) relata os seguintes dados sobre o tempo de queima (em horas) para amostras de carvalho e pinheiro. Teste ao nível 0,05 para verificar se há qualquer diferença no tempo de queima médio real para os dois tipos de madeira.

<b>Carvalho</b>	1,72	0,67	1,55	1,56	1,42	1,23	1,77	0,48
<b>Pinheiro</b>	0,98	1,40	1,33	1,52	0,73	1,20		

**12.** O processo de produção de um determinado tipo de filme “tempo-zero” (filme que inicia a revelação assim que uma foto é tirada) sofreu uma modificação. Devido ao fato de a modificação envolver custos extras, ela será incorporada ao processo somente se os dados de amostra indicarem fortemente uma diminuição no tempo médio real de revelação maior do que 1 segundo. Assumindo que as distribuições dos tempos de revelação diferem somente no que diz respeito à localização, utilize o nível 0,05 para o teste de soma de postos de

Wilcoxon nos dados correspondentes para testar as hipóteses apropriadas.

<b>Processo Original</b>	8,6	5,1	4,5	5,4	6,3	6,6	5,7	8,5
<b>Processo Modificado</b>	5,5	4,0	3,8	6,0	5,8	4,9	7,0	5,7

**13.** Os dados a seguir são resultantes de um experimento, realizado em porquinhos-da-Índia, durante o período de seis semanas, a fim de comparar os efeitos da vitamina C presente em suco de laranja e em ácido ascórbico sintético no comprimento dos odontoblastos. (“The Growth of the Odontoblasts of the Incisor Tooth as a Criterion of the Vitamin C Intake of the Guinea Pig”, *J. Nutrition*, 1947, p. 491-504). Utilize o teste de soma de postos de Wilcoxon ao nível 0,01 para decidir se o comprimento médio real difere para os dois tipos de vitamina C ingeridos. Calcule também um valor  $P$  aproximado.

<b>Suco de Laranja</b>	8,2	9,4	9,6	9,7	10,0	14,5
	15,2	16,1	17,6	21,5		
<b>Ácido Ascórbico</b>	4,2	5,2	5,8	6,4	7,0	7,3
	10,1	11,2	11,3	11,5		

**14.** Teste as hipóteses sugeridas no Exercício 13 utilizando os seguintes dados:

<b>Suco de Laranja</b>	8,2	9,5	9,5	9,7	10,0	14,5
	15,2	16,1	17,6	21,5		
<b>Ácido Ascórbico</b>	4,2	5,2	5,8	6,4	7,0	7,3
	9,5	10,0	11,5	11,5		

**15.** O artigo “Measuring the Exposure of Infants to Tobacco Smoke” (*N. Engl. J. Med.*, 1984, p. 1075-1078)



relata um estudo no qual várias medidas foram tomadas com base em duas amostras aleatórias: uma de crianças que foram expostas a fumantes dentro de casa; e outra, de crianças não-expostas. Os dados a seguir consistem em observações da concentração de cotanina, um dos principais metabólitos da nicotina, na urina (os valores correspondem a um subconjunto dos dados originais e foram obtidos a partir de um gráfico apresentado no artigo). Os dados sugerem que o nível médio real de cotanina em crianças expostas, comparado àquele em crianças não-expostas, é superior a 25? Realize um teste ao nível de significância 0,05.

<b>Não-expostas</b>	8	11	12	14	20	43	111	
<b>Expostas</b>	35	56	83	92	128	150	176	208

16. Considere a situação descrita no Exercício 79 do Capítulo 9 e o resultado do MINITAB correspondente (a letra grega eta é utilizada para indicar uma mediana).

**Intervalo de Confiança e Teste de Mann-Whitney**  
 bom N = 8 Mediana = 0,540  
 ruim N = 8 Mediana = 2,400  
 Estimativa Pontual para ETA1-ETA2 é -1,155  
 IC 95,9% para ETA1-ETA2 é (-3,160, -0,409)  
 W = 41,0  
 Teste de ETA1 = ETA2 vs ETA1 < ETA2 é significativo para 0,0027

- Verifique se o valor do teste estatístico do MINITAB está correto.
- Realize um teste de hipóteses apropriado utilizando um nível de significância de 0,01.

## 15.3 | Intervalos de Confiança Livres de Distribuição

O método que utilizamos anteriormente para construir um intervalo de confiança (IC) foi o seguinte: iniciar com uma variável aleatória ( $Z$ ,  $T$ ,  $\chi^2$ ,  $F$  ou semelhante) que depende do parâmetro de interesse e da expressão para a probabilidade em função da variável. Manipular as desigualdades da expressão para isolar os parâmetros entre as extremidades aleatórias e finalmente substituir os valores calculados por variáveis aleatórias. Outro método geral para obter IC's tira vantagem da relação entre os procedimentos de teste e os ICs. Um IC de  $100(1 - \alpha)\%$  para um parâmetro  $\theta$  pode ser obtido a partir de um teste de nível  $\alpha$  para  $H_0: \theta = \theta_0$  versus  $H_a: \theta \neq \theta_0$ . Este método será utilizado para obter intervalos associados com o teste de postos com sinais de Wilcoxon e o teste de soma de postos de Wilcoxon.

Antes de utilizar o método para obter novos intervalos, considere o teste  $t$  e o intervalo  $t$ . Suponha que uma amostra aleatória de  $n = 25$  observações a partir de uma população normal produz as estatísticas  $\bar{x} = 100$ ,  $s = 20$ . Então um IC de 90% para  $\mu$  é

$$\left( \bar{x} - t_{0,05,24} \cdot \frac{s}{\sqrt{25}}, \bar{x} + t_{0,05,24} \cdot \frac{s}{\sqrt{25}} \right) = (93,16, 106,84) \quad (15.7)$$

Suponha que em vez de um IC, quiséssemos testar uma hipótese sobre  $\mu$ . Para  $H_0: \mu = \mu_0$  versus  $H_a: \mu \neq \mu_0$ , o teste  $t$  ao nível 0,10 especifica que  $H_0$  deve ser rejeitada se  $t$  for  $\geq 1,711$  ou  $\leq -1,711$ , onde

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{25}} = \frac{100 - \mu_0}{20/\sqrt{25}} = \frac{100 - \mu_0}{4} \quad (15.8)$$

Considere agora o valor nulo  $\mu_0 = 95$ . Então  $t = 1,25$  e, portanto,  $H_0$  não é rejeitada. Deste modo, se  $\mu_0 = 104$ , então  $t = -1$ , portanto,  $H_0$  não é rejeitada novamente. Entretanto, se  $\mu_0 = 90$ , então  $t = 2,5$ , portanto  $H_0$  é rejeitada, e se  $\mu_0 = 108$ , então  $t = -2$ , portanto  $H_0$  é novamente rejeitada. Considerando outros valores de  $\mu_0$  e a decisão resultante de cada um, surge a seguinte questão: *Cada número contido no intervalo (15.7) especifica um valor de  $\mu_0$  para o qual o valor  $t$  de (15.8) leva à não rejeição de  $H_0$ , enquanto todo número fora do intervalo (15.7) corresponde a um  $t$  para o qual  $H_0$  é rejeitada.* Ou seja, para os valores fixados de  $n$ ,  $\bar{x}$  e  $s$ , o conjunto de todos os valores de  $\mu_0$  para os quais o teste  $H_0: \mu = \mu_0$  versus  $H_a: \mu \neq \mu_0$  resulta na não-rejeição de  $H_0$  é exatamente o intervalo (15.7).

**PROPOSIÇÃO**

Considere um procedimento de teste ao nível  $\alpha$  para testar  $H_0: \theta = \theta_0$  versus  $H_a: \theta \neq \theta_0$ . Para valores da amostra fixos, seja  $A$  o conjunto de todos os valores  $\theta_0$  para os quais  $H_0$  não é rejeitada. Então  $A$  é um IC  $100(1 - \alpha)\%$  para  $\theta$ .

Existem certos exemplos patológicos nos quais o conjunto  $A$  definido na proposição não é um intervalo de valores  $\theta$  mas, em vez disso, o complemento de um intervalo ou algo ainda mais estranho. Mais precisamente, devemos substituir a noção de um IC por um conjunto de confiança. Nos casos de interesse, o conjunto  $A$  deixa de ser um intervalo.

**O Intervalo de Postos com Sinais de Wilcoxon**

Para testar  $H_0: \mu = \mu_0$  versus  $H_a: \mu \neq \mu_0$ , utilizando o teste de postos com sinais de Wilcoxon, onde  $\mu$  é a média de uma distribuição simétrica contínua, os valores absolutos  $|x_1 - \mu_0|, \dots, |x_n - \mu_0|$  são ordenados do menor para o maior, com o menor recebendo a classificação 1 e o maior, a classificação  $n$ . Cada classificação é, então, dada pelo sinal de seu  $x_i - \mu_0$  associado e a estatística de teste é a soma dos postos com sinal positivo. O teste bi-caudal rejeita  $H_0$  se  $s_+$  for  $\geq c$  ou  $\leq n(n + 1)/2 - c$ , onde  $c$  é obtido da Tabela A.13 do Apêndice, visto que o nível de significância  $\alpha$  desejado é especificado. Para  $x_1, \dots, x_n$ , fixados, o intervalo de classificação com sinal  $100(1 - \alpha)\%$  consistirá de todos  $\mu_0$  para os quais  $H_0: \mu = \mu_0$  não é rejeitada ao nível  $\alpha$ . Para identificar este intervalo, é conveniente expressar o teste estatístico  $S_+$  de outra forma.

$$S_+ = \text{número de médias dos pares } (X_i + X_j)/2 \text{ com } i \leq j \text{ que são } \geq \mu_0 \tag{15.9}$$

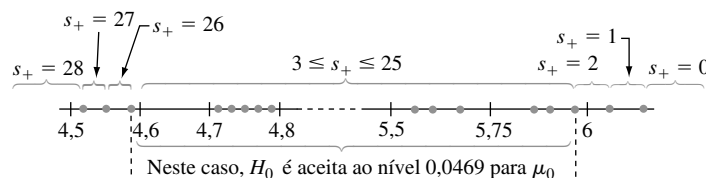
Ou seja, calculando a média entre cada  $x_j$ , na lista, e cada  $x_i$  à sua esquerda, inclusive  $(x_j + x_j)/2$  (que vale exatamente  $x_j$ ) e, contando o número destas médias que são  $\geq \mu_0$ , obtemos o valor de  $s_+$ . Ao caminhar da esquerda para a direita, na lista de valores amostrados, estamos simplesmente calculando a média de cada par de observações na amostra [novamente, incluindo  $(x_j + x_j)/2$ ] exatamente uma vez, portanto a ordem na qual as observações são listadas antes do cálculo da média não é importante. Não é difícil verificar a equivalência entre os dois métodos para calcular  $s_+$ . O número de médias de pares é  $\binom{n}{2} + n$  (o primeiro termo é devido ao cálculo das médias de diferentes observações e o segundo devido ao cálculo das médias de cada  $x_i$  com ele mesmo), que é igual a  $n(n + 1)/2$ . Se um grande número ou muito pouco dessas médias de pares são  $\geq \mu_0$ ,  $H_0$  é rejeitada.

**Exemplo 15.6**

As observações a seguir são valores de taxas de metabolismo cerebral para macacos rhesus:  $x_1 = 4,51, x_2 = 4,59, x_3 = 4,90, x_4 = 4,93, x_5 = 6,80, x_6 = 5,08, x_7 = 5,67$ . As 28 médias de pares são, em ordem crescente, valem

- 4,51 4,55 4,59 4,705 4,72 4,745 4,76 4,795 4,835 4,90
- 4,915 4,93 4,99 5,005 5,08 5,09 5,13 5,285 5,30 5,375
- 5,655 5,67 5,695 5,85 5,865 5,94 6,235 6,80

Os primeiros e os últimos destes são representados em um eixo de medidas na Figura 15.3.



**Figura 15.3** Gráfico dos dados para o Exemplo 15.6

Devido ao fato de a distribuição de  $S_+$  ser discreta,  $\alpha = 0,05$  não pode ser obtido com exatidão. A região de rejeição  $\{0, 1, 2, 26, 27, 28\}$  apresenta  $\alpha = 0,046$  que está o mais próximo possível de 0,05, de forma que o nível seja de, aproximadamente, 0,05. Assim, se o número de médias dos pares  $\geq \mu_0$  estiver entre 3 e 25, inclusive,  $H_0$  não é rejeitada. A partir da Figura 15.3 o IC 95% (aproximado) para  $\mu$  é (4,59; 5,94). ■

Em geral, uma vez que as médias dos pares são ordenadas da menor para a maior, as extremidades do intervalo de Wilcoxon são as duas médias “extremas”. Mais precisamente, seja  $\bar{x}_{(1)}$  a menor média dos pares,  $\bar{x}_{(2)}, \dots$ , a segunda menor,  $\dots$ , e  $\bar{x}_{(n(n+1)/2)}$  a maior delas.

#### PROPOSIÇÃO

Se o teste de postos com sinais de Wilcoxon, ao nível  $\alpha$ , para  $H_0: \mu = \mu_0$  versus  $H_a: \mu \neq \mu_0$  rejeitar  $H_0$  para  $s_+ \geq c$  ou  $s_+ \leq n(n+1)/2 - c$ , então um IC  $100(1 - \alpha)\%$  para  $\mu$  é

$$(\bar{x}_{(n(n+1)/2-c+1)}, \bar{x}_{(c)}) \quad (15.10)$$

Em outras palavras, o intervalo estende-se da  $d$ -ésima menor média dos pares até a  $d$ -ésima maior média, onde  $d = n(n+1)/2 - c + 1$ . A Tabela A.15 do Apêndice fornece os valores de  $c$  que correspondem aos níveis de confiança usuais para  $n = 5, 6, \dots, 25$ .

**Exemplo 15.7** (Continuação do Exemplo 15.6)

Para  $n = 7$ , um intervalo 89,1% (aproximadamente 90%) é obtido utilizando  $c = 24$  (já que a região de rejeição  $\{0, 1, 2, 3, 4, 24, 25, 26, 27, 28\}$  tem  $\alpha = 0,109$ ). O intervalo é  $(\bar{x}_{(28-24+1)}, \bar{x}_{(24)}) = (\bar{x}_{(5)}, \bar{x}_{(24)}) = (4,72; 5,85)$ , se estendendo da quinta menor até a quinta maior média dos pares. ■

## Outros Usos e Propriedades do Intervalo de Postos com Sinais de Wilcoxon

A obtenção de um intervalo depende da existência de uma única amostra de uma distribuição simétrica contínua com média (mediana)  $\mu$ . Quando os dados são emparelhados, o intervalo construído a partir das diferenças  $d_1, d_2, \dots, d_n$  é um IC para a média (mediana) das diferenças, ente  $\mu_D$ . Neste caso, não é necessária a condição de simetria das distribuições de  $X$  e  $Y$ ; desde que as distribuições de  $X$  e  $Y$  apresentem a mesma forma, a distribuição  $X - Y$  será simétrica e, portanto, é necessária apenas a condição de continuidade.

Para  $n > 20$ , a aproximação de amostra grande para o teste de Wilcoxon, baseado na padronização de  $S_+$ , resulta em uma aproximação para  $c$  em (15.10). O resultado [para um intervalo  $100(1 - \alpha)\%$ ] é

$$c \approx \frac{n(n+1)}{4} + z_{\alpha/2} \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

A eficiência do intervalo de Wilcoxon, relativa ao intervalo  $t$ , é comparável à do teste de Wilcoxon, relativamente ao teste  $t$ . Em particular, para amostras grandes, quando a população considerada for normal, o intervalo de Wilcoxon tenderá a ser ligeiramente maior do que o intervalo  $t$ , mas se a população for aproximadamente não-normal (simétrica, mas com caudas pesadas), então o intervalo de Wilcoxon tenderá a ser muito mais curto do que o intervalo  $t$ .

## O Teste de Soma de Postos de Wilcoxon

O teste de soma de postos de Wilcoxon para testar  $H_0: \mu_1 - \mu_2 = \Delta_0$  é realizado combinando, primeiramente, os  $(X_i - \Delta_0)$ s e  $Y_j$ s em uma amostra de tamanho  $m + n$  e, em seguida, classificando-os do menor (classificação 1) para o maior (classificação  $m + n$ ). A estatística de teste  $W$  é, então, a soma dos postos dos  $(X_i - \Delta_0)$ s. Para a alternativa bi-caudal,  $H_0$  é rejeitada se  $w$  for muito pequeno ou muito grande.

Para obter o IC associado, para  $x_i$ s e  $y_j$ s fixados, devemos determinar o conjunto de todos os valores de  $\Delta_0$  para os quais  $H_0$  não é rejeitada. Isto pode ser mais facilmente realizado se primeiro expressarmos a estatística de teste de uma maneira um pouco diferente. O menor valor possível de  $W$  é  $m(m + 1)/2$ , correspondente a todo  $(X_i - \Delta_0)$  menor do que todo  $Y_j$ , e existem  $mn$  diferenças do tipo  $(X_i - \Delta_0) - Y_j$ . Através de manipulações, obtemos

$$\begin{aligned}
 W &= [\text{número de } (X_i - Y_j - \Delta_0)s \geq 0] + \frac{m(m + 1)}{2} \\
 &= [\text{número de } (X_i - Y_j)s \geq \Delta_0] + \frac{m(m + 1)}{2}
 \end{aligned}
 \tag{15.11}$$

Assim, rejeitar  $H_0$  se o número de  $(x_i - y_j)s \geq \Delta_0$  for muito pequeno ou muito grande é equivalente a rejeitar  $H_0$  para  $w$  pequeno ou grande.

A expressão (15.11) sugere que calculemos  $x_i - y_j$  para cada  $i$  e  $j$  e ordenemos estas  $mn$  diferenças da menor para a maior. Então, se o valor nulo  $\Delta_0$  não for nem menor do que a maioria das diferenças, nem maior do que a maioria delas,  $H_0: \mu_1 - \mu_2 = \Delta_0$  não é rejeitada. Variando  $\Delta_0$ , é possível mostrar que um IC para  $\mu_1 - \mu_2$  terá como sua menor extremidade um dos  $(x_i - y_j)s$  ordenados, da mesma forma que para a extremidade superior.

**PROPOSIÇÃO**

Sejam  $x_1, \dots, x_m$  e  $y_1, \dots, y_n$ , os valores observados em duas amostras independentes de distribuições contínuas, que diferem somente na localização (e não na forma). Com  $d_{ij} = x_i - y_j$  e as diferenças ordenadas indicadas por  $d_{ij(1)}, d_{ij(2)}, \dots, d_{ij(mn)}$ , a forma geral de um IC  $100(1 - \alpha)\%$  para  $\mu_1 - \mu_2$  é

$$(d_{ij(mn-c+1)}, d_{ij(c)})
 \tag{15.12}$$

onde  $c$  é a constante crítica para um teste bi-caudal de soma de postos de Wilcoxon, ao nível  $\alpha$ .

Observe que o tipo do intervalo da soma de postos de Wilcoxon (15.12) é muito semelhante ao intervalo de postos com sinais de Wilcoxon (15.10); (15.10) utiliza as médias dos pares de uma amostra simples, enquanto (15.12) utiliza a diferença dos pares de duas amostras. A Tabela A.16 do Apêndice fornece os valores de  $c$  para valores selecionados de  $m$  e  $n$ .

**Exemplo 15.8**

O artigo “Some Mechanical Properties of Impregnated Bark Board” (*Forest Products J.*, 1977, p. 31-38) relata os seguintes dados para máxima resistência à compressão (psi), para uma amostra de quadros de madeira impregnados com epóxi e para uma amostra de quadros de madeira impregnados com outro polímero:

<b>Epóxi (xs)</b>	10.860	11.120	11.340	12.130	14.380	13.070
<b>Outro (ys)</b>	4.590	4.850	6.510	5.640	6.390	

Obtenha um IC 95% para a média real das diferenças entre as resistências à compressão para os quadros impregnados de epóxi e para os outros tipos de quadros.

A partir da Tabela A.16 do Apêndice, já que o tamanho da menor amostra é 5 e o tamanho da maior amostra é 6,  $c = 26$  para um nível de confiança de aproximadamente 95%. Os  $d_{ij}$ s aparecem na Tabela 15.5. Os cinco menores  $d_{ij}$ s [ $d_{ij(1)}, \dots, d_{ij(5)}$ ] valem 4.350, 4.470, 4.610, 4.730 e 4.830; e os cinco maiores  $d_{ij}$ s valem (em ordem decrescente) 9.790, 9.530, 8.740, 8.480 e 8.220. Assim, o IC é  $(d_{ij(5)}, d_{ij(26)}) = (4.830, 8.220)$ .

**Tabela 15.5** Diferenças para o intervalo de soma de postos no Exemplo 15.8

$d_{ij}$	4590	4850	$y_j$ 5640	6390	6510
$x_i$					
10.860	6.270	6.010	5.220	4.470	4.350
11.120	6.530	6.270	5.480	4.730	4.610
11.340	6.750	6.490	5.700	4.950	4.830
12.130	7.540	7.280	6.490	5.740	5.620
13.070	8.480	8.220	7.430	6.680	6.560
14.380	9.790	9.530	8.740	7.990	7.870

Quando  $m$  e  $n$  são ambos grandes, a estatística de teste de Wilcoxon tem uma distribuição aproximadamente normal. Isto pode ser utilizado para obter uma aproximação de amostra grande para o valor  $c$  no intervalo (15.12). O resultado é

$$c \approx \frac{mn}{2} + z_{\alpha/2} \sqrt{\frac{mn(m+n+1)}{12}} \quad (15.13)$$

Assim como para o intervalo de postos com sinais, o intervalo de soma de postos (15.12) é quase tão eficiente quanto o intervalo  $t$ ; em amostras grandes, (15.12) tenderá a ser apenas um pouco maior do que o intervalo  $t$  quando as populações em estudo forem normais e pode ser consideravelmente menor do que o intervalo  $t$ , se as populações analisadas apresentarem caudas mais pesadas do que as populações normais.

## Exercícios | Seção 15.3 (17–22)

- O artigo “The Lead Content and Acidity of Christchurch Precipitation” (*New Zealand J. Science*, 1980, p. 311-312) relata os dados a seguir sobre a concentração de chumbo ( $\mu\text{g/L}$ ) em amostras coletadas durante oito chuvas de verão: 17,0; 21,4; 30,6; 5,0; 12,2; 11,8; 17,3 e 18,8. Assumindo que a distribuição do conteúdo de chumbo seja simétrica, utilize o intervalo de postos com sinais de Wilcoxon para obter um IC 95% para  $\mu$ .
- Calcule o intervalo de postos com sinais de 99% para o valor médio real de pH  $\mu$  (sob a hipótese de simetria), utilizando os dados do Exercício 3. [Sugestão: Tente calcular apenas aquelas médias dos pares que apresentam valores relativamente pequenos ou grandes (em vez de todas as 105 médias).]
- Calcule um IC para  $\mu_D$  do Exemplo 15.2, utilizando os dados ali fornecidos; seu nível de confiança deve ser próximo de 95%.
- As seguintes observações correspondem a quantidades de emissões de hidrocarbonetos, resultantes do uso de pneus descalibrado sem rodovias, submetidos a uma carga de 522 kg, sob pressão de 228 kPa e na velocidade de 64 km/h, durante 6 horas (“Characterization of Tire Emissions Using an Indoor Test Facility”, *Rubber Chemistry and Technology*, 1978 p. 7-25): 0,045; 0,117; 0,062 e 0,072. Que intervalos de confiança podem ser obtidos para este tamanho de amostra, utilizando o intervalo de postos com sinais? Selecione um nível de confiança apropriado e calcule o intervalo.
- Calcule o IC de soma de postos de 90% para  $\mu_1 - \mu_2$ , utilizando os dados no Exercício 10.
- Calcule um IC 99% para  $\mu_1 - \mu_2$ , utilizando os dados do Exercício 11.

## 15.4 | ANOVA Livre de Distribuição

O modelo ANOVA de fator único do Capítulo 10 para comparar  $I$  populações ou tratamentos, assumido que para  $i = 1, 2, \dots, I$ , uma amostra aleatória de tamanho  $J_i$  foi selecionada a partir de uma população normal com média  $\mu_i$  e variância  $\sigma^2$ . O enunciado pode ser escrito como

$$X_{ij} = \mu_i + \epsilon_{ij} \quad j = 1, \dots, J_i; i = 1, \dots, I \tag{15.14}$$

onde os  $\epsilon_{ij}$ s são independentes e normalmente distribuídos, com média zero e variância  $\sigma^2$ . Embora a hipótese de normalidade fosse necessária para a validade do teste  $F$ , descrito no Capítulo 10, a validade do teste de Kruskal-Wallis para testar a igualdade dos  $\mu_i$ s depende apenas dos  $\epsilon_{ij}$ s que apresentam a mesma distribuição contínua.

### O Teste de Kruskal-Wallis

Seja  $N = \sum J_i$ , o número total de observações em um conjunto de dados e suponha que classificamos todas as  $N$  observações de 1 (o menor  $X_{ij}$ ) até  $N$  (o maior  $X_{ij}$ ). Quando  $H_0: \mu_1 = \mu_2 = \dots = \mu_I$  é verdadeira, todas as  $N$  observações originam-se da mesma distribuição, de forma que todos os postos 1, 2, ...,  $N$  para as  $I$  amostras são equiprováveis e esperamos que os postos sejam mesclados nestas amostras. Se, entretanto,  $H_0$  é falsa, então algumas amostras consistirão, na sua maioria, de observações com postos pequenos na amostra combinada, enquanto outras consistirão, na sua maioria, de observações com postos elevados. Mais especificamente, se  $R_{ij}$  indica a classificação de  $X_{ij}$ , entre as  $N$  observações, e  $R_i$  e  $\bar{R}_i$  indicam, respectivamente, o total e a média dos postos na  $i$ -ésima amostra, então quando  $H_0$  é verdadeira

$$E(R_{ij}) = \frac{N + 1}{2} \quad E(\bar{R}_i) = \frac{1}{J_i} \sum_j E(R_{ij}) = \frac{N + 1}{2}$$

A estatística de teste de Kruskal-Wallis é uma medida da extensão na qual  $\bar{R}_i$ s desviam de seus valores esperados comuns  $(N + 1)/2$  e  $H_0$  é rejeitada se o valor calculado da estatística indicar uma discrepância muito grande entre as médias dos postos observados e dos esperados.

ESTATÍSTICA DE TESTE

$$\begin{aligned} K &= \frac{12}{N(N + 1)} \sum_{j=1}^I J_i \left( \bar{R}_i - \frac{N + 1}{2} \right)^2 \\ &= \frac{12}{N(N + 1)} \sum_{i=1}^I \frac{R_i^2}{J_i} - 3(N + 1) \end{aligned} \tag{15.15}$$

A segunda expressão para  $K$  é uma fórmula computacional. Ela considera os totais classificados ( $R_i$ s), no lugar das médias e requer apenas uma subtração.

Se  $H_0$  é rejeitada quando  $k \geq c$ , então  $c$  deve ser escolhido tal que o teste apresente nível  $\alpha$ . Ou seja,  $c$  deve ser o valor crítico da cauda superior da distribuição de  $K$  quando  $H_0$  é verdadeira. Sob  $H_0$ , cada atribuição possível de postos para as  $I$  amostras é equiprovável, portanto, teoricamente, todas as atribuições podem ser enumeradas, o valor de  $K$  determinado para cada uma e a distribuição nula obtida, contando-se o número de vezes que cada valor de  $K$  ocorre. Claramente, este cálculo é tedioso, portanto, embora existam tabelas de distribuição nula exata e valores críticos para pequenos valores dos  $J_i$ s, utilizaremos a seguinte aproximação para “amostra grande”.

PROPOSIÇÃO

Quando  $H_0$  é verdadeira e se

$$I = 3 \quad J_i \geq 6 \quad (i = 1, 2, 3)$$

ou

$$I > 3 \quad J_i \geq 5 \quad (i = 1, \dots, I)$$

então  $K$  tem uma distribuição aproximadamente qui-quadrada com  $I - 1$  gl. Isto implica que em um teste com nível de significância de aproximadamente  $\alpha$  rejeita  $H_0$  se  $k \geq \chi_{\alpha, I-1}^2$ .

**Exemplo 15.9**

As observações correspondentes (Tabela 15.6) em um índice de firmeza axial resultado de um estudo de estrutura de chapas de metal conectadas em cinco comprimentos de chapas diferentes — 4 pol., 6 pol., 8 pol., 10 pol. e 12 pol. — foram usadas (“Modeling Joints Made with Light-Gauge Metal Connector Plates”, *Forest Products J.*, 1979, p. 39-44).

**Tabela 15.6** Dados e postos para o Exemplo 15.9

	$i = 1$ (4"):	309,2	309,7	311,0	316,8	326,5	349,8	409,5		
	$i = 2$ (6"):	331,0	347,2	348,9	361,0	381,7	402,1	404,5		
	$i = 3$ (8"):	351,0	357,1	366,2	367,3	382,0	392,4	409,9		
	$i = 4$ (10"):	346,7	362,6	384,2	410,6	433,1	452,9	461,4		
	$i = 5$ (12"):	407,4	410,7	419,9	441,2	441,8	465,8	473,4		
							$r_i$	$\bar{r}_i$		
<b>Postos</b>	$i = 1$ :	1	2	3	4	5	10	24	49	7,00
	$i = 2$ :	6	8	9	13	17	21	22	96	13,71
	$i = 3$ :	11	12	15	16	18	20	25	117	16,71
	$i = 4$ :	7	14	19	26	29	32	33	160	22,86
	$i = 5$ :	23	27	28	30	31	34	35	208	29,71

O valor calculado de  $K$  é

$$k = \frac{12}{35(36)} \left[ \frac{(49)^2}{7} + \frac{(96)^2}{7} + \frac{(117)^2}{7} + \frac{(160)^2}{7} + \frac{(208)^2}{7} \right] - 3(36)$$

$$= 20,12$$

Ao nível 0,01,  $\chi_{0,01,4}^2 = 13,277$  e como  $20,12 \geq 13,277$ ,  $H_0$  é rejeitada e concluímos que a rigidez axial esperada depende do comprimento do suporte. ■

### Teste de Friedman para um Experimento em Bloco Aleatório

Suponha que  $X_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}$ , onde  $\alpha_i$  é o efeito do  $i$ -ésimo tratamento,  $\beta_j$  é o  $j$ -ésimo efeito de bloco e os  $\epsilon_{ij}$  são desenvolvidos independentemente a partir da mesma distribuição contínua (mas não necessariamente normal). Então para testar  $H_0: \alpha_1 = \alpha_2 = \dots = \alpha_I = 0$ , a hipótese nula dos efeitos do não tratamento, as observações são primeiro classificadas separadamente de 1 até  $I$  dentro de cada bloco, e então a média de classificação  $\bar{r}_i$  é calculada para cada um dos  $I$  tratamentos. Quando  $H_0$  é verdadeira, o  $\bar{r}_i$ s devem ser fechados perto de outro, como dentro de cada bloco todas as atribuições  $I!$  de postos para tratamentos são igualmente prováveis. O teste estatístico de Friedman mede a discrepância entre o valor esperado  $(I + 1)/2$  de cada média de classificação e os  $\bar{r}_i$ s

TESTE ESTATÍSTICO

$$F_r = \frac{12J}{I(I+1)} \sum_{i=1}^I \left( \bar{R}_i - \frac{I+1}{2} \right)^2 = \frac{12}{IJ(I+1)} \sum R_i^2 - 3J(I+1)$$

Como no teste de Kruskal-Wallis, o teste de Friedman rejeita  $H_0$  quando o valor calculado do teste estatístico é muito grande. Para os casos  $I = 3, J = 2, \dots, 15$  e  $I = 4, J = 2, \dots, 8$ , o livro Lehmann (veja a Bibliografia do capítulo) fornece os valores críticos da cauda superior para o teste. Por outro lado, mesmo para os valores moderados de  $J$ , o teste estatístico  $F_r$  tem aproximadamente uma distribuição qui-quadrado com  $I - 1$  gl quando  $H_0$  é verdadeira, podendo ser rejeitada se  $f_r \geq \chi_{\alpha, I-1}^2$ .

**Exemplo 15.10**

O artigo “Physiological Effects During Hypnotically Requested Emotions” (*Psychosomatic Med.*, 1963, p. 334-343) informa os seguintes dados (Tabela 15.7) sobre o potencial da pele (mV) quando as emoções de medo, felicidade, depressão e calma são solicitadas a partir de cada um dos oito indivíduos.

**Tabela 15.7** Dados e postos para o Exemplo 15.10

$x_{ij}$	Bloco (Indivíduos)									
	1	2	3	4	5	6	7	8		
<b>Medo</b>	23,1	57,6	10,5	23,6	11,9	54,6	21,0	20,3		
<b>Felicidade</b>	22,7	53,2	9,7	19,6	13,8	47,1	13,6	23,6		
<b>Depressão</b>	22,5	53,7	10,8	21,1	13,7	39,2	13,7	16,3		
<b>Calma</b>	22,6	53,1	8,3	21,6	13,3	37,0	14,8	14,8		
<b>Postos</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	$r_i$	$r_i^2$
<b>Medo</b>	4	4	3	4	1	4	4	3	27	729
<b>Felicidade</b>	3	2	2	1	4	3	1	4	20	400
<b>Depressão</b>	1	3	4	2	3	2	2	2	19	361
<b>Calma</b>	2	1	1	3	2	1	3	1	14	196
										1686

Desta forma

$$f_r = \frac{12}{4(8)(5)} (1.686) - 3(8)(5) = 6,45$$

Ao nível 0,05,  $\chi^2_{0,05,3} = 7,815$  e por que 6,45 não é  $\geq 7,815$ ,  $H_0$  não é rejeitada. Não há evidência de que o potencial médio da pele depende de qual emoção é solicitada. ■

O livro de Myles Hollander e Douglas Wolfe (veja Bibliografia do capítulo) discute vários procedimentos de comparação associados com os testes de Kruskal-Wallis e Friedman, assim como outros aspectos da livre distribuição ANOVA.

**Exercícios** | Seção 15.4 (23–27)

**23.** Os dados correspondentes referem-se à concentração de isótopos radioativos de estrôncio-90 nas amostras de leite obtidas a partir de cinco indústrias de laticínios selecionadas aleatoriamente em cada uma das quatro diferentes regiões.

Região	1	2	3	4	5
1	6,4	5,8	6,5	7,7	6,1
2	7,1	9,9	11,2	10,5	8,8
3	5,7	5,9	8,2	6,6	5,1
4	9,5	12,1	10,3	12,4	11,7

Testar ao nível 0,10 para verificar se a média verdadeira das concentrações de estrôncio-90 difere em pelo menos duas regiões.

**24.** O artigo “Production of Gaseous Nitrogen in Human Steady-State Conditions” (*J. Applied Physiology*, 1972,

p. 155-159) informa as seguintes observações sobre a quantidade de nitrogênio expirado (em litros) sob quatro regimes de alimentação: (1) jejum, (2) 23% de proteína, (3) 32% de proteína e (4) 67% de proteína. Utilize o teste de Kruskal-Wallis ao nível 0,05 para testar a igualdade dos correspondentes  $\mu_i$ s.

1.	4,079	4,859	3,540	5,047	3,298
2.	4,368	5,668	3,752	5,848	3,802
3.	4,169	5,709	4,416	5,666	4,123
4.	4,928	5,608	4,940	5,291	4,674
1.	4,679	2,870	4,648	3,847	
2.	4,844	3,578	5,393	4,374	
3.	5,059	4,403	4,496	4,688	
4.	5,038	4,905	5,208	4,806	



25. Os dados correspondentes sobre o nível de cortisol foram informados no artigo "Cortisol, Cortisone, and 11-Deoxycortisol Levels in Human Umbilical and Maternal Plasma in Relation to the Onset Labor" (*J. Obstetric Gynaecology British Commonwealth*, 1974, p. 737-745, ). As pessoas do experimento foram mulheres grávidas cujos bebês nasceram entre 38 e 42 semanas de gestação. Os indivíduos do grupo 1 escolheram ter seus bebês por cesariana antes do início do trabalho de parto, as do grupo 2 tiveram seus bebês por cesariana de emergência durante o parto induzido e o grupo 3 teve seus bebês por parto normal. Utilize o teste de Kruskal-Wallis ao nível 0,05 para testar a igualdade das médias das três populações.

<b>Grupo 1</b>	262	307	211	323	454	339
	304	154	287	356		
<b>Grupo 2</b>	465	501	455	355	468	362
<b>Grupo 3</b>	343	772	207	1.048	838	687

26. Em um teste para determinar se o solo pré-tratado com pequenas quantidades de Básico-H torna o solo mais permeável à água, amostras de solo foram divididas em blocos, e cada bloco recebeu um dos quatro tratamentos sob estudo. Os tratamentos foram (A) água com 0,001% de Básico-H derramada sobre o solo de controle, (B) água sem Básico-H sobre o solo de controle, (C) água com Básico-H derramada sobre o solo pré-tratado com Básico-H e (D) água sem Básico-H sobre o solo pré-tratado com Básico-H. Testar ao nível 0,01 para verificar se há qualquer efeito devido aos diferentes tratamentos.

	Blocos				
	1	2	3	4	5
<b>A</b>	37,1	31,8	28,0	25,9	25,5
<b>B</b>	33,2	25,3	20,2	20,3	18,3
<b>C</b>	58,9	54,2	49,2	47,9	38,2
<b>D</b>	56,7	49,6	46,4	40,9	39,4
	6	7	8	9	10
<b>A</b>	25,3	23,7	24,4	21,7	26,2
<b>B</b>	19,3	17,3	17,0	16,7	18,3
<b>C</b>	48,8	47,8	40,2	44,0	46,4
<b>D</b>	37,1	37,5	39,6	35,1	36,5

27. Em um experimento para estudar a maneira pela qual diferentes anestésicos afetam a concentração de epinefrina no sangue, dez cães foram selecionados e a concentração foi medida enquanto eles estavam sob a influência dos anestésicos isoflurano, halotano e cicloprano ("Sympathoadrenal and Hemodynamic Effects of Isoflurane, Halothane, and Cyclopropane in Dogs", *Anesthesiology*, 1974, p. 465-470). Testar ao nível 0,05 para verificar se há um efeito anestésico na concentração.

	Cães				
	1	2	3	4	5
<b>Isoflurano</b>	0,28	0,51	1,00	0,39	0,29
<b>Halotano</b>	0,30	0,39	0,63	0,38	0,21
<b>Ciclopropano</b>	1,07	1,35	0,69	0,28	1,24
	6	7	8	9	10
<b>Isoflurano</b>	0,36	0,32	0,69	0,17	0,33
<b>Halotano</b>	0,88	0,39	0,51	0,32	0,42
<b>Ciclopropano</b>	1,53	0,49	0,56	1,02	0,30

### Exercícios Suplementares (28–36)

28. O artigo "Effects of a Rice-Rich Versus Potato-Rich Diet on Glucose, Lipoprotein, and Cholesterol Metabolism in Noninsulin-Dependent Diabetics" (*Amer.: J. Clinical Nutr.*, 1984, p. 598-600) fornece os dados correspondentes à taxa da síntese de colesterol para oito indivíduos diabéticos. Os indivíduos foram alimentados com uma dieta padronizada de batata ou arroz como as maiores fontes de carboidratos. Os participantes receberam ambas as dietas por períodos específicos de tempo, com a taxa de síntese de colesterol (mmol/dia) medida no final de cada período de dieta. A análise apresentada neste artigo utilizou um teste de livre distribuição. Utilize um teste com nível de significância 0,05 para determinar se a média verdadeira da taxa de síntese de colesterol difere significativamente para as duas fontes de carboidratos.

Indivíduo	Taxa de Síntese de Colesterol							
	1	2	3	4	5	6	7	8
<b>Batata</b>	1,88	2,60	1,38	4,41	1,87	2,89	3,96	2,31
<b>Arroz</b>	1,70	3,84	1,13	4,97	0,86	1,93	3,36	2,15

29. Táticas de venda de alta pressão ou vendedores de porta em porta podem ser bastante desagradáveis. Muitas pessoas sucumbem a tais táticas, assinam um contrato de compra e depois se arrependem de suas ações. Em meados da década de 1970, a Federal Trade Commission implementou regulamentos esclarecendo e estendendo os direitos dos compradores de cancelar tais contratos. Os dados correspondentes são um subconjunto daqueles fornecidos no artigo "Evaluating the FTC Cooling-Off Rule" (*J. Consumer Affairs*, 1977, p. 101-106). Observações individuais são as taxas de cancelamento para cada um dos nove vendedores durante 4 anos. Utilize um teste apropriado ao nível 0,05 para verificar se a média verdadeira da taxa de cancelamento depende do ano.

	Vendedor								
	1	2	3	4	5	6	7	8	9
<b>1973</b>	2,8	5,9	3,3	4,4	1,7	3,8	6,6	3,1	0,0
<b>1974</b>	3,6	1,7	5,1	2,2	2,1	4,1	4,7	2,7	1,3
<b>1975</b>	1,4	0,9	1,1	3,2	0,8	1,5	2,8	1,4	0,5
<b>1976</b>	2,0	2,2	0,9	1,1	0,5	1,2	1,4	3,5	1,2

30. Os dados correspondentes sobre a concentração de fósforo na superfície do solo para quatro diferentes tratamentos de solo apareceram no artigo “Fertilisers for Lotus and Clover Establishment on a Sequence of Acid Soils on the East Otago Uplands” (*N. Zeal. J. Exper. Ag.*, 1984, p. 119-129). Utilize um procedimento livre de distribuição para testar a hipótese nula da não-diferença na média real da concentração de fósforo (mg/g) para os quatro tratamentos de solo.

	I	8,1	5,9	7,0	8,0	9,0
Tratamento	II	11,5	10,9	12,1	10,3	11,9
	III	15,3	17,4	16,4	15,8	16,0
	IV	23,0	33,0	28,4	24,6	27,7

31. Consulte os dados do Exercício 30 e calcule um IC de 95% para a diferença entre a média verdadeira das concentrações aos tratamentos II e III.

32. O estudo informado em “Gait Patterns During Choice Ladder Ascents” (*Human Movement Sci.*, 1983, p. 187-195) foi motivado para a conscientização pública do aumento da taxa de acidentes para indivíduos que sobem escadas. Um número de padrões de passos diferentes foi utilizado pelos indivíduos subindo em uma escada reta portátil de acordo com as instruções especificadas. Os tempos de subida para sete indivíduos que utilizaram um modo de andar lateral e seis indivíduos que utilizaram um modo de andar diagonal são fornecidos.

Lateral	0,86	1,31	1,64	1,51	1,53	1,39	1,09
Diagonal	1,27	1,82	1,66	0,85	1,45	1,24	

- a. Realize um teste utilizando  $\alpha = 0,05$  para verificar se os dados sugerem qualquer diferença na média verdadeira dos tempos de subida para os dois modos de andar.
- b. Calcule um IC de 95% para a diferença entre a média verdadeira dos tempos dos modos de andar.

33. O teste de sinal é um procedimento simples para testar a hipótese sobre uma população mediana assumindo apenas que a distribuição subjacente é contínua. Para ilustrar, considere a seguinte amostra de 20 observações do componente tempo de vida (h):

1,7	3,3	5,1	6,9	12,6	14,4	16,4
24,6	26,0	26,5	32,1	37,4	40,1	40,5
41,5	72,4	80,1	86,4	87,5	100,2	

Desejamos testar  $H_0: \tilde{\mu} = 25,0$  versus  $H_a: \tilde{\mu} > 25,0$ . O teste estatístico é  $Y =$  o número de observações que excedem 25.

- a. Considere a rejeição de  $H_0$  se  $Y \geq 15$ . Qual é o valor de  $\alpha$  (a probabilidade de um erro do tipo I) para este teste? *Sugestão:* Entenda como “sucesso” o tempo de vida que excede 25,0. Então,  $Y$  é o número de sucessos na amostra. Qual o tipo de uma distribuição  $Y$  quando  $\tilde{\mu} = 25,0$ ?

- b. Qual região de rejeição do tipo  $Y \geq c$  especifica um teste com nível de significância tão próximo de 0,05 quanto possível? Use essa região para realizar o teste para os dados fornecidos.

Nota: O teste estatístico é o número de diferenças  $X_i - 25,0$  que têm sinais positivos, por isso o nome teste de sinal.

34. Consulte o Exercício 33 e considere um intervalo de confiança associado ao teste de sinal, o **intervalo de sinal**. As hipóteses relevantes são agora  $H_0: \tilde{\mu} = \tilde{\mu}_0$  versus  $H_a: \tilde{\mu} \neq \tilde{\mu}_0$ . Vamos usar a seguinte região de rejeição:  $Y \geq 15$  ou  $Y \leq 5$ .

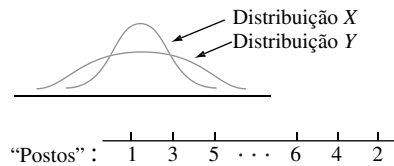
- a. Qual é o nível de significância para este teste?
- b. O intervalo de confiança consistirá de valores  $\tilde{\mu}_0$  para os quais  $H_0$  não é rejeitado. Determine o IC para os dados fornecidos e defina o nível de confiança.

35. Suponha que desejamos testar

$H_0$ : as distribuições  $X$  e  $Y$  são idênticas versus

$H_a$ : a distribuição  $X$  é menos dispersa do que a distribuição  $Y$

A figura correspondente ilustra as distribuições  $X$  e  $Y$  para as quais  $H_a$  é verdadeira. O teste de soma de postos de Wilcoxon não é apropriado nesta situação porque  $H_a$  é verdadeira como ilustrado, os  $Y$ s tenderão a ser as extremidades da amostra combinada (resultando nos menores e maiores postos de  $Y$ ), portanto a soma dos postos  $X$  resultará em um valor  $W$  que não é grande nem pequeno.



Considere modificar o procedimento para a atribuição de postos como segue. Após a amostra combinada de  $m + n$  observações ser ordenada, à menor observação é dado o posto 1, à maior observação é dado o posto 2, à segunda menor é dado o posto 3, à segunda maior é dado o posto 4 e assim por diante. Então, se  $H_a$  é verdadeira como mostrado, os valores de  $X$  tenderão a ser médias da amostra e assim receber postos altos. Seja  $W'$  a soma dos  $X$  postos e considere a rejeição de  $H_0$  em favor de  $H_a$  quando  $w' \geq c$ . Quando  $H_0$  é verdadeira, todo o conjunto possível de postos de  $X$  tem a mesma probabilidade, portanto  $W'$  tem a mesma distribuição de  $W$  quando  $H_0$  é verdadeira. Assim  $c$  pode ser escolhido da Tabela A.14 do Apêndice para produzir um teste de nível  $\alpha$ . Os dados correspondentes se referem à espessura do músculo mediano para arteríolas a partir de pulmões de crianças que morreram da síndrome de morte infantil repentina ( $x$ s) e um grupo de crianças de controle ( $y$ s). Realize o teste de  $H_0$  versus  $H_a$  ao nível 0,05.

<b>SMIR</b>	4,0	4,4	4,8	4,9	
<b>Controle</b>	3,7	4,1	4,3	5,1	5,6

Consulte o livro de Lehmann (na Bibliografia do capítulo) para obter mais informações sobre este teste, chamado de *teste Siegel-Tukey*.

36. O procedimento de classificação descrito no Exercício 35 é algo assimétrico, porque a menor observação recebe posto 1 enquanto a maior recebe posto 2 e assim por diante. Suponha que ambas, a menor e a maior, recebam o posto 1, a segunda menor e a segunda maior recebam posto 2 e assim sucessivamente, e seja  $W''$  a

soma dos  $X$  postos. A distribuição nula de  $W''$  não é idêntica à distribuição nula de  $W$ , portanto tabelas diferentes são necessárias. Considere o caso  $m = 3$  e  $n = 4$ . Liste todas as 35 possibilidades de ordenação dos três valores de  $X$  entre as sete observações (por exemplo, 1, 3, 7 ou 4, 5, 6), atribua postos da maneira descrita, calcule o valor de  $W''$  para cada possibilidade e então tabele a distribuição nula de  $W''$ . Para o teste que rejeita se  $w'' \geq c$ , qual valor de  $c$  está associado a um nível de teste de aproximadamente 0,10? Este é o *teste de Ansari-Bradley*; para obter informações adicionais, veja o livro de Hollander e Wolfe na Bibliografia do capítulo.

---

## Bibliografia

HOLLANDER, Myles e DOUGLAS Wolfe. *Nonparametric Statistical Methods*, Wiley, Nova York, 1999 (2. ed.). Uma referência muito boa sobre os métodos de livre distribuição com uma excelente coleção de tabelas.

LEHMANN, Erich. *Nonparametrics: Statistical Methods Based on Ranks*, Holden-Day, São Francisco, 1975. Uma discussão excelente dos mais importantes métodos de livre distribuição, apresentados com uma grande quantidade de comentários pertinentes.

# Métodos de Controle de Qualidade

## Introdução

As características de qualidade dos produtos manufaturados receberam muita atenção dos engenheiros de projeto e pessoal de produção, assim como daqueles envolvidos com gerenciamento financeiro. Uma crença desenvolvida ao longo dos anos foi a de que níveis muito altos de qualidade e bem-estar econômico eram objetivos incompatíveis. Recentemente, entretanto, tem-se tornado aparente que níveis de qualidade crescente podem levar a um decréscimo dos custos, a um maior grau de satisfação do consumidor e, assim, a um aumento de lucratividade. Isto resultou na ênfase renovada nas técnicas de estatística para o desenvolvimento de qualidades nos produtos e para a identificação de problemas de qualidade em vários estágios da produção e distribuição.

Gráficos de controle agora são utilizados extensivamente na indústria como uma técnica de diagnóstico para o monitoramento dos processos de produção a fim de identificar instabilidade e circunstâncias não-usuais. Depois de apresentar as idéias básicas na Seção 16.1, vários gráficos de controle diferentes são expostos nas próximas quatro seções. A base para a maioria deles recai em nosso trabalho anterior relacionado às distribuições de probabilidade de várias estatísticas, tais como a média  $\bar{X}$  de amostra e a proporção da amostra  $\hat{p} = X/n$ .

Outra situação comumente encontrada no ambiente industrial envolve o cliente decidir se um lote de itens oferecido pelo fornecedor é de qualidade aceitável. Na última seção do capítulo, faremos um resumo breve de alguns métodos de amostragem aceitáveis para decisão, baseados nos dados de amostra, sobre a aceitação ou rejeição de um lote.

Além de gráficos de controle e planos de amostragem aceitáveis, os quais foram primeiramente desenvolvidos nas décadas de 1920 e 1930, os estatísticos e engenheiros têm apresentado recentemente vários métodos estatísticos novos para a identificação de tipos e níveis de insumos de produção que garantirão produtos de alta qualidade. Pesquisadores japoneses, e em particular o engenheiro/estatístico G. Taguchi e seus discípulos, tiveram muita influência neste aspecto, e há agora uma grande quantidade de material conhecido como "métodos Taguchi". As idéias de planejamento de experimento e em

particular experimentos fatoriais fracionários, são os ingredientes-chave. Há ainda muita controvérsia na comunidade estatística sobre quais as formas de planejamento de experimento e métodos de análise mais adequadas às tarefas manuais. Uma crítica recente é encontrada no artigo exposto por George Box et. al., citado na Bibliografia do capítulo; no livro de Thomas Ryan mencionado há também uma boa fonte de informação.

## 16.1 Comentários Gerais sobre Gráficos de Controle

Uma mensagem central em todo este livro tem sido a que ocorre naturalmente em quaisquer características ou atributos de diferentes indivíduos ou assuntos. Em um contexto de produção, não importa que as máquinas sejam cuidadosamente calibradas, fatores ambientais sejam controlados, materiais e outros insumos sejam monitorados e trabalhadores treinados, o diâmetro variará de porca para porca, algumas folhas de plástico serão mais fortes do que outras e alguns fusos serão danificados enquanto outros não, e assim por diante. Podemos pensar sobre tal variação aleatória natural como um ruído não-controlável que faz parte natural do processo.

Há, entretanto, outras fontes de variação que podem ter um impacto pernicioso sobre a qualidade dos itens produzidos por alguns processos. Tais variações podem ser atribuídas a material contaminado, ajuste de máquinas incorreto, uso de ferramentas estranhas e coisas deste tipo. Estas fontes de variação têm sido chamadas de *causas assinaláveis* na literatura sobre controle de qualidade. Os **gráficos de controle** fornecem um mecanismo para reconhecer situações em que as causas assinaláveis podem afetar adversamente a qualidade do produto. Uma vez que o gráfico indica uma situação fora de controle, uma investigação pode ser iniciada para identificar as causas e providenciar ações corretivas.

Um elemento básico dos gráficos de controle é que as amostras do processo de interesse foram selecionadas em uma seqüência de pontos no tempo. Dependendo do aspecto do processo sob investigação, alguma estatística, tal como a média amostral ou a proporção amostral dos itens com defeito, é escolhida. O valor desta estatística então é calculado para cada amostra em resposta. Um gráfico de controle tradicional resulta da marcação destes valores calculados ao longo do tempo, como ilustrado na Figura 16.1, a seguir.

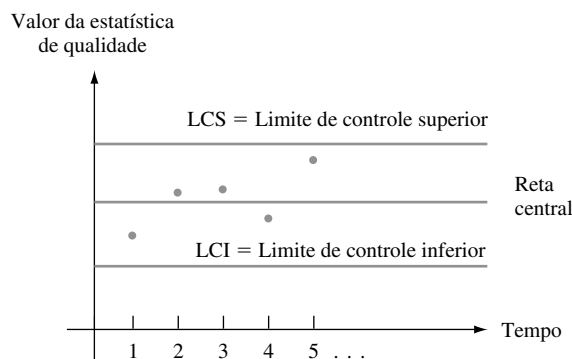


Figura 16.1 Um protótipo de gráfico de controle

Observe que além dos próprios pontos marcados, o gráfico tem uma reta central e dois limites de controle. A base para a escolha de uma reta central é algumas vezes um valor-alvo ou especificação de projeto, por exemplo, um valor desejado do diâmetro do mancal. Em outros casos, a altura da reta central é estimada a partir dos dados. Se todos os pontos sobre o gráfico caem entre os dois limites de controle, o processo é considerado sob controle. Ou seja, o processo é tido como operacional de maneira estável refletindo apenas a variação aleatória natural. Um “sinal” de estar fora de controle acontece se um ponto marcado cai fora dos limites. Isto é

entendido como sendo provocado por alguma causa assinalável e a busca por tais causas é iniciada. Os limites são projetados tal que um processo sob controle gere uma quantidade mínima de alarmes falsos, enquanto um processo fora do controle mostre rapidamente um ponto fora dos limites.

Há uma forte analogia entre a lógica do gráfico de controle e nosso trabalho anterior em testes de hipótese. A hipótese nula aqui é que o processo está sob controle. Quando um processo sob controle produz um ponto fora dos limites de controle (um sinal descontrolado), ocorreu um erro tipo I. Por outro lado, um erro do tipo II acontece quando um processo fora de controle produz um ponto dentro dos limites de controle. A escolha apropriada do tamanho da amostra e os limites de controle (o último corresponde à especificação de uma região de rejeição no teste de hipótese) tornarão as probabilidades de erro adequadamente pequenas.

Enfatizamos que “sob controle” não é sinônimo de “alcança as especificações de projeto ou tolerância”. A extensão da variação natural pode ser tal que a porcentagem de itens em não-conformidade com a especificação seja muito maior do que pode ser tolerado. Em tais casos, uma reestruturação importante do processo será necessária para melhorar a capacitação do processo. Um processo sob controle é simplesmente aquele cujo comportamento em relação à variação é estável ao longo do tempo, não mostrando indicações de causas estranhas não-usuais.

Como um comentário introdutório final, softwares para gráficos de controle são agora facilmente encontrados. O jornal *Quality Progress* contém muitos anúncios de programas de computador para controle de qualidade estatísticos. Além disso, o SAS e a mais recente versão do MINITAB, entre outros programas de propósito geral, têm recursos de controle de qualidade atrativos.

---

## Exercícios | Seção 16.1 (1–3)

- Um gráfico de controle para a espessura de folhas de aço enroladas é baseado em um limite de controle superior de 0,0520 polegada e um limite inferior de 0,0475 polegada. Os primeiros dez valores da estatística de qualidade (neste caso  $\bar{X}$ , a média amostral de espessura de  $n = 5$  folhas de amostra) são 0,0506; 0,0493; 0,0502; 0,0501; 0,0512; 0,0498; 0,0485; 0,0500; 0,0505 e 0,0483. Construa a parte inicial do gráfico de controle de qualidade e comente sobre sua aparência.
- Refira-se ao Exercício 1 e suponha que os dez valores mais recentes da estatística de qualidade são 0,0493; 0,0485; 0,0490; 0,0503; 0,0492; 0,0486; 0,0494; 0,0493 e 0,0488. Construa a porção relevante do gráfico de controle correspondente e comente sobre sua aparência.
- Suponha que um gráfico de controle seja construído tal que a probabilidade de um ponto cair fora do limite quando o processo está realmente sob controle é 0,002. Qual a probabilidade de dez pontos sucessivos (baseado na independência das amostras selecionadas) estarem dentro dos limites de controle? Qual a probabilidade de que todos os 25 pontos sucessivos estejam dentro dos limites de controle? Qual é o menor número de pontos sucessivos para os quais a probabilidade de observar pelo menos um ponto fora dos limites de controle exceda 0,10?

---

## 16.2 | Gráficos de Controle para Localização do Processo

Suponha que as características de qualidade de interesse estejam associadas com uma variável cujos valores observados resultem de medidas realizadas. Por exemplo, a característica pode ser resistência do fio elétrico (ohms), diâmetro interno das juntas de expansão de borracha moldada (cm) ou dureza de uma determinada liga (unidades de Brinell). Um uso importante dos gráficos de controle é verificar se alguma medida de localização da distribuição de variáveis permanece estável ao longo do tempo. O gráfico mais popular para esta finalidade é o gráfico  $\bar{X}$ .

### O Gráfico $\bar{X}$ Baseado em Valores de Parâmetro Conhecidos

Como há incerteza em relação ao valor da variável para qualquer item particular ou espécime, indicamos esse valor como uma variável *aleatória* (va)  $X$ . Faça a hipótese de que para um processo sob controle,  $X$  tem uma

distribuição normal com valor médio  $\mu$  e desvio padrão  $\sigma$ . Então se  $\bar{X}$  indica a média amostral para uma amostra aleatória de tamanho  $n$  selecionada em um instante de tempo particular, sabemos que

1.  $E(\bar{X}) = \mu$
2.  $\sigma_{\bar{X}} = \sigma/\sqrt{n}$
3.  $\bar{X}$  tem uma distribuição normal.

Segue que

$$P(\mu - 3\sigma_{\bar{X}} \leq \bar{X} \leq \mu + 3\sigma_{\bar{X}}) = P(-3,00 \leq Z \leq 3,00) = 0,9974$$

onde  $Z$  é um va normal padronizada.<sup>1</sup> Então, é altamente provável que para um processo sob controle, a média amostral caia dentro de 3 desvios padrão ( $3\sigma_{\bar{X}}$ ) da média do processo  $\mu$ .

Considere primeiro o caso no qual os valores de ambos  $\mu$  e  $\sigma$  são conhecidos. Suponha que em cada um dos pontos de tempo 1, 2, 3, ..., uma amostra aleatória de tamanho  $n$  esteja disponível. Sejam  $\bar{x}_1, \bar{x}_2, \bar{x}_3, \dots$  os valores calculados das médias amostrais correspondentes. Um gráfico  $\bar{X}$  resulta da marcação destes  $\bar{x}_i$ s ao longo do tempo, ou seja, marcar no gráfico os pontos (1,  $\bar{x}_1$ ), (2,  $\bar{x}_2$ ), (3,  $\bar{x}_3$ ) e assim por diante, e então desenhar retas horizontais ao longo de todo o gráfico sobre os limites

$$\text{LCI} = \text{controle inferior} = \mu - 3 \cdot \frac{\sigma}{\sqrt{n}}$$

$$\text{LCS} = \text{limite de controle superior} = \mu + 3 \cdot \frac{\sigma}{\sqrt{n}}$$

Tal gráfico é freqüentemente chamado de carta de controle 3-sigma. Qualquer ponto fora dos limites de controle sugere que o processo pode ter ficado fora de controle naquele momento, portanto, uma busca por causas assinaláveis deve ser iniciada.

**Tabela 16.1** Dados de viscosidade para o Exemplo 16.1

Dia	Observações de Viscosidade			$\bar{x}$	$s$	Variação
1	10,37	10,19	10,36	10,307	0,101	0,18
2	10,48	10,24	10,58	10,433	0,175	0,34
3	10,77	10,22	10,54	10,510	0,276	0,55
4	10,47	10,26	10,31	10,347	0,110	0,21
5	10,84	10,75	10,53	10,707	0,159	0,31
6	10,48	10,53	10,50	10,503	0,025	0,05
7	10,41	10,52	10,46	10,463	0,055	0,11
8	10,40	10,38	10,69	10,490	0,173	0,31
9	10,33	10,35	10,49	10,390	0,087	0,16
10	10,73	10,45	10,30	10,493	0,218	0,43
11	10,41	10,68	10,25	10,447	0,217	0,43
12	10,00	10,60	10,71	10,437	0,382	0,71
13	10,37	10,50	10,34	10,403	0,085	0,16
14	10,47	10,60	10,75	10,607	0,140	0,28
15	10,46	10,46	10,56	10,493	0,058	0,10
16	10,44	10,68	10,32	10,480	0,183	0,36
17	10,65	10,42	10,26	10,443	0,196	0,39
18	10,73	10,72	10,83	10,760	0,061	0,11
19	10,39	10,75	10,27	10,470	0,250	0,48
20	10,59	10,23	10,35	10,390	0,183	0,36
21	10,47	10,67	10,64	10,593	0,108	0,20
22	10,40	10,55	10,38	10,443	0,093	0,17
23	10,24	10,71	10,27	10,407	0,263	0,47
24	10,37	10,69	10,40	10,487	0,177	0,32
25	10,46	10,35	10,37	10,393	0,059	0,11

<sup>1</sup> O uso de gráficos baseados em três limites de desvio padrão é tradicional, mas as tradições não são invioláveis.

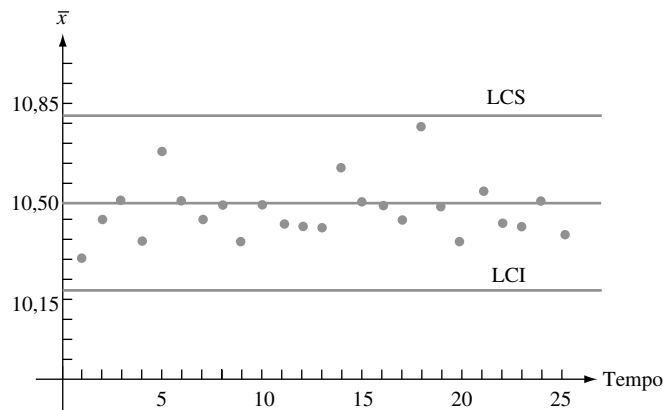
**Exemplo 16.1**

Uma vez por dia, três tipos de óleo de motor são selecionados aleatoriamente a partir do processo de produção, e cada um é analisado para determinar a viscosidade. Os dados correspondentes (Tabela 16.1) são para um período de 25 dias. Experimentos contínuos com este processo sugerem que quando o processo está sob controle, a viscosidade de uma espécie é normalmente distribuída com média 10,5 e desvio padrão de 0,18. Assim  $\sigma_{\bar{x}} = \sigma/\sqrt{n} = 0,18/\sqrt{3} = 0,104$ , portanto os limites de controle do tipo 3 DP são

$$LCI = \mu - 3 \cdot \frac{\sigma}{\sqrt{n}} = 10,5 - 3(0,104) = 10,188$$

$$LCS = \mu + 3 \cdot \frac{\sigma}{\sqrt{n}} = 10,5 + 3(0,104) = 10,812$$

Todos os pontos no gráfico de controle mostrados na Figura 16.2 estão entre os limites de controle, indicando o comportamento estável da média do processo ao longo do período de tempo (o desvio padrão e a amplitude de cada amostra serão utilizados na próxima subseção).



**Figura 16.2** Gráfico  $\bar{X}$  para os dados de viscosidade do Exemplo 16.1

**Gráficos  $\bar{X}$  Baseados em Parâmetros Estimados**

Na prática acontece frequentemente que os valores de  $\mu$  e  $\sigma$  são desconhecidos, portanto eles devem ser estimados a partir dos dados de amostra antes de determinar os limites de controle. Isto é especialmente verdadeiro quando um processo está sujeito à análise do controle de qualidade pela primeira vez. Novamente, seja  $n$  o número de observações em cada amostra e  $k$  o número de amostras disponíveis. Valores típicos de  $n$  são 3, 4, 5 ou 6, e recomenda-se que o de  $k$  seja pelo menos 20. Fazemos a hipótese de que as amostras  $k$  foram coletadas durante um período quando se acreditou que o processo estava sob controle. Mais será dito sobre esta hipótese brevemente.

Com  $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k$  indicando as médias amostrais  $k$  calculadas, a estimativa usual de  $\mu$  é simplesmente a média dessas médias:

$$\hat{\mu} = \bar{\bar{x}} = \frac{\sum_{i=1}^k \bar{x}_i}{k}$$

Apresentamos dois métodos diferentes comumente utilizados para estimar  $\sigma$ , um baseado nos  $k$  desvios padrão da amostra e o outro sobre as  $k$  amplitudes amostrais (lembre-se de que a amplitude amostral é a diferença entre a maior e a menor observações da amostra). Antes da grande disponibilidade de boas calculadoras e softwares estatísticos para computadores, a facilidade em cálculos manuais era uma consideração importante, portanto, as variações de métodos predominavam. Entretanto, no caso de uma população com uma distribuição normal, o estimador não-tendencioso de  $\sigma$  baseado em  $S$  é sabido como tendo variância menor do que aquela baseada na amplitude amostral. Os estatísticos dizem que o estimador anterior é mais *eficiente* do que este último. A perda em eficiência para o estimador é pequena quando  $n$  é muito pequeno mas torna-se importante para  $n > 4$ .



Lembre-se de que o desvio padrão amostral não é um estimador imparcial para  $\sigma$ . Quando  $X_1, \dots, X_n$  é uma amostra aleatória de uma distribuição normal, pode ser mostrado (verificar o Exercício 6.37) que

$$E(S) = a_n \cdot \sigma$$

onde

$$a_n = \frac{\sqrt{2}\Gamma(n/2)}{\sqrt{n-1}\Gamma[(n-1)/2]}$$

e  $\Gamma(\cdot)$  indica a função gama (veja a Seção 4.4). Uma tabulação de  $a_n$  para  $n$  selecionados segue:

$n$	3	4	5	6	7	8
$a_n$	0,886	0,921	0,940	0,952	0,959	0,965

Seja

$$\bar{S} = \frac{\sum_{i=1}^k S_i}{k}$$

onde  $S_1, S_2, \dots, S_k$  são os desvios padrão amostrais para as  $k$  amostras. Então

$$E(\bar{S}) = \frac{1}{k}E\left(\sum_{i=1}^k S_i\right) = \frac{1}{k}\sum_{i=1}^k E(S_i) = \frac{1}{k}\sum_{i=1}^k a_n \cdot \sigma = a_n \cdot \sigma$$

Desta maneira

$$E\left(\frac{\bar{S}}{a_n}\right) = \frac{1}{a_n}E(\bar{S}) = \frac{1}{a_n} \cdot a_n \cdot \sigma = \sigma$$

portanto  $\hat{\sigma} = \bar{S}/a_n$  é um estimador não-tendencioso de  $\sigma$ .

#### Limites de Controle Baseados nos Desvios padrão Amostrais

$$LCI = \bar{\bar{x}} - 3 \cdot \frac{\bar{s}}{a_n \sqrt{n}}$$

$$LCS = \bar{\bar{x}} + 3 \cdot \frac{\bar{s}}{a_n \sqrt{n}}$$

onde

$$\bar{\bar{x}} = \frac{\sum_{i=1}^k \bar{x}_i}{k} \quad \bar{s} = \frac{\sum_{i=1}^k s_i}{k}$$

#### Exemplo 16.2

Em relação aos dados de viscosidade do Exemplo 16.1, temos  $n = 3$  e  $k = 25$ . Os valores de  $\bar{x}_i$  e  $s_i$  ( $i = 1, \dots, 25$ ) aparecem na Tabela 16.1, a partir da qual segue-se que  $\bar{\bar{x}} = 261,896/25 = 10,476$  e  $\bar{s} = 3,834/25 = 0,153$ . Com  $a_3 = 0,886$ , temos

$$LCI = 10,476 - 3 \cdot \frac{0,153}{0,886 \sqrt{3}} = 10,476 - 0,299 = 10,177$$

$$LCS = 10,476 + 3 \cdot \frac{0,153}{0,886 \sqrt{3}} = 10,476 + 0,299 = 10,775$$

Estes limites diferem um pouco dos limites anteriores baseados em  $\mu = 10,5$  e  $\sigma = 0,18$ , pois agora  $\hat{\mu} = 10,476$  e  $\hat{\sigma} = \bar{s}/a_3 = 0,173$ . A observação da Tabela 16.1 mostra que cada  $\bar{x}_i$  está entre esses novos limites, portanto novamente nenhuma situação fora de controle é evidenciada. ■

Para obter uma estimativa de  $\sigma$  baseada na amplitude amostral, observe que se  $X_1, \dots, X_n$  formam uma amostra aleatória de uma distribuição normal, então

$$\begin{aligned} R &= \text{amplitude}(X_1, \dots, X_n) \\ &= \text{máx}(X_1, \dots, X_n) - \text{mín}(X_1, \dots, X_n) \\ &= \text{máx}(X_1 - \mu, \dots, X_n - \mu) - \text{mín}(X_1 - \mu, \dots, X_n - \mu) \\ &= \sigma \left\{ \text{máx} \left( \frac{X_1 - \mu}{\sigma}, \dots, \frac{X_n - \mu}{\sigma} \right) - \text{mín} \left( \frac{X_1 - \mu}{\sigma}, \dots, \frac{X_n - \mu}{\sigma} \right) \right\} \\ &= \sigma \cdot \{ \text{máx}(Z_1, \dots, Z_n) - \text{mín}(Z_1, \dots, Z_n) \} \end{aligned}$$

onde  $Z_1, \dots, Z_n$  são variáveis normais padronizadas independentes. Assim

$$\begin{aligned} E(R) &= \sigma \cdot E(\text{amplitude de uma amostra normal padronizada}) \\ &= \sigma \cdot b_n \end{aligned}$$

tal que  $R/b_n$  é um estimador não-tendencioso de  $\sigma$ .

Agora indique as amplitudes para as  $k$  amostras no conjunto de dados de controle de qualidade por  $r_1, r_2, \dots, r_k$ . O argumento dado implica que a estimativa

$$\hat{\sigma} = \frac{\frac{1}{k} \sum_{i=1}^k r_i}{b_n} = \frac{\bar{r}}{b_n}$$

provém de um estimador não-tendencioso para  $\sigma$ . Valores selecionados de  $b_n$  aparecem na tabela a seguir [seu cálculo é baseado na teoria estatística e na integração numérica para determinar  $E(\text{mín}(Z_1, \dots, Z_n))$  e  $E(\text{máx}(Z_1, \dots, Z_n))$ ].

$n$	3	4	5	6	7	8
$b_n$	1,693	2,058	2,325	2,536	2,706	2,844

#### Limites de Controle Baseados nas Amplitudes da Amostra

$$\text{LCI} = \bar{\bar{x}} - 3 \cdot \frac{\bar{r}}{b_n \sqrt{n}}$$

$$\text{LCS} = \bar{\bar{x}} + 3 \cdot \frac{\bar{r}}{b_n \sqrt{n}}$$

onde  $\bar{r} = \sum_{i=1}^k r_i/k$  e  $r_1, \dots, r_k$  são as  $k$  amplitudes amostrais individuais.

#### Exemplo 16.3 (Continuação do Exemplo 16.2)

É facilmente verificado pela Tabela 16.1 que  $\bar{r} = 0,292$ , portanto  $\hat{\sigma} = 0,292/b_3 = 0,292/1,693 = 0,172$  e

$$\text{LCI} = 10,476 - 3 \cdot \frac{0,292}{1,693 \sqrt{3}} = 10,476 - 0,299 = 10,177$$

$$\text{LCS} = 10,476 + 3 \cdot \frac{0,292}{1,693 \sqrt{3}} = 10,476 + 0,299 = 10,775$$

Estes limites são idênticos àqueles baseados em  $\bar{s}$  e novamente cada  $\bar{x}_i$  fica entre os limites. ■

## Recalculando os Limites de Controle

Fizemos a hipótese de que os dados amostrais utilizados para estimar  $\mu$  e  $\sigma$  foram obtidos a partir de um processo sob controle. Suponha, entretanto, que um dos pontos sobre o gráfico de controle resultante fique fora dos limites de controle. Então se uma causa assinalável para esta situação fora de controle puder ser encontrada e verificada, é recomendado que novos limites de controle sejam calculados depois de remover a amostra correspondente do conjunto de dados. Da mesma maneira, se mais de um ponto cair fora dos limites originais, novos limites devem ser determinados depois de eliminar qualquer ponto para o qual uma causa assinalável pode ser identificada e tratada. Pode acontecer que um ou mais pontos fiquem fora dos novos limites, neste caso, o processo de remoção/recálculo deve ser repetido.

## Características de Desempenho dos Gráficos de Controle

Genericamente falando, um gráfico de controle será efetivo se ele fornecer pouquíssimos sinais fora de controle quando o processo estiver sob controle, mas mostrar um ponto fora dos limites de controle assim que o processo saia de controle. Uma avaliação da eficácia de um gráfico é baseada na noção de “probabilidades de erro”. Suponha que a variável de interesse seja normalmente distribuída com  $\sigma$  conhecido (o mesmo valor para um processo sob controle ou fora de controle). Além disso, considere um gráfico 3-sigma baseado no valor-alvo  $\mu_0$  com  $\mu = \mu_0$  quando o processo está sob controle. Uma probabilidade de erro é

$$\begin{aligned}\alpha &= P(\text{uma única amostra fornece um ponto fora dos limites de controle quando } \mu = \mu_0) \\ &= P(\bar{X} > \mu_0 + 3\sigma/\sqrt{n} \text{ ou } \bar{X} < \mu_0 - 3\sigma/\sqrt{n} \text{ quando } \mu = \mu_0) \\ &= P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > 3 \text{ ou } \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < -3 \text{ quando } \mu = \mu_0\right)\end{aligned}$$

A variável padronizada  $Z = (\bar{X} - \mu_0)/(\sigma/\sqrt{n})$  tem uma distribuição normal quando  $\mu = \mu_0$ , então

$$\alpha = P(Z > 3 \text{ ou } Z < -3) = \Phi(-3,00) + 1 - \Phi(3,00) = 0,0026$$

Se 3,09 em vez de 3 fosse utilizado para determinar os limites de controle (isto é comum na Inglaterra), então

$$\alpha = P(Z > 3,09 \text{ ou } Z < -3,09) = 0,0020$$

O uso de limites 3-sigma torna altamente improvável que um sinal fora de controle resulte de um processo sob controle.

Agora suponha que o processo saia de controle porque  $\mu$  foi deslocado para  $\mu + \Delta\sigma$  ( $\Delta$  pode ser positivo ou negativo);  $\Delta$  é o número de desvios padrão pelo qual  $\mu$  mudou. Uma segunda probabilidade de erro é

$$\begin{aligned}\beta &= P\left(\text{Uma amostra única resulta em um ponto dentro} \right. \\ &\quad \left. \text{dos limites de controle quando } \mu = \mu_0 + \Delta\sigma\right) \\ &= P(\mu_0 - 3\sigma/\sqrt{n} < \bar{X} < \mu_0 + 3\sigma/\sqrt{n} \text{ quando } \mu = \mu_0 + \Delta\sigma)\end{aligned}$$

Agora, padronizamos primeiramente subtraindo  $\mu_0 + \Delta\sigma$  de cada termo dentro dos parênteses e então dividindo por  $\sigma/\sqrt{n}$ . Isso fornece

$$\begin{aligned}\beta &= P(-3 - \sqrt{n}\Delta < \text{desvio padrão va} < 3 - \sqrt{n}\Delta) \\ &= \Phi(3 - \sqrt{n}\Delta) - \Phi(-3 - \sqrt{n}\Delta)\end{aligned}$$

Esta probabilidade de erro depende de  $\Delta$ , o que determina o tamanho do deslocamento, e do tamanho da amostra  $n$ . Em particular, para  $\Delta$  fixado,  $\beta$  decresce quando  $n$  cresce (quanto maior a amostra, mais provável que resulte um sinal fora de controle), e para  $n$  fixo,  $\beta$  decresce quando  $|\Delta|$  cresce (quanto maior a magnitude do deslocamento, mais provável que resulte um sinal fora de controle). A tabela a seguir fornece  $\beta$  para valores selecionados de  $\Delta$  quando  $n = 4$ .

$\Delta$	0,25	0,50	0,75	1,00	1,50	2,00	2,50	3,00
$\beta$ quando $n = 4$	0,9936	0,9772	0,9332	0,8413	0,5000	0,1587	0,0668	0,0013

Fica claro que um deslocamento pequeno provavelmente passe despercebido em uma única amostra.

Se 3 for substituído por 3,09 nos limites de controle, então  $\alpha$  decresce de 0,0026 para 0,002, mas para qualquer  $n$  e  $\sigma$  fixos,  $\beta$  crescerá. Esta é uma manifestação da relação inversa entre os dois tipos de probabilidades de erro no teste de hipóteses. Por exemplo, alterar 3 para 2,5 aumentará  $\alpha$  e diminuirá  $\beta$ .

As probabilidades de erro discutidas anteriormente foram calculadas sob a hipótese de que a variável de interesse fosse normalmente distribuída. Se a distribuição é apenas ligeiramente não-normal, o efeito do Teorema do Limite Central implica que  $\bar{X}$  terá aproximadamente uma distribuição normal ainda quando  $n$  for pequeno, neste caso, as probabilidades de erro afirmadas serão aproximadamente corretas. Isso, é claro, não é o caso quando a distribuição da variável desvia consideravelmente da normalidade.

Uma segunda avaliação de desempenho utiliza o comprimento esperado ou médio do período necessário para observar um sinal fora de controle. Quando o processo está sob controle, esperaríamos observar várias amostras antes de ver uma cujo  $\bar{x}$  fique fora dos limites de controle. Por outro lado, se um processo sai do controle, o número esperado de amostras necessárias para detectar isso deveria ser pequeno.

Seja  $p$  a probabilidade de que uma única amostra produza um valor  $\bar{x}$  fora dos limites de controle, ou seja,

$$p = P(\bar{X} < \mu_0 - 3\sigma/\sqrt{n} \text{ ou } \bar{X} > \mu_0 + 3\sigma/\sqrt{n})$$

Considere primeiro o processo sob controle, portanto  $\bar{X}_1, \bar{X}_2, \bar{X}_3, \dots$  são todos normalmente distribuídos com valor médio  $\mu_0$  e desvio padrão  $\sigma/\sqrt{n}$ . Defina um  $Y$  por

$$Y = \text{o primeiro } i \text{ para o qual } \bar{X}_i \text{ fique fora dos limites de controle}$$

Se pensarmos em cada número de amostra como um teste e uma amostra fora de controle como um sucesso, então  $Y$  é o número de testes (independentes) necessários para observar um sucesso. Este  $Y$  tem uma distribuição geométrica e mostramos no Exemplo 3.18 que  $E(Y) = 1/p$ . O acrônimo CM (*para comprimento médio*) é frequentemente usado no lugar de  $E(Y)$ . Como  $p = \alpha$  para um processo sob controle, temos

$$CM = E(Y) = \frac{1}{p} = \frac{1}{\alpha} = \frac{1}{0,0026} = 384,62$$

Substituindo 3 nos limites de controle por 3,09 resultará em  $CM = 1/0,002 = 500$ .

Agora suponha que, em um instante de tempo particular, a média do processo seja deslocada para  $\mu = \mu_0 + \Delta\sigma$ . Se definirmos  $Y$  como o primeiro  $i$  subsequente ao deslocamento para o qual uma amostra gera um sinal fora de controle, é novamente verdade que  $CM = E(Y) = 1/p$ , mas agora  $p = 1 - \beta$ . A tabela a seguir resulta nos CMs selecionados para um gráfico 3-sigma quando  $n = 4$ . Estes resultados mostram novamente a eficácia do gráfico em detectar deslocamentos grandes, mas também sua falta de habilidade de identificar rapidamente pequenos deslocamentos. Quando a amostragem é feita com pouca frequência, um grande número de itens é provavelmente produzido antes de um pequeno deslocamento em  $\mu$  ser detectado. Os procedimentos CUSUM discutidos na Seção 16.5 foram desenvolvidos para corrigir esta deficiência.

$\Delta$	0,25	0,50	0,75	1,00	1,50	2,00	2,50	3,00
CM quando $n = 4$	156,25	43,86	14,97	6,30	2,00	1,19	1,07	1,0013

## Regras Suplementares para Gráficos $\bar{X}$

A falta de habilidade dos gráficos  $\bar{X}$  com limites 3-sigma para detectar rapidamente pequenos deslocamentos na média de processo tem alertado os pesquisadores para desenvolverem procedimentos que forneçam melhores comportamentos a este respeito. Uma abordagem envolve a introdução de condições adicionais que façam com que um sinal fora de controle seja gerado. As condições a seguir foram recomendadas pela Western Electric

(então uma subsidiária da AT&T). Uma intervenção para tomar ações corretivas é apropriada quando uma dessas condições for satisfeita:

1. Dois de três pontos sucessivos ficam fora dos limites 2-sigma do mesmo lado da reta central.
2. Quatro de cinco pontos sucessivos ficam fora dos limites 1-sigma do mesmo lado da reta central.
3. Oito pontos sucessivos ficam do mesmo lado da reta central.

Um texto sobre Controle de Qualidade deve ser consultado para uma discussão sobre estas e outras regras suplementares. Na Seção 16.5 apresentamos um tipo diferente de procedimento para superar a deficiência dos gráficos  $\bar{X}$ .

## Gráficos de Controle Robustos

A presença de outliers nos dados da amostra tende a reduzir a sensibilidade dos procedimentos de gráfico de controle quando os parâmetros devem ser estimados. Isto porque os limites de controle são movidos no sentido para fora da reta central, tornando mais difícil a identificação de pontos não-usuais. Não queremos que as estatísticas, cujos valores são colocados no gráfico de controle, sejam resistentes a outliers porque isso mascararia qualquer sinal fora de controle. Por exemplo, colocar as medianas amostrais no gráfico de controle seria menos eficiente do que colocar as  $\bar{x}_1, \bar{x}_2, \dots$  como é feito no gráfico  $\bar{X}$ .

O artigo “Robust Control Charts” por David M. Rocke (*Technometrics*, 1989, p.173-184) apresenta um estudo de procedimentos para os quais os limites de controle estão baseados em estatísticas resistentes aos efeitos de outliers. Rocke recomenda limites de controle calculados a partir do *intervalo interquartil* (IQR), que é muito semelhante à quarta variação apresentada no Capítulo 1. Em particular,

$$\text{IQR} = \begin{cases} (2^{\text{o}} \text{ maior } x_i) - (2^{\text{o}} \text{ menor } x_i) & n = 4, 5, 6, 7 \\ (3^{\text{o}} \text{ maior } x_i) - (3^{\text{o}} \text{ menor } x_i) & n = 8, 9, 10, 11 \end{cases}$$

Para uma amostra aleatória de uma distribuição normal,  $E(\text{IQR}) = k_n \sigma$ ; os valores de  $k_n$  são dados na tabela a seguir.

$n$	4	5	6	7	8
$k_n$	0,596	0,990	1,282	1,512	0,942

Os limites de controle sugeridos são

$$\text{LCI} = \bar{x} - 3 \cdot \frac{\text{IQR}}{k_n \sqrt{n}} \quad \text{LCS} = \bar{x} + 3 \cdot \frac{\text{IQR}}{k_n \sqrt{n}}$$

Os valores de  $\bar{x}_1, \bar{x}_2, \bar{x}_3, \dots$  são colocados no gráfico. As simulações informadas no artigo indicam que o desempenho do gráfico com estes limites é superior ao do gráfico  $\bar{X}$  tradicional.

## Exercícios | Seção 16.2 (4–13)

4. No caso de conhecermos  $\mu$  e  $\sigma$ , quais limites de controle são necessários para a probabilidade de um único ponto estar fora dos limites para um processo sob controle ser 0,005?
  - a.  $\mu_0 + 0,5\sigma$
  - b.  $\mu_0 - \sigma$
  - c.  $\mu_0 + 2\sigma$
5. Considere um gráfico de controle 3-sigma com reta central em  $\mu_0$ , baseado em  $n = 5$ . Sob a hipótese de normalidade, calcule a probabilidade de um único ponto cair fora dos limites de controle quando a média do processo real é
  - a.  $\mu_0 + 0,5\sigma$
  - b.  $\mu_0 - \sigma$
  - c.  $\mu_0 + 2\sigma$
6. A tabela a seguir fornece os dados do conteúdo de uma mistura para amostras de um determinado tipo de tecido. Determine os limites de controle para um gráfico com reta central na altura 13,00 baseado em  $\sigma = 0,600$ , construa o gráfico de controle e comente sobre sua aparência.

## Dados para o Exercício 6

Nº da amostra	Observações de conteúdo da mistura					$\bar{x}$	$s$	Intervalo
1	12,2	12,1	13,3	13,0	13,0	12,72	0,536	1,2
2	12,4	13,3	12,8	12,6	12,9	12,80	0,339	0,9
3	12,9	12,7	14,2	12,5	12,9	13,04	0,669	1,7
4	13,2	13,0	13,0	12,6	13,9	13,14	0,477	1,3
5	12,8	12,3	12,2	13,3	12,0	12,52	0,526	1,3
6	13,9	13,4	13,1	12,4	13,2	13,20	0,543	1,5
7	12,2	14,4	12,4	12,4	12,5	12,78	0,912	2,2
8	12,6	12,8	13,5	13,9	13,1	13,18	0,526	1,3
9	14,6	13,4	12,2	13,7	12,5	13,28	0,963	2,4
10	12,8	12,3	12,6	13,2	12,8	12,74	0,329	0,9
11	12,6	13,1	12,7	13,2	12,3	12,78	0,370	0,9
12	13,5	12,3	12,8	13,1	12,9	12,92	0,438	1,2
13	13,4	13,3	12,0	12,9	13,1	12,94	0,559	1,4
14	13,5	12,4	13,0	13,6	13,4	13,18	0,492	1,2
15	12,3	12,8	13,0	12,8	13,5	12,88	0,432	1,2
16	12,6	13,4	12,1	13,2	13,3	12,92	0,554	1,3
17	12,1	12,7	13,4	13,0	13,9	13,02	0,683	1,8
18	13,0	12,8	13,0	13,3	13,1	13,04	0,182	0,5
19	12,4	13,2	13,0	14,0	13,1	13,14	0,573	1,6
20	12,7	12,4	12,4	13,9	12,8	12,84	0,619	1,5
21	12,6	12,8	12,7	13,4	13,0	12,90	0,316	0,8
22	12,7	13,4	12,1	13,2	13,3	12,94	0,541	1,3

7. Consulte os dados fornecidos no Exercício 6 e construa um gráfico de controle com uma reta central estimada e limites baseados no uso de desvios padrão amostrais para estimar  $\sigma$ . Há alguma evidência que o processo esteja fora de controle?

8. Consulte os exercícios 6 e 7 e agora aplique os limites de controle baseados no uso das amplitudes amostrais para estimar  $\sigma$ . O processo parece estar sob controle?

9. A tabela a seguir fornece médias amostrais e desvios padrão, cada um baseado em  $n = 6$  observações do índice refratário do cabo de fibra óptica. Construa um gráfico de controle e comente sobre sua aparência. [Sugestão:  $\sum \bar{x}_i = 2317,07$  e  $\sum s_i = 30,34$ .]

Dia	$\bar{x}$	$s$	Dia	$\bar{x}$	$s$
1	95,47	1,30	9	96,63	1,48
2	97,38	0,88	10	96,50	0,80
3	96,85	1,43	11	97,22	1,42
4	96,64	1,59	12	96,55	1,65
5	96,87	1,52	13	97,02	1,28
6	96,52	1,27	14	95,55	1,14
7	96,08	1,16	15	96,29	1,37
8	96,48	0,79	16	96,80	1,40

(continua)

Dia	$\bar{x}$	$s$	Dia	$\bar{x}$	$s$
17	96,01	1,58	21	97,06	1,34
18	95,39	0,98	22	98,34	1,60
19	96,58	1,21	23	96,42	1,22
20	96,43	0,75	24	95,99	1,18

10. Consulte o Exercício 9. Uma causa assinalável foi encontrada para a média amostral não usualmente alta do índice de refração do dia 22. Recalcule os limites de controle depois de remover os dados deste dia. O que você poderia concluir?

11. Considere o gráfico de controle baseado nos limites de controle  $\mu_0 \pm 2,81 \sigma/\sqrt{n}$ .

a. Qual é o CM quando o processo está sob controle?

b. Qual é o CM quando  $n = 4$  e a média do processo está deslocada para  $\mu = \mu_0 + \sigma$ ?

c. Como os valores das partes (a) e (b) podem ser comparados com os valores correspondentes para um gráfico 3-sigma?

12. Aplique as regras suplementares sugeridas no texto aos dados do Exercício 6. Há quaisquer sinais fora de controle?

13. Calcule os limites de controle para os dados do Exercício 6 utilizando o procedimento robusto apresentado nesta seção.

## 16.3 | Gráficos de Controle de Variações de Processo

Os gráficos de controle discutidos na seção anterior foram elaborados a fim de controlar a localização — ou seja, a tendência central, — de um processo, sendo a média, em particular, uma medida de localização. É igualmente importante assegurar que um processo está sob controle em relação à variação. Na verdade, recomenda-se que o controle seja estabelecido sobre a variação antes de construir um gráfico  $\bar{X}$  ou qualquer outro gráfico para controle de localização. Nesta seção, consideramos os gráficos para variação baseados no desvio padrão da amostra  $S$  e também gráficos baseados na amplitude da amostra  $R$ . O primeiro é, em geral, preferido pois o desvio padrão é um método mais eficiente de avaliação da variação do que a amplitude, mas o gráfico de  $R$  foi utilizado primeiro e tradições são difíceis de acabar.

### O Gráfico $S$

Supomos, novamente, que  $k$  amostras independentemente selecionadas estão disponíveis, cada uma consistindo de  $n$  observações de uma variável distribuída normalmente. Represente os desvios padrão da amostra por  $s_1, s_2, \dots, s_k$ , com  $\bar{s} = \sum s_i/k$ . Os valores de  $s_1, s_2, s_3, \dots$  são marcados em seqüência em um gráfico  $S$ . A reta central do gráfico terá altura  $\bar{s}$  e os limites 3-sigma requerem que sejam determinados  $3\sigma_S$  (assim como os limites 3-sigma de um gráfico  $\bar{X}$  necessitavam de  $3\sigma_{\bar{X}} = 3\sigma/\sqrt{n}$ , com  $\sigma$  então estimado a partir dos dados).

Lembre-se de que para qualquer  $Y$ ,  $V(Y) = E(Y^2) - [E(Y)]^2$ , e que a variância amostral  $S^2$  é um estimador não-tendencioso de  $\sigma^2$ , ou seja  $E(S^2) = \sigma^2$ . Assim,

$$V(S) = E(S^2) - [E(S)]^2 = \sigma^2 - (a_n\sigma)^2 = \sigma^2(1 - a_n^2)$$

onde os valores  $a_n$  para  $n = 3, \dots, 8$  estão tabulados na seção anterior. O desvio padrão de  $S$  é então

$$\sigma_S = \sqrt{V(S)} = \sigma\sqrt{1 - a_n^2}$$

É natural estimar  $\sigma$  utilizando  $s_1, \dots, s_k$  como foi feito na seção anterior, isto é,  $\hat{\sigma} = \bar{s}/a_n$ . Substituir  $\hat{\sigma}$  por  $\sigma$  na expressão para  $\sigma_S$  fornece a quantidade utilizada para calcular os limites 3-sigma.

Os limites de controle 3-sigma para um gráfico de controle  $S$  são

$$\text{LCI} = \bar{s} - 3\bar{s}\sqrt{1 - a_n^2}/a_n$$

$$\text{LCS} = \bar{s} + 3\bar{s}\sqrt{1 - a_n^2}/a_n$$

A expressão para LCI será negativa se  $n \leq 5$ , neste caso, é comum utilizar  $\text{LCI} = 0$ .

#### Exemplo 16.4

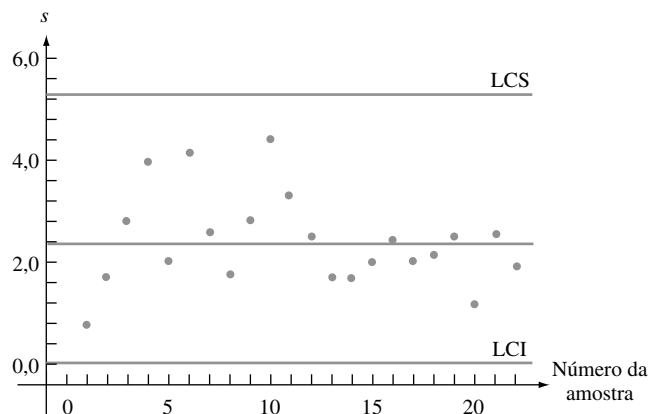
A Tabela 16.2 mostra as observações da resistência à tensão de folhas de plástico (a força, em psi, necessária para quebrar uma folha). Há  $k = 22$  amostras, obtidas para instantes de tempo igualmente espaçados e  $n = 4$  observações em cada exemplo. É fácil verificar que  $\sum s_i = 51,10$  e  $\bar{s} = 2,32$ , portanto o centro do gráfico  $S$  será em 2,32 (no entanto, já que  $n = 4$ ,  $\text{LCI} = 0$  e a reta central não estará equidistante aos limites de controle). Da seção anterior,  $a_4 = 0,921$ , a partir da qual o LCS é

$$\begin{aligned} \text{LCS} &= 2,32 + 3(2,32)(\sqrt{1 - (0,921)^2})/0,921 \\ &= 2,32 + 2,94 \\ &= 5,26 \end{aligned}$$

**Tabela 16.2** Dados de resistência à tensão para o Exemplo 16.4

Nº da amostra	Observações				DP	Intervalo
1	29,7	29,0	28,8	30,2	0,64	1,4
2	32,2	29,3	32,2	32,9	1,60	3,6
3	35,9	29,1	32,1	31,3	2,83	6,8
4	28,8	27,2	28,5	35,7	3,83	8,5
5	30,9	32,6	28,3	28,3	2,11	4,3
6	30,6	34,3	34,8	26,3	3,94	8,5
7	32,3	27,7	30,9	27,8	2,30	4,6
8	32,0	27,9	31,0	30,8	1,76	4,1
9	24,2	27,5	28,5	31,1	2,85	6,9
10	33,7	24,4	34,3	31,0	4,53	9,9
11	35,3	33,2	31,4	28,0	3,09	7,3
12	28,1	34,0	31,0	30,8	2,41	5,9
13	28,7	28,9	25,8	29,7	1,71	3,9
14	29,0	33,0	30,2	30,1	1,71	4,0
15	33,5	32,6	33,6	29,2	2,07	4,4
16	26,9	27,3	32,1	28,5	2,37	5,2
17	30,4	29,6	31,0	33,8	1,83	4,2
18	29,0	28,9	31,8	26,7	2,09	5,1
19	33,8	30,9	31,7	28,2	2,32	5,6
20	29,7	27,9	29,1	30,1	0,96	2,2
21	27,9	27,7	30,2	32,9	2,43	5,2
22	30,0	31,4	27,7	28,1	1,72	3,7

O gráfico de controle resultante é mostrado na Figura 16.3. Todos os pontos marcados estão dentro dos limites de controle, sugerindo um comportamento de processo estável, em relação à variação.

**Figura 16.3** O gráfico S para os dados de resistência à tensão para o Exemplo 16.4

## O Gráfico R

Sejam  $r_1, r_2, \dots, r_k$  a amplitude da amostra  $k$  e  $\bar{r} = \sum r_i/k$ . A reta central do gráfico R estará à altura  $\bar{r}$ . A determinação dos limites de controle requer  $\sigma_R$ , onde R indica a amplitude (antes de fazer as observações como uma variável aleatória) de uma amostra aleatória de tamanho  $n$  a partir de uma distribuição normal com valor médio  $\mu$  e desvio padrão  $\sigma$ . Como

$$\begin{aligned} R &= \text{máx}(X_1, \dots, X_n) - \text{mín}(X_1, \dots, X_n) \\ &= \sigma\{\text{máx}(Z_1, \dots, Z_n) - \text{mín}(Z_1, \dots, Z_n)\} \end{aligned}$$

onde  $Z_i = (X_i - \mu)/\sigma$ , e os  $Z_i$ s são os vas normais padronizadas, segue que



$$\begin{aligned} \sigma_R &= \sigma \cdot \left( \text{desvio padrão da amplitude de uma amostra aleatória de} \right. \\ &\quad \left. \text{tamanho } n \text{ a partir de uma distribuição normal padronizada} \right) \\ &= \sigma \cdot c_n \end{aligned}$$

Os valores de  $c_n$  para  $n = 3, \dots, 8$  são exibidos na tabela a seguir.

$n$	3	4	5	6	7	8
$c_n$	0,888	0,880	0,864	0,848	0,833	0,820

É comum estimar  $\sigma$  por  $\hat{\sigma} = \bar{r}/b_n$ , como discutido na seção anterior. Isto resulta em  $\hat{\sigma}_R = c_n \bar{r}/b_n$  como o desvio padrão estimado de  $R$ .

Os limites 3-sigma para um gráfico  $R$  são

$$LCI = \bar{r} - 3c_n \bar{r}/b_n$$

$$LCS = \bar{r} + 3c_n \bar{r}/b_n$$

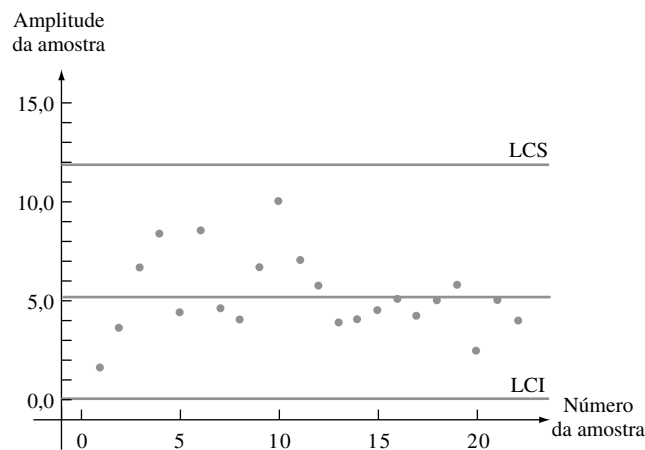
A expressão para LCI será negativa se  $n \leq 6$ , neste caso  $LCI = 0$  deverá ser utilizado.

**Exemplo 16.5** (Continuação do Exemplo 16.4)

A Tabela 16.2 produz  $\sum r_i = 115,3$  e  $\bar{r} = 5,24$ . Como  $n = 4$ ,  $LCI = 0$ . Com  $b_4 = 2,058$  e  $c_4 = 0,880$ ,

$$LCS = 5,24 + 3(0,880)(5,24)/2,058 = 11,96$$

O gráfico  $R$  aparece na Figura 16.4. Assim como com o gráfico  $S$ , todos os pontos encontram-se entre os limites, indicando um processo sob controle quanto à variação admitida.



**Figura 16.4** O gráfico  $R$  para dados de resistência à tensão do Exemplo 16.5

**Gráficos Baseados em Limites de Probabilidade**

Considere um gráfico  $\bar{X}$  baseado no valor  $\mu_0$  sob controle (alvo) e  $\sigma$  conhecido. Quando a variável de interesse é normalmente distribuída e o processo está sob controle,

$$P(\bar{X}_i > \mu_0 + 3\sigma/\sqrt{n}) = 0,0013 = P(\bar{X}_i < \mu_0 - 3\sigma/\sqrt{n})$$

Ou seja, a probabilidade que um ponto sobre o gráfico tem de estar acima do LCS é de 0,0013, assim como é a probabilidade de aquele ponto estar abaixo de LCI (utilizando 3,09 no lugar de 3 resulta em 0,001 para cada probabilidade). Quando os limites de controle são baseados em estimativas de  $\mu$  e  $\sigma$ , estas probabilidades serão aproximadamente corretas desde que  $n$  não seja tão pequeno e  $k$  seja, pelo menos, maior do que 20.

Por outro lado, este não é o caso para o gráfico 3-sigma  $S$  em que  $P(S_i > \text{LCS}) = P(S_i < \text{LCI}) = 0,0013$ , nem é verdadeiro para um gráfico 3-sigma  $R$  em que  $P(R_i > \text{LCS}) = P(R_i < \text{LCI}) = 0,0013$ . Isto ocorre pois nem o desvio padrão amostral  $S$  nem a amplitude da amostra  $R$  têm uma distribuição normal, mesmo quando a população é normalmente distribuída. Em vez disso, ambos  $S$  e  $R$  têm distribuições distorcidas. O melhor que pode ser dito para os gráficos 3-sigma  $S$  e  $R$  é que em um processo sob controle é pouco provável produzir um ponto em qualquer instante particular que esteja fora dos limites de controle. Alguns autores afirmam que o uso de limites de controle para os quais a “probabilidade em excesso” de cada limite é de aproximadamente 0,001. O livro *Statistical Methods for Quality Improvement* (veja a Bibliografia do capítulo) contém mais informações sobre este assunto.

## Exercícios | Seção 16.3 (14–18)

14. Um fabricante de giz antialérgico instituiu um programa de controle de qualidade para monitorar a densidade do giz. Os desvios padrão amostrais da densidade para 24 subgrupos diferentes, cada um consistindo de  $n = 8$  espécimes de giz, foram os seguintes:

0,204 0,315 0,096 0,184 0,230 0,212 0,322 0,287  
0,145 0,211 0,053 0,145 0,272 0,351 0,159 0,214  
0,388 0,187 0,150 0,229 0,276 0,118 0,091 0,056

Calcule os limites para um gráfico  $S$ , construa o gráfico e identifique os pontos fora de controle. Se há um ponto fora de controle, elimine-o e repita o processo.

15. Subgrupos de unidades de fornecimento de energia são selecionados a cada hora de uma linha de montagem e a saída de alta voltagem de cada unidade é determinada.
- Suponha que a soma dessas amplitudes amostrais resultantes para 30 subgrupos, cada um consistindo de quatro unidades, seja de 85,2. Calcule os limites de controle para um gráfico  $R$ .
  - Repita o item (a) se cada subgrupo consiste de oito unidades e a soma é 106,2.
16. Calcule os limites de controle para um gráfico  $S$  e um gráfico  $R$  utilizando os dados de conteúdo de mistura do Exercício 6. Verifique a presença de quaisquer sinais fora de controle.

17. Calcule os limites de controle para um gráfico  $S$  a partir dos dados de índice de refração do Exercício 9. O processo parece estar sob controle em relação à variabilidade? Por quê?

18. Quando  $S^2$  é a variância amostral de uma amostra aleatória normalmente distribuída,  $(n - 1)S^2/\sigma^2$  tem uma distribuição qui-quadrado com  $n - 1$  gl, portanto

$$P\left(\chi_{0,999,n-1}^2 < \frac{(n-1)S^2}{\sigma^2} < \chi_{0,001,n-1}^2\right) = 0,998$$

A partir do qual

$$P\left(\frac{\sigma^2 \chi_{0,999,n-1}^2}{n-1} < S^2 < \frac{\sigma^2 \chi_{0,001,n-1}^2}{n-1}\right) = 0,998$$

Isto sugere que um gráfico alternativo para o controle da variação do processo compreende na representação gráfica das variâncias amostrais e na utilização de limites de controle

$$\text{LCI} = \frac{\bar{s}^2 \chi_{0,999,n-1}^2}{n-1}$$

$$\text{LCS} = \frac{\bar{s}^2 \chi_{0,001,n-1}^2}{n-1}$$

Construa o gráfico correspondente para os dados do Exercício 9. (*Sugestão:* Os valores críticos para as caudas inferior e superior da distribuição qui-quadrado para 5 gl são 0,210 e 20,515, respectivamente.)

## 16.4 | Gráficos de Controle de Atributos

A expressão *dados de atributo* é utilizada na literatura de controle de qualidade para descrever duas situações:

- Cada item produzido é defeituoso ou não-defeituoso (conforme esteja ou não de acordo com as especificações).
- Um único item pode ter um ou mais defeitos, e o número de defeitos é determinado.

No primeiro caso, um gráfico de controle está baseado na distribuição binomial; no segundo caso, a distribuição de Poisson é a base para o gráfico.

### O Gráfico $p$ para Defeitos de Fração

Suponha que quando um processo está sob controle, a probabilidade de que qualquer item particular tenha defeito é  $p$  (de maneira equivalente,  $p$  é a proporção a longo prazo dos itens com defeito para um processo sob controle) e que itens diferentes são independentes um do outro em relação às suas condições. Considere uma amostra de  $n$  itens obtidos em um instante particular, e seja  $X$  o número de defeitos e  $\hat{p} = X/n$ . Como  $X$  tem uma distribuição binomial,  $E(X) = np$  e  $V(X) = np(1 - p)$ , portanto

$$E(\hat{p}) = p \quad V(\hat{p}) = \frac{p(1-p)}{n}$$

Além disso, se  $np \geq 10$  e  $n(1 - p) \geq 10$ ,  $\hat{p}$  tem uma distribuição aproximadamente normal.

No caso de  $p$  ser conhecido (ou um gráfico baseado no valor-alvo), os limites de controle são

$$\text{LCL} = p - 3\sqrt{\frac{p(1-p)}{n}} \quad \text{UCL} = p + 3\sqrt{\frac{p(1-p)}{n}}$$

Se cada amostra consiste de  $n$  itens, o número de itens com defeito na  $i$ -ésima amostra é  $x_i$ , e  $\hat{p}_i = x_i/n$ , então  $\hat{p}_1, \hat{p}_2, \hat{p}_3, \dots$  são marcados no gráfico de controle.

Normalmente o valor de  $p$  deve ser estimado a partir dos dados. Suponha que  $k$  amostras, a partir das quais acredita-se que esteja um processo sob controle, estão disponíveis, e seja

$$\bar{p} = \frac{\sum_{i=1}^k \hat{p}_i}{k}$$

O  $\bar{p}$  estimado é então usado no lugar de  $p$  nos limites de controle anteriormente mencionados.

O gráfico de  $p$  para a fração de itens com defeito tem sua reta central na altura  $\bar{p}$  e limites de controle

$$\text{LCL} = \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

$$\text{UCL} = \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

Se LCI é negativo, ele é substituído por 0.

#### Exemplo 16.6

Uma amostra de 100 xícaras de chá de um determinado conjunto de jantar foi selecionada a cada 25 dias sucessivos, e cada uma foi examinada em busca de defeitos. Os números resultantes de xícaras defeituosas e as proporções amostrais correspondentes são os seguintes:

<i>Dia (i)</i>	1	2	3	4	5	6	7	8	9	10	11	12	13
$x_i$	7	4	3	6	4	9	6	7	5	3	7	8	4
$\hat{p}_i$	0,07	0,04	0,03	0,06	0,04	0,09	0,06	0,07	0,05	0,03	0,07	0,08	0,04
<i>Dia (i)</i>	14	15	16	17	18	19	20	21	22	23	24	25	
$x_i$	6	2	9	7	6	7	11	6	7	4	8	6	
$\hat{p}_i$	0,06	0,02	0,09	0,07	0,06	0,07	0,11	0,06	0,07	0,04	0,08	0,06	

Fazendo a hipótese de que o processo estava sob controle durante este período, vamos estabelecer os limites de controle e construir um gráfico  $p$ . Temos que  $\sum \hat{p}_i = 1,52$ , que leva a  $\bar{p} = 1,52/25 = 0,0608$  e

$$\begin{aligned} \text{LCI} &= 0,0608 - 3\sqrt{(0,0608)(0,9392)/100} = 0,0608 - 0,0717 = -0,0109 \\ \text{LCS} &= 0,0608 + 3\sqrt{(0,0608)(0,9392)/100} = 0,0608 + 0,0717 = 0,1325 \end{aligned}$$

O LCI é, portanto, definido como 0. O gráfico de controle exibido na Figura 16.5 mostra que todos os pontos estão dentro dos limites de controle. Isto é consistente com a hipótese de um processo sob controle.

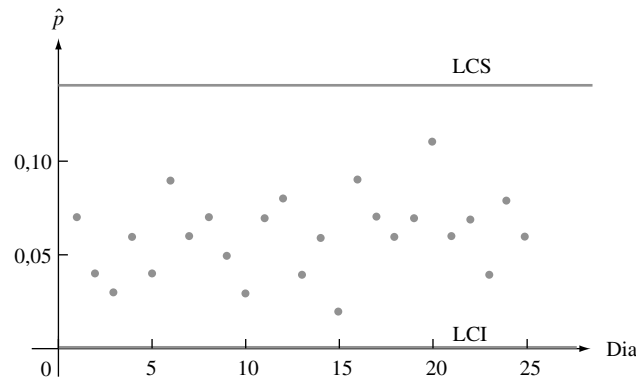


Figura 16.5 Gráfico de controle para os dados de fração de defeito do Exemplo 16.6

## O Gráfico $c$ para o Número de Defeitos

Consideraremos, agora, situações nas quais a observação em cada instante no tempo é o número de defeitos em uma unidade de algum tipo. A unidade pode consistir de um único item (por exemplo, um automóvel) ou um grupo de itens (por exemplo, rasgos em um conjunto de quatro pneus). No segundo caso, o tamanho do grupo é, por hipótese, o mesmo em cada instante do tempo.

O gráfico de controle para o número de defeitos está baseado na distribuição de probabilidades de Poisson. Lembre-se de que se  $Y$  é uma variável aleatória de Poisson com parâmetro  $\theta$ , então

$$E(Y) = \theta \quad V(Y) = \theta \quad \sigma_Y = \sqrt{\theta}$$

Além disso,  $Y$  tem uma distribuição aproximadamente normal quando  $\theta$  é grande ( $\theta \geq 10$  será suficiente para a maioria dos casos). Além disso, se  $Y_1, Y_2, \dots, Y_n$  são variáveis de Poisson independentes com parâmetros  $\theta_1, \theta_2, \dots, \theta_n$ , pode-se mostrar que  $Y_1 + \dots + Y_n$  tem uma distribuição de Poisson com parâmetro  $\theta_1 + \dots + \theta_n$ . Em particular, se  $\theta_1 = \dots = \theta_n = \theta$  (a distribuição do número de defeitos por item é o mesmo para cada item), então o parâmetro de Poisson é  $\lambda = n\theta$ .

Seja  $\lambda$  o parâmetro de Poisson para o número de defeitos em uma unidade (é o número esperado de defeitos por unidade). No caso de  $\lambda$  conhecido (ou de um gráfico baseado em um valor-alvo),

$$\text{LCI} = \lambda - 3\sqrt{\lambda} \quad \text{LCS} = \lambda + 3\sqrt{\lambda}$$

Com  $x_i$  indicando o número total de defeitos na  $i$ -ésima unidade ( $i = 1, 2, 3, \dots$ ), então pontos com alturas  $x_1, x_2, x_3, \dots$  são marcados sobre o gráfico. Normalmente, o valor de  $\lambda$  deve ser estimado a partir dos dados. Como  $E(X_i) = \lambda$ , é natural utilizar o  $\lambda = \bar{x}$  estimado (baseado em  $x_1, x_2, \dots, x_k$ ).

O gráfico  $c$  para o número de defeitos em uma unidade tem reta central na altura  $\bar{x}$  e

$$\begin{aligned} \text{LCI} &= \bar{x} - 3\sqrt{\bar{x}} \\ \text{LCS} &= \bar{x} + 3\sqrt{\bar{x}} \end{aligned}$$

Se LCI for negativo, ele será substituído por 0.

**Exemplo 16.7**

Uma companhia produz painéis de metal que são conformados após terem sido inicialmente revestidas com uma mistura contendo cerâmica em pó. Algumas vezes aparecem defeitos na finalização destas painéis e a companhia deseja estabelecer um gráfico de controle para o número de defeitos. Este, em cada uma das 24 painéis amostradas em intervalos de tempo iguais são como segue:

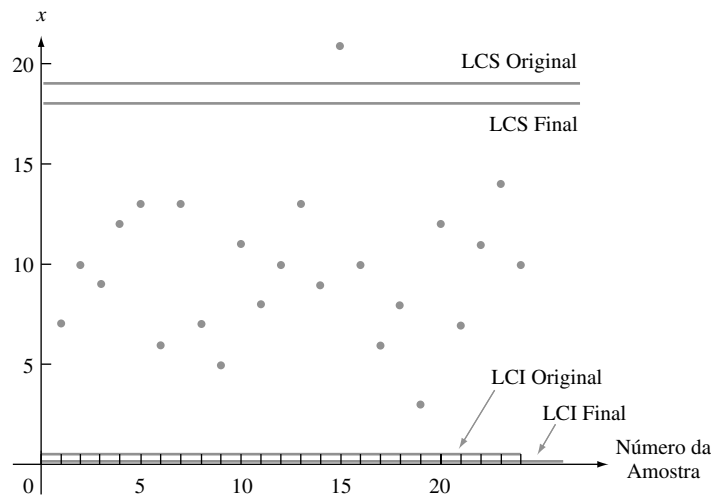
7 10 9 12 13 6 13 7 5 11 8 10  
13 9 21 10 6 8 3 12 7 11 14 10

com  $\sum x_i = 235$  e  $\hat{\lambda} = \bar{x} = 235/24 = 9,79$ . Os limites de controle são

$$LCI = 9,79 - 3\sqrt{9,79} = 0,40 \quad LCS = 9,79 + 3\sqrt{9,79} = 19,18$$

O gráfico de controle está na Figura 16.6. O ponto correspondente à painél de número quinze está acima do LCS. Após investigação, descobriu-se que o revestimento utilizado sobre aquela painél possuía viscosidade baixa (uma possível causa). Eliminando esta observação do conjunto de dados, resulta em  $\bar{x} = 214/23 = 9,30$  e os novos limites de controle

$$LCI = 9,30 - 3\sqrt{9,30} = 0,15 \quad LCS = 9,30 + 3\sqrt{9,30} = 18,45$$



**Figura 16.6** O gráfico de controle para o número de defeitos dos dados do Exemplo 16.7

As 23 observações remanescentes ficam todas entre estes limites, indicando um processo sob controle. ■

### Gráficos de Controle Baseados em Dados Transformados

Presume-se que o uso de limites de controle 3-sigma resulte em  $P(\text{estatística} < LCI) \approx P(\text{estatística} > LCS) \approx 0,0013$  quando o processo está sob controle. Entretanto, quando  $p$  é pequeno, a aproximação normal para a distribuição de  $\hat{p} = X/n$ , em geral, não será muito precisa nas caudas extremas. A Tabela 16.3 fornece a evidência de que este comportamento para valores selecionados de  $p$  e  $n$  (o valor de  $p$  é utilizado para calcular os limites de controle). Em muitos casos, a probabilidade de que um único ponto caia fora dos limites de controle é muito diferente da probabilidade nominal de 0,0026.

**Tabela 16.3** Probabilidades sob controle para um gráfico  $p$ 

$p$	$n$	$P(\hat{p} < \text{LCI})$	$P(\hat{p} > \text{LCS})$	$P(\text{ponto fora de controle})$
0,10	100	0,00003	0,00198	0,00201
0,10	200	0,00048	0,00299	0,00347
0,10	400	0,00044	0,00171	0,00215
0,05	200	0,00004	0,00266	0,00270
0,05	400	0,00020	0,00207	0,00227
0,05	600	0,00031	0,00189	0,00220
0,02	600	0,00007	0,00275	0,00282
0,02	800	0,00036	0,00374	0,00410
0,02	1.000	0,00023	0,00243	0,00266

Esse problema pode ser remediado com a aplicação de uma transformação nos dados. Seja  $h(X)$  uma função aplicada para transformar a variável binomial  $X$ . Então  $h(\cdot)$  deve ser escolhida tal que  $h(X)$  tenha uma distribuição aproximadamente normal e que esta aproximação seja precisa nas caudas. Uma transformação recomendada é baseada na função arcsen (ou seja,  $\text{sen}^{-1}$ ):

$$Y = h(X) = \text{sen}^{-1}(\sqrt{X/n})$$

Então  $Y$  é aproximadamente normal com média  $\text{sen}^{-1}(\sqrt{p})$  e variância  $1/(4n)$ ; observe que a variância é independente de  $p$ . Seja  $y_i = \text{sen}^{-1}(\sqrt{x_i/n})$ . Então os pontos sobre o gráfico de controle são marcados nas alturas  $y_1, y_2, \dots$ . Para  $n$  conhecido, os limites de controle são

$$\text{LCI} = \text{sen}^{-1}(\sqrt{p}) - 3\sqrt{1/(4n)} \quad \text{LCS} = \text{sen}^{-1}(\sqrt{p}) + 3\sqrt{1/(4n)}$$

Quando  $p$  não é conhecido,  $\text{sen}^{-1}(\sqrt{p})$  é substituído por  $\bar{y}$ .

Comentários semelhantes se aplicam à distribuição de Poisson quando  $\lambda$  é pequeno. A transformação sugerida é  $Y = h(X) = 2\sqrt{X}$  que tem média  $2\sqrt{\lambda}$  e variância 1. Os limites de controle resultantes são,  $2\sqrt{\lambda} \pm 3$  quando  $\lambda$  é conhecido e  $\bar{y} \pm 3$ , caso contrário. O livro *Statistical Methods for Quality Improvement*, listado na Bibliografia do capítulo, discute estas questões com mais detalhes.

## Exercícios | Seção 16.4 (19–26)

- Em cada um dos 25 dias anteriores, 100 dispositivos eletrônicos de um tipo determinado foram selecionados aleatoriamente e sujeitos a um severo teste de *stress* ao calor. O número total de itens que falharam no teste foi 578.
  - Determine os limites de controle para um gráfico  $p$  3-sigma.
  - O número mais alto de itens falhos em um determinado dia foi 39 e o menor número foi 13. Algum destes corresponde a um ponto fora de controle? Explique.
- Uma amostra de 200 chips ROM de computador foi selecionada em cada um dos 30 dias consecutivos, e o número de chips em não-conformidade em cada dia foi o seguinte: 10, 18, 24, 17, 37, 19, 7, 25, 11, 24, 29, 15, 16, 21, 18, 17, 15, 22, 12, 20, 17, 18, 12, 24, 30, 16, 11, 20, 14, 28. Construa um gráfico  $p$  e examine-o para quaisquer pontos fora de controle.
- Quando  $n = 150$ , qual é o menor valor de  $\bar{p}$  para o qual o LCI em um gráfico  $p$  é positivo?
- Consulte os dados do Exercício 20 e construa um gráfico de controle utilizando a transformação  $\text{sen}^{-1}$  como sugerido no texto.
- As observações correspondentes são números de defeitos em 25 amostras de 1 jarda quadrada de tecido fabricado de um determinado tipo: 3, 7, 5, 3, 4, 2, 8, 4, 3, 3, 6, 7, 2, 3, 2, 4, 7, 3, 2, 4, 4, 1, 5, 4, 6. Construa um gráfico  $c$  para o número de defeitos.
- Para quais valores de  $\bar{x}$  o LCI em um gráfico  $c$  é negativo?

25. Em algumas situações, os tamanhos de espécies amostradas varia e espera-se que espécies maiores tenham mais defeitos dos que as menores. Por exemplo, por um lado tamanhos das amostras de tecidos inspecionados para detectar defeitos podem variar ao longo do tempo. Por outro lado, o número de itens inspecionados pode mudar com o tempo. Seja

$$u_i = \frac{\text{o número de defeitos observados no momento } i}{\text{o tamanho do item inspecionado no momento } i}$$

$$= \frac{x_i}{g_i}$$

onde “tamanho” pode referir-se à área, comprimento, volume ou simplesmente ao número de itens inspecionados. Então um **gráfico  $u$**  com plotes  $u_1, u_2, \dots$ , tem reta central  $\bar{u}$  e os limites de controle para as  $i$ -ésimas observações são  $\bar{u} \pm 3\sqrt{\bar{u}/g_i}$ .

Os painéis pintados foram examinados seqüencialmente no tempo e para cada um, o número de manchas em uma região de amostragem especificada foi determinado. A área de superfície (ft<sup>2</sup> — pés quadrados) da região examinada variou de painel para painel. Os resultados são dados a seguir. Construa um gráfico  $u$ .

Painel	Área examinada	Nº de Manchas
1	0,8	3
2	0,6	2
3	0,8	3
4	0,8	2
5	1,0	5
6	1,0	5
7	0,8	10
8	1,0	12
9	0,6	4
10	0,6	2
11	0,6	1
12	0,8	3
13	0,8	5
14	1,0	4
15	1,0	6
16	1,0	12
17	0,8	3
18	0,6	3
19	0,6	5
20	0,6	1

26. Construa um gráfico de controle para os dados do Exercício 23 utilizando a transformação sugerida no texto.

## 16.5 | Procedimentos de CUSUM

Um defeito do gráfico  $X$  tradicional é sua falta de habilidade para detectar uma mudança relativamente pequena em uma média de processo. Isto é com certeza uma consequência do fato de que para um processo ser julgado como fora de controle em um momento específico, depende somente da amostra naquele instante, e não da história passada do processo. Os gráficos de controle e procedimentos de soma acumulada (CUSUM) foram projetados para corrigir este defeito.

Há duas versões equivalentes de um procedimento CUSUM para uma média de processo, uma gráfica e a outra computacional. A versão computacional é utilizada quase que exclusivamente na prática, mas a lógica por trás do procedimento é mais fácil de compreender considerando primeiro a forma gráfica.

### A Máscara V

Seja  $\mu_0$  um valor-alvo ou objetivo para a média de processo e sejam definidas as somas acumuladas por:

$$S_1 = \bar{x}_1 - \mu_0$$

$$S_2 = (\bar{x}_1 - \mu_0) + (\bar{x}_2 - \mu_0) = \sum_{i=1}^2 (\bar{x}_i - \mu_0)$$

$$\vdots$$

$$\vdots$$

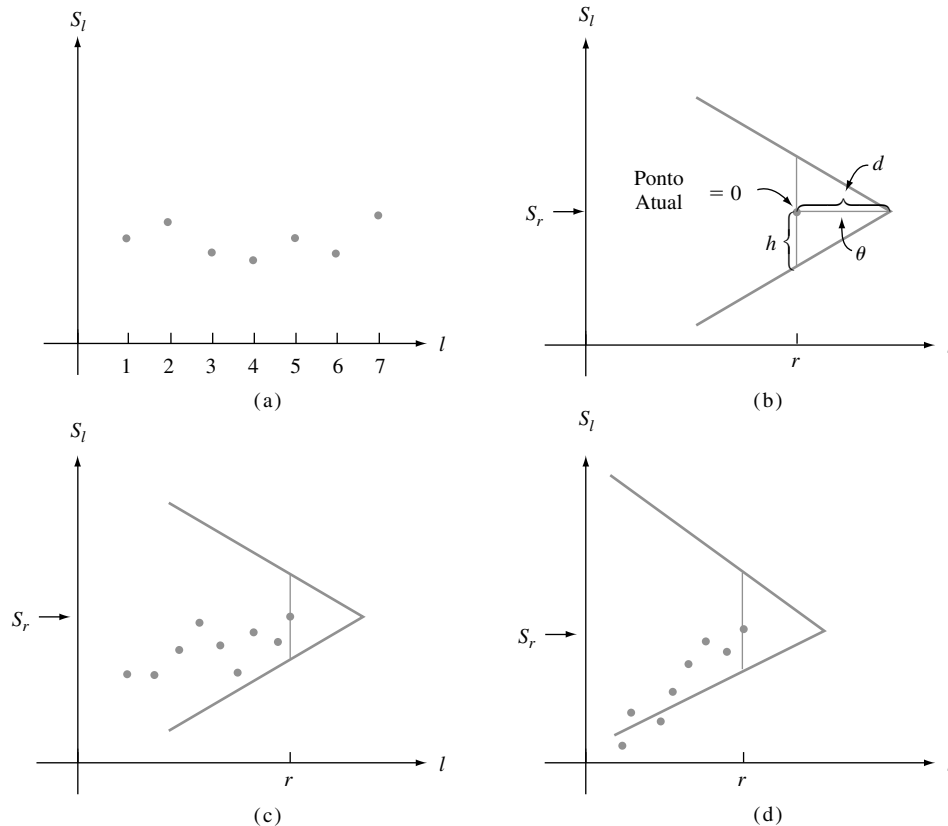
$$\vdots$$

$$S_l = (\bar{x}_1 - \mu_0) + \dots + (\bar{x}_l - \mu_0) = \sum_{i=1}^l (\bar{x}_i - \mu_0)$$

(na ausência de um valor-alvo,  $\bar{x}$  é utilizado no lugar de  $\mu_0$ ). Essas somas acumuladas são marcadas ao longo do tempo. Ou seja, no momento 1 (“ele”), marcamos um ponto na altura  $S_1$ . No ponto do tempo atual  $r$ , os pontos marcados são  $(1, S_1), (2, S_2), (3, S_3), \dots, (r, S_r)$ .

Agora, uma “máscara” com formato de V é sobreposta sobre a carta de controle, como mostra a Figura 16.7. O ponto 0, que está a uma distância  $d$  atrás do ponto no qual os dois braços da máscara se cruzam, é posicionado no ponto CUSUM atual  $(r, S_r)$ . No momento  $r$ , o processo é julgado como fora de controle se qualquer um dos pontos

marcados ficar fora da máscara V, tanto acima do braço superior quanto abaixo do braço inferior. Quando o processo está sob controle, os  $\bar{x}_i$ s variarão em torno do valor-alvo  $\mu_0$ , portanto  $S_i$ s sucessivos deverão variar em torno de 0. Suponha, entretanto, que em um tempo determinado, a média do processo se desloca para um valor maior do que o alvo. Daquele ponto em diante, as diferenças  $\bar{x}_i - \mu_0$  tenderão a ser positivas, tal que  $S_i$ s sucessivos crescerão e os pontos marcados migrarão para cima. Se um deslocamento tiver ocorrido primeiro para o ponto  $r$  do tempo atual, há uma boa chance de que  $(r, S_r)$  seja substancialmente maior do que alguns outros pontos na carta de controle, neste caso esses outros pontos estarão abaixo do braço inferior da máscara. Da mesma maneira, um deslocamento para um valor menor do que o alvo resultará subseqüentemente em pontos acima do braço superior da máscara.



**Figura 16.7** Marcações CUSUM: (a) pontos sucessivos  $(l, S_l)$  em uma carta de controle CUSUM, (b) uma máscara V com  $0 = (r, S_r)$ ; (c) um processo sob controle; (d) um processo fora de controle

Qualquer máscara V particular é determinada pela especificação da “distância”  $d$  e o “arco metade”  $\theta$ , ou, de maneira equivalente, especificando  $d$  e a altura  $h$  do segmento reto vertical de 0 até o braço inferior (ou superior) da máscara. Um método para decidir qual máscara utilizar envolve especificar o tamanho de um deslocamento na média do processo, que é de decisão particular do pesquisador. Então, os parâmetros da máscara são escolhidos para resultar nos valores desejados de  $\alpha$  e  $\beta$ , a probabilidade de falso alarme e a probabilidade de não detectar o deslocamento especificado, respectivamente. Um método alternativo compreende selecionar a máscara que produza os valores especificados de CM (comprimento de execução da média) ambos para um processo sob controle e para um processo no qual a média tenha se deslocado em uma quantidade especificada. Depois do desenvolvimento computacional do procedimento CUSUM, ilustraremos o segundo método de construção.

### Exemplo 16.8

Uma empresa de produtos de madeira produz carvão vegetal para churrasqueiras. O carvão é empacotado em sacos de vários tamanhos, o maior dos quais supõe-se que contenha 40 lb. A Tabela 16.4 mostra os pesos dos sacos de 16 amostras diferentes, cada uma delas com tamanho  $n = 4$ . Os 10 primeiros pesos foram obtidos de uma distribuição normal com  $\mu = \mu_0 = 40$  e  $\sigma = 0,5$ . Iniciando com a amostra de número 11, a média foi deslocada para cima para  $\mu = 40,3$ .

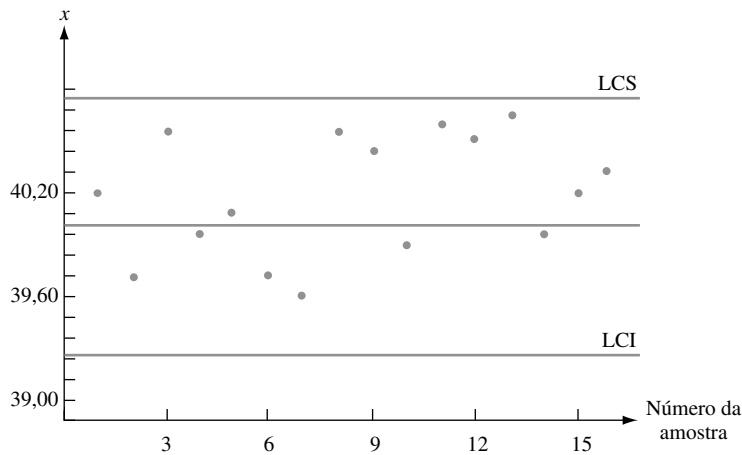


**Tabela 16.4** Observações,  $\bar{x}$ 's, e somas acumuladas para o Exemplo 16.8

Número da Amostra	Observações				$\bar{x}$	$\sum(\bar{x}_i - 40)$
1	40,77	39,95	40,86	39,21	40,20	0,20
2	38,94	39,70	40,37	39,88	39,72	-0,08
3	40,43	40,27	40,91	40,05	40,42	0,34
4	39,55	40,10	39,39	40,89	39,98	0,32
5	41,01	39,07	39,85	40,32	40,06	0,38
6	39,06	39,90	39,84	40,22	39,76	0,14
7	39,63	39,42	40,04	39,50	39,65	-0,21
8	41,05	40,74	40,43	39,40	40,41	0,20
9	40,28	40,89	39,61	40,48	40,32	0,52
10	39,28	40,49	38,88	40,72	39,84	0,36
11	40,57	40,04	40,85	40,51	40,49	0,85
12	39,90	40,67	40,51	40,53	40,40	1,25
13	40,70	40,54	40,73	40,45	40,61	1,86
14	39,58	40,90	39,62	39,83	39,98	1,84
15	40,16	40,69	40,37	39,69	40,23	2,07
16	40,46	40,21	40,09	40,58	40,34	2,41

A Figura 16.8 mostra um gráfico  $\bar{X}$  com reta central na altura 40 e limites de controle em

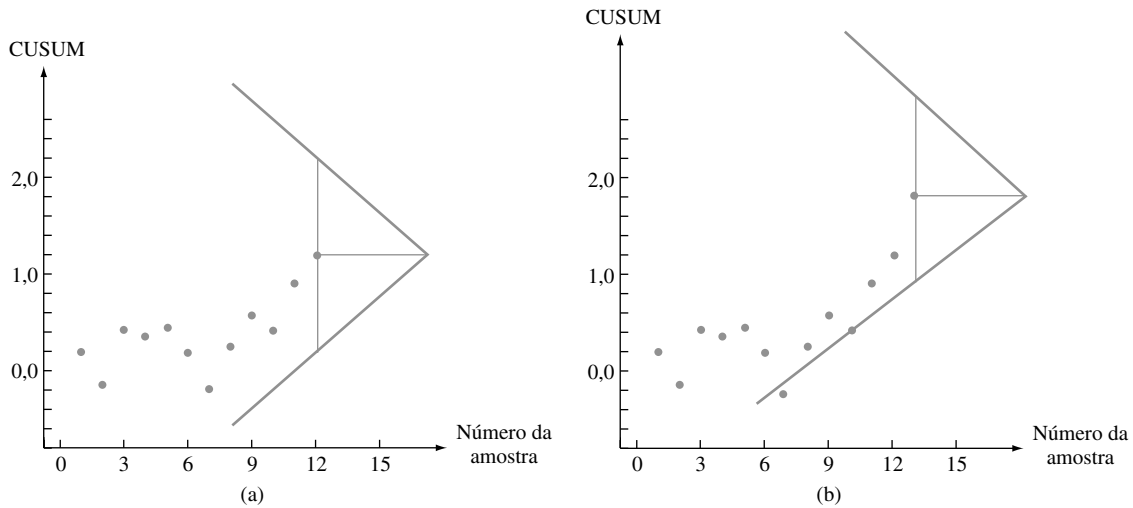
$$\mu_0 \pm 3\sigma_{\bar{x}} = 40 \pm 3 \cdot (0,5/\sqrt{4}) = 40 \pm 0,75$$



**Figura 16.8** Carta de controle  $\bar{X}$  para os dados do Exemplo 16.8

Nenhum ponto do gráfico fica fora dos limites de controle. Esse gráfico sugere um processo estável para o qual a média permaneceu no alvo.

A Figura 16.9 mostra as marcações CUSUM com uma máscara V particular sobreposta. A marcação na Figura 16.9(a) é o tempo atual  $r = 12$ . Todos os pontos nesta carta de controle ficam dentro dos braços da máscara. Entretanto, a marcação para  $r = 13$  mostrada na Figura 16.9(b) fornece um sinal fora de controle. O ponto que cai abaixo do braço inferior da máscara sugere um aumento no valor da média do processo. A máscara em  $r = 16$  é ainda mais enfática em sua mensagem fora de controle. Isto está em contraste marcante com o gráfico  $\bar{X}$  comum.



**Figura 16.9** Marcações CUSUM e máscaras V para dados do Exemplo 16.8: (a) máscara V no instante  $r = 12$ , processo sob controle; (b) máscara V no instante  $r = 13$ , sinal fora de controle

### Forma Computacional do Procedimento CUSUM

Descrevemos primeiro a versão computacional do procedimento CUSUM, então mostramos sua equivalência em relação à forma gráfica e finalmente discutimos como desenvolver um procedimento para alcançar o critério especificado.

Faça a hipótese de que  $d_0 = e_0 = 0$  e calcule  $d_1, d_2, d_3, \dots$  e  $e_1, e_2, e_3, \dots$  recursivamente utilizando as relações

$$d_l = \text{máx}[0, d_{l-1} + (\bar{x}_l - (\mu_0 + k))] \quad (l = 1, 2, 3, \dots)$$

$$e_l = \text{máx}[0, e_{l-1} - (\bar{x}_l - (\mu_0 - k))]$$

Nesse caso, o símbolo  $k$  indica o coeficiente angular do braço inferior da máscara V e seu valor é normalmente tomado como  $\Delta/2$  (onde  $\Delta$  é o tamanho de um deslocamento  $\mu$  no qual a atenção é concentrada).

Se no instante atual  $r$ , tanto  $d_r > h$  ou  $e_r > h$ , o processo é julgado como fora de controle. A primeira desigualdade sugere que a média do processo deslocou-se para um valor maior do que o alvo, ao passo que  $e_r > h$  indica um deslocamento para um valor menor.

#### Exemplo 16.9

Reconsidere os dados de carvão vegetal mostrados na Tabela 16.4 do Exemplo 16.8. O valor-alvo é  $\mu_0 = 40$  e o tamanho de um deslocamento para que seja rapidamente detectado é  $\Delta = 0,3$ . Assim

$$k = \frac{\Delta}{2} = 0,15 \quad \mu_0 + k = 40,15 \quad \mu_0 - k = 39,85$$

assim

$$d_l = \text{máx}[0, d_{l-1} + (\bar{x}_l - 40,15)]$$

$$e_l = \text{máx}[0, e_{l-1} - (\bar{x}_l - 39,85)]$$

Os cálculos dos primeiros  $d_i$ s ocorrem como a seguir:

$$\begin{aligned}
 d_0 &= 0 \\
 d_1 &= \text{máx}[0, d_0 + (\bar{x}_1 - 40,15)] \\
 &= \text{máx}[0, 0 + (40,20 - 40,15)] \\
 &= 0,05 \\
 d_2 &= \text{máx}[0, d_1 + (\bar{x}_2 - 40,15)] \\
 &= \text{máx}[0, 0,05 + (39,72 - 40,15)] \\
 &= 0 \\
 d_3 &= \text{máx}[0, d_2 + (\bar{x}_3 - 40,15)] \\
 &= \text{máx}[0, 0 + (40,42 - 40,15)] \\
 &= 0,27
 \end{aligned}$$

Os cálculos remanescentes estão resumidos na Tabela 16.5.

**Tabela 16.5** Cálculos de CUSUM para o Exemplo 16.9

Número da Amostra	$\bar{x}_i$	$\bar{x}_i - 40,15$	$d_i$	$\bar{x}_i - 39,85$	$e_i$
1	40,20	0,05	0,05	0,35	0
2	39,72	-0,43	0	-0,13	0,13
3	40,42	0,27	0,27	0,57	0
4	39,98	-0,17	0,10	0,13	0
5	40,06	-0,09	0,01	0,21	0
6	39,76	-0,39	0	-0,09	0,09
7	39,65	-0,50	0	-0,20	0,29
8	40,41	0,26	0,26	0,56	0
9	40,32	0,17	0,43	0,47	0
10	39,84	-0,31	0,12	-0,01	0,01
11	40,49	0,34	0,46	0,64	0
12	40,40	0,25	0,71	0,55	0
13	40,61	0,46	1,17	0,76	0
14	39,98	-0,17	1,00	0,13	0
15	40,23	0,08	1,08	0,38	0
16	40,34	0,19	1,27	0,49	0

O valor  $h = 0,95$  resulta em um procedimento CUSUM com propriedades desejáveis — alarmes falsos (sinais fora de controle incorretos) raramente ocorrem; contudo, um deslocamento de  $\Delta = 0,3$  será normalmente detectado de forma bastante rápida. Com este valor de  $h$ , o primeiro sinal fora de controle inicia-se depois da décima terceira amostra disponível. Como  $d_{13} = 1,17 > 0,95$ , parece que a média deslocou-se para um valor maior do que o alvo. Esta é a mesma mensagem que aquela fornecida pela máscara V na Figura 16.9(b). ■

### Equivalência da Máscara V e a Forma Computacional

Novamente, seja  $r$  o ponto no tempo atual, tal que  $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_r$  estejam disponíveis. A Figura 16.10 mostra uma máscara V com o ponto denominado 0 em  $(r, S_r)$ . O coeficiente angular do braço inferior, o qual indicamos por  $k$  é  $h/d$ . Assim, os pontos sobre o braço inferior acima de  $r, r - 1, r - 2, \dots$  são alturas  $S_r - h, S_r - h - k, S_r - h - 2k$  e assim por diante.

O processo está sob controle se todos os pontos estiverem sobre ou entre os braços da máscara. Desejamos descrever esta condição algebricamente. Considere

$$T_l = \sum_{i=1}^l [\bar{x}_i - (\mu_0 + k)] \quad l = 1, 2, 3, \dots, r$$

As condições sob as quais todos os pontos estão sobre ou acima do braço inferior são

$$\begin{array}{ll}
 S_r - h \leq S_r & \text{(trivialidade satisfeita)} \quad \text{ou seja, } S_r \leq S_r + h \\
 S_r - h - k \leq S_{r-1} & \text{ou seja, } S_r \leq S_{r-1} + h + k \\
 S_r - h - 2k \leq S_{r-2} & \text{ou seja, } S_r \leq S_{r-2} + h + 2k \\
 \cdot & \cdot \\
 \cdot & \cdot \\
 \cdot & \cdot
 \end{array}$$

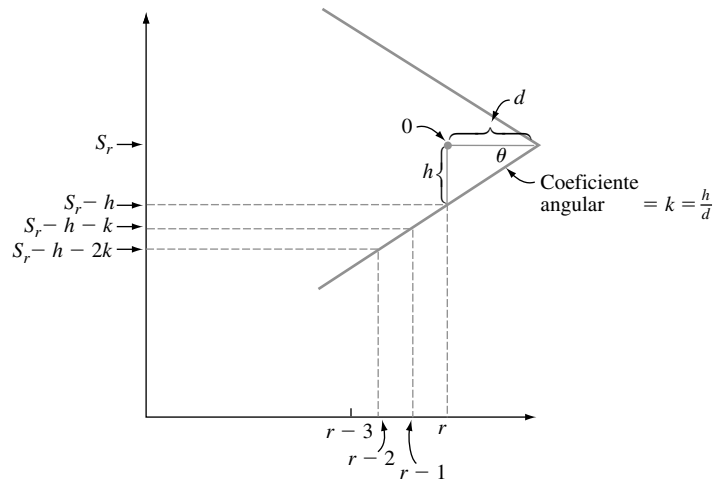


Figura 16.10 Uma máscara V com coeficiente angular de braço inferior = k

Agora, subtraia  $rk$  de ambos os lados de cada desigualdade para obter

$$\begin{array}{ll}
 S_r - rk \leq S_r - rk + h & \text{ou seja, } T_r \leq T_r + h \\
 S_r - rk \leq S_{r-1} - (r-1)k + h & \text{ou seja, } T_r \leq T_{r-1} + h \\
 S_r - rk \leq S_{r-2} - (r-2)k + h & \text{ou seja, } T_r \leq T_{r-2} + h \\
 \cdot & \cdot \\
 \cdot & \cdot \\
 \cdot & \cdot
 \end{array}$$

Assim, todos os pontos marcados ficam sobre ou acima do braço inferior se, e somente se  $T_r - T_r \leq h$ ,  $T_r - T_{r-1} \leq h$ ,  $T_r - T_{r-2} \leq h$ , e assim por diante. Isto é equivalente a

$$T_r - \text{mín}(T_1, T_2, \dots, T_r) \leq h$$

De maneira semelhante, se fizermos a hipótese

$$V_r = \sum_{i=1}^r [\bar{x}_i - (\mu_0 - k)] = S_r + rk$$

pode ser mostrado que todos os pontos ficam sobre ou abaixo do braço superior se e somente se máx

$$\text{máx}(V_1, \dots, V_r) - V_r \leq h$$

Se agora considerarmos

$$\begin{aligned}
 d_r &= T_r - \text{mín}(T_1, \dots, T_r) \\
 e_r &= \text{máx}(V_1, \dots, V_r) - V_r
 \end{aligned}$$

vemos facilmente que  $d_1, d_2, \dots$  e  $e_1, e_2, \dots$  podem ser calculados recursivamente como previamente ilustrado. Por exemplo, a expressão para  $d_r$ , segue da consideração de dois casos:

1.  $\text{mín}(T_1, \dots, T_r) = T_r$ , por isso  $d_r = 0$

2.  $\text{mín}(T_1, \dots, T_r) = \text{mín}(T_1, \dots, T_{r-1})$ , tal que

$$\begin{aligned} d_r &= T_r - \text{mín}(T_1, \dots, T_{r-1}) \\ &= \bar{x}_r - (\mu_0 + k) + T_{r-1} - \text{mín}(T_1, \dots, T_{r-1}) \\ &= \bar{x}_r - (\mu_0 + k) + d_{r-1} \end{aligned}$$

Como  $d_r$  não pode ser negativo, é a maior destas duas quantidades.

### Projetando um Procedimento CUSUM

Seja  $\Delta$  o tamanho de um deslocamento em  $\mu$  que seja rapidamente detectado utilizando um procedimento CUSUM.<sup>2</sup> É prática comum adotar  $k = \Delta/2$ . Suponha agora que um prático de controle de qualidade especifique valores desejados de dois comprimentos de execução de médias:

1. CM quando o processo está sob controle ( $\mu = \mu_0$ )
2. CM quando o processo está fora de controle porque a média foi deslocada de  $\Delta$  ( $\mu = \mu_0 + \Delta$  ou  $\mu = \mu_0 - \Delta$ )

Um gráfico desenvolvido por Kenneth Kemp (“The Use of Cumulative Sums for Sampling Inspection Schemes”, *Applied Statistics*, 1962, p. 23), chamado *nomograma*, pode então ser utilizado para determinar valores de  $h$  e  $n$  que alcancem os CMs<sup>3</sup> especificados. Este gráfico é mostrado na Figura 16.11. O método para sua utilização é descrito na caixa correspondente. O valor de  $\sigma$  deve ser conhecido ou uma estimativa é utilizada em seu lugar.

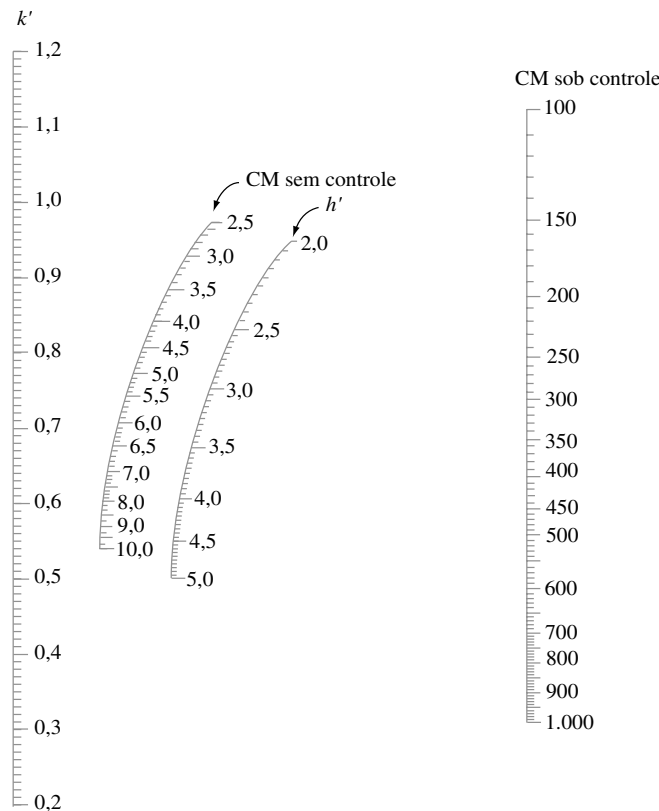


Figura 16.11 Nomograma de Kemp<sup>†</sup>

<sup>2</sup> Isto contrasta com a notação prévia, onde  $\Delta$  representou o número de desvios padrão pelo qual  $\mu$  foi modificado.

<sup>3</sup> A palavra *nomograma* não é específica para este gráfico; nomogramas são utilizados para várias outras finalidades.

<sup>†</sup> FONTE: KEMP, Kenneth. W., “The Use of Cumulative Sums for Sampling Inspection Schemes”, *Applied Statistics*, Vol. XI, p. 23, 1962. Com a permissão de Blackwell Publishing.

## Utilizando o Nomograma de Kemp

1. Localize os CMs desejados nas escalas sob controle e fora de controle. Conecte estes dois pontos com uma reta.
2. Observe onde a reta cruza a escala  $k'$  e resolva para  $n$  utilizando a equação

$$k' = \frac{\Delta/2}{\sigma/\sqrt{n}}$$

Então arredonde  $n$  para cima para o inteiro mais próximo.

3. Conecte o ponto sobre a escala  $k'$  com o ponto sobre a escala CM sob controle utilizando uma segunda reta e observe onde esta reta cruza a escala  $h'$ . Então  $h = (\sigma/\sqrt{n}) \cdot h'$ .

O valor  $h = 0,95$  foi utilizado no Exemplo 16.9. Naquela situação, segue-se que o CM sob controle é 500 e o CM fora de controle (para  $\Delta = 0,3$ ) é 7.

## Exemplo 16.10

O valor-alvo para o diâmetro do núcleo interior de uma bomba hidráulica é 2.250 pol. Se o desvio padrão do diâmetro do núcleo é  $\sigma = 0,004$ , qual procedimento CUSUM produzirá um CM sob controle de 500 e um CM de 5 quando o diâmetro do núcleo médio se desloca em uma quantidade igual a 0,003 pol?

Conectando o ponto 500 sobre uma escala CM sob controle ao ponto 5 sobre uma escala CM fora de controle e estendendo a reta para a escala  $k'$  sobre a esquerda na Figura 16.11 resulta em  $k' = 0,74$ . Assim

$$k' = 0,74 = \frac{\Delta/2}{\sigma/\sqrt{n}} = \frac{0,0015}{0,004/\sqrt{n}} = 0,375\sqrt{n}$$

de modo que

$$\sqrt{n} = \frac{0,74}{0,375} = 1,973 \quad n = (1,973)^2 = 3,894$$

O procedimento CUSUM deve portanto ser baseado em um tamanho da amostra  $n = 4$ . Conectando agora 0,74 sobre a escala  $k'$  a 500 sobre a escala CM sob controle resulta em  $h' = 3,2$ , de onde

$$h = (\sigma/\sqrt{n}) \cdot (3,2) = (0,004/\sqrt{4})(3,2) = 0,0064$$

Um sinal fora de controle resulta tanto em  $d_r > 0,0064$  como em  $e_r > 0,0064$ . ■

Discutimos os procedimentos CUSUM para a localização do processo de controle. Há também procedimentos CUSUM para a variação de processo de controle e para dados de atributo. As referências do Capítulo devem ser consultadas para obter-se informações sobre estes procedimentos.

---

**Exercícios** | Seção 16.5 (27–30)

27. Contêineres de um tratamento determinado para tanques sépticos devem conter 16 oz de líquido. Uma amostra de cinco contêineres é selecionada a partir de uma linha de produção uma vez a cada hora e o conteúdo da média amostral é determinado. Considere os seguintes resultados: 15,992; 16,051; 16,066; 15,912; 16,030; 16,060; 15,982; 15,899; 16,038; 16,074; 16,029; 15,935; 16,032; 15,960; 16,055. Utilizando  $\Delta = 0,10$  e  $h = 0,20$ , empregue a forma computacional do procedimento CUSUM para investigar o comportamento deste processo.
28. O valor-alvo para o diâmetro de um determinado tipo de eixo de motor é de 0,75 in. O tamanho do deslocamento

na média do diâmetro considerado importante para ser detectado é 0,002 in. Os diâmetros da média amostral para grupos sucessivos de  $n = 4$  eixos são os seguintes: 0,7507; 0,7504; 0,7492; 0,7501; 0,7503; 0,7510; 0,7490; 0,7497; 0,7488; 0,7504; 0,7516; 0,7472; 0,7489; 0,7483; 0,7471; 0,7498; 0,7460; 0,7482; 0,7470; 0,7493; 0,7462; 0,7481. Utilize a forma computacional do procedimento CUSUM com  $h = 0,003$  para verificar se a média do processo remanescente sobre o alvo atravessa o tempo de observação.

29. O desvio padrão de uma determinada dimensão de uma parte de um avião é 0,005 cm. Qual procedimento CUSUM fornece um CM sob controle de 600 e um CM fora de controle de 4 quando o valor médio da dimensão varia em 0,004 cm?

30. Quando o CM fora de controle corresponde a um deslocamento de 1 desvio padrão na média do processo, quais são as características do procedimento CUSUM que tem CMs de 250 e 4,8 para as condições sob controle e fora de controle, respectivamente?

## 16.6 | Amostragem de Aceitação

Itens procedentes de um processo de produção são freqüentemente enviados em grupos a outra empresa ou estabelecimento comercial. Um grupo poderá consistir de todas as unidades de uma produção particular executada ou deslocada, em um contêiner de navio de algum tipo, enviado em resposta a um pedido particular e assim por diante. O grupo de itens é normalmente chamado de *lote*, o remetente é denominado *produtor* e o destinatário do lote é o *consumidor*. Nosso objetivo será observar situações nas quais cada item é tanto defeituoso como não-defeituoso, com  $p$  indicando a proporção das unidades com defeito em um lote. O consumidor deverá naturalmente querer aceitar o lote somente se o valor de  $p$  for aceitavelmente pequeno. A amostragem de aceitação significa que parte da estatística aplicada lida com métodos para decidir se o consumidor deve aceitar ou rejeitar um lote.

Até pouco tempo, os procedimentos de gráfico de controle e técnicas de amostragem de aceitação eram considerados pelos praticantes como partes igualmente importantes da metodologia de controle de qualidade. Este não é o caso. A razão é que o uso de gráficos de controle e outras estratégias desenvolvidas recentemente oferecem a oportunidade de projetar a qualidade em um produto, enquanto a amostragem de aceitação está de acordo com o que já foi produzido e assim não fornece qualquer controle direto sobre a qualidade do processo. Isto remete ao conhecedor de controle de qualidade norte-americana, W. E. Deming, a principal força em persuadir os japoneses a fazerem uso substancial da metodologia de controle de qualidade, a questionar fortemente contra o uso da amostragem de aceitação em várias situações. Em uma tendência semelhante, o livro recente escrito por Ryan (veja a Bibliografia do capítulo) dedica vários capítulos aos gráficos de controle e menciona a amostragem de aceitação somente de passagem. Como um reflexo desta ênfase, nos contentaremos aqui com uma breve apresentação dos conceitos básicos. Mais informações podem ser encontradas em várias referências da Bibliografia do capítulo.

### Planos de Amostragem Única

O tipo mais direto de plano de amostragem de aceitação compreende a seleção de uma única amostra aleatória de tamanho  $n$  e a rejeição do lote se o número de defeitos na amostra exceder um valor crítico especificado  $c$ . Seja a va  $X$  o número de itens com defeito em um lote e por  $A$  o evento em que o lote é aceito. Então  $P(A) = P(X \leq c)$  é uma função de  $p$ ; quanto maior o valor de  $p$ , menor será a probabilidade de aceitação do lote.

Se o tamanho da amostra  $n$  é relativamente grande para  $N$ ,  $P(A)$  é calculado utilizando a distribuição hipergeométrica (o número de defeitos no lote é  $Np$ ):

$$P(X \leq c) = \sum_{x=0}^c h(x; n, Np, N) = \sum_{x=0}^c \frac{\binom{Np}{x} \cdot \binom{N(1-p)}{n-x}}{\binom{N}{n}}$$

Quando  $n$  é pequeno em relação a  $N$  (a regra prática sugerida previamente foi  $n \leq 0,05N$ , mas alguns autores empregam a regra menos conservadora  $n \leq 0,10N$ ), a distribuição binomial pode ser utilizada:

$$P(X \leq c) = \sum_{x=0}^c b(x; n, p) = \sum_{x=0}^c \binom{n}{x} p^x (1-p)^{n-x}$$

Finalmente, se  $P(A)$  é grande somente quando  $p$  é pequeno (isto depende do valor de  $c$ ), a aproximação de Poisson para a distribuição binomial é justificada:

$$P(X \leq c) \approx \sum_{x=0}^c p(x; np) = \sum_{x=0}^c \frac{e^{-np}(np)^x}{x!}$$

O comportamento de um plano de amostragem pode ser melhor resumido construindo-se o gráfico  $P(A)$  como uma função de  $p$ . Este gráfico é chamado de **curva característica da operação (CO)** para o plano.

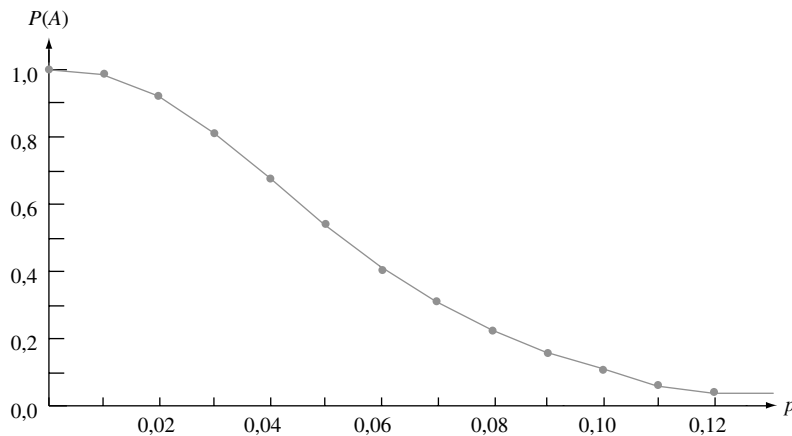
**Exemplo 16.11**

Considere o plano de amostragem com valor crítico  $c = 2$  e tamanho da amostra  $n = 50$  e suponha que o tamanho do lote  $N$  exceda 1.000, portanto, a distribuição binomial pode ser utilizada. Isso fornece

$$P(A) = P(X \leq 2) = (1 - p)^{50} + 50p(1 - p)^{49} + 1.255p^2(1 - p)^{48}$$

A tabela a seguir mostra  $P(A)$  para os valores selecionados de  $p$  e a curva característica de operação (CO) correspondente é mostrada na Figura 16.12.

$p$	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09	0,10	0,12	0,15
$P(A)$	0,986	0,922	0,811	0,677	0,541	0,416	0,311	0,226	0,161	0,112	0,051	0,014



**Figura 16.12** A Curva CO para plano de amostragem com  $c = 2$  e  $n = 50$

A curva CO para o plano do Exemplo 16.11 tem  $P(A)$  próxima a 1 para  $p$  muito próximo a 0. Entretanto, em muitas aplicações, uma taxa de defeito de 8% [para a qual  $P(A) = 0,226$ ] ou ainda 5% [ $P(A) = 0,541$ ] deveria ser considerada excessiva, caso em que as probabilidades de aceitação são muito altas. Aumentando-se o valor crítico  $c$  enquanto se mantém  $n$  fixo resulta em um plano para o qual  $P(A)$  aumenta em cada  $p$  (exceto 0 e 1), portanto a nova curva CO localiza-se acima da anterior. Isto é desejável para  $p$  próximo a 0, mas não para valores de  $p$  maiores. Manter  $c$  constante enquanto aumenta-se  $n$  fornece uma curva CO mais baixa, o que é bom para  $p$  grande, mas não para  $p$  próximo a 0. Queremos uma curva CO que seja mais alta para  $p$  muito pequeno e mais baixa para  $p$  grande. Isto pode ser alcançado aumentando-se  $n$  e ajustando  $c$ .



## Projetando um Plano de Amostra Única

Um plano de amostragem efetivo tem as seguintes características:

1. Tem uma probabilidade alta especificada de aceitar lotes que o produtor considera de boa qualidade.
2. Tem uma probabilidade baixa especificada de aceitar lotes que o consumidor considera de baixa qualidade.

Um plano deste tipo pode ser desenvolvido procedendo-se como a seguir. Vamos designar dois valores diferentes de  $p$ , um para o qual  $P(A)$  é um valor especificado próximo a 1, e o outro para o qual  $P(A)$  é um valor especificado próximo a 0. Esses dois valores de  $p$ , digamos  $p_1$  e  $p_2$ , são freqüentemente chamados **nível de qualidade aceitável (NQA)** e a **porcentagem de tolerância a defeitos do lote (PTDL)**. Ou seja, exigimos um plano para o qual

1.  $P(A) = 1 - \alpha$  quando  $p = p_1 = \text{NQA}$  ( $\alpha$  pequeno)
2.  $P(A) = \beta$  quando  $p = p_2 = \text{PTDL}$  ( $\beta$  pequeno)

Isto é análogo a verificar um procedimento de teste com a probabilidade de erro tipo I especificado  $\alpha$ , e a probabilidade de erro tipo II especificado  $\beta$  quando testada a hipótese. Por exemplo, devemos ter

$$\begin{array}{lll} \text{NQA} = 0,01 & \alpha = 0,05 & (P(A) = 0,95) \\ \text{PTDL} = 0,045 & \beta = 0,10 & (P(A) = 0,10) \end{array}$$

Como  $X$  é discreto, devemos normalmente estar satisfeitos com valores de  $n$  e  $c$  que atendam aproximadamente estas condições.

A Tabela 16.6 fornece informações a partir das quais  $n$  e  $c$  podem ser determinados no caso  $\alpha = 0,05$ ,  $\beta = 0,10$ .

**Tabela 16.6** Fatores para determinação de  $n$  e  $c$  para um plano de amostra única com  $\alpha = 0,05$ ,  $\beta = 0,10$

$c$	$np_1$	$np_2$	$p_2/p_1$	$c$	$np_1$	$np_2$	$p_2/p_1$
0	0,051	2,30	45,10	8	4,695	12,99	2,77
1	0,355	3,89	10,96	9	5,425	14,21	2,62
2	0,818	5,32	6,50	10	6,169	15,41	2,50
3	1,366	6,68	4,89	11	6,924	16,60	2,40
4	1,970	7,99	4,06	12	7,690	17,78	2,31
5	2,613	9,28	3,55	13	8,464	18,86	2,24
6	3,285	10,53	3,21	14	9,246	20,13	2,18
7	3,981	11,77	2,96	15	10,040	21,29	2,12

### Exemplo 16.12

Vamos determinar um plano para o qual  $\text{NQA} = p_1 = 0,01$  e  $\text{PTDL} = p_2 = 0,045$ . A razão de  $p_2$  para  $p_1$  é

$$\frac{\text{PTDL}}{\text{NQA}} = \frac{p_2}{p_1} = \frac{0,045}{0,01} = 4,50$$

Este valor localiza-se entre a razão 4,89, dada na Tabela 16.6, para a qual  $c = 3$ , e 4,06, para a qual  $c = 4$ . Uma vez que um desses valores de  $c$  seja escolhido,  $n$  pode ser determinado tanto dividindo o valor de  $np_1$  na Tabela 16.6 por  $p_1$  quanto por  $np_2/p_2$ . Assim, quatro planos diferentes (dois valores de  $c$  e para cada dois valores de  $n$ ) resultam em aproximadamente o valor especificado de  $\alpha$  e  $\beta$ . Considere, por exemplo, o uso de  $c = 3$  e

$$n = \frac{np_1}{p_1} = \frac{1,366}{0,01} = 136,6 \approx 137$$

Então

$$\begin{aligned}\alpha &= 1 - P(X \leq 3 \text{ quando } p = p_1) \\ &= 1 - \sum_{x=0}^3 \binom{137}{x} (0,01)^x (0,99)^{137-x} = 0,050\end{aligned}$$

(a aproximação de Poisson com  $\lambda = 1,37$  também resulta em 0,050) e

$$\beta = P(X \leq 3 \text{ quando } p = p_2) = 0,131$$

O plano com  $c = 4$  e  $n$  determinado a partir de  $np_2 = 7,99$  tem  $n = 178$ ,  $\alpha = 0,034$  e  $\beta = 0,094$ . O tamanho maior da amostra resulta em um plano com ambos  $\alpha$  e  $\beta$  menores do que os valores especificados correspondentes. ■

O livro de Douglas Montgomery citado na Bibliografia do capítulo contém um gráfico a partir do qual  $c$  e  $n$  podem ser determinados para quaisquer  $\alpha$  e  $\beta$  especificados.

Pode acontecer que o número de itens defeituosos na amostra chegue a  $c + 1$  antes de todos os itens terem sido examinados. Por exemplo, no caso  $c = 3$  e  $n = 137$ , pode ser que o item de número 125 examinado seja o quarto item com defeito, portanto, os 12 remanescentes não precisam ser examinados. Entretanto, geralmente recomenda-se que todos os itens sejam examinados mesmo quando isto ocorre de modo a fornecer um histórico de qualidade lote a lote e estimativas de  $p$  todo o tempo.

## Planos de Amostragem Dupla

Em um plano de amostragem dupla, o número de itens com defeito,  $x$ , em uma amostra inicial de tamanho  $n$  é determinado. Há então três possíveis cursos de ação: imediatamente aceitar o lote, imediatamente rejeitar o lote ou tirar uma segunda amostra de  $n_2$  itens e rejeitar ou aceitar o lote dependendo do número total  $x_1 + x_2$  de itens com defeito nas duas amostras. Apesar dos dois tamanhos da amostra, um plano específico é caracterizado por mais três números:  $c_1$ ,  $r_1$  e  $c_2$ , como segue:

1. Rejeita o lote se  $x_1 \geq r_1$ .
2. Aceita o lote se  $x_1 \leq c_1$ .
3. Se  $c_1 < x_1 < r_1$ , tirar uma segunda amostra; então aceita o lote se  $x_1 + x_2 \leq c_2$  e o rejeita, caso contrário.

### Exemplo 16.13

Considere o plano de amostragem dupla com  $n_1 = 80$ ,  $n_2 = 80$ ,  $c_1 = 2$ ,  $r_1 = 5$  e  $c_2 = 6$ . Assim o lote será aceito se (1)  $x_1 = 0, 1$  ou  $2$ ; (2)  $x_1 = 3$  e  $x_2 = 0, 1, 2$  ou  $3$ ; ou (3)  $x_1 = 4$  e  $x_2 = 0, 1$  ou  $2$ .

Sob a hipótese de que o tamanho do lote seja grande o suficiente para a aproximação binomial ser aplicada, a probabilidade  $P(A)$  de aceitação do lote é

$$\begin{aligned}P(A) &= P(X_1 = 0, 1 \text{ ou } 2) + P(X_1 = 3, X_2 = 0, 1, 2 \text{ ou } 3) \\ &\quad + P(X_1 = 4, X_2 = 0, 1 \text{ ou } 2) \\ &= \sum_{x_1=0}^2 b(x_1; 80, p) + b(3; 80, p) \sum_{x_2=0}^3 b(x_2; 80, p) \\ &\quad + b(4; 80, p) \sum_{x_2=0}^2 b(x_2; 80, p)\end{aligned}$$

Novamente, o gráfico de  $P(A)$  versus  $p$  é a curva CO do plano. A curva CO para este plano aparece na Figura 16.13.

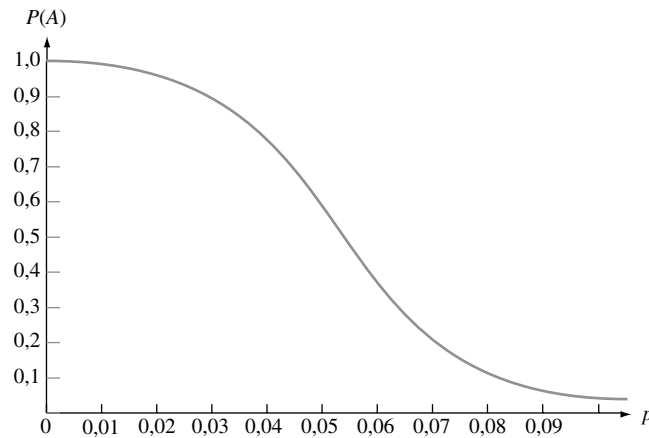


Figura 16.13 Curva CO para o plano de amostragem dupla do Exemplo 16.13

Um método-padrão para projetar um plano de amostragem dupla envolve proceder como sugerido anteriormente para planos de amostragem simples. Especifique valores  $p_1$  e  $p_2$  com as probabilidades de aceitação correspondentes  $1 - \alpha$  e  $\beta$ . Então encontre um plano que satisfaça estas condições. O livro de Montgomery fornece tabelas semelhantes à Tabela 16.6 para este propósito nos casos  $n_2 = n_1$  e  $n_2 = 2n_1$  com  $1 - \alpha = 0,95$ ,  $\beta = 0,10$ . Tabulações de plano muito mais extensas estão disponíveis em outras fontes.

Da mesma forma que na prática-padrão com planos de amostragem simples, recomenda-se que todos os itens na primeira amostra sejam examinados mesmo quando o item defeituoso de número  $(r_1 + 1)$  seja descoberto antes da inspeção do  $n_1$ -ésimo item. Entretanto, é comum terminar a inspeção da segunda amostra se o número de defeitos for suficiente para justificar a rejeição antes de todos os itens terem sido examinados. Isto é chamado de redução na segunda amostra. Sob *redução*, pode ser mostrado que o número esperado de itens inspecionados em um plano de amostragem dupla é menor do que o número de itens examinados em um plano de amostragem única quando as curvas CO dos dois planos são muito parecidas. Essa é a maior virtude dos planos de amostragem dupla. Para obter mais informações sobre essas questões, assim como uma discussão de vários planos de amostragem múltiplos e sequenciais (o que compreende a seleção de itens para inspeção um por um em vez de em grupos), um livro sobre controle de qualidade deve ser consultado.

### Inspeção de Retificação e Outros Critérios de Projeto

Em algumas situações, a inspeção da amostragem é realizada utilizando *retificação*. Para planos de amostragem simples, isso significa que cada item com defeito na amostra é substituído por um bom, e se o número de defeitos na amostra exceder o limite  $c$  de aceitação, *todos* os itens no lote são examinados e os itens bons são substituídos por quaisquer que estejam defeituosos. Seja  $N$  o tamanho do lote. Uma característica importante de um plano de amostragem com inspeção de retificação é a **qualidade média resultante**, indicada por **QMR**. Esta é a proporção de execução longa de itens defeituosos entre aqueles enviados depois de o plano de amostragem ser empregado. Agora, os defeitos ocorrerão somente entre os  $N - n$  itens não inspecionados em um lote julgado aceitável sobre as bases de uma amostra. Suponha, por exemplo, que  $P(A) = P(X \leq c) = 0,985$  quando  $p = 0,01$ . Então na execução longa, 98,5% dos  $N - n$  itens que não estão na amostra não serão inspecionados, dos quais esperamos que 1% esteja com defeito. Isto implica que o número esperado de defeitos em um lote selecionado aleatoriamente seja  $(N - n) \cdot P(A) \cdot p = 0,00985(N - n)$ . Dividindo essa expressão pelo número de itens em um lote resulta na qualidade média resultante:

$$\begin{aligned} \text{QMR} &= \frac{(N - n) \cdot P(A) \cdot p}{N} \\ &\approx P(A) \cdot p \quad \text{se } N \gg n \end{aligned}$$

Como  $\text{QMR} = 0$  quando tanto  $p = 0$  quanto  $p = 1$  [ $P(A) = 0$  no caso anterior], a seguir há um valor de  $p$  entre 0 e 1 para o qual o QMR é um máximo. O valor máximo de QMR é chamado **limite de qualidade**

**média resultante, LQMR.** Por exemplo, para o plano com  $n = 137$  e  $c = 3$  discutido anteriormente, o LQMR = 0,0142, o valor de QMR em  $p \approx 0,02$ .

Escolhas apropriadas de  $n$  e  $c$  produzirão um plano de amostragem para o qual o LQMR é um número pequeno especificado. Tal plano não é, entretanto, único, portanto outra condição pode ser imposta. Frequentemente, esta segunda condição compreenderá o **número médio do total inspecionado** (ou seja, esperada), indicado por **MTI**. O número de itens inspecionados em um lote escolhido aleatoriamente é uma variável aleatória que assume o valor  $n$  com probabilidade  $P(A)$ , e  $N$  com probabilidade  $1 - P(A)$ . Assim, o número esperado de itens inspecionados em um lote selecionado aleatoriamente é

$$\text{MTI} = n \cdot P(A) + N \cdot (1 - P(A))$$

É uma prática comum selecionar um plano de amostragem que tem um LQMR especificado e, além disso, um MTI mínimo em um nível de qualidade particular  $p$ .

## Planos de Amostragem-Padrão

Pode parecer que a determinação de um plano de amostragem que satisfaça simultaneamente vários critérios seja um tanto difícil. Felizmente, outros já deixaram a maior parte do trabalho pronto na forma de tabulações extensivas de tais planos. MIL STD 105D desenvolvido pelos militares depois da Segunda Guerra Mundial é o mais utilizado conjunto de planos. Uma versão civil, ANSI/ASQC Z1.4, é bastante semelhante à versão militar. Um terceiro conjunto de planos que é bastante popular foi desenvolvido pela Bell Laboratories antes da Segunda Guerra Mundial por dois estatísticos aplicados chamados Dodge e Romig. O livro de Montgomery (veja a Bibliografia do capítulo) contém uma introdução legível para o uso destes planos.

---

### Exercícios | Seção 16.6 (31–38)

31. Considere o plano de amostragem única com  $c = 2$  e  $n = 50$  como discutido no Exemplo 16.11, mas agora suponha que o tamanho do lote seja  $N = 500$ . Calcule  $P(A)$ , a probabilidade de aceitação do lote, para  $p = 0,01; 0,02; \dots; 0,10$  utilizando a distribuição hipergeométrica. A aproximação binomial fornece resultados satisfatórios neste caso?
32. Uma amostra de 50 itens deve ser selecionada de um lote que consiste de 5.000 itens. O lote será aceito se a amostra contiver no máximo um item com defeito. Calcule a probabilidade de o lote ser aceito para  $p = 0,01, 0,02, \dots, 0,10$  e esboce a curva CO.
33. Consulte o Exercício 32 e considere o plano com  $n = 100$  e  $c = 2$ . Calcule  $P(A)$  para  $p = 0,01, 0,02, \dots, 0,05$ , e esboce as duas curvas CO no mesmo conjunto de eixos. Qual dos dois planos é preferível (desconsiderando o custo da amostragem) e por quê?
34. Desenvolva um plano de amostragem única para o qual QMR = 0,02 e LQMR = 0,07 no caso  $\alpha = 0,05$ ,  $\beta = 0,10$ . Uma vez determinados os valores de  $n$  e  $c$ , calcule os valores encontrados de  $\alpha$  e  $\beta$  para o plano.
35. Considere o plano de amostragem dupla para o qual ambos os tamanhos da amostra são 50. O lote é aceito depois da primeira amostra se o número de defeitos é no máximo 1, rejeitado se o número de defeitos for no mínimo 4 e rejeitado depois da segunda amostra se o número total de defeitos for 6 ou mais. Calcule a probabilidade de aceitação do lote quando  $p = 0,01, 0,05$  e  $0,10$ .
36. Algumas fontes afirmam que um tipo ainda mais restritivo de plano de amostragem dupla no qual  $r_1 = c_2 + 1$ ; ou seja, o lote é rejeitado se em qualquer estágio o número (total) de defeitos é pelo menos  $r_1$  (veja o livro de Montgomery). Considere este tipo de plano de amostragem com  $n_1 = 50$ ,  $n_2 = 100$ ,  $c_1 = 1$  e  $r_1 = 4$ . Calcule a probabilidade de aceitação do lote quando  $p = 0,02, 0,05$  e  $0,10$ .
37. Consulte o Exemplo 16.11 no qual um plano de amostragem única com  $n = 50$  e  $c = 2$  foi empregado.
  - a. Calcule QMR para  $p = 0,01, 0,02, \dots, 10$ . O que isto sugere sobre o valor de  $p$  para o qual o QMR é um máximo e o LQMR correspondente?
  - b. Determine o valor de  $p$  para o qual o QMR é um máximo e o valor corresponde de LQMR. (*Sugestão:* Utilize o cálculo.)
  - c. Para  $N = 2.000$ , calcule o MTI para os valores de  $p$  dados na parte (a).
38. Considere um plano de amostra única que utiliza  $n = 50$  e  $c = 1$  quando  $N = 2.000$ . Determine os valores de QMR e MTI para valores selecionados de  $p$  e construa o gráfico de cada um deles contra  $p$ . Determine também o valor de LQMR.

Exercícios Suplementares (39–44)

- 39.** As observações sobre a força de atrito para 26 grupos de teste de pontos de solda, cada um consistindo de seis soldas, produz  $\sum \bar{x}_i = 10.980$ ,  $\sum s_i = 402$ , e  $\sum r_i = 1.074$ . Calcule os limites de controle para quaisquer gráficos de controle relevantes.
- 40.** O número de arranhões na superfície de cada uma das 24 chapas de metal retangulares é determinado, produzindo os seguintes dados: 8, 1, 7, 5, 2, 0, 2, 3, 4, 3, 1, 2, 5, 7, 3, 4, 6, 5, 2, 4, 0, 10, 2, 6. Construa um gráfico de controle apropriado e comente.
- 41.** Os números seguintes são observações da tensão de amostras de tecido sintético selecionados de um processo de produção a intervalos de tempo iguais. Construa os gráficos de controle apropriados e comente (faça a hipótese de que uma causa assinalável é identificável para quaisquer observações fora de controle).
- |                |      |      |                |      |      |
|----------------|------|------|----------------|------|------|
| <b>a.</b> 51,3 | 51,7 | 49,5 | <b>m.</b> 49,6 | 48,4 | 50,0 |
| <b>b.</b> 51,0 | 50,0 | 49,3 | <b>n.</b> 49,8 | 51,2 | 49,7 |
| <b>c.</b> 50,8 | 51,1 | 49,0 | <b>o.</b> 50,4 | 49,9 | 50,7 |
| <b>d.</b> 50,6 | 51,1 | 49,0 | <b>p.</b> 49,4 | 49,5 | 49,0 |
| <b>e.</b> 49,6 | 50,5 | 50,9 | <b>q.</b> 50,7 | 49,0 | 50,0 |
| <b>f.</b> 51,3 | 52,0 | 50,3 | <b>r.</b> 50,8 | 49,5 | 50,9 |
| <b>g.</b> 49,7 | 50,5 | 50,3 | <b>s.</b> 48,5 | 50,3 | 49,3 |
| <b>h.</b> 51,8 | 50,3 | 50,0 | <b>t.</b> 49,6 | 50,6 | 49,4 |
| <b>i.</b> 48,6 | 50,5 | 50,7 | <b>u.</b> 50,9 | 49,4 | 49,7 |
| <b>j.</b> 49,6 | 49,8 | 50,5 | <b>v.</b> 54,1 | 49,8 | 48,5 |
| <b>l.</b> 49,9 | 50,7 | 49,8 | <b>x.</b> 50,2 | 49,6 | 51,5 |
- 42.** Uma alternativa ao gráfico  $p$  para um defeito de fração é o gráfico  $np$  para o número de defeitos. Este gráfico tem  $LCS = n\bar{p} + 3\sqrt{n\bar{p}(1 - \bar{p})}$ ,  $LCI = n\bar{p} - 3\sqrt{n\bar{p}(1 - \bar{p})}$ , e o número de defeitos de cada amostra é marcado no gráfico. Construa um gráfico para os dados do Exemplo 16.6. O uso do gráfico  $np$  fornecerá sempre a mesma mensagem que a do uso de um gráfico  $p$  (ou seja, os dois gráficos são equivalentes)?
- 43.** Observações de resistência (ohms) para subgrupos de um determinado tipo de registro forneceram o seguinte resumo de quantidades:

$i$	$n_i$	$\bar{x}_i$	$s_i$	$i$	$n_i$	$\bar{x}_i$	$s_i$
1	4	430,0	22,5	7	4	420,8	25,4
2	4	418,2	20,6	8	4	431,4	24,0
3	3	435,5	25,1	9	4	428,7	21,2
4	4	427,6	22,3	10	4	440,1	25,8
5	4	444,0	21,5	11	4	445,2	27,3
6	3	431,4	28,9	12	4	430,1	22,2
13	4	427,2	24,0	17	3	447,0	19,8
14	4	439,6	23,3	18	4	434,4	23,7
15	3	415,9	31,2	19	4	422,2	25,1
16	4	419,8	27,5	20	4	425,7	24,4

Construa limites de controle apropriados. [Sugestão: Utilize  $\bar{x} = \sum n_i \bar{x}_i / \sum n_i$  e  $s^2 = \sum (n_i - 1) s_i^2 / \sum (n_i - 1)$ .

- 44.** Seja  $\alpha$  um número entre 0 e 1 e defina a seqüência  $W_1, W_2, W_3, \dots$  por  $W_0 = \mu$  e  $W_t = \alpha \bar{X}_t + (1 - \alpha)W_{t-1}$  para  $t = 1, 2, \dots$  Substituindo por  $W_{t-1}$  sua representação em termos de  $\bar{X}_{t-1}$  e  $W_{t-2}$  então substituindo por  $W_{t-2}$  e assim por diante, resulta em

$$W_t = \alpha \bar{X}_t + \alpha(1 - \alpha) \bar{X}_{t-1} + \dots + \alpha(1 - \alpha)^{t-1} \bar{X}_1 + (1 - \alpha)^t \mu$$

O fato de que  $W_t$  depende não apenas de  $\bar{X}_t$ , mas também das médias de pontos anteriores de tempo, apesar de que com decréscimos de peso (exponencialmente), sugere que mudanças na média do processo serão mais rapidamente refletidas nos  $W_t$ s do que nos  $\bar{X}_t$ s individuais.

- a.** Mostre que  $E(W_t) = \mu$ .  
**b.** Considere  $\sigma_t^2 = V(W_t)$ , e mostre que

$$\sigma_t^2 = \frac{\alpha[1 - (1 - \alpha)^{2t}]}{2 - \alpha} \cdot \frac{\sigma^2}{n}$$

- c.** Um gráfico de controle de média móvel ponderada exponencialmente marca os  $W_t$ s e utiliza limites de controle  $\mu_0 \pm 3\sigma_t$  (ou  $\bar{x}$  no lugar de  $\mu_0$ ). Construa um gráfico para os dados do Exemplo 16.9 utilizando  $\mu_0 = 40$ .

Bibliografia

BOX, George, Soren Bisgaard e Conrad Fung, "An Explanation and Critique of Taguchi's Contributions to Quality Engineering", *Quality and Reliability Engineering*. International, 1988, p. 123-131.

MONTGOMERY, Douglas C., *Introduction to Statistical Quality Control*. Wiley, Nova York: Wiley 1996 (3. ed.). Esta é uma introdução abrangente para os vários aspectos do controle de qualidade no mesmo nível deste livro.

RYAN, Thomas P., *Statistical Methods for Quality Improvement*. Nova York: Wiley 1989. Capta muito bem o sabor moderno do controle de qualidade, exigindo pré-requisitos mínimos dos leitores.

VARDEMAN, Stephen B. e J. Marcus Jobe, *Statistical Quality Assurance Methods for Engineers*. Nova York: Wiley, 1999. Inclui tópicos tradicionais de qualidade e também material de planejamento de experimento alemão para questões de qualidade; informal e conceitualmente correto.

# **Tabelas do Apêndice**

**Tabela A.1** Probabilidades Binomiais Acumuladas

$$B(x; n, p) = \sum_{y=0}^x b(y; n, p)$$

**a. n = 5**

		<i>p</i>														
		0,01	0,05	0,10	0,20	0,25	0,30	0,40	0,50	0,60	0,70	0,75	0,80	0,90	0,95	0,99
<i>x</i>	0	0,951	0,774	0,590	0,328	0,237	0,168	0,078	0,031	0,010	0,002	0,001	0,000	0,000	0,000	0,000
	1	0,999	0,977	0,919	0,737	0,633	0,528	0,337	0,188	0,087	0,031	0,016	0,007	0,000	0,000	0,000
	2	1,000	0,999	0,991	0,942	0,896	0,837	0,683	0,500	0,317	0,163	0,104	0,058	0,009	0,001	0,000
	3	1,000	1,000	1,000	0,993	0,984	0,969	0,913	0,812	0,663	0,472	0,367	0,263	0,081	0,023	0,001
	4	1,000	1,000	1,000	1,000	0,999	0,998	0,990	0,969	0,922	0,832	0,763	0,672	0,410	0,226	0,049

**b. n = 10**

		<i>p</i>														
		0,01	0,05	0,10	0,20	0,25	0,30	0,40	0,50	0,60	0,70	0,75	0,80	0,90	0,95	0,99
<i>x</i>	0	0,904	0,599	0,349	0,107	0,056	0,028	0,006	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	1	0,996	0,914	0,736	0,376	0,244	0,149	0,046	0,011	0,002	0,000	0,000	0,000	0,000	0,000	0,000
	2	1,000	0,988	0,930	0,678	0,526	0,383	0,167	0,055	0,012	0,002	0,000	0,000	0,000	0,000	0,000
	3	1,000	0,999	0,987	0,879	0,776	0,650	0,382	0,172	0,055	0,011	0,004	0,001	0,000	0,000	0,000
	4	1,000	1,000	0,998	0,967	0,922	0,850	0,633	0,377	0,166	0,047	0,020	0,006	0,000	0,000	0,000
	5	1,000	1,000	1,000	0,994	0,980	0,953	0,834	0,623	0,367	0,150	0,078	0,033	0,002	0,000	0,000
	6	1,000	1,000	1,000	0,999	0,996	0,989	0,945	0,828	0,618	0,350	0,224	0,121	0,013	0,001	0,000
	7	1,000	1,000	1,000	1,000	1,000	0,998	0,988	0,945	0,833	0,617	0,474	0,322	0,070	0,012	0,000
	8	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,989	0,954	0,851	0,756	0,624	0,264	0,086	0,004
	9	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,994	0,972	0,944	0,893	0,651	0,401	0,096

**c. n = 15**

		<i>p</i>														
		0,01	0,05	0,10	0,20	0,25	0,30	0,40	0,50	0,60	0,70	0,75	0,80	0,90	0,95	0,99
<i>x</i>	0	0,860	0,463	0,206	0,035	0,013	0,005	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	1	0,990	0,829	0,549	0,167	0,080	0,035	0,005	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	2	1,000	0,964	0,816	0,398	0,236	0,127	0,027	0,004	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	3	1,000	0,995	0,944	0,648	0,461	0,297	0,091	0,018	0,002	0,000	0,000	0,000	0,000	0,000	0,000
	4	1,000	0,999	0,987	0,836	0,686	0,515	0,217	0,059	0,009	0,001	0,000	0,000	0,000	0,000	0,000
	5	1,000	1,000	0,998	0,939	0,852	0,722	0,402	0,151	0,034	0,004	0,001	0,000	0,000	0,000	0,000
	6	1,000	1,000	1,000	0,982	0,943	0,869	0,610	0,304	0,095	0,015	0,004	0,001	0,000	0,000	0,000
	7	1,000	1,000	1,000	0,996	0,983	0,950	0,787	0,500	0,213	0,050	0,017	0,004	0,000	0,000	0,000
	8	1,000	1,000	1,000	0,999	0,996	0,985	0,905	0,696	0,390	0,131	0,057	0,018	0,000	0,000	0,000
	9	1,000	1,000	1,000	1,000	0,999	0,996	0,966	0,849	0,597	0,278	0,148	0,061	0,002	0,000	0,000
	10	1,000	1,000	1,000	1,000	1,000	0,999	0,991	0,941	0,783	0,485	0,314	0,164	0,013	0,001	0,000
	11	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,982	0,909	0,703	0,539	0,352	0,056	0,005	0,000
	12	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,996	0,973	0,873	0,764	0,602	0,184	0,036	0,000
	13	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,995	0,965	0,920	0,833	0,451	0,171	0,010
	14	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,995	0,987	0,965	0,794	0,537	0,140

(continuação)

**Tabela A.1** Probabilidades Binomiais Acumuladas (cont.)

$$B(x; n, p) = \sum_{y=0}^x b(y; n, p)$$

**d.  $n = 20$**

		<i>p</i>														
		0,01	0,05	0,10	0,20	0,25	0,30	0,40	0,50	0,60	0,70	0,75	0,80	0,90	0,95	0,99
<i>x</i>	0	0,818	0,358	0,122	0,012	0,003	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	1	0,983	0,736	0,392	0,069	0,024	0,008	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,999	0,925	0,677	0,206	0,091	0,035	0,004	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	3	1,000	0,984	0,867	0,411	0,225	0,107	0,016	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	4	1,000	0,997	0,957	0,630	0,415	0,238	0,051	0,006	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	5	1,000	1,000	0,989	0,804	0,617	0,416	0,126	0,021	0,002	0,000	0,000	0,000	0,000	0,000	0,000
	6	1,000	1,000	0,998	0,913	0,786	0,608	0,250	0,058	0,006	0,000	0,000	0,000	0,000	0,000	0,000
	7	1,000	1,000	1,000	0,968	0,898	0,772	0,416	0,132	0,021	0,001	0,000	0,000	0,000	0,000	0,000
	8	1,000	1,000	1,000	0,990	0,959	0,887	0,596	0,252	0,057	0,005	0,001	0,000	0,000	0,000	0,000
	9	1,000	1,000	1,000	0,997	0,986	0,952	0,755	0,412	0,128	0,017	0,004	0,001	0,000	0,000	0,000
	10	1,000	1,000	1,000	0,999	0,996	0,983	0,872	0,588	0,245	0,048	0,014	0,003	0,000	0,000	0,000
	11	1,000	1,000	1,000	1,000	0,999	0,995	0,943	0,748	0,404	0,113	0,041	0,010	0,000	0,000	0,000
	12	1,000	1,000	1,000	1,000	1,000	0,999	0,979	0,868	0,584	0,228	0,102	0,032	0,000	0,000	0,000
	13	1,000	1,000	1,000	1,000	1,000	1,000	0,994	0,942	0,750	0,392	0,214	0,087	0,002	0,000	0,000
	14	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,979	0,874	0,584	0,383	0,196	0,011	0,000	0,000
	15	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,994	0,949	0,762	0,585	0,370	0,043	0,003	0,000
	16	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,984	0,893	0,775	0,589	0,133	0,016	0,000
	17	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,996	0,965	0,909	0,794	0,323	0,075	0,001
	18	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,992	0,976	0,931	0,608	0,264	0,017
	19	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,997	0,988	0,878	0,642	0,182

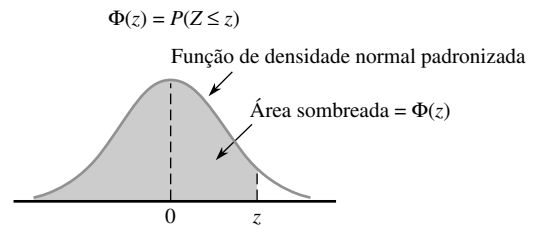
**e.  $n = 25$**

		<i>p</i>														
		0,01	0,05	0,10	0,20	0,25	0,30	0,40	0,50	0,60	0,70	0,75	0,80	0,90	0,95	0,99
<i>x</i>	0	0,778	0,277	0,072	0,004	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	1	0,974	0,642	0,271	0,027	0,007	0,002	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,998	0,873	0,537	0,098	0,032	0,009	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	3	1,000	0,966	0,764	0,234	0,096	0,033	0,002	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	4	1,000	0,993	0,902	0,421	0,214	0,090	0,009	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	5	1,000	0,999	0,967	0,617	0,378	0,193	0,029	0,002	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	6	1,000	1,000	0,991	0,780	0,561	0,341	0,074	0,007	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	7	1,000	1,000	0,998	0,891	0,727	0,512	0,154	0,022	0,001	0,000	0,000	0,000	0,000	0,000	0,000
	8	1,000	1,000	1,000	0,953	0,851	0,677	0,274	0,054	0,004	0,000	0,000	0,000	0,000	0,000	0,000
	9	1,000	1,000	1,000	0,983	0,929	0,811	0,425	0,115	0,013	0,000	0,000	0,000	0,000	0,000	0,000
	10	1,000	1,000	1,000	0,994	0,970	0,902	0,586	0,212	0,034	0,002	0,000	0,000	0,000	0,000	0,000
	11	1,000	1,000	1,000	0,998	0,980	0,956	0,732	0,345	0,078	0,006	0,001	0,000	0,000	0,000	0,000
	12	1,000	1,000	1,000	1,000	0,997	0,983	0,846	0,500	0,154	0,017	0,003	0,000	0,000	0,000	0,000
	13	1,000	1,000	1,000	1,000	0,999	0,994	0,922	0,655	0,268	0,044	0,020	0,002	0,000	0,000	0,000
	14	1,000	1,000	1,000	1,000	1,000	0,998	0,966	0,788	0,414	0,098	0,030	0,006	0,000	0,000	0,000
	15	1,000	1,000	1,000	1,000	1,000	1,000	0,987	0,885	0,575	0,189	0,071	0,017	0,000	0,000	0,000
	16	1,000	1,000	1,000	1,000	1,000	1,000	0,996	0,946	0,726	0,323	0,149	0,047	0,000	0,000	0,000
	17	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,978	0,846	0,488	0,273	0,109	0,002	0,000	0,000
	18	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,993	0,926	0,659	0,439	0,220	0,009	0,000	0,000
	19	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,971	0,807	0,622	0,383	0,033	0,001	0,000
	20	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,991	0,910	0,786	0,579	0,098	0,007	0,000
	21	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,967	0,904	0,766	0,236	0,034	0,000
	22	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,991	0,968	0,902	0,463	0,127	0,002
	23	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,993	0,973	0,729	0,358	0,026
	24	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,996	0,928	0,723	0,222





**Tabela A.3** Área sob a Curva Normal Padronizada



<i>z</i>	<b>0,00</b>	<b>0,01</b>	<b>0,02</b>	<b>0,03</b>	<b>0,04</b>	<b>0,05</b>	<b>0,06</b>	<b>0,07</b>	<b>0,08</b>	<b>0,09</b>
-3,4	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0002
-3,3	0,0005	0,0005	0,0005	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004	0,0003
-3,2	0,0007	0,0007	0,0006	0,0006	0,0006	0,0006	0,0006	0,0005	0,0005	0,0005
-3,1	0,0010	0,0009	0,0009	0,0009	0,0008	0,0008	0,0008	0,0008	0,0007	0,0007
-3,0	0,0013	0,0013	0,0013	0,0012	0,0012	0,0011	0,0011	0,0011	0,0010	0,0010
-2,9	0,0019	0,0018	0,0017	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014
-2,8	0,0026	0,0025	0,0024	0,0023	0,0023	0,0022	0,0021	0,0021	0,0020	0,0019
-2,7	0,0035	0,0034	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026
-2,6	0,0047	0,0045	0,0044	0,0043	0,0041	0,0040	0,0039	0,0038	0,0037	0,0036
-2,5	0,0062	0,0060	0,0059	0,0057	0,0055	0,0054	0,0052	0,0051	0,0049	0,0038
-2,4	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0064
-2,3	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0084
-2,2	0,0139	0,0136	0,0132	0,0129	0,0125	0,0122	0,0119	0,0116	0,0113	0,0110
-2,1	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143
-2,0	0,0228	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0183
-1,9	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0239	0,0233
-1,8	0,0359	0,0352	0,0344	0,0336	0,0329	0,0322	0,0314	0,0307	0,0301	0,0294
-1,7	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367
-1,6	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455
-1,5	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0571	0,0559
-1,4	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0722	0,0708	0,0694	0,0681
-1,3	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823
-1,2	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0985
-1,1	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1170
-1,0	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379
-0,9	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611
-0,8	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	0,1867
-0,7	0,2420	0,2389	0,2358	0,2327	0,2296	0,2266	0,2236	0,2206	0,2177	0,2148
-0,6	0,2743	0,2709	0,2676	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2451
-0,5	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2843	0,2810	0,2776
-0,4	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121
-0,3	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3482
-0,2	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859
-0,1	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247
-0,0	0,5000	0,4960	0,4920	0,4880	0,4840	0,4801	0,4761	0,4721	0,4681	0,4641
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879

(continuação)

**Tabela A.3** Área sob a Curva Normal Padronizada (cont.)

$$\Phi(z) = P(Z \leq z)$$

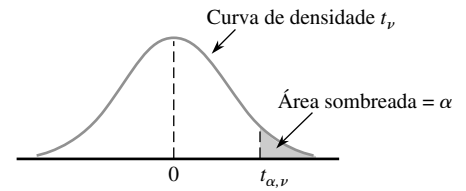
z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9278	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998

**Tabela A.4** Função Gama Incompleta

$$F(x; \alpha) = \int_0^x \frac{1}{\Gamma(\alpha)} y^{\alpha-1} e^{-y} dy$$

x \ α	1	2	3	4	5	6	7	8	9	10
1	0,632	0,264	0,080	0,019	0,004	0,001	0,000	0,000	0,000	0,000
2	0,865	0,594	0,323	0,143	0,053	0,017	0,005	0,001	0,000	0,000
3	0,950	0,801	0,577	0,353	0,185	0,084	0,034	0,012	0,004	0,001
4	0,982	0,908	0,762	0,567	0,371	0,215	0,111	0,051	0,021	0,008
5	0,993	0,960	0,875	0,735	0,560	0,384	0,238	0,133	0,068	0,032
6	0,998	0,983	0,938	0,849	0,715	0,554	0,394	0,256	0,153	0,084
7	0,999	0,993	0,970	0,918	0,827	0,699	0,550	0,401	0,271	0,170
8	1,000	0,997	0,986	0,958	0,900	0,809	0,687	0,547	0,407	0,283
9		0,999	0,994	0,979	0,945	0,884	0,793	0,676	0,544	0,413
10		1,000	0,997	0,990	0,971	0,933	0,870	0,780	0,667	0,542
11			0,999	0,995	0,985	0,962	0,921	0,857	0,768	0,659
12			1,000	0,998	0,992	0,980	0,954	0,911	0,845	0,758
13				0,999	0,996	0,989	0,974	0,946	0,900	0,834
14				1,000	0,998	0,994	0,986	0,968	0,938	0,891
15					0,999	0,997	0,992	0,982	0,963	0,930

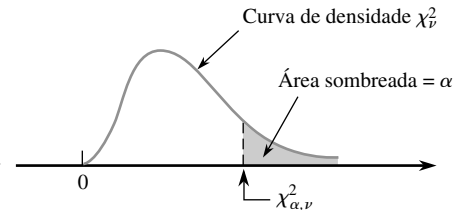
**Tabela A.5** Valores Críticos para as Distribuições  $t$



$\nu$	$\alpha$						
	0,10	0,05	0,025	0,01	0,005	0,001	0,0005
1	3,078	6,314	12,706	31,821	63,657	318,31	636,62
2	1,886	2,920	4,303	6,965	9,925	22,326	31,598
3	1,638	2,353	3,182	4,541	5,841	10,213	12,924
4	1,533	2,132	2,776	3,747	4,604	7,173	8,610
5	1,476	2,015	2,571	3,365	4,032	5,893	6,869
6	1,440	1,943	2,447	3,143	3,707	5,208	5,959
7	1,415	1,895	2,365	2,998	3,499	4,785	5,408
8	1,397	1,860	2,306	2,896	3,355	4,501	5,041
9	1,383	1,833	2,262	2,821	3,250	4,297	4,781
10	1,372	1,812	2,228	2,764	3,169	4,144	4,587
11	1,363	1,796	2,201	2,718	3,106	4,025	4,437
12	1,356	1,782	2,179	2,681	3,055	3,930	4,318
13	1,350	1,771	2,160	2,650	3,012	3,852	4,221
14	1,345	1,761	2,145	2,624	2,977	3,787	4,140
15	1,341	1,753	2,131	2,602	2,947	3,733	4,073
16	1,337	1,746	2,120	2,583	2,921	3,686	4,015
17	1,333	1,740	2,110	2,567	2,898	3,646	3,965
18	1,330	1,734	2,101	2,552	2,878	3,610	3,922
19	1,328	1,729	2,093	2,539	2,861	3,579	3,883
20	1,325	1,725	2,086	2,528	2,845	3,552	3,850
21	1,323	1,721	2,080	2,518	2,831	3,527	3,819
22	1,321	1,717	2,074	2,508	2,819	3,505	3,792
23	1,319	1,714	2,069	2,500	2,807	3,485	3,767
24	1,318	1,711	2,064	2,492	2,797	3,467	3,745
25	1,316	1,708	2,060	2,485	2,787	3,450	3,725
26	1,315	1,706	2,056	2,479	2,779	3,435	3,707
27	1,314	1,703	2,052	2,473	2,771	3,421	3,690
28	1,313	1,701	2,048	2,467	2,763	3,408	3,674
29	1,311	1,699	2,045	2,462	2,756	3,396	3,659
30	1,310	1,697	2,042	2,457	2,750	3,385	3,646
32	1,309	1,694	2,037	2,449	2,738	3,365	3,622
34	1,307	1,691	2,032	2,441	2,728	3,348	3,601
36	1,306	1,688	2,028	2,434	2,719	3,333	3,582
38	1,304	1,686	2,024	2,429	2,712	3,319	3,566
40	1,303	1,684	2,021	2,423	2,704	3,307	3,551
50	1,299	1,676	2,009	2,403	2,678	3,262	3,496
60	1,296	1,671	2,000	2,390	2,660	3,232	3,460
120	1,289	1,658	1,980	2,358	2,617	3,160	3,373
$\infty$	1,282	1,645	1,960	2,326	2,576	3,090	3,291



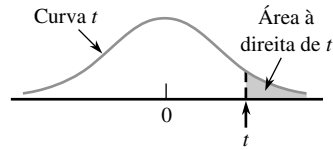
**Tabela A.7** Valores Críticos para Distribuições Qui-Quadrado



$\nu$	$\alpha$									
	0,995	0,99	0,975	0,95	0,90	0,10	0,05	0,025	0,01	0,005
1	0,000	0,000	0,001	0,004	0,016	2,706	3,843	5,025	6,637	7,882
2	0,010	0,020	0,051	0,103	0,211	4,605	5,992	7,378	9,210	10,597
3	0,072	0,115	0,216	0,352	0,584	6,251	7,815	9,348	11,344	12,837
4	0,207	0,297	0,484	0,711	1,064	7,779	9,488	11,143	13,277	14,860
5	0,412	0,554	0,831	1,145	1,610	9,236	11,070	12,832	15,085	16,748
6	0,676	0,872	1,237	1,635	2,204	10,645	12,592	14,440	16,812	18,548
7	0,989	1,239	1,690	2,167	2,833	12,017	14,067	16,012	18,474	20,276
8	1,344	1,646	2,180	2,733	3,490	13,362	15,507	17,534	20,090	21,954
9	1,735	2,088	2,700	3,325	4,168	14,684	16,919	19,022	21,665	23,587
10	2,156	2,558	3,247	3,940	4,865	15,987	18,307	20,483	23,209	25,188
11	2,603	3,053	3,816	4,575	5,578	17,275	19,675	21,920	24,724	26,755
12	3,074	3,571	4,404	5,226	6,304	18,549	21,026	23,337	26,217	28,300
13	3,565	4,107	5,009	5,892	7,041	19,812	22,362	24,735	27,687	29,817
14	4,075	4,660	5,629	6,571	7,790	21,064	23,685	26,119	29,141	31,319
15	4,600	5,229	6,262	7,261	8,547	22,307	24,996	27,488	30,577	32,799
16	5,142	5,812	6,908	7,962	9,312	23,542	26,296	28,845	32,000	34,267
17	5,697	6,407	7,564	8,682	10,085	24,769	27,587	30,190	33,408	35,716
18	6,265	7,015	8,231	9,390	10,865	25,989	28,869	31,526	34,805	37,156
19	6,843	7,632	8,906	10,117	11,651	27,203	30,143	32,852	36,190	38,580
20	7,434	8,260	9,591	10,851	12,443	28,412	31,410	34,170	37,566	39,997
21	8,033	8,897	10,283	11,591	13,240	29,615	32,670	35,478	38,930	41,399
22	8,643	9,542	10,982	12,338	14,042	30,813	33,924	36,781	40,289	42,796
23	9,260	10,195	11,688	13,090	14,848	32,007	35,172	38,075	41,637	44,179
24	9,886	10,856	12,401	13,848	15,659	33,196	36,415	39,364	42,980	45,558
25	10,519	11,523	13,120	14,611	16,473	34,381	37,652	40,646	44,313	46,925
26	11,160	12,198	13,844	15,379	17,292	35,563	38,885	41,923	45,642	48,290
27	11,807	12,878	14,573	16,151	18,114	36,741	40,113	43,194	46,962	49,642
28	12,461	13,565	15,308	16,928	18,939	37,916	41,337	44,461	48,278	50,993
29	13,120	14,256	16,147	17,708	19,768	39,087	42,557	45,772	49,586	52,333
30	13,787	14,954	16,791	18,493	20,599	40,256	43,773	46,979	50,892	53,672
31	14,457	15,655	17,538	19,280	21,433	41,422	44,985	48,231	52,190	55,000
32	15,134	16,362	18,291	20,072	22,271	42,585	46,194	49,480	53,486	56,328
33	15,814	17,073	19,046	20,866	23,110	43,745	47,400	50,724	54,774	57,646
34	16,501	17,789	19,806	21,664	23,952	44,903	48,602	51,966	56,061	58,964
35	17,191	18,508	20,569	22,465	24,796	46,059	49,802	53,203	57,340	60,272
36	17,887	19,233	21,336	23,269	25,643	47,212	50,998	54,437	58,619	61,581
37	18,584	19,960	22,105	24,075	26,492	48,363	52,192	55,667	59,891	62,880
38	19,289	20,691	22,878	24,884	27,343	49,513	53,384	56,896	61,162	64,181
39	19,994	21,425	23,654	25,695	28,196	50,660	54,572	58,119	62,426	65,473
40	20,706	22,164	24,433	26,509	29,050	51,805	55,758	59,342	63,691	66,766

Para  $\nu > 40$ ,  $\chi^2_{\alpha,\nu} \approx \nu \left( 1 - \frac{2}{9\nu} + z_\alpha \sqrt{\frac{2}{9\nu}} \right)^3$

**Tabela A.8** Áreas de Cauda da Curva  $t$



$t$	$\nu$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
0.0		0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,500	0,500
0.1		0,468	0,465	0,463	0,463	0,462	0,462	0,462	0,461	0,461	0,461	0,461	0,461	0,461	0,461	0,461	0,461	0,461	0,461
0.2		0,437	0,430	0,427	0,426	0,425	0,424	0,424	0,423	0,423	0,423	0,423	0,422	0,422	0,422	0,422	0,422	0,422	0,422
0.3		0,407	0,396	0,392	0,390	0,388	0,387	0,386	0,386	0,386	0,385	0,385	0,385	0,384	0,384	0,384	0,384	0,384	0,384
0.4		0,379	0,364	0,358	0,355	0,353	0,352	0,351	0,350	0,349	0,349	0,348	0,348	0,348	0,347	0,347	0,347	0,347	0,347
0.5		0,352	0,333	0,326	0,322	0,319	0,317	0,316	0,315	0,315	0,314	0,313	0,313	0,313	0,312	0,312	0,312	0,312	0,312
0.6		0,328	0,305	0,295	0,290	0,287	0,285	0,284	0,283	0,282	0,281	0,280	0,280	0,279	0,279	0,279	0,278	0,278	0,278
0.7		0,306	0,278	0,267	0,261	0,258	0,255	0,253	0,252	0,251	0,250	0,249	0,249	0,248	0,247	0,247	0,247	0,247	0,246
0.8		0,285	0,254	0,241	0,234	0,230	0,227	0,225	0,223	0,222	0,221	0,220	0,220	0,219	0,218	0,218	0,218	0,217	0,217
0.9		0,267	0,232	0,217	0,210	0,205	0,201	0,199	0,197	0,196	0,195	0,194	0,193	0,192	0,191	0,191	0,191	0,190	0,190
1.0		0,250	0,211	0,196	0,187	0,182	0,178	0,175	0,173	0,172	0,170	0,169	0,169	0,168	0,167	0,167	0,166	0,166	0,165
1.1		0,235	0,193	0,176	0,167	0,162	0,157	0,154	0,152	0,150	0,149	0,147	0,146	0,146	0,144	0,144	0,144	0,143	0,143
1.2		0,221	0,177	0,158	0,148	0,142	0,138	0,135	0,132	0,130	0,129	0,128	0,127	0,126	0,124	0,124	0,124	0,123	0,123
1.3		0,209	0,162	0,142	0,132	0,125	0,121	0,117	0,115	0,113	0,111	0,110	0,109	0,108	0,107	0,107	0,106	0,105	0,105
1.4		0,197	0,148	0,128	0,117	0,110	0,106	0,102	0,100	0,098	0,096	0,095	0,093	0,092	0,091	0,091	0,090	0,090	0,089
1.5		0,187	0,136	0,115	0,104	0,097	0,092	0,089	0,086	0,084	0,082	0,081	0,080	0,079	0,077	0,077	0,077	0,076	0,075
1.6		0,178	0,125	0,104	0,092	0,085	0,080	0,077	0,074	0,072	0,070	0,069	0,068	0,067	0,065	0,065	0,065	0,064	0,064
1.7		0,169	0,116	0,094	0,082	0,075	0,070	0,065	0,064	0,062	0,060	0,059	0,057	0,056	0,055	0,055	0,054	0,054	0,053
1.8		0,161	0,107	0,085	0,073	0,066	0,061	0,057	0,055	0,053	0,051	0,050	0,049	0,048	0,046	0,046	0,045	0,045	0,044
1.9		0,154	0,099	0,077	0,065	0,058	0,053	0,050	0,047	0,045	0,043	0,042	0,041	0,040	0,038	0,038	0,038	0,037	0,037
2.0		0,148	0,092	0,070	0,058	0,051	0,046	0,043	0,040	0,038	0,037	0,035	0,034	0,033	0,032	0,032	0,031	0,031	0,030
2.1		0,141	0,085	0,063	0,052	0,045	0,040	0,037	0,034	0,033	0,031	0,030	0,029	0,028	0,027	0,027	0,026	0,025	0,025
2.2		0,136	0,079	0,058	0,046	0,040	0,035	0,032	0,029	0,028	0,026	0,025	0,024	0,023	0,022	0,022	0,021	0,021	0,021
2.3		0,131	0,074	0,052	0,041	0,035	0,031	0,027	0,025	0,023	0,022	0,021	0,020	0,019	0,018	0,018	0,018	0,017	0,017
2.4		0,126	0,069	0,048	0,037	0,031	0,027	0,024	0,022	0,020	0,019	0,018	0,017	0,016	0,015	0,015	0,014	0,014	0,014
2.5		0,121	0,065	0,044	0,033	0,027	0,023	0,020	0,018	0,017	0,016	0,015	0,014	0,013	0,012	0,012	0,012	0,011	0,011
2.6		0,117	0,061	0,040	0,030	0,024	0,020	0,018	0,016	0,014	0,013	0,012	0,012	0,011	0,010	0,010	0,010	0,009	0,009
2.7		0,113	0,057	0,037	0,027	0,021	0,018	0,015	0,014	0,012	0,011	0,010	0,010	0,009	0,008	0,008	0,008	0,008	0,007
2.8		0,109	0,054	0,034	0,024	0,019	0,016	0,013	0,012	0,010	0,009	0,009	0,008	0,008	0,007	0,007	0,006	0,006	0,006
2.9		0,106	0,051	0,031	0,022	0,017	0,014	0,011	0,010	0,009	0,008	0,007	0,007	0,006	0,005	0,005	0,005	0,005	0,005
3.0		0,102	0,048	0,029	0,020	0,015	0,012	0,010	0,009	0,007	0,007	0,006	0,006	0,005	0,004	0,004	0,004	0,004	0,004
3.1		0,099	0,045	0,027	0,018	0,013	0,011	0,009	0,007	0,006	0,006	0,005	0,005	0,004	0,004	0,004	0,003	0,003	0,003
3.2		0,096	0,043	0,025	0,016	0,012	0,009	0,008	0,006	0,005	0,005	0,004	0,004	0,003	0,003	0,003	0,003	0,003	0,002
3.3		0,094	0,040	0,023	0,015	0,011	0,008	0,007	0,005	0,005	0,004	0,004	0,003	0,003	0,002	0,002	0,002	0,002	0,002
3.4		0,091	0,038	0,021	0,014	0,010	0,007	0,006	0,005	0,004	0,003	0,003	0,003	0,002	0,002	0,002	0,002	0,002	0,002
3.5		0,089	0,036	0,020	0,012	0,009	0,006	0,005	0,004	0,003	0,003	0,002	0,002	0,002	0,002	0,002	0,001	0,001	0,001
3.6		0,086	0,035	0,018	0,011	0,008	0,006	0,004	0,004	0,003	0,002	0,002	0,002	0,002	0,001	0,001	0,001	0,001	0,001
3.7		0,084	0,033	0,017	0,010	0,007	0,005	0,004	0,003	0,002	0,002	0,002	0,002	0,001	0,001	0,001	0,001	0,001	0,001
3.8		0,082	0,031	0,016	0,010	0,006	0,004	0,003	0,003	0,002	0,002	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001
3.9		0,080	0,030	0,015	0,009	0,006	0,004	0,003	0,002	0,002	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001
4.0		0,078	0,029	0,014	0,008	0,005	0,004	0,003	0,002	0,002	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,000	0,000

(continuação)





Tabela A.9 Valores Críticos para Distribuições F

$\alpha$		$\nu_1 = \text{gl do numerador}$								
		1	2	3	4	5	6	7	8	9
1	0,100	39,86	49,50	53,59	55,83	57,24	58,20	58,91	59,44	59,86
	0,050	161,45	199,50	215,71	224,58	230,16	233,99	236,77	238,88	240,54
	0,010	4052,2	4999,5	5403,4	5624,6	5763,6	5859,0	5928,4	5981,1	6022,5
	0,001	405284	500000	540379	562500	576405	585937	592873	598144	602284
2	0,100	8,53	9,00	9,16	9,24	9,29	9,33	9,35	9,37	9,38
	0,050	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38
	0,010	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39
	0,001	998,50	999,00	999,17	999,25	999,30	999,33	999,36	999,37	999,39
3	0,100	5,54	5,46	5,39	5,34	5,31	5,28	5,27	5,25	5,24
	0,050	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81
	0,010	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35
	0,001	167,03	148,50	141,11	137,10	134,58	132,85	131,58	130,62	129,86
4	0,100	4,54	4,32	4,19	4,11	4,05	4,01	3,98	3,95	3,94
	0,050	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00
	0,010	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66
	0,001	74,14	61,25	56,18	53,44	51,71	50,53	49,66	49,00	48,47
5	0,100	4,06	3,78	3,62	3,52	3,45	3,40	3,37	3,34	3,32
	0,050	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77
	0,010	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16
	0,001	47,18	37,12	33,20	31,09	29,75	28,83	28,16	27,65	27,24
6	0,100	3,78	3,46	3,29	3,18	3,11	3,05	3,01	2,98	2,96
	0,050	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10
	0,010	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98
	0,001	35,51	27,00	23,70	21,92	20,80	20,03	19,46	19,03	18,69
7	0,100	3,59	3,26	3,07	2,96	2,88	2,83	2,78	2,75	2,72
	0,050	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68
	0,010	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72
	0,001	29,25	21,69	18,77	17,20	16,21	15,52	15,02	14,63	14,33
8	0,100	3,46	3,11	2,92	2,81	2,73	2,67	2,62	2,59	2,56
	0,050	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39
	0,010	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91
	0,001	25,41	18,49	15,83	14,39	13,48	12,86	12,40	12,05	11,77
9	0,100	3,36	3,01	2,81	2,69	2,61	2,55	2,51	2,47	2,44
	0,050	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18
	0,010	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35
	0,001	22,86	16,39	13,90	12,56	11,71	11,13	10,70	10,37	10,11
10	0,100	3,29	2,92	2,73	2,61	2,52	2,46	2,41	2,38	2,35
	0,050	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02
	0,010	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94
	0,001	21,04	14,91	12,55	11,28	10,48	9,93	9,52	9,20	8,96
11	0,100	3,23	2,86	2,66	2,54	2,45	2,39	2,34	2,30	2,27
	0,050	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90
	0,010	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63
	0,001	19,69	13,81	11,56	10,35	9,58	9,05	8,66	8,35	8,12
12	0,100	3,18	2,81	2,61	2,48	2,39	2,33	2,28	2,24	2,21
	0,050	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80
	0,010	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39
	0,001	18,64	12,97	10,80	9,63	8,89	8,38	8,00	7,71	7,48

(continuação)

**Tabela A.9** Valores Críticos para Distribuições  $F$  (cont.)

$\nu_1 = \text{gl do numerador}$										
10	12	15	20	25	30	40	50	60	120	1000
60,19	60,71	61,22	61,74	62,05	62,26	62,53	62,69	62,79	63,06	63,30
241,88	243,91	245,95	248,01	249,26	250,10	251,14	251,77	252,20	253,25	254,19
6055,8	6106,3	6157,3	6208,7	6239,8	6260,6	6286,8	6302,5	6313,0	6339,4	6362,7
605621	610668	615764	620908	624017	626099	628712	630285	631337	633972	636301
9,39	9,41	9,42	9,44	9,45	9,46	9,47	9,47	9,47	9,48	9,49
19,40	19,41	19,43	19,45	19,46	19,46	19,47	19,48	19,48	19,49	19,49
99,40	99,42	99,43	99,45	99,46	99,47	99,47	99,48	99,48	99,49	99,50
999,40	999,42	999,43	999,45	999,46	999,47	999,47	999,48	999,48	999,49	999,50
5,23	5,22	5,20	5,18	5,17	5,17	5,16	5,15	5,15	5,14	5,13
8,79	8,74	8,70	8,66	8,63	8,62	8,59	8,58	8,57	8,55	8,53
27,23	27,05	26,87	26,69	26,58	26,50	26,41	26,35	26,32	26,22	26,14
129,25	128,32	127,37	126,42	125,84	125,45	124,96	124,66	124,47	123,97	123,53
3,92	3,90	3,87	3,84	3,83	3,82	3,80	3,80	3,79	3,78	3,76
5,96	5,91	5,86	5,80	5,77	5,75	5,72	5,70	5,69	5,66	5,63
14,55	14,37	14,20	14,02	13,91	13,84	13,75	13,69	13,65	13,56	13,47
48,05	47,41	46,76	46,10	45,70	45,43	45,09	44,88	44,75	44,40	44,09
3,30	3,27	3,24	3,21	3,19	3,17	3,16	3,15	3,14	3,12	3,11
4,74	4,68	4,62	4,56	4,52	4,50	4,46	4,44	4,43	4,40	4,37
10,05	9,89	9,72	9,55	9,45	9,38	9,29	9,24	9,20	9,11	9,03
26,92	26,42	25,91	25,39	25,08	24,87	24,60	24,44	24,33	24,06	23,82
2,94	2,90	2,87	2,84	2,81	2,80	2,78	2,77	2,76	2,74	2,72
4,06	4,00	3,94	3,87	3,83	3,81	3,77	3,75	3,74	3,70	3,67
7,87	7,72	7,56	7,40	7,30	7,23	7,14	7,09	7,06	6,97	6,89
18,41	17,99	17,56	17,12	16,85	16,67	16,44	16,31	16,21	15,98	15,77
2,70	2,67	2,63	2,59	2,57	2,56	2,54	2,52	2,51	2,49	2,47
3,64	3,57	3,51	3,44	3,40	3,38	3,34	3,32	3,30	3,27	3,23
6,62	6,47	6,31	6,16	6,06	5,99	5,91	5,86	5,82	5,74	5,66
14,08	13,71	13,32	12,93	12,69	12,53	12,33	12,20	12,12	11,91	11,72
2,54	2,50	2,46	2,42	2,40	2,38	2,36	2,35	2,34	2,32	2,30
3,35	3,28	3,22	3,15	3,11	3,08	3,04	3,02	3,01	2,97	2,93
5,81	5,67	5,52	5,36	5,26	5,20	5,12	5,07	5,03	4,95	4,87
11,54	11,19	10,84	10,48	10,26	10,11	9,92	9,80	9,73	9,53	9,36
2,42	2,38	2,34	2,30	2,27	2,25	2,23	2,22	2,21	2,18	2,16
3,14	3,07	3,01	2,94	2,89	2,86	2,83	2,80	2,79	2,75	2,71
5,26	5,11	4,96	4,81	4,71	4,65	4,57	4,52	4,48	4,40	4,32
9,89	9,57	9,24	8,90	8,69	8,55	8,37	8,26	8,19	8,00	7,84
2,32	2,28	2,24	2,20	2,17	2,16	2,13	2,12	2,11	2,08	2,06
2,98	2,91	2,85	2,77	2,73	2,70	2,66	2,64	2,62	2,58	2,54
4,85	4,71	4,56	4,41	4,31	4,25	4,17	4,12	4,08	4,00	3,92
8,75	8,45	8,13	7,80	7,60	7,47	7,30	7,19	7,12	6,94	6,78
2,25	2,21	2,17	2,12	2,10	2,08	2,05	2,04	2,03	2,00	1,98
2,85	2,79	2,72	2,65	2,60	2,57	2,53	2,51	2,49	2,45	2,41
4,54	4,40	4,25	4,10	4,01	3,94	3,86	3,81	3,78	3,69	3,61
7,92	7,63	7,32	7,01	6,81	6,68	6,52	6,42	6,35	6,18	6,02
2,19	2,15	2,10	2,06	2,03	2,01	1,99	1,97	1,96	1,93	1,91
2,75	2,69	2,62	2,54	2,50	2,47	2,43	2,40	2,38	2,34	2,30
4,30	4,16	4,01	3,86	3,76	3,70	3,62	3,57	3,54	3,45	3,37
7,29	7,00	6,71	6,40	6,22	6,09	5,93	5,83	5,76	5,59	5,44

(continuação)

Tabela A.9 Valores Críticos para Distribuições  $F$  (cont.)

		$\nu_1 = \text{gl do numerador}$									
		1	2	3	4	5	6	7	8	9	
$\nu_2 = \text{graus de liberdade do denominador}$	13	0,100	3,14	2,76	2,56	2,43	2,35	2,28	2,23	2,20	2,16
		0,050	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71
		0,010	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19
		0,001	17,82	12,31	10,21	9,07	8,35	7,86	7,49	7,21	6,98
	14	0,100	3,10	2,73	2,52	2,39	2,31	2,24	2,19	2,15	2,12
		0,050	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65
		0,010	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03
		0,001	17,14	11,78	9,73	8,62	7,92	7,44	7,08	6,80	6,58
	15	0,100	3,07	2,70	2,49	2,36	2,27	2,21	2,16	2,12	2,09
		0,050	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59
		0,010	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89
		0,001	16,59	11,34	9,34	8,25	7,57	7,09	6,74	6,47	6,26
	16	0,100	3,05	2,67	2,46	2,33	2,24	2,18	2,13	2,09	2,06
		0,050	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54
		0,010	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78
		0,001	16,12	10,97	9,01	7,94	7,27	6,80	6,46	6,19	5,98
	17	0,100	3,03	2,64	2,44	2,31	2,22	2,15	2,10	2,06	2,03
		0,050	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49
		0,010	8,40	6,11	5,19	4,67	4,34	4,10	3,93	3,79	3,68
		0,001	15,72	10,66	8,73	7,68	7,02	6,56	6,22	5,96	5,75
	18	0,100	3,01	2,62	2,42	2,29	2,20	2,13	2,08	2,04	2,00
		0,050	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46
		0,010	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60
		0,001	15,38	10,39	8,49	7,46	6,81	6,35	6,02	5,76	5,56
	19	0,100	2,99	2,61	2,40	2,27	2,18	2,11	2,06	2,02	1,98
		0,050	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42
		0,010	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52
		0,001	15,08	10,16	8,28	7,27	6,62	6,18	5,85	5,59	5,39
	20	0,100	2,97	2,59	2,38	2,25	2,16	2,09	2,04	2,00	1,96
		0,050	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39
		0,010	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46
		0,001	14,82	9,95	8,10	7,10	6,46	6,02	5,69	5,44	5,24
	21	0,100	2,96	2,57	2,36	2,23	2,14	2,08	2,02	1,98	1,95
		0,050	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37
		0,010	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40
		0,001	14,59	9,77	7,94	6,95	6,32	5,88	5,56	5,31	5,11
	22	0,100	2,95	2,56	2,35	2,22	2,13	2,06	2,01	1,97	1,93
		0,050	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34
		0,010	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35
		0,001	14,38	9,61	7,80	6,81	6,19	5,76	5,44	5,19	4,99
	23	0,100	2,94	2,55	2,34	2,21	2,11	2,05	1,99	1,95	1,92
		0,050	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32
		0,010	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30
		0,001	14,20	9,47	7,67	6,70	6,08	5,65	5,33	5,09	4,89
	24	0,100	2,93	2,54	2,33	2,19	2,10	2,04	1,98	1,94	1,91
		0,050	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30
		0,010	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26
		0,001	14,03	9,34	7,55	6,59	5,98	5,55	5,23	4,99	4,80

(continuação)

**Tabela A.9** Valores Críticos para Distribuições  $F$  (cont.)

$\nu_1 = \text{gl do numerador}$										
10	12	15	20	25	30	40	50	60	120	1000
2,14	2,10	2,05	2,01	1,98	1,96	1,93	1,92	1,90	1,88	1,85
2,67	2,60	2,53	2,46	2,41	2,38	2,34	2,31	2,30	2,25	2,21
4,10	3,96	3,82	3,66	3,57	3,51	3,43	3,38	3,34	3,25	3,18
6,80	6,52	6,23	5,93	5,75	5,63	5,47	5,37	5,30	5,14	4,99
2,10	2,05	2,01	1,96	1,93	1,91	1,89	1,87	1,86	1,83	1,80
2,60	2,53	2,46	2,39	2,34	2,31	2,27	2,24	2,22	2,18	2,14
3,94	3,80	3,66	3,51	3,41	3,35	3,27	3,22	3,18	3,09	3,02
6,40	6,13	5,85	5,56	5,38	5,25	5,10	5,00	4,94	4,77	4,62
2,06	2,02	1,97	1,92	1,89	1,87	1,85	1,83	1,82	1,79	1,76
2,54	2,48	2,40	2,33	2,28	2,25	2,20	2,18	2,16	2,11	2,07
3,80	3,67	3,52	3,37	3,28	3,21	3,13	3,08	3,05	2,96	2,88
6,08	5,81	5,54	5,25	5,07	4,95	4,80	4,70	4,64	4,47	4,33
2,03	1,99	1,94	1,89	1,86	1,84	1,81	1,79	1,78	1,75	1,72
2,49	2,42	2,35	2,28	2,23	2,19	2,15	2,12	2,11	2,06	2,02
3,69	3,55	3,41	3,26	3,16	3,10	3,02	2,97	2,93	2,84	2,76
5,81	5,55	5,27	4,99	4,82	4,70	4,54	4,45	4,39	4,23	4,08
2,00	1,96	1,91	1,86	1,83	1,81	1,78	1,76	1,75	1,72	1,69
2,45	2,38	2,31	2,23	2,18	2,15	2,10	2,08	2,06	2,01	1,97
3,59	3,46	3,31	3,16	3,07	3,00	2,92	2,87	2,83	2,75	2,66
5,58	5,32	5,05	4,78	4,60	4,48	4,33	4,24	4,18	4,02	3,87
1,98	1,93	1,89	1,84	1,80	1,78	1,75	1,74	1,72	1,69	1,66
2,41	2,34	2,27	2,19	2,14	2,11	2,06	2,04	2,02	1,97	1,92
3,51	3,37	3,23	3,08	2,98	2,92	2,84	2,78	2,75	2,66	2,58
5,39	5,13	4,87	4,59	4,42	4,30	4,15	4,06	4,00	3,84	3,69
1,96	1,91	1,86	1,81	1,78	1,76	1,73	1,71	1,70	1,67	1,64
2,38	2,31	2,23	2,16	2,11	2,07	2,03	2,00	1,98	1,93	1,88
3,43	3,30	3,15	3,00	2,91	2,84	2,76	2,71	2,67	2,58	2,50
5,22	4,97	4,70	4,43	4,26	4,14	3,99	3,90	3,84	3,68	3,53
1,94	1,89	1,84	1,79	1,76	1,74	1,71	1,69	1,68	1,64	1,61
2,35	2,28	2,20	2,12	2,07	2,04	1,99	1,97	1,95	1,90	1,85
3,37	3,23	3,09	2,94	2,84	2,78	2,69	2,64	2,61	2,52	2,43
5,08	4,82	4,56	4,29	4,12	4,00	3,86	3,77	3,70	3,54	3,40
1,92	1,87	1,83	1,78	1,74	1,72	1,69	1,67	1,66	1,62	1,59
2,32	2,25	2,18	2,10	2,05	2,01	1,96	1,94	1,92	1,87	1,82
3,31	3,17	3,03	2,88	2,79	2,72	2,64	2,58	2,55	2,46	2,37
4,95	4,70	4,44	4,17	4,00	3,88	3,74	3,64	3,58	3,42	3,28
1,90	1,86	1,81	1,76	1,73	1,70	1,67	1,65	1,64	1,60	1,57
2,30	2,23	2,15	2,07	2,02	1,98	1,94	1,91	1,89	1,84	1,79
3,26	3,12	2,98	2,83	2,73	2,67	2,58	2,53	2,50	2,40	2,32
4,83	4,58	4,33	4,06	3,89	3,78	3,63	3,54	3,48	3,32	3,17
1,89	1,84	1,80	1,74	1,71	1,69	1,66	1,64	1,62	1,59	1,55
2,27	2,20	2,13	2,05	2,00	1,96	1,91	1,88	1,86	1,81	1,76
3,21	3,07	2,93	2,78	2,69	2,62	2,54	2,48	2,45	2,35	2,27
4,73	4,48	4,23	3,96	3,79	3,68	3,53	3,44	3,38	3,22	3,08
1,88	1,83	1,78	1,73	1,70	1,67	1,64	1,62	1,61	1,57	1,54
2,25	2,18	2,11	2,03	1,97	1,94	1,89	1,86	1,84	1,79	1,74
3,17	3,03	2,89	2,74	2,64	2,58	2,49	2,44	2,40	2,31	2,22
4,64	4,39	4,14	3,87	3,71	3,59	3,45	3,36	3,29	3,14	2,99

(continuação)

Tabela A.9 Valores Críticos para Distribuições  $F$  (cont.)

		$\nu_1 = \text{gl do numerador}$									
		1	2	3	4	5	6	7	8	9	
$\nu_2 = \text{graus de liberdade do denominador}$	25	0,100	2,92	2,53	2,32	2,18	2,09	2,02	1,97	1,93	1,89
		0,050	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28
		0,010	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22
		0,001	13,88	9,22	7,45	6,49	5,89	5,46	5,15	4,91	4,71
	26	0,100	2,91	2,52	2,31	2,17	2,08	2,01	1,96	1,92	1,88
		0,050	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27
		0,010	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18
		0,001	13,74	9,12	7,36	6,41	5,80	5,38	5,07	4,83	4,64
	27	0,100	2,90	2,51	2,30	2,17	2,07	2,00	1,95	1,91	1,87
		0,050	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25
		0,010	7,68	5,49	4,60	4,11	3,78	3,56	3,39	3,26	3,15
		0,001	13,61	9,02	7,27	6,33	5,73	5,31	5,00	4,76	4,57
	28	0,100	2,89	2,50	2,29	2,16	2,06	2,00	1,94	1,90	1,87
		0,050	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24
		0,010	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12
		0,001	13,50	8,93	7,19	6,25	5,66	5,24	4,93	4,69	4,50
	29	0,100	2,89	2,50	2,28	2,15	2,06	1,99	1,93	1,89	1,86
		0,050	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22
		0,010	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,09
		0,001	13,39	8,85	7,12	6,19	5,59	5,18	4,87	4,64	4,45
	30	0,100	2,88	2,49	2,28	2,14	2,05	1,98	1,93	1,88	1,85
		0,050	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21
		0,010	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07
		0,001	13,29	8,77	7,05	6,12	5,53	5,12	4,82	4,58	4,39
	40	0,100	2,84	2,44	2,23	2,09	2,00	1,93	1,87	1,83	1,79
		0,050	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12
		0,010	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89
		0,001	12,61	8,25	6,59	5,70	5,13	4,73	4,44	4,21	4,02
	50	0,100	2,81	2,41	2,20	2,06	1,97	1,90	1,84	1,80	1,76
		0,050	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07
		0,010	7,17	5,06	4,20	3,72	3,41	3,19	3,02	2,89	2,78
		0,001	12,22	7,96	6,34	5,46	4,90	4,51	4,22	4,00	3,82
	60	0,100	2,79	2,39	2,18	2,04	1,95	1,87	1,82	1,77	1,74
		0,050	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04
		0,010	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72
		0,001	11,97	7,77	6,17	5,31	4,76	4,37	4,09	3,86	3,69
	100	0,100	2,76	2,36	2,14	2,00	1,91	1,83	1,78	1,73	1,69
		0,050	3,94	3,09	2,70	2,46	2,31	2,19	2,10	2,03	1,97
		0,010	6,90	4,82	3,98	3,51	3,21	2,99	2,82	2,69	2,59
		0,001	11,50	7,41	5,86	5,02	4,48	4,11	3,83	3,61	3,44
200	0,100	2,73	2,33	2,11	1,97	1,88	1,80	1,75	1,70	1,66	
	0,050	3,89	3,04	2,65	2,42	2,26	2,14	2,06	1,98	1,93	
	0,010	6,76	4,71	3,88	3,41	3,11	2,89	2,73	2,60	2,50	
	0,001	11,15	7,15	5,63	4,81	4,29	3,92	3,65	3,43	3,26	
1000	0,100	2,71	2,31	2,09	1,95	1,85	1,78	1,72	1,68	1,64	
	0,050	3,85	3,00	2,61	2,38	2,22	2,11	2,02	1,95	1,89	
	0,010	6,66	4,63	3,80	3,34	3,04	2,82	2,66	2,53	2,43	
	0,001	10,89	6,96	5,46	4,65	4,14	3,78	3,51	3,30	3,13	

(continuação)

**Tabela A.9** Valores Críticos para Distribuições  $F$  (cont.)

$\nu_1 = \text{gl do numerador}$										
10	12	15	20	25	30	40	50	60	120	1000
1,87	1,82	1,77	1,72	1,68	1,66	1,63	1,61	1,59	1,56	1,52
2,24	2,16	2,09	2,01	1,96	1,92	1,87	1,84	1,82	1,77	1,72
3,13	2,99	2,85	2,70	2,60	2,54	2,45	2,40	2,36	2,27	2,18
4,56	4,31	4,06	3,79	3,63	3,52	3,37	3,28	3,22	3,06	2,91
1,86	1,81	1,76	1,71	1,67	1,65	1,61	1,59	1,58	1,54	1,51
2,22	2,15	2,07	1,99	1,94	1,90	1,85	1,82	1,80	1,75	1,70
3,09	2,96	2,81	2,66	2,57	2,50	2,42	2,36	2,33	2,23	2,14
4,48	4,24	3,99	3,72	3,56	3,44	3,30	3,21	3,15	2,99	2,84
1,85	1,80	1,75	1,70	1,66	1,64	1,60	1,58	1,57	1,53	1,50
2,20	2,13	2,06	1,97	1,92	1,88	1,84	1,81	1,79	1,73	1,68
3,06	2,93	2,78	2,63	2,54	2,47	2,38	2,33	2,29	2,20	2,11
4,41	4,17	3,92	3,66	3,49	3,38	3,23	3,14	3,08	2,92	2,78
1,84	1,79	1,74	1,69	1,65	1,63	1,59	1,57	1,56	1,52	1,48
2,19	2,12	2,04	1,96	1,91	1,87	1,82	1,79	1,77	1,71	1,66
3,03	2,90	2,75	2,60	2,51	2,44	2,35	2,30	2,26	2,17	2,08
4,35	4,11	3,86	3,60	3,43	3,32	3,18	3,09	3,02	2,86	2,72
1,83	1,78	1,73	1,68	1,64	1,62	1,58	1,56	1,55	1,51	1,47
2,18	2,10	2,03	1,94	1,89	1,85	1,81	1,77	1,75	1,70	1,65
3,00	2,87	2,73	2,57	2,48	2,41	2,33	2,27	2,23	2,14	2,05
4,29	4,05	3,80	3,54	3,38	3,27	3,12	3,03	2,97	2,81	2,66
1,82	1,77	1,72	1,67	1,63	1,61	1,57	1,55	1,54	1,50	1,46
2,16	2,09	2,01	1,93	1,88	1,84	1,79	1,76	1,74	1,68	1,63
2,98	2,84	2,70	2,55	2,45	2,39	2,30	2,25	2,21	2,11	2,02
4,24	4,00	3,75	3,49	3,33	3,22	3,07	2,98	2,92	2,76	2,61
1,76	1,71	1,66	1,61	1,57	1,54	1,51	1,48	1,47	1,42	1,38
2,08	2,00	1,92	1,84	1,78	1,74	1,69	1,66	1,64	1,58	1,52
2,80	2,66	2,52	2,37	2,27	2,20	2,11	2,06	2,02	1,92	1,82
3,87	3,64	3,40	3,14	2,98	2,87	2,73	2,64	2,57	2,41	2,25
1,73	1,68	1,63	1,57	1,53	1,50	1,46	1,44	1,42	1,38	1,33
2,03	1,95	1,87	1,78	1,73	1,69	1,63	1,60	1,58	1,51	1,45
2,70	2,56	2,42	2,27	2,17	2,10	2,01	1,95	1,91	1,80	1,70
3,67	3,44	3,20	2,95	2,79	2,68	2,53	2,44	2,38	2,21	2,05
1,71	1,66	1,60	1,54	1,50	1,48	1,44	1,41	1,40	1,35	1,30
1,99	1,92	1,84	1,75	1,69	1,65	1,59	1,56	1,53	1,47	1,40
2,63	2,50	2,35	2,20	2,10	2,03	1,94	1,88	1,84	1,73	1,62
3,54	3,32	3,08	2,83	2,67	2,55	2,41	2,32	2,25	2,08	1,92
1,66	1,61	1,56	1,49	1,45	1,42	1,38	1,35	1,34	1,28	1,22
1,93	1,85	1,77	1,68	1,62	1,57	1,52	1,48	1,45	1,38	1,30
2,50	2,37	2,22	2,07	1,97	1,89	1,80	1,74	1,69	1,57	1,45
3,30	3,07	2,84	2,59	2,43	2,32	2,17	2,08	2,01	1,83	1,64
1,63	1,58	1,52	1,46	1,41	1,38	1,34	1,31	1,29	1,23	1,16
1,88	1,80	1,72	1,62	1,56	1,52	1,46	1,41	1,39	1,30	1,21
2,41	2,27	2,13	1,97	1,87	1,79	1,69	1,63	1,58	1,45	1,30
3,12	2,90	2,67	2,42	2,26	2,15	2,00	1,90	1,83	1,64	1,43
1,61	1,55	1,49	1,43	1,38	1,35	1,30	1,27	1,25	1,18	1,08
1,84	1,76	1,68	1,58	1,52	1,47	1,41	1,36	1,33	1,24	1,11
2,34	2,20	2,06	1,90	1,79	1,72	1,61	1,54	1,50	1,35	1,16
2,99	2,77	2,54	2,30	2,14	2,02	1,87	1,77	1,69	1,49	1,22

**Tabela A.10** Valores Críticos para Distribuições de Amplitude Studentizada

$\nu$	$\alpha$	$m$										
		2	3	4	5	6	7	8	9	10	11	12
5	0,05	3,64	4,60	5,22	5,67	6,03	6,33	6,58	6,80	6,99	7,17	7,32
	0,01	5,70	6,98	7,80	8,42	8,91	9,32	9,67	9,97	10,24	10,48	10,70
6	0,05	3,46	4,34	4,90	5,30	5,63	5,90	6,12	6,32	6,49	6,65	6,79
	0,01	5,24	6,33	7,03	7,56	7,97	8,32	8,61	8,87	9,10	9,30	9,48
7	0,05	3,34	4,16	4,68	5,06	5,36	5,61	5,82	6,00	6,16	6,30	6,43
	0,01	4,95	5,92	6,54	7,01	7,37	7,68	7,94	8,17	8,37	8,55	8,71
8	0,05	3,26	4,04	4,53	4,89	5,17	5,40	5,60	5,77	5,92	6,05	6,18
	0,01	4,75	5,64	6,20	6,62	6,96	7,24	7,47	7,68	7,86	8,03	8,18
9	0,05	3,20	3,95	4,41	4,76	5,02	5,24	5,43	5,59	5,74	5,87	5,98
	0,01	4,60	5,43	5,96	6,35	6,66	6,91	7,13	7,33	7,49	7,65	7,78
10	0,05	3,15	3,88	4,33	4,65	4,91	5,12	5,30	5,46	5,60	5,72	5,83
	0,01	4,48	5,27	5,77	6,14	6,43	6,67	6,87	7,05	7,21	7,36	7,49
11	0,05	3,11	3,82	4,26	4,57	4,82	5,03	5,20	5,35	5,49	5,61	5,71
	0,01	4,39	5,15	5,62	5,97	6,25	6,48	6,67	6,84	6,99	7,13	7,25
12	0,05	3,08	3,77	4,20	4,51	4,75	4,95	5,12	5,27	5,39	5,51	5,61
	0,01	4,32	5,05	5,50	5,84	6,10	6,32	6,51	6,67	6,81	6,94	7,06
13	0,05	3,06	3,73	4,15	4,45	4,69	4,88	5,05	5,19	5,32	5,43	5,53
	0,01	4,26	4,96	5,40	5,73	5,98	6,19	6,37	6,53	6,67	6,79	6,90
14	0,05	3,03	3,70	4,11	4,41	4,64	4,83	4,99	5,13	5,25	5,36	5,46
	0,01	4,21	4,89	5,32	5,63	5,88	6,08	6,26	6,41	6,54	6,66	6,77
15	0,05	3,01	3,67	4,08	4,37	4,59	4,78	4,94	5,08	5,20	5,31	5,40
	0,01	4,17	4,84	5,25	5,56	5,80	5,99	6,16	6,31	6,44	6,55	6,66
16	0,05	3,00	3,65	4,05	4,33	4,56	4,74	4,90	5,03	5,15	5,26	5,35
	0,01	4,13	4,79	5,19	5,49	5,72	5,92	6,08	6,22	6,35	6,46	6,56
17	0,05	2,98	3,63	4,02	4,30	4,52	4,70	4,86	4,99	5,11	5,21	5,31
	0,01	4,10	4,74	5,14	5,43	5,66	5,85	6,01	6,15	6,27	6,38	6,48
18	0,05	2,97	3,61	4,00	4,28	4,49	4,67	4,82	4,96	5,07	5,17	5,27
	0,01	4,07	4,70	5,09	5,38	5,60	5,79	5,94	6,08	6,20	6,31	6,41
19	0,05	2,96	3,59	3,98	4,25	4,47	4,65	4,79	4,92	5,04	5,14	5,23
	0,01	4,05	4,67	5,05	5,33	5,55	5,73	5,89	6,02	6,14	6,25	6,34
20	0,05	2,95	3,58	3,96	4,23	4,45	4,62	4,77	4,90	5,01	5,11	5,20
	0,01	4,02	4,64	5,02	5,29	5,51	5,69	5,84	5,97	6,09	6,19	6,28
24	0,05	2,92	3,53	3,90	4,17	4,37	4,54	4,68	4,81	4,92	5,01	5,10
	0,01	3,96	4,55	4,91	5,17	5,37	5,54	5,69	5,81	5,92	6,02	6,11
30	0,05	2,89	3,49	3,85	4,10	4,30	4,46	4,60	4,72	4,82	4,92	5,00
	0,01	3,89	4,45	4,80	5,05	5,24	5,40	5,54	5,65	5,76	5,85	5,93
40	0,05	2,86	3,44	3,79	4,04	4,23	4,39	4,52	4,63	4,73	4,82	4,90
	0,01	3,82	4,37	4,70	4,93	5,11	5,26	5,39	5,50	5,60	5,69	5,76
60	0,05	2,83	3,40	3,74	3,98	4,16	4,31	4,44	4,55	4,65	4,73	4,81
	0,01	3,76	4,28	4,59	4,82	4,99	5,13	5,25	5,36	5,45	5,53	5,60
120	0,05	2,80	3,36	3,68	3,92	4,10	4,24	4,36	4,47	4,56	4,64	4,71
	0,01	3,70	4,20	4,50	4,71	4,87	5,01	5,12	5,21	5,30	5,37	5,44
$\infty$	0,05	2,77	3,31	3,63	3,86	4,03	4,17	4,29	4,39	4,47	4,55	4,62
	0,01	3,64	4,12	4,40	4,60	4,76	4,88	4,99	5,08	5,16	5,23	5,29

**Tabela A.11** Áreas de Cauda da Curva Qui-Quadrado

Área da Cauda Superior	$\nu = 1$	$\nu = 2$	$\nu = 3$	$\nu = 4$	$\nu = 5$
> 0,100	< 2,70	< 4,60	< 6,25	< 7,77	< 9,23
0,100	2,70	4,60	6,25	7,77	9,23
0,095	2,78	4,70	6,36	7,90	9,37
0,090	2,87	4,81	6,49	8,04	9,52
0,085	2,96	4,93	6,62	8,18	9,67
0,080	3,06	5,05	6,75	8,33	9,83
0,075	3,17	5,18	6,90	8,49	10,00
0,070	3,28	5,31	7,06	8,66	10,19
0,065	3,40	5,46	7,22	8,84	10,38
0,060	3,53	5,62	7,40	9,04	10,59
0,055	3,68	5,80	7,60	9,25	10,82
0,050	3,84	5,99	7,81	9,48	11,07
0,045	4,01	6,20	8,04	9,74	11,34
0,040	4,21	6,43	8,31	10,02	11,64
0,035	4,44	6,70	8,60	10,34	11,98
0,030	4,70	7,01	8,94	10,71	12,37
0,025	5,02	7,37	9,34	11,14	12,83
0,020	5,41	7,82	9,83	11,66	13,38
0,015	5,91	8,39	10,46	12,33	14,09
0,010	6,63	9,21	11,34	13,27	15,08
0,005	7,87	10,59	12,83	14,86	16,74
0,001	10,82	13,81	16,26	18,46	20,51
< 0,001	> 10,82	> 13,81	> 16,26	> 18,46	> 20,51
Área da Cauda Superior	$\nu = 6$	$\nu = 7$	$\nu = 8$	$\nu = 9$	$\nu = 10$
> 0,100	< 10,64	< 12,01	< 13,36	< 14,68	< 15,98
0,100	10,64	12,01	13,36	14,68	15,98
0,095	10,79	12,17	13,52	14,85	16,16
0,090	10,94	12,33	13,69	15,03	16,35
0,085	11,11	12,50	13,87	15,22	16,54
0,080	11,28	12,69	14,06	15,42	16,75
0,075	11,46	12,88	14,26	15,63	16,97
0,070	11,65	13,08	14,48	15,85	17,20
0,065	11,86	13,30	14,71	16,09	17,44
0,060	12,08	13,53	14,95	16,34	17,71
0,055	12,33	13,79	15,22	16,62	17,99
0,050	12,59	14,06	15,50	16,91	18,30
0,045	12,87	14,36	15,82	17,24	18,64
0,040	13,19	14,70	16,17	17,60	19,02
0,035	13,55	15,07	16,56	18,01	19,44
0,030	13,96	15,50	17,01	18,47	19,92
0,025	14,44	16,01	17,53	19,02	20,48
0,020	15,03	16,62	18,16	19,67	21,16
0,015	15,77	17,39	18,97	20,51	22,02
0,010	16,81	18,47	20,09	21,66	23,20
0,005	18,54	20,27	21,95	23,58	25,18
0,001	22,45	24,32	26,12	27,87	29,58
< 0,001	> 22,45	> 24,32	> 26,12	> 27,87	> 29,58

(continuação)



**Tabela A.11** Áreas de Cauda da Curva Qui-Quadrado (cont.)

Área da Cauda Superior	$\nu = 11$	$\nu = 12$	$\nu = 13$	$\nu = 14$	$\nu = 15$
> 0,100	< 17,27	< 18,54	< 19,81	< 21,06	< 22,30
0,100	17,27	18,54	19,81	21,06	22,30
0,095	17,45	18,74	20,00	21,26	22,51
0,090	17,65	18,93	20,21	21,47	22,73
0,085	17,85	19,14	20,42	21,69	22,95
0,080	18,06	19,36	20,65	21,93	23,19
0,075	18,29	19,60	20,89	22,17	23,45
0,070	18,53	19,84	21,15	22,44	23,72
0,065	18,78	20,11	21,42	22,71	24,00
0,060	19,06	20,39	21,71	23,01	24,31
0,055	19,35	20,69	22,02	23,33	24,63
0,050	19,67	21,02	22,36	23,68	24,99
0,045	20,02	21,38	22,73	24,06	25,38
0,040	20,41	21,78	23,14	24,48	25,81
0,035	20,84	22,23	23,60	24,95	26,29
0,030	21,34	22,74	24,12	25,49	26,84
0,025	21,92	23,33	24,73	26,11	27,48
0,020	22,61	24,05	25,47	26,87	28,25
0,015	23,50	24,96	26,40	27,82	29,23
0,010	24,72	26,21	27,68	29,14	30,57
0,005	26,75	28,29	29,81	31,31	32,80
0,001	31,26	32,90	34,52	36,12	37,69
< 0,001	> 31,26	> 32,90	> 34,52	> 36,12	> 37,69
Área da Cauda Superior	$\nu = 16$	$\nu = 17$	$\nu = 18$	$\nu = 19$	$\nu = 20$
> 0,100	< 23,54	< 24,77	< 25,98	< 27,20	< 28,41
0,100	23,54	24,76	25,98	27,20	28,41
0,095	23,75	24,98	26,21	27,43	28,64
0,090	23,97	25,21	26,44	27,66	28,88
0,085	24,21	25,45	26,68	27,91	29,14
0,080	24,45	25,70	26,94	28,18	29,40
0,075	24,71	25,97	27,21	28,45	29,69
0,070	24,99	26,25	27,50	28,75	29,99
0,065	25,28	26,55	27,81	29,06	30,30
0,060	25,59	26,87	28,13	29,39	30,64
0,055	25,93	27,21	28,48	29,75	31,01
0,050	26,29	27,58	28,86	30,14	31,41
0,045	26,69	27,99	29,28	30,56	31,84
0,040	27,13	28,44	29,74	31,03	32,32
0,035	27,62	28,94	30,25	31,56	32,85
0,030	28,19	29,52	30,84	32,15	33,46
0,025	28,84	30,19	31,52	32,85	34,16
0,020	29,63	30,99	32,34	33,68	35,01
0,015	30,62	32,01	33,38	34,74	36,09
0,010	32,00	33,40	34,80	36,19	37,56
0,005	34,26	35,71	37,15	38,58	39,99
0,001	39,25	40,78	42,31	43,81	45,31
< 0,001	> 39,25	> 40,78	> 42,31	> 43,81	> 45,31

**Tabela A.12** Valores Críticos para o teste de Ryan-Joiner de Normalidade

		$\alpha$		
		<b>0,10</b>	<b>0,05</b>	<b>0,01</b>
<i>n</i>	<b>5</b>	0,9033	0,8804	0,8320
	<b>10</b>	0,9347	0,9180	0,8804
	<b>15</b>	0,9506	0,9383	0,9110
	<b>20</b>	0,9600	0,9503	0,9290
	<b>25</b>	0,9662	0,9582	0,9408
	<b>30</b>	0,9707	0,9639	0,9490
	<b>40</b>	0,9767	0,9715	0,9597
	<b>50</b>	0,9807	0,9764	0,9664
	<b>60</b>	0,9835	0,9799	0,9710
	<b>75</b>	0,9865	0,9835	0,9757

**Tabela A.13** Valores Críticos para o Teste de Wilcoxon com Pares de Observações

$$P_0(S_+ \geq c_1) = P(S_+ \geq c_1 \text{ quando } H_0 \text{ é verdadeiro})$$

$n$	$c_1$	$P_0(S_+ \geq c_1)$	$n$	$c_1$	$P_0(S_+ \geq c_1)$
3	6	0,125		78	0,011
4	9	0,125		79	0,009
	10	0,062		81	0,005
5	13	0,094	14	73	0,108
	14	0,062		74	0,097
	15	0,031		79	0,052
6	17	0,109		84	0,025
	19	0,047		89	0,010
	20	0,031		92	0,005
	21	0,016	15	83	0,104
7	22	0,109		84	0,094
	24	0,055		89	0,053
	26	0,023		90	0,047
	28	0,008		95	0,024
8	28	0,098		100	0,011
	30	0,055		101	0,009
	32	0,027		104	0,005
	34	0,012	16	93	0,106
	35	0,008		94	0,096
	36	0,004		100	0,052
9	34	0,102		106	0,025
	37	0,049		112	0,011
	39	0,027		113	0,009
	42	0,010		116	0,005
	44	0,004	17	104	0,103
10	41	0,097		105	0,095
	44	0,053		112	0,049
	47	0,024		118	0,025
	50	0,010		125	0,010
	52	0,005		129	0,005
11	48	0,103	18	116	0,098
	52	0,051		124	0,049
	55	0,027		131	0,024
	59	0,009		138	0,010
	61	0,005		143	0,005
12	56	0,102	19	128	0,098
	60	0,055		136	0,052
	61	0,046		137	0,048
	64	0,026		144	0,025
	68	0,010		152	0,010
	71	0,005		157	0,005
13	64	0,108	20	140	0,101
	65	0,095		150	0,049
	69	0,055		158	0,024
	70	0,047		167	0,010
	74	0,024		172	0,005

**Tabela A.14** Valores Críticos do Teste de Soma de Postos de Wilcoxon

$$P_0(W \geq c) = P(W \geq c \text{ quando } H_0 \text{ é verdadeiro})$$

<i>m</i>	<i>n</i>	<i>c</i>	$P_0(W \geq c)$	<i>m</i>	<i>n</i>	<i>c</i>	$P_0(W \geq c)$	
3	3	15	0,05	6	6	40	0,004	
		4	0,057			40	0,041	
	5	18	0,029			41	0,026	
		20	0,036			43	0,009	
		21	0,018			44	0,004	
	6	22	0,048			43	0,053	
		23	0,024		45	0,024		
		24	0,012		47	0,009		
	7	24	0,058		48	0,005		
		26	0,017		8	47	0,047	
		27	0,008			49	0,023	
	8	27	0,042			51	0,009	
28		0,024	52	0,005				
29		0,012	50	0,047				
4	4	24	0,057	6	6	52	0,021	
		25	0,029			54	0,008	
		26	0,014			55	0,004	
	5	27	0,056			7	54	0,051
		28	0,032				56	0,026
		29	0,016				58	0,011
	6	30	0,008	8	60	0,004		
		30	0,057		58	0,054		
		32	0,019		61	0,021		
	7	7	33	0,010	7	7	63	0,01
			34	0,005			65	0,004
			33	0,055			66	0,049
35		0,021	68	0,027				
36		0,012	71	0,009				
37		0,006	72	0,006				
8	8	36	0,055	8	8	71	0,047	
		38	0,024			73	0,027	
		40	0,008			76	0,01	
5	5	41	0,004	8	8	78	0,005	
		36	0,048			84	0,052	
		37	0,028			87	0,025	
		39	0,008			90	0,01	
						92	0,005	

**Tabela A.15** Valores Críticos para Intervalo de Pares de Observações de Wilcoxon $(\bar{x}_{(n(n+1)/2-c+1)}, \bar{x}_{(c)})$ 

<i>n</i>	Nível de Confiança (%)	<i>c</i>	<i>n</i>	Nível de Confiança (%)	<i>c</i>	<i>n</i>	Nível de Confiança (%)	<i>c</i>
5	93,8	15	13	99,0	81	20	99,1	173
	87,5	14		95,2	74		95,2	158
6	96,9	21	14	90,6	70	21	90,3	150
	93,7	20		99,1	93		99,0	188
7	90,6	19	15	95,1	84	22	95,0	172
	98,4	28		89,6	79		89,7	163
8	95,3	26	16	99,0	104	23	99,0	204
	89,1	24		95,2	95		95,0	187
9	99,2	36	17	90,5	90	24	90,2	178
	94,5	32		99,1	117		99,0	221
10	89,1	30	18	94,9	106	25	95,2	203
	99,2	44		89,5	100		90,2	193
11	94,5	39	19	99,1	130		99,0	239
	90,2	37		94,9	118		95,1	219
12	99,0	52		90,2	112		89,9	208
	95,1	47		99,0	143		99,0	257
	89,5	44		95,2	131		95,2	236
	99,0	61		90,1	124		89,9	224
	94,6	55		99,1	158			
	89,8	52		95,1	144			
	99,1	71		90,4	137			
	94,8	64						
	90,8	61						

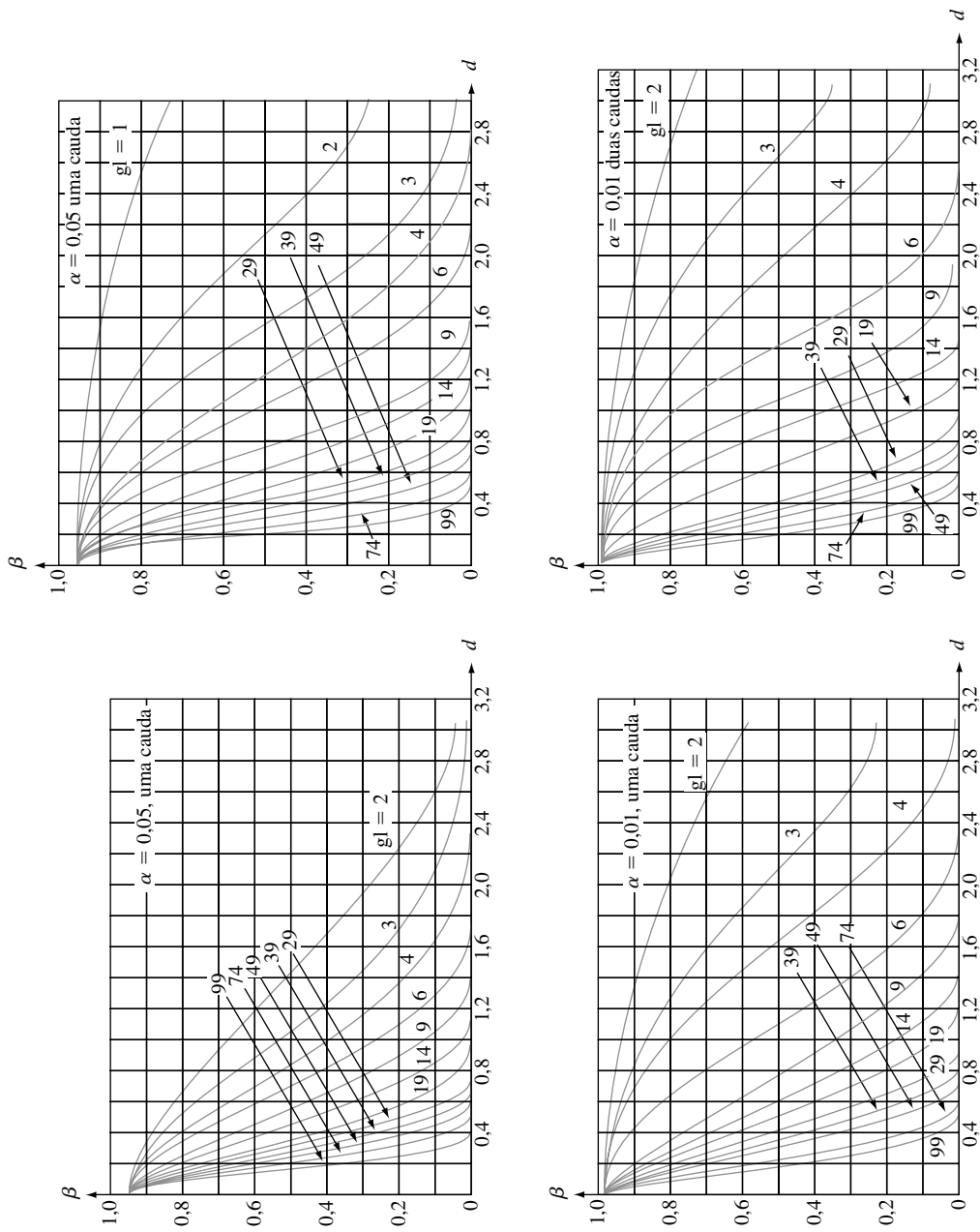
**Tabela A.16** Valores Críticos para o Intervalo de Soma de Postos de Wilcoxon $(d_{ij(mn-c+1)}, d_{ij(c)})$ 

		<b>Tamanho da Amostra Menor</b>							
		<b>5</b>		<b>6</b>		<b>7</b>		<b>8</b>	
<b>Tamanho da Amostra Maior</b>		<b>Nível de Confiança (%)</b>	<b><i>c</i></b>	<b>Nível de Confiança (%)</b>	<b><i>c</i></b>	<b>Nível de Confiança (%)</b>	<b><i>c</i></b>	<b>Nível de Confiança (%)</b>	<b><i>c</i></b>
5		99,2	25						
		94,4	22						
		90,5	21						
6		99,1	29	99,1	34				
		94,8	26	95,9	31				
		91,8	25	90,7	29				
7		99,0	33	99,2	39	98,9	44		
		95,2	30	94,9	35	94,7	40		
		89,4	28	89,9	33	90,3	38		
8		98,9	37	99,2	44	99,1	50	99,0	56
		95,5	34	95,7	40	94,6	45	95,0	51
		90,7	32	89,2	37	90,6	43	89,5	48
9		98,8	41	99,2	49	99,2	56	98,9	62
		95,8	38	95,0	44	94,5	50	95,4	57
		88,8	35	91,2	42	90,9	48	90,7	54
10		99,2	46	98,9	53	99,0	61	99,1	69
		94,5	41	94,4	48	94,5	55	94,5	62
		90,1	39	90,7	46	89,1	52	89,9	59
11		99,1	50	99,0	58	98,9	66	99,1	75
		94,8	45	95,2	53	95,6	61	94,9	68
		91,0	43	90,2	50	89,6	57	90,9	65
12		99,1	54	99,0	63	99,0	72	99,0	81
		95,2	49	94,7	57	95,5	66	95,3	74
		89,6	46	89,8	54	90,0	62	90,2	70

		<b>Tamanho da Amostra Menor</b>							
		<b>9</b>		<b>10</b>		<b>11</b>		<b>12</b>	
<b>Tamanho da Amostra Maior</b>		<b>Nível de Confiança (%)</b>	<b><i>c</i></b>	<b>Nível de Confiança (%)</b>	<b><i>c</i></b>	<b>Nível de Confiança (%)</b>	<b><i>c</i></b>	<b>Nível de Confiança (%)</b>	<b><i>c</i></b>
9		98,9	69						
		95,0	63						
		90,6	60						
10		99,0	76	99,1	84				
		94,7	69	94,8	76				
		90,5	66	89,5	72				
11		99,0	83	99,0	91	98,9	99		
		95,4	76	94,9	83	95,3	91		
		90,5	72	90,1	79	89,9	86		
12		99,1	90	99,1	99	99,1	108	99,0	116
		95,1	82	95,0	90	94,9	98	94,8	106
		90,5	78	90,7	86	89,6	93	89,9	101

Tabela A.17 Curvas  $\beta$  para Testes  $t$







Ambos os conjuntos de contagens são muito dispersos. Parece não haver outliers. A distribuição de contagens de crocantes parece estar deslocada para a direita (na direção dos valores maiores) do que para contagens de cremosas por algo na ordem de 10.

17. a. # Não-conformidade	Freqüência	Freq. Rel.
0	7	0,117
1	12	0,200
2	13	0,217
3	14	0,233
4	6	0,100
5	3	0,050
6	3	0,050
7	1	0,017
8	1	0,017
	<u>60</u>	<u>1,001</u>

b. 0,917, 0,867,  $1 - 0,867 = 0,133$

c. O histograma tem uma inclinação substancialmente positiva. Está centrado em algum lugar entre 2 e 3 e dispersa-se um pouco em seu centro.

19. a. 0,99 (99%), 0,71 (71%)

b. 0,64 (64%), 0,44 (44%)

c. Estritamente falando, o histograma não é unimodal, mas está próximo de ser, portanto com uma inclinação positiva moderada. Um tamanho de amostra muito maior resultaria provavelmente em uma figura mais plana.

21. a. y	Freq.	Freq. Rel.	b. z	Freq.	Freq. Rel.
0	17	0,362	0	13	0,277
1	22	0,468	1	11	0,234
2	6	0,128	2	3	0,064
3	1	0,021	3	7	0,149
4	0	0,000	4	5	0,106
5	<u>1</u>	<u>0,021</u>	5	3	0,064
	47	1,000	6	3	0,064
			7	0	0,000
			8	<u>2</u>	<u>0,043</u>
				47	1,001
					0,894; 0,830

23. a. Classe	Freq.	Freq. Rel.
0-<100	21	0,21
100-<200	32	0,32
200-<300	26	0,26
300-<400	12	0,12
400-<500	4	0,04
500-<600	3	0,03
600-<700	1	0,01
700-<800	0	0,00
800-<900	<u>1</u>	<u>0,01</u>
	100	1,00

b. Classe	Freq.	Freq. Rel.	Densidade
0-<50	8	0,08	0,0016
50-<100	13	0,13	0,0026
100-<150	11	0,11	0,0022
150-<200	21	0,21	0,0042
200-<300	26	0,26	0,0026
300-<400	12	0,12	0,0012
400-<500	4	0,04	0,0004
500-<600	3	0,03	0,0003
600-<900	<u>2</u>	<u>0,02</u>	0,00007
	100	1,00	

c. 0,79

25. Classe	Freq.	Classe	Freq.
10-<20	8	1,1-<1,2	2
20-<30	14	1,2-<1,3	6
30-<40	8	1,3-<1,4	7
40-<50	4	1,4-<1,5	9
50-<60	3	1,5-<1,6	6
60-<70	2	1,6-<1,7	4
70-<80	<u>1</u>	1,7-<1,8	5
	40	1,8-<1,9	<u>1</u>
			40

Original: inclinada positivamente;

Transformada: muito mais simétrica, não distante da forma de sino.

27. a. A observação 50 cai em uma classe limite.

b. Classe	Freq.	Freq. Rel.
0-<50	9	0,18
50-<100	19	0,38
100-<150	11	0,22
150-<200	4	0,08
200-<300	4	0,08
300-<400	2	0,04
400-<500	0	0,00
500-<600	<u>1</u>	<u>0,02</u>
	50	1,00

Um valor (central) representativo é tanto um pouco abaixo como um pouco acima de 100, dependendo de como se mede o centro. Há grande variação na vida útil, especialmente nos valores na extremidade superior dos dados. Há vários candidatos à outliers.

c. Classe	Freq.	Freq. Rel.
2,25-<2,75	2	0,04
2,75-<3,25	2	0,04
3,25-<3,75	3	0,06
3,75-<4,25	8	0,16
4,25-<4,75	18	0,36
4,75-<5,25	10	0,20
5,25-<5,75	4	0,08
5,75-<6,25	<u>3</u>	<u>0,06</u>
	50	1,00

Há muito mais simetria na distribuição dos valores de  $\ln(x)$  do que nos valores  $x$  em si mesmos, e menos variabilidade. Não há lacunas grandes ou óbvios outliers.

d. 0,38; 0,14.

29. Reclamação	Freq.	Freq. Rel.
J	10	0,1667
F	9	0,1500
B	7	0,1167
M	4	0,0667
C	3	0,0500
N	6	0,1000
O	21	0,3500
	60	1,0001

31. Classe	Freq.	Freq. Acum.	Freq. Rel. Acum.
0-<4	2	2	0,050
4-<8	14	16	0,400
8-<12	11	27	0,675
12-<16	8	35	0,875
16-<20	4	39	0,975
20-<24	0	39	0,975
24-<28	1	40	1,000

- parece haver uma quantidade relativamente pequena de variabilidade nos dados.
59. a. ED: 0,4; 0,10; 2,75; 2,65;  
 Não-Ed: 1,60; 0,30; 7,90; 7,60  
 b. ED: 8,9 e 9,2 são outliers moderados e 11,7 e 21,0 são outliers extremos.  
 Não há outliers na amostra não-ED.  
 c. Quatro outliers para ED, nenhum para não-ED. Inclinação positiva significativa em ambas as amostras; menos variabilidade em ED (o menor  $f_s$ ) e observações não-ED tendem a ser algo maiores do que as observações ED.
61. Outliers, ambos moderados e extremos, apenas às 6 da manhã. As distribuições em outros horários são bem simétricas. A variabilidade aumenta consideravelmente até 14h e então diminui ligeiramente, e o mesmo acontece para os valores de coeficiente gasolina-vapor “típicos”.
63. Baseadas em um boxplot comparativa, as três amostras parecem significativamente diferentes em relação ao centro. Há muito menos variabilidade para a taxa de fluxo de 160 do que para as duas outras taxas. Nenhum outlier presente. Há evidência de inclinação positiva moderada nos 50% do meio de cada amostra, mas não para a amostra da taxa de fluxo de 160 ou a amostra da taxa de fluxo de 200 como um todo.
65. a. 9,59; 59,41  
 b.  $CV = 0,396$  para HC e  $= 0,323$  para CO, portanto há maior variabilidade relativa nos dados de HC ainda que seu DP seja muito menor.
67. 10,65
69. a.  $\bar{y} = a\bar{x} + b$ ,  $s_y^2 = a^2s_x^2$     b. 189,14; 1,87
71. a. A média, a mediana e a média ordenada são virtualmente idênticas, sugerindo uma quantidade substancial de simetria nos dados: o fato de os quartis terem aproximadamente as mesmas distâncias da mediana e de as menores e maiores observações serem aproximadamente equidistantes do centro fornece suporte adicional para a simetria. O desvio padrão é bem pequeno em relação à média e à mediana.  
 b. Veja os comentários de (a). Além disso, utilizando  $1,5(Q3 - Q1)$  como padrão de medida, as duas maiores e as três menores observações são outliers moderados.
73.  $\bar{x} = 0,9255$ ,  $s = 0,0809$ ,  $\tilde{x} = 0,93$ , quantidade pequena de variabilidade, um pouco de inclinação.
75. a. Os “resumos de cinco números” ( $\tilde{x}$ , os dois quartos e a menor e a maior observações) são idênticos e não há outliers, portanto os três *boxplots* individuais são idênticos.  
 b. Diferenças na variabilidade, natureza das lacunas e existência de agrupamentos para as três amostras.  
 c. Não. Detalhe foi perdido.
33. a.  $\bar{x} = 192,57$ ,  $\tilde{x} = 189,0$   
 b. Novo  $\bar{x} = 189,71$ ;  $\tilde{x}$  inalterado  
 c. 191,0; 7,14%    d. 122,6
35. a.  $\bar{x} = 12,55$ ,  $\tilde{x} = 12,50$ ,  $\bar{x}_{tr(12,5)} = 12,40$ . Deleção da maior observação (18,0) faz com que  $\tilde{x}$  e  $\bar{x}_{tr}$  sejam um pouco menores do que  $\bar{x}$ .  
 b. Por no máximo 4.0  
 c. Não; multiplique os valores de  $\bar{x}$  e  $\tilde{x}$  pelo fator de conversão 1/2,2.
37.  $\bar{x}_{tr(10)} = 11,46$
39. a.  $\bar{x} = 1,0297$ ,  $\tilde{x} = 1,009$     b. 0,383
41. a. 0,7    b. Também 0,7    c. 13
43.  $\tilde{x} = 68,0$ ,  $\bar{x}_{tr(20)} = 66,2$ ,  $\bar{x}_{tr(30)} = 67,5$
45. a.  $\bar{x} = 115,58$ ; os desvios são 0,82; 0,32; -0,98; -0,38; 0,22  
 b. 0,482; 0,694    c. 0,482    d. 0,482
47.  $\bar{x} = 116,2$ ,  $s = 25,75$ . A magnitude de  $s$  indica uma quantidade substancial de variação em torno do centro (um desvio “representativo” de aproximadamente 25).
49. a. 0,56,80; 197,8040    b. 0,5016; 0,708
51. a. 1264,766; 35,564    b. 0,351; 0,593
53. a. 2,74; 3,88    b. 1,14    c. Inalterado  
 d. No máximo 0,40    e. 1,19
55. a. 33    b. Não  
 c. Inclinação positiva leve na metade do meio, mas simétrica no todo. A extensão da variabilidade parece significativa  
 d. No máximo 32
57. a. Sim. 125,8 é um outlier extremo e 250,2 é uma outlier moderado.  
 b. Além da presença de outliers, há inclinações positivas ambas em 50% do meio dos dados e, excetuando os outliers, em todo ele. Exceto para os dois outliers,

77. a.	0	2355566777888	
	1	0000135555	
	2	00257	
	3	0033	
	4	0057	caule: unidades
	5	044	folha: dezenas
	6		
	7	05	
	8	8	
	9	0	
	10	3	
	HI	22,0; 24,5	

b. Classe	Freq.	Freq. Rel.	Densidade
0-<2	23	0,500	0,250
2-<4	9	0,196	0,098
4-<6	7	0,152	0,076
6-<10	4	0,087	0,022
10-<20	1	0,022	0,002
20-<30	2	0,043	0,004

79. a.  $\bar{x}_{n+1} = (n\bar{x}_n + x_{n+1})/(n + 1)$   
 c. 12,53; 0,532

81. Uma inclinação substancial positiva (assumindo unimodalidade)

83. a. Todos os pontos caem em uma reta de 45°. Os pontos caem abaixo de uma reta de 45°.

b. Os pontos caem bem abaixo de uma reta de 45° indicando uma inclinação substancial positiva.

## Capítulo 2

1. a.  $\mathcal{S} = \{1324, 3124, 1342, 3142, 1423, 1432, 4123, 4132, 2314, 2341, 3214, 3241, 2413, 2431, 4213, 4231\}$   
 b.  $A = \{1324, 1342, 1423, 1432\}$   
 c.  $B = \{2314, 2341, 3214, 3241, 2413, 2431, 4213, 4231\}$

d.  $A \cup B = \{1324, 1342, 1423, 1432, 2314, 2341, 3214, 3241, 2413, 2431, 4213, 4231\}$ ,  
 $A \cap B$  não contém saídas ( $A$  e  $B$  são disjuntos),  
 $A' = \{3124, 3142, 4123, 4132, 2314, 2341, 3214, 3241, 2413, 2431, 4213, 4231\}$

3. a.  $A = \{SSF, SFS, FSS\}$   
 b.  $B = \{SSF, SFS, FSS, SSS\}$   
 c.  $C = \{SFS, SSF, SSS\}$   
 d.  $C' = \{FFF, FSF, FFS, FSS, SFF\}$ ,  
 $A \cup C = \{SSF, SFS, FSS, SSS\}$ ,  
 $A \cap C = \{SSF, SFS\}$ ,  
 $B \cup C = \{SSF, SFS, FSS, SSS\} = B$ ,  
 $B \cap C = \{SSF, SFS, SSS\} = C$

5. a.  $\mathcal{S} = \{(1, 1, 1), (1, 1, 2), (1, 1, 3), (1, 2, 1), (1, 2, 2), (1, 2, 3), (1, 3, 1), (1, 3, 2), (1, 3, 3), (2, 1, 1), (2, 1, 2), (2, 1, 3), (2, 2, 1), (2, 2, 2), (2, 2, 3), (2, 3, 1), (2, 3, 2), (2, 3, 3), (3, 1, 1), (3, 1, 2), (3, 1, 3), (3, 2, 1), (3, 2, 2), (3, 2, 3), (3, 3, 1), (3, 3, 2), (3, 3, 3)\}$  b.  $\{(1, 1, 1), (2, 2, 2), (3, 3, 3)\}$  c.  $\{(1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1)\}$  d.  $\{(1, 1, 1), (1, 1, 3), (1, 3, 1), (1, 3, 3), (3, 1, 1), (3, 1, 3), (3, 3, 1), (3, 3, 3)\}$

7. a. Há 35 resultados de  $\mathcal{S}$ . b.  $\{AABABAB, AABAABB, AAABBAB, AAABABB, AAAABBB\}$

11. a. 0,07 b. 0,30 c. 0,57

13. a. 0,36 b. 0,64 c. 0,53  
 d. 0,47 e. 0,17 f. 0,75

15. a. 0,572 b. 0,879

17. a. Há pacotes de software estatísticos além do SPSS e do SAS.  
 b. 0,70 c. 0,80 d. 0,20

19. a. 0,8841 b. 0,0435

21. a. 0,10 b. 0,18; 0,19 c. 0,41 d. 0,59  
 e. 0,31 f. 0,69

23. a. 0,067 b. 0,400 c. 0,933 d. 0,533

25. a. 0,98 b. 0,02 c. 0,03 d. 0,24

27. a. 0,1 b. 0,7 c. 0,6

29. a. 20 b. 60 c. 10

31. a. 243 b. 3.645 dias (aproximadamente 10 anos)

33. a. 53,130 b. 1.190 c. 0,022 d. 0,023

35. 0,20

37. 0,0456

39. a. 0,0839 b. 0,24975

41. a. 0,929 b. 0,0714 c. 0,99997520

43. 0,000394; 0,00394; 0,00001539

45. a. 0,447; 0,500; 0,200 b. 0,400; 0,447 c. 0,211

47. a. 0,50 b. 0,50 c. 0,625  
 d. 0,375 e. 0,769

49. 0,217; 0,178

51. 0,436; 0,581

53. 0,083

55. a. 0,0111 b. 0,333 c. 0,2

59. a. 0,21 b. 0,455 c. 0,264; 0,274

61. a. 0,578; 0,278; 0,144 b. 0; 0,457; 0,543

63. b. 0,54 c. 0,68 d. 0,74 e. 0,7941

65.  $P(\text{Média} | S) = 0,3922$ ,  $P(\text{Mediana} | S) = 0,2941$ , portanto Média e Mediana são mais e menos prováveis, respectivamente.

67. a. 0,126 b. 0,05 c. 0,1125  
 d. 0,2725 e. 0,5325 f. 0,2113

69. a. 0,300    b. 0,820    c. 0,146  
 73. 0,401; 0,722  
 75. a. 0,06235    b. 0,00421  
 77. 0,0059  
 79. a. 0,95  
 81. a. 0,10; 0,20    b. 0  
 83. a.  $p(2-p)$     b.  $1 - (1-p)^n$     c.  $(1-p)^3$   
 d.  $0,9 + (1-p)^3(0,1)$   
 e.  $0,1(1-p)^3/[0,9 + 0,1(1-p)^3] = 0,0137$  para  $p = 0,5$   
 85. 0,8588; 0,9897  
 87.  $[2\pi(1-\pi)]/(1-\pi^2)$   
 89. a. 0,333; 0,444    b. 0,150    c. 0,291  
 91. 0,45; 0,32
93. a. 0,0083    b. 0,2    c. 0,2  
 95. 0,905  
 97. a. 0,956    b. 0,994  
 99. 0,926  
 101. a. 0,018    b. 0,601  
 103. a. 0,883; 0,117    b. 23    c. 0,156  
 105.  $1 - (1-p_1)(1-p_2) \cdots (1-p_n)$   
 107. a. 0,0417    b. 0,375  
 109.  $P(\text{contratação do nº } 1) = 6/24$  para  $s = 0$ ,  $= 11/24$  para  $s = 1$ ,  $= 10/24$  para  $s = 2$  e  $= 6/24$  para  $s = 3$ , portanto  $s = 1$  é melhor.  
 111.  $1/4 = P(A_1 \cap A_2 \cap A_3) \neq P(A_1) \cdot P(A_2) \cdot P(A_3) = 1/8$

## Capítulo 3

1.  $x = 0$  para *FFF*;  $x = 1$  para *SFF*, *FSF* e *FFS*;  $x = 2$  para *SSF*, *SFS* e *FSS*; e  $x = 3$  para *SSS*  
 3.  $Z =$  média de dois números, com valores possíveis  $2/2, 3/2, \dots, 12/2$ ;  $W =$  valor absoluto da diferença, com valores possíveis 0, 1, 2, 3, 4, 5  
 5. Não. No Exemplo 3.4 assumimos  $Y = 1$  se no máximo três baterias são examinadas e assumimos  $Y = 0$  caso contrário. Então  $Y$  tem apenas dois valores.  
 7. a.  $\{0, 1, \dots, 12\}$ ; discreto    c.  $\{1, 2, 3, \dots\}$ ; discreto  
 e.  $\{0, c, 2c, \dots, 10.000c\}$ , onde  $c$  é o direito autoral por livro; discreto    g.  $\{x: m \leq x \leq M\}$  onde  $m$  ( $M$ ) é a tensão mínima (máxima) possível; contínua  
 9. a.  $\{2, 4, 6, 8, \dots\}$ , ou seja,  $\{2(1), 2(2), 2(3), 2(4), \dots\}$ , uma seqüência infinita; discreta  
 b.  $\{2, 3, 4, 5, 6, \dots\}$ , ou seja,  $\{1 + 1, 1 + 2, 1 + 3, 1 + 4, \dots\}$ , uma seqüência infinita; discreta  
 11. a.  $p(4) = 0,45, p(6) = 0,40, p(8) = 0,15, p(x) = 0$  para  $x \neq 4, 6$  ou  $8$     c. 0,55; 0,15  
 13. a. 0,70    b. 0,45    c. 0,55  
 d. 0,71    e. 0,65    f. 0,45  
 15. a. (1, 2), (1, 3), (1, 4), (1, 5), (2, 3), (2, 4), (2, 5), (3, 4), (3, 5), (4, 5)    b.  $p(0) = 0,3, p(1) = 0,6, p(2) = 0,1$     c.  $F(x) = 0$  para  $x < 0$ ,  $= 0,3$  para  $0 \leq x < 1$ ,  $= 0,9$  para  $1 \leq x < 2$  e  $= 1$  para  $2 \leq x$   
 17. a. 0,81    b. 0,162    c. É A; AUUUA, UAUUA, UUAUA, UUUAA; 0,00405  
 19.  $p(1) = p(2) = p(3) = p(4) = 0,25$   
 21.  $F(x) = 0$  para  $x < 0$ ,  $= 0,10$  para  $0 \leq x < 1$ ,  $= 0,25$  para  $1 \leq x < 2$ ,  $= 0,45$  para  $2 \leq x < 3$ ,  $= 0,70$  para  $3 \leq x < 4$ ,  $= 0,90$  para  $4 \leq x < 5$ ,  $= 0,96$  para  $5 \leq x < 6$  e  $= 1$  para  $6 \leq x$   
 23. a.  $p(1) = 0,30, p(3) = 0,10, p(4) = 0,05, p(6) = 0,15, p(12) = 0,40$     b. 0,30, 0,60
25. a.  $p(x) = (\frac{2}{3})^{x-1} \cdot (\frac{1}{3})$  para  $x = 1, 2, 3, \dots$   
 b.  $p(y) = (\frac{2}{3})^{y-2} \cdot (\frac{1}{3})$  para  $y = 2, 3, 4, 5, \dots$   
 c.  $p(z) = (\frac{25}{34}) \cdot (\frac{2}{3})^{2z-2}$  para  $z = 1, 2, 3, \dots$   
 29. a. 0,60    b. \$110  
 31. a. 16,38; 272,298; 3,9936    b. 401    c. 2.496  
 d. 13,66  
 33. Sim, porque  $\sum_{x=1}^{\infty} 1/x^2 < \infty$ .  
 35. \$700  
 37.  $E(h(X)) = 0,408$ , portanto joga.  
 39.  $V(-X) = V(X)$   
 41. a. 32,5    b. 7,5  
 c.  $V(X) = E[X(X-1)] + E(X) - [E(X)]^2$   
 43. a. 0,25; 0,11; 0,06; 0,04; 0,01    b. 2,64; 1,54; 0,04, conservativo    c. Eles são iguais    d.  $p(-1) = 0,02 = p(1), p(0) = 0,96$   
 45. a. 0,850    b. 0,200    c. 0,200    d. 0,701  
 e. 0,851    f. 0,000    g. 0,570  
 47. a. 0,354    b. 0,115    c. 0,918  
 49. a. 0,403    b. 0,787    c. 0,774  
 51. 0,1478  
 53. 0,407, independência  
 55. a. 0,017    b. 0,811; 0,425    c. 0,006; 0,902; 0,586  
 57. Quando  $p = 0,9$ , a probabilidade é 0,99 para A e 0,9963 para B. Se  $p = 0,5$ , estas probabilidades são 0,75 e 0,6875, respectivamente  
 59. A tabulação para  $p > 0,5$  é desnecessária  
 61. a. 20, 16    b. 70, 21  
 63.  $P(|X - \mu| \geq 2\sigma) = 0,042$  quando  $p = 0,5$  e  $= 0,065$  quando  $p = 0,75$ , comparado ao limite superior de 0,25. Usando  $k = 3$  no lugar de  $k = 2$ , estas probabilidades

são 0,002 e 0,004, respectivamente, considerando que o limite superior é 0,11.

65. a. 0,114    b. 0,879    c. 0,121    d. Use a distribuição com  $n = 15, p = 0,10$
67. a.  $h(x; 15, 10, 20)$  para  $x = 5, \dots, 10$   
b. 0,0325    c. 0,697
69. a.  $h(x; 10, 10, 20)$     b. 0,033
71. a.  $nb(x; 2, 0,5)$     b. 0,188    c. 0,688    d. 2, 4
73.  $nb(x; 6; 0,5), 6$
75. a. 0,932    b. 0,065    c. 0,068  
d. 0,492    e. 0,251
77. a. 0,011    b. 0,441    c. 0,554; 0,459    d. 0,945
79. Poisson(5)    a. 0,492    b. 0,133
81. a. 0,313; 0,809; 0,283    b. 12; 3,464  
c. 0,530; 0,011
83. a. 0,099    b. 0,135    c. 2
85. a. 4    b. 0,215    c. Pelo menos  $-\ln(0,1)/2 \approx 1,1513$  anos
87. a. 0,221    b. 6.800,000    c.  $p(x; 20,106)$
91. b. 3,114; 0,405; 0,636
93. a.  $b(x; 15; 0,75)$     b. 0,686  
c. 0,313    d. 11,25; 2,81    e. 0,310
95. 0,991
97. a.  $p(x; 2,5)$     b. 0,067    c. 0,109
99. 1,813; 3,05
101.  $p(2) = p^2, p(3) = (1 - p)p^2, p(4) = (1 - p)p^2, p(x) = [1 - p(2) - \dots - p(x - 3)](1 - p)p^2$  para  $x = 5, 6, 7, \dots; 0,99950841$
103. a. 0,0029    b. 0,0767; 0,9702
105. a. 0,135    b. 0,00144    c.  $\sum_{x=0}^{\infty} [p(x; 2)]^5$
107. 3,590
109. a. Não    b. 0,0273
111. b.  $0,6p(x; \lambda) + 0,4p(x; \mu)$     c.  $(\lambda + \mu)/2$   
d.  $(\lambda - \mu)^2/4 + (\lambda + \mu)/2$
113.  $\sum_{i=1}^{10} (p_{i+j+1} + p_{i-j-1})p_i$ , onde  $p_k = 0$  se  $k < 0$  ou  $k > 10$ .

## Capítulo 4

1. a. 0,25    b. 0,50    c. 0,4375
3. b. 0,5    c. 0,6875    d. 0,6328
5. a. 0,375    b. 0,125    c. 0,297    d. 0,578
7. a.  $f(x) = 0,1$  para  $25 \leq x \leq 35$  e 0 caso contrário  
b. 0,20    c. 0,40    d. 0,20
9. a. 0,562    b. 0,438; 0,438    c. 0,071
11. a. 0,25    b. 0,1875    c. 0,9375    d. 1,4142  
e.  $f(x) = x/2$  para  $0 < x < 2$   
f. 1,33    g. 0,222; 0,471    h. 2
13. a. 3    b. 0 para  $x \leq 1, 1 - x^3$  para  $x > 1$   
c. 0,125; 0,088    d. 1,5; 0,866    e. 0,924
15. a.  $F(x) = 0$  para  $x \leq 0, = 90[\frac{x^9}{9} - \frac{x^{10}}{10}]$  para  $0 < x < 1, = 1$  para  $x \geq 1$     b. 0,0107    c. 0,0107; 0,0107  
d. 0,9036    e. 0,818; 0,111    f. 0,3137
17. a. Para  $2 \leq x \leq 4, F(x) = 0,25[3x - 7 - (x - 3)^3]$   
b. 3    c. 3; 0,2
19. a. 0,597    b. 0,369  
c.  $f(x) = 0,3466 - 0,25 \ln(x)$  para  $0 < x < 4$
21. 314,79
23. 248; 3,60
25. b. 1,8(o nonagésimo percentil para  $X$ ) + 32  
c.  $a(X \text{ percentil}) + b$
27. a. 2,14    b. 0,81    c. 1,17  
d. 0,97    e. 2,41
29. a. 2,54    b. 1,34    c. -0,42
31. a. 0,9918    b. 0,0082    c. 0,9544
33. a. 0,3336    b. Aproximadamente 0  
c. 0,5795    d. 6,524    e. 0,8028
35. a. 36,7    b. 22,225    c. 3,179
37. 0,002
39. 10; 0,2
41. 7,3%
43. 21,155
45. a. 0,1190; 0,6969    b. 0,0021    c. 0,7054  
d.  $>5020$  ou  $<1844$  (usando  $z_{0,0005} = 3,295$ )  
e. Normal,  $\mu = 7,576, \sigma = 1,064; 0,7054$
47. 0,3174 para  $k = 1; 0,0456$  para  $k = 2; 0,0026$  para  $k = 3$ , quando comparado com os limites de 1,25 e 0,111, respectivamente.
49. a. Exato: 0,212, 0,577, 0,573; Aproximado: 0,211, 0,567; 0,596    b. Exato: 0,885; 0,575; 0,017; Aproximado: 0,885; 0,579; 0,012    c. Exato: 0,002; 0,029; 0,617; Aproximado: 0,003; 0,033; 0,599
51. a. 0,9666    b. 0,0099; 0
53. b. Normal,  $\mu = 239, \sigma^2 = 12,96$
55. a. 120    b. 1,329    c. 0,371  
d. 0,735    e. 0
57. a. 5, 4    b. 0,715    c. 0,411
59. a. 1    b. 1    c. 0,982    d. 0,129
61. a. 0,449; 0,699; 0,148    b. 0,05; 0,018
63. a.  $\cap A_i$     b. Exponencial com  $\lambda = 0,05$   
c. Exponencial com parâmetro  $n\lambda$

67. a. 0,826; 0,826; 0,0636    b. 0,664    c. 172,727
71. a. 0,930    b. 0,298    c. 98,18
73. a. 68,0; 122,1    b. 0,3204  
c. 0,7257, inclinação
75. a. 149,157, 223,595    b. 0,9573    c. 0,0414  
d. 148,41    e. 9,57    f. 125,90
77.  $\alpha = \beta$
79. b.  $[\Gamma(\alpha + \beta) \cdot \Gamma(m + \beta)] / [\Gamma(\alpha + \beta + m) \cdot \Gamma(\beta)]$ ,  $\beta / (\alpha + \beta)$
81. Sim, visto que o padrão na marcação é bem linear.
83. Sim
85. Sim
87. Marcação  $\ln(x)$  versus percentil  $z$ . O padrão é retilíneo, portanto uma distribuição de população lognormal é plausível.
89. O padrão na marcação é bem linear; é muito plausível que a força seja distribuída normalmente.
91. Há curvatura considerável na marcação.  $\lambda$  é um parâmetro de escala (como é  $\sigma$  para a família normal).
93. a.  $F(y) = \frac{1}{48}(y^2 - y^3/18)$  para  $0 \leq y \leq 12$   
b. 0,259; 0,5; 0,241    c. 6; 43,2; 7,2  
d. 0,518    e. 3,75
95. a.  $f(x) = x^2$  para  $0 \leq x < 1$  e  $e = \frac{7}{4} - \frac{3}{4}x$  para  $1 \leq x \leq \frac{7}{4}$   
b. 0,917    c. 1,213
97. a. 0,9162    b. 0,9549    c. 1,3374
99. a. 0,3859    b. 0,0663    c. (72,97; 119,03)
101. b.  $F(x) = 0$  para  $x < -1$ ,  $= (4x - x^3/3)/9 + \frac{11}{27}$  para  $-1 \leq x \leq 2$  e  $= 1$  para  $x > 2$   
c. Não.  $F(0) < 0,5 \Rightarrow \tilde{\mu} > 0$   
d.  $Y \sim \text{Bin}(10, \frac{5}{27})$
103. a. 0,368; 0,828; 0,460  
b. 352,53  
c.  $1/\beta \cdot \exp[-\exp(-(x - \alpha)/\beta)] \cdot \exp(-(x - \alpha)/\beta)$   
d.  $\alpha$   
e.  $\mu = 201,95$ , moda = 150,  $\tilde{\mu} = 182,99$
105. a.  $\mu$     b. Não    c. 0  
d.  $(\alpha - 1)\beta$     e.  $\nu - 2$
107. b.  $p(1 - \exp(-\lambda_1 x)) + (1 - p)(1 - \exp(-\lambda_2 x))$  para  $x \geq 0$     c.  $p/\lambda_1 + (1 - p)/\lambda_2$   
d.  $V(X) = 2p/\lambda_1^2 + 2(1 - p)/\lambda_2^2 - \mu^2$   
e. 1,  $CV > 1$     f.  $CV < 1$
109. a. Lognormal    b. 1    c. 2,72; 0,0185
113. a. Exponencial com  $\lambda = 1$   
c. Gama com parâmetros  $\alpha$  e  $c\beta$
115. a.  $(1/365)^3$     b.  $(1/365)^2$     c. 0,000002145
117. b. Sejam  $u_1, u_2, u_3, \dots$  uma seqüência de observações de uma distribuição Unif[0,1] (uma seqüência de números aleatórios). Então com  $x_i = (-0,1)\ln(1 - u_i)$ , os  $x_i$ s são observações de uma distribuição exponencial com  $\lambda = 10$ .
119.  $g(E(X)) \leq E(g(X))$

## Capítulo 5

1. a. 0,20    b. 0,42    c. Pelo menos uma mangueira é utilizada em cada bomba: 0,70.    d.  $p_x(x) = 0,16; 0,34; 0,50$  para  $x = 0, 1, 2$ , respectivamente;  $p_y(y) = 0,24; 0,38; 0,38$  para  $y = 0, 1, 2$ , respectivamente; 0,50  
e. Não;  $p(0, 0) \neq p_x(0) \cdot p_y(0)$
3. a. 0,15    b. 0,40    c. 0,22    d. 0,17; 0,46
5. a. 0,054    b. 0,00018
7. a. 0,030    b. 0,120    c. 0,300  
d. 0,380    e. Sim
9. a. 3/380,000    b. 0,3024    c. 0,3593  
d.  $10Kx^2 + 0,05$  para  $20 \leq x \leq 30$     e. Não
11. a.  $e^{-\lambda-\mu} \cdot \lambda^x \cdot \mu^y / x!y!$     b.  $e^{-\lambda-\mu} \cdot [1 + \lambda + \mu]$   
c.  $e^{-(\lambda+\mu)} \cdot (\lambda + \mu)^m / m!$ ; Poisson ( $\lambda + \mu$ )
13. a.  $e^{-x-y}$  para  $x \geq 0, y \geq 0$     b. 0,400    c. 0,594  
d. 0,330
15. a.  $F(y) = 1 - e^{-\lambda y} + (1 - e^{-\lambda y})^2 - (1 - e^{-\lambda y})^3$  para  $y \geq 0$   
b.  $2/3\lambda$
17. a. 0,25    b. 0,318    c. 0,637  
d.  $f_X(x) = 2\sqrt{R^2 - x^2} / \pi R^2$  para  $-R \leq x \leq R$ ; não
19. a.  $K(x^2 + y^2) / (10Kx^2 + 0,05)$ ;  $K(x^2 + y^2) / (10Ky^2 + 0,05)$   
b. 0,556; 0,549    c. 25,37; 2,87
21.  $= f_Y(y)$  para todo  $x$
23. 0,15
25.  $L^2$
27.  $\frac{1}{3}$  h
29.  $-\frac{2}{3}$
31. a. -0,1082    b. -0,0131
37. a. 

$\bar{x}$	25	32,5	40	45	52,5	65
$p(\bar{x})$	0,04	0,20	0,25	0,12	0,30	0,09

,  $E(\bar{X}) = \mu = 44,5$   
b. 

$s^2$	0	112,5	312,5	800
$p(s^2)$	0,38	0,20	0,30	0,12

,  $E(S^2) = 212,25 = \sigma^2$
39. 

Proporção	0	0,1	0,2	0,3	0,4	0,5
Probabilidade	0,000	0,000	0,000	0,001	0,005	0,027
Proporção	0,6	0,7	0,8	0,9	1,0	
Probabilidade	0,088	0,201	0,302	0,269	0,107	
41. a. 

$\bar{x}$	1	1,5	2	2,5	3	3,5	4
$p(\bar{x})$	0,16	0,24	0,25	0,20	0,10	0,04	0,01

  
b. 0,85    c. 

$r$	0	1	2	3
$p(r)$	0,30	0,40	0,22	0,08

47. a. 0,6826    b. 0,1056  
 49. a. 0,6026    b. 0,2981  
 51. 0,7720  
 53. a. 0,0062    b. 0  
 55. a. 0,0968    b. 0,8882  
 57. 0,9616  
 59. a. 0,9986; 0,9986    b. 0,9015; 0,3970  
     c. 0,8357    d. 0,9525; 0,0003  
 61. a. 3,5; 2,27; 1,51    b. 15,4; 75,94; 8,71  
 63. a. 0,695    b.  $4,0675 > 2,6775$   
 65. a. 0,9232    b. 0,9660  
 67. 0,1588  
 69. a. 2,400    b. 1,205; independência  
     c. 2,400; 41,77  
 71. a. 158; 430,25    b. 0,9788  
 73. a. Aproximadamente normal com média = 105, DP = 1,2649; Aproximadamente normal com média = 100, DP = 1,0142  
     b. Aproximadamente normal com média = 5, DP = 1,6213  
     c. 0,0068    d. 0,0010, sim  
 75. a. 0,2; 0,5; 0,3 para  $x = 12, 15, 20$ ; 0,10; 0,35; 0,55 para  $y = 12, 15, 20$     b. 0,25    c. Não    d. 33,35    e. 3,85  
 77. a.  $3/81,250$     b.  $f_x(x) = k(250x - 10x^2)$  para  $0 \leq x \leq 20$  e  $k(450x - 30x^2 + \frac{1}{2}x^3)$  para  $20 < x \leq 30$ ;  $f_y(y)$  resulta da substituição de  $y$  por  $x$  em  $f_x(x)$ . Eles não são independentes.  
     c. 0,355    d. 25,969  
     e. 204,6154; -0,894    f. 7,66  
 79.  $\approx 1$   
 81. a. 400 min    b. 70  
 83. 97  
 85. 0,9973  
 89. b, c. Qui-quadrado com  $\nu = n$ .  
 91. a.  $\sigma_w^2/(\sigma_w^2 + \sigma_E^2)$     b. 0,9999  
 93. 26; 1,64

---

## Capítulo 6

1. a. 8,14,  $\bar{X}$     b. 0,77,  $\bar{X}$     c. 1,66,  $S$   
     d. 0,148    e. 0,204,  $S/\bar{X}$   
 3. a. 1,348,  $\bar{X}$     b. 1,348,  $\bar{X}$   
     c. 1,781,  $\bar{X} + 1,28S$   
     d. 0,6736    e. 0,0905  
 5. 1,703,000; 1,591,300; 1,601,438,281  
 7. a. 120,6    b. 1.206.000    c. 0,80    d. 120,0  
 9. a. 2,11    b. 0,119  
 11. b.  $\left[ \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2} \right]^{1/2}$   
     c. Use  $\hat{p}_i = x_i/n_i$  e  $\hat{q}_i = 1 - \hat{p}_i$  no lugar de  $p_i$  e  $q_i$  na parte (b) para  $i = 1, 2$ .  
     d. -0,245    e. 0,041  
 15. a.  $\hat{\theta} = \sum X_i^2/2n$     b. 74,505  
 17. b. 0,444  
 19. a.  $\hat{p} = 2\hat{\lambda} - 0,30 = 0,20$     b.  $\hat{p} = (100\hat{\lambda} - 9)/70$   
 21. b.  $\hat{\alpha} = 5, \hat{\beta} = 28,0/\Gamma(1,2)$   
 23.  $\hat{\lambda}_1 = \bar{x}, \hat{\lambda}_2 = \bar{y}$ , estimado de  $(\lambda_1 - \lambda_2)$  é  $\bar{x} - \bar{y}$ .  
 25. a. 384,4; 18,86    b. 415,42  
 29. a.  $\hat{\theta} = \text{mín}(X_i), \hat{\lambda} = n/\sum [X_i - \text{mín}(X_i)]$   
     b. 0,64; 0,202  
 33. Com  $x_i$  = tempo entre o nascimento  $i - 1$  e o nascimento  $i$ ,  $\hat{\lambda} = 6/\sum_{i=1}^6 x_i = 0,0436$ .  
 35. 29,5  
 37. 1,0132

---

## Capítulo 7

1. a. 99,5%    b. 85%    c. 2,96    d. 1,15  
 3. a. Mais estreito    b. Não    c. Não    d. Não  
 5. a. (4,52; 5,18)    b. (4,12; 5,00)  
     c. 0,55    d. 94  
 7. Por um fator de 4; a largura é decrescente por um fator de 5.  
 9. a.  $(\bar{x} - 1,645\sigma/\sqrt{n}, \infty)$ ; (4,57,  $\infty$ )  
     b.  $(\bar{x} - z_\alpha \cdot \sigma/\sqrt{n}, \infty)$     c.  $(-\infty, \bar{x} + z_\alpha \cdot \sigma/\sqrt{n})$ ;  
      $(-\infty, 59,7)$   
 11. 950; 0,8714  
 13. a. (0,990; 1,066)    b. 158  
 15. a. 80%    b. 98%    c. 75%  
 17. 134,53  
 19. (0,513; 0,615)  
 21. 0,218  
 23. a. (0,438; 0,814)    b. 659  
 25. a. 381    b. 339

29. a. 2,228    b. 2,086    c. 2,845    d. 2,680  
e. 2,485    f. 2,571
31. a. 1,812    b. 1,753    c. 2,602    d. 3,747  
e. 2,1716 (do MINITAB)    f. Aproximadamente 2,43
33. a. Quantidade razoável de simetria, sem outliers  
b. Sim (baseado em uma marcação de probabilidade normal)  
c. (430,5; 446,1), sim, não
35. a. 95% CI: (2783,3; 2991,9)  
b. 95% PI: (2632,1; 3143,1), aproximadamente 2,5 vezes mais largo
37. a. (0,888; 0,964)    b. (0,752; 1,100)  
c. (0,634; 1,218)
39. a. Sim    b. (6,45; 98,01)  
c. (18,63; 85,83)
41. Todos os 70%; (c), porque é o mais curto
43. a. 18,307    b. 3,940    c. 0,95    d. 0,10
45. (3,6; 8,1); não
47. a. 95% IC: (6,702; 9,456)    b. (0,166; 0,410)
49. a. Parece haver uma ligeira inclinação positiva na metade do meio da amostra, mas as suíças inferiores são mais longas do que as superiores. A extensão da variabilidade é muito substancial, embora não haja outliers.  
b. Sim. O padrão dos pontos em uma marcação de probabilidade normal é razoavelmente linear.  
c. (33,53; 43,79)
51. a. (0,624; 0,732)    b. 1.080    c. Não
53. (-0,84, -0,16)
55. 246
57.  $(2t_r/\chi_{1-\alpha/2,2r}^2, 2t_r/\chi_{\alpha/2,2r}^2) = (65,3; 232,5)$
59. a.  $(\max(x_i)/(1-\alpha/2)^{1/n}, \max(x_i)/(\alpha/2)^{1/n})$   
b.  $(\max(x_i), \max(x_i)/\alpha^{1/n})$     c. (b); (4,2; 7,65)
61. (73,6; 78,8) versus (75,1; 79,6)

## Capítulo 8

1. a. Sim    b. Não    c. Não  
d. Sim    e. Não    f. Sim
5.  $H_0: \sigma = 0,05$  versus  $H_a: \sigma < 0,05$ . I: conclui que a variabilidade na espessura é satisfatória quando de fato não é. II: conclui que a variabilidade na espessura não é satisfatória quando de fato é.
7. I: concluindo que a usina não está em conformidade quando está; II: concluindo que a usina está em conformidade quando não está.
9. a.  $R_1$     b. I: julgando que uma das duas companhias é favorecida em relação à outra quando este não é o caso; II: julgando que nenhuma companhia é favorecida em relação a outra quando na realidade uma das duas é realmente a preferida.    c. 0,044  
d.  $\beta(0,3) = \beta(0,7) = 0,488$ ,  $\beta(0,4) = \beta(0,6) = 0,845$   
e. Rejeitar  $H_0$  em favor de  $H_a$ .
11. a.  $H_0: \mu = 10$  versus  $H_a: \mu \neq 10$     b. 0,01  
c. 0,5319; 0,0078    d. 2,58  
e. 10,1032 é substituído por 10,124, e 9,8968 é substituído por 9,876.    f.  $\bar{x} = 10,020$ , portanto  $H_0$  não deve ser rejeitada.    g.  $z \geq 2,58$  ou  $z \leq -2,58$
13. b. 0,0004; 0; menor do que 0,01
15. a. 0,0301    b. 0,003    c. 0,004
17. a.  $z = 2,56 \geq 2,33$ ; portanto rejeita  $H_0$ .  
b. 0,8413    c. 143    d. 0,0052
19. a.  $z = -2,27$ ; portanto não rejeita  $H_0$ .    b. 0,2266  
c. 22
21. a.  $t_{0,025,12} = 2,179 > 1,6$ ; portanto não rejeita  $H_0: \mu = 0,5$ .  
b.  $-1,6 > -2,179$ ; portanto não rejeita  $H_0$ .  
c. Não rejeita  $H_0$ .  
d. Rejeita  $H_0$  em favor de  $H_a: \mu \neq 0,5$ .
23.  $t = 2,24 \geq 1,708$ ; portanto  $H_0$  deve ser rejeitada. Os dados sugerem uma contradição da primeira hipótese.
25. a.  $z = -3,33 \leq -2,58$ ; portanto rejeita  $H_0$ .  
b. 0,1056    c. 217
27.  $-2,09 > -2,33$ ; portanto não rejeita  $H_0$ .
29. a.  $0,498 < 1,895$ ; portanto não rejeita  $H_0$ .    b. 0,72
31.  $-1,24 > -1,397$ ; portanto a primeira hipótese não parece ter sido contradita
35. Sim, porque  $-2,47 \leq -1,96$ .
37.  $z = 3,67 \geq 2,58$ ; portanto rejeita  $H_0: p = 0,40$ . Não.
39.  $-6,1$  é muito menor do que  $-2,33$ ; portanto os dados suportam fortemente esta conclusão.
41. a. {15, ..., 20}    b. 0,021, sim    c. 0,874; 0,196  
d. Usando  $R = \{14, \dots, 20\}$ , não.
43.  $z = -0,61 > -2,33$ ; não
45. a. Não rejeita  $H_0$ .    b. Não rejeita  $H_0$ .  
c. Não rejeita  $H_0$ .    d. Rejeita  $H_0$ .  
e. Não rejeita  $H_0$ .    f. Não rejeita  $H_0$ .
47. a. 0,0358    b. 0,0802    c. 0,5824  
d. 0,1586    e. 0
49. a. Valor  $P = 0,003 < 0,05$ ; portanto rejeita  $H_0$ .  
b. Valor  $P = 0,055 > 0,01$ ; portanto  $H_0$  não pode ser rejeitada.  
c. Valor  $P > 0,5$ ; portanto  $H_0$  não pode ser rejeitada em qualquer  $\alpha$  razoável.



51. Valor  $P < 0,0004 < 0,01$ ; portanto  $H_0: \mu = 5$  deve ser rejeitada em favor de  $H_a: \mu \neq 5$ .
53. Não; Valor  $P \approx 0,2$
55.  $t \approx 1,9$ , portanto valor  $P \approx 0,041$ . Com o valor  $P \leq \alpha$ ,  $H_0: \mu = 25$  deve ser rejeitada em favor de  $H_a: \mu > 25$ .
57.  $t \approx 1,9$ , portanto valor  $P \approx 0,116$ .  $H_0$  não deve, portanto ser rejeitada.
59. a. 0,8980; 0,1049; 0,0014    b. valor  $P \approx 0$ . Sim.  
c. Não
61.  $z = -3,12 \leq -1,96$ , portanto  $H_0$  deve ser rejeitada.
63. a.  $H_0: \mu = 0,85$  versus  $H_a: \mu \neq 0,85$   
b.  $H_0$  não pode ser rejeitada para qualquer  $\alpha$ .
65. a. Sim, porque  $t = 12,9 \geq 2,228$ .  
b. Distribuição de população normal
67. a. Não; não  
b. Não. Porque  $z = 0,44$  e valor  $P = 0,33 > 0,10$ .
69. a. Aproximadamente 0,6; aproximadamente 0,2 (pela Tabela A.17 do Apêndice)    b.  $n = 28$
71. a.  $z = 1,64 < 1,96$ , portanto  $H_0$  não pode ser rejeitada; Tipo II    b. 0,10. Sim.
73. Sim.  $z = -3,32 \leq -3,08$ , portanto  $H_0$  deve ser rejeitada.
75. Não, visto que  $z = 1,33 < 2,05$ .
77. Não, visto que  $12,11 < 21,665$ .
79. a.  $\bar{X} + 2,33S$     b.  $3,7145\sigma^2/n$ ;  $1,927s/\sqrt{n}$   
c.  $z = -1,22$ , portanto  $H_0: \mu + 2,33\sigma = 6,75$  não pode ser rejeitada em favor de  $H_a: \mu + 2,33\sigma < 6,75$ .
81. b.  $\Phi(z_\gamma + (\mu_0 - \mu)/(\sigma/\sqrt{n})) - \Phi(-z_{\alpha-\gamma} + (\mu_0 - \mu)/(\sigma/\sqrt{n}))$     c.  $\gamma > \alpha/2$

## Capítulo 9

1. a. -0,4 h; não    b. 0,0724; 0,2691    c. Não
3.  $z = 1,76 < 2,33$ , portanto não rejeita  $H_0$ .
5. a.  $z = -2,90$ , portanto rejeita  $H_0$ .    b. 0,0019  
c. 0,8212    d. 66
7. Sim, visto que  $z = 1,83 \geq 1,645$ .
9. a. 6,2; sim    b.  $z = 1,14$ , valor  $P \approx 0,25$ ; não  
c. Não    d. Um IC de 95% é (10,0; 21,8).
11. Um IC de 95% é (0,99; 2,41).
13. 50
15. b. Cresce.
17. a. 17    b. 21    c. 18    d. 26
19.  $t = -1,20 > -t_{0,01,9} = -2,821$ , portanto não rejeita  $H_0$ .
21. Sim;  $-2,64 \leq -2,602$ , portanto rejeita  $H_0$ .
23. b. Não    c.  $t = -0,38 > -t_{\alpha/2,10}$  para qualquer  $\alpha$ , razoável, portanto não rejeita  $H_0$  (valor  $P \approx 0,7$ ).
25. (0,3; 6,1), sim, sim
27. (6,5; 31,3) baseado em 9 gl; sim, sim
29.  $t = -2,10$ , gl = 25, valor  $P = 0,023$ . No nível de significância 0,05 concluiremos que a cola resulta em força média maior, mas não no nível de significância 0,01.
31. a. Centros virtualmente idênticos, substancialmente mais variabilidade na observação da variação média do que em observação de variação mais alta.  
b. (-7,9; 9,6), baseado em 23 gl; não
33.  $t = -2,2$ , valor  $P = 0,022$ , não rejeita  $H_0$ .
35.  $t = -2,2$ , gl = 16, valor  $P = 0,021 > 0,01 = \alpha$ , portanto não rejeita  $H_0$ .
37. a. (-0,561, -0,287)    b. Entre -1,224 e 0,376
39. a. Sim  
b.  $t = 2,7$ , valor  $P = 0,018 < 0,05 = \alpha$ , portanto  $H_0$  deve ser rejeitada.
41.  $t = 1,9$ , valor  $P = 0,047$ .  $H_0$  não pode ser rejeitada no nível de significância 0,01, mas é quase rejeitada em  $\alpha = 0,05$ .
43. a. Não    b. -49,1    c. 49,1
45. (-10,499; -4,589)
47.  $H_0$  é rejeitada porque  $-4,18 \leq -2,33$
49. Valor  $P = 0,4247$ , portanto  $H_0$  não pode ser rejeitada.
51. a.  $z = 1,48 < 1,645$ , portanto não rejeita  $H_0$ .  
b.  $n = 6582$
53. a. O IC para  $\ln(\theta)$  é  $\ln(\hat{\theta}) \pm z_{\alpha/2}[(m-x)/(mx) + (n-y)/(ny)]^{1/2}$ . O cálculo dos antilogs dos limites inferiores e superiores fornece um IC de  $\theta$  para si mesmo.  
b. (1,43; 2,31); a aspirina parece ser benéfica.
55. (-0,35; 0,07)
57. a. 3,69    b. 4,82    c. 0,207    d. 0,271  
e. 4,30    f. 0,212    g. 0,95    h. 0,94
59.  $f = 0,384$ ; como  $0,167 < 0,384 < 3,63$ , não rejeita  $H_0$ .
61.  $f = 2,85 \geq 2,08$ , portanto rejeita  $H_0$ ; parece haver mais variabilidade no ganho de peso de baixa dosagem.
63. ( $s_2^2 F_{1-\alpha/2}/s_1^2$ ,  $s_2^2 F_{\alpha/2}/s_1^2$ ); (0,023, 1,99)
65. Não.  $t = 3,2$ , gl = 15, valor  $P = 0,006$ , portanto rejeita  $H_0: \mu_1 - \mu_2 = 0$  usando tanto  $\alpha = 0,05$  como 0,01.
67.  $z > 0 \Rightarrow$  valor  $P > 0,5$ , portanto  $H_0: p_1 - p_2 = 0$  não pode ser rejeitada.
69. (-299,3; 1.517,9)
71. (1.024,0; 1.336,0), sim

73. Sim.  $t = -2,25$ ,  $gl = 57$ , valor  $P \approx 0,028$
75. a. Não.  $t = -2,84$ ,  $gl = 18$ , valor  $P \approx 0,012$   
 b. Não.  $t = -0,56$ , Valor  $P \approx 0,29$
77. Não em um nível de significância 0,05.  $t = -1,76 > -t_{0,05;4} = -2,015$
79. Não, nem o teste  $t$  de amostra dupla deve ser utilizado, pois uma marcação de probabilidade normal sugere que a distribuição de boa visibilidade não é normal.
81. Desagrupado:  $gl = 15$ ,  $t = -1,8$ , valor  $P \approx 0,092$   
 Agrupado:  $gl = 24$ ,  $t = -1,9$ , valor  $P \approx 0,070$

83. a.  $m = 141$ ,  $n = 47$     b.  $m = 240$ ,  $n = 160$
85. a. Não;  $-t_{0,0005,11} = -4,437 < 3,3 < 4,437$ , portanto não rejeita  $H_0$ .    b. valor  $P = 0,008$ , portanto  $H_0$  seria rejeitada se  $\alpha = 0,05$  ou 0,01.
87. 0,9015; 0,8264; 0,0294; 0,0000; média verdadeira dos QIs; não
89. Sim;  $z = 4,2$ , valor  $P \approx 0$
91. a. Sim.  $t = -6,4$ ,  $gl = 57$ , e valor  $P \approx 0$   
 b.  $t = 1,1$ , valor  $P = 0,14$ , portanto não rejeita  $H_0$ .
93.  $(-1,29, -0,59)$

## Capítulo 10

1. a.  $f = 1,85 < 3,06 = F_{0,05;4,15}$ , portanto não rejeita  $H_0$ .  
 b. valor  $P > 0,10$
3.  $f = 1,30 < 2,57 = F_{0,10;2,21}$ , portanto valor  $P > 0,10$ .  $H_0$  não pode ser rejeitada em qualquer nível de significância razoável.
5.  $f = 1,73 < 5,49 = F_{0,01;2,27}$ , portanto as três categorias não parecem diferentes.
7.  $f = 1,70 < 2,46 = F_{0,10;3,16}$ , portanto valor  $P > 0,10$ .  $H_0$  não pode ser rejeitada em qualquer nível de significância razoável.
9.  $f = 3,96$  e  $F_{0,05;3,20} = 3,10 < 3,96 < 4,94 = F_{0,01;3,20}$ , portanto  $0,01 < \text{valor } P < 0,05$ . Assim  $H_0$  pode ser rejeitada no nível de significância 0,05; parece haver diferenças na quantidade de grãos.
11.  $w = 36,09$     3    1    4    2    5  
                   437,5   462,0   469,3   512,8   532,1  
 As marcas 2 e 5 não parecem diferir, nem parece haver quaisquer diferenças entre as marcas 1, 3 e 4, mas cada marca no primeiro grupo parece diferir significativamente de todas as marcas no segundo grupo.
13.            3            1            4            2            5  
               427,5    462,0    469,3    502,8    532,1
15.  $w = 5,94$             2            1            3            4  
                           24,69    26,08    29,95    33,84  
 As únicas diferenças significativas estão entre 4 e ambas 1 e 2.
17.  $(-0,029; 0,379)$
19. Qualquer valor de SQE entre 422,16 e 431,88 funcionará.
21. a.  $f = 22,6$  e  $F_{0,01;5,78} \approx 3,3$ , portanto rejeita  $H_0$ .  
 b.  $(-99,16; -35,64)$ ,  $(29,34; 94,16)$

23.

	1	2	3	4
1	–	$2,88 \pm 5,81$	$7,43 \pm 5,81$	$12,78 \pm 5,48$
2	–	–	$4,55 \pm 6,13$	$9,90 \pm 5,81$
3	–	–	–	$5,35 \pm 5,81$
4	–	–	–	–
		<u>4</u>	<u>3</u>	<u>2</u> <u>1</u>

25. a. Normal, variâncias iguais  
 b.  $SQTr = 8,33$ ,  $SQE = 77,79$ ,  $f = 1,7$ ,  $H_0$  não deve ser rejeitada (valor  $P > 0,10$ )
27. a.  $f = 3,75 \geq 3,10 = F_{0,05;3,20}$ , portanto as marcas parecem diferir.    b. Normalidade é bastante plausível (uma marcação de probabilidade normal dos resíduos  $x_{ij} - \bar{x}_i$  mostram um padrão linear).  
 c. 4 3 2 1    Somente as marcas 1 e 4 parecem diferir significativamente.
31. Aproximadamente 0,62
33.  $\arcsin(\sqrt{x/n})$
35. a.  $3,68 < 4,94$ , portanto  $H_0$  não é rejeitada.  
 b.  $0,029 > 0,01$ , portanto  $H_0$  não é rejeitada.
37.  $f = 8,44 > 6,49 = F_{0,001}$ , portanto o valor  $P < 0,001$  e  $H_0$  devem ser rejeitados.  
           5 3 1 4 2    Este padrão de subpontuação é um pouco complicado de ser interpretado.
39. O IC é  $(-0,144; 0,474)$ , o que inclui 0.
41.  $f = 3,96 < 4,07$ , portanto  $H_0: \sigma_A^2 = 0$  não pode ser rejeitada.
43.  $(-3,70; 1,04)$ ,  $(-4,83; -0,33)$ ,  $(-3,77; 1,27)$ ,  $(-3,99; 0,15)$ . Apenas  $\mu_1 - \mu_3$  entre estes quatro contrastes parece diferir significativamente de 0.
45. Eles são idênticos.

## Capítulo 11

1. **a.**  $f_A = 1,55$ , portanto não rejeita  $H_{0A}$ .  
**b.**  $f_B = 2,98$ , portanto não rejeita  $H_{0B}$ .
3. **a.**  $f_A = 12,987 \geq F_{0,01;3,9}$ , portanto conclui que há um efeito da taxa de gás:  $f_B = 105,31$ , portanto conclua que há efeito da taxa de líquido efetiva. **b.**  $w = 95,44$ ; 231,75 325,25 441,0 613,25, portanto somente as duas taxas mais baixas não diferem significativamente uma da outra.  
**c.** 336,75 382,25 419,25 473, portanto somente as taxas menores e maiores parecem diferir significativamente uma da outra.
5.  $f_A = 2,56$ ,  $F_{0,01;3,12} = 5,95$ , portanto parece não haver efeito devido ao ângulo de tração.

7. **a.**

Fonte	gl	SQ	QM	$f$
Tratamentos	2	28,78	14,39	1,04
Blocos	17	2.977,67	175,16	12,68
Erros	34	469,55	13,81	
Total	53	3.476,00		

A pontuação de adaptação da média verdadeira não parece depender de qual tratamento é dado.

**b.** Sim,  $f_B$  é bastante amplo, sugerindo uma grande variabilidade entre os indivíduos.

9.

Fonte	gl	SQ	QM	$f$	$F_{0,05}$
Tratamentos	3	81,19	27,06	22,4	3,01
Blocos	8	66,50	8,31		
Erros	24	29,06	1,21		
Total	35	176,75			
1	4	3	2		
8,56	9,22	10,78	12,44		

11. Os resíduos são 0,0350, 0,0117, -0,0750, 0,0283, 0,0875, -0,0758, -0,0825, 0,0708, -0,1225, 0,0642, 0,1575 e -0,0992. O padrão na marcação da probabilidade normal é bem linear.

13. **b.** Cada SQ é multiplicado por  $c^2$ , mas  $f_A$  e  $f_B$  não são alterados.

15. **a.** Aproximadamente 0,20; 0,43 **b.** Aproximadamente 0,30

17. **a.**  $f_A = 3,76$ ,  $f_B = 6,82$ ,  $f_{AB} = 0,74$  e  $F_{0,05;2,9} = 4,26$ , portanto a quantidade de fibra de carbono adicionada parece significativa. **b.**  $f_A = 6,54$ ,  $f_B = 5,33$ ,  $f_{AB} = 0,27$

19. **a.**

Fonte	gl	SQ	QM	$f$
Carvão	2	1,00241	0,50121	29,49
NaOH	2	0,12431	0,06216	3,66
Interação	4	0,01456	0,00364	0,21
Erros	9	0,15295	0,01699	
Total	17	1,29423		

O tipo de carvão parece afetar a acidez total.

- b.** Carvões 1 e 3 não diferem significativamente um do outro, mas ambos diferem significativamente do carvão 2.

21. **a, b.**

Fonte	gl	SQ	QM	$f$
A	2	22.941,80	11.470,90	22,98
B	4	22.765,53	5.691,38	5,60
AB	8	3.993,87	4.99,23	0,49
Erros	15	15.253,50	1.016,90	
Total	29	64.954,70		

$H_{0A}$  e  $H_{0B}$  são ambas rejeitadas.

23.

Fonte	gl	SQ	QM	$f$
A	2	11.573,38	5.786,69	$\frac{MSA}{MSAB} = 26,70$
B	4	17.930,09	4.482,52	$\frac{MSB}{MSE} = 28,51$
AB	8	1.734,17	216,77	$\frac{MSAB}{MSE} = 1,38$
Erros	30	4.716,67	157,22	
Total	44	35.954,31		

Como  $F_{0,01;8,30} = 3,17$ ,  $F_{0,01;2,8} = 8,65$  e  $F_{0,01;4,30} = 4,02$ ,  $H_{0G}$  não é rejeitada, mas ambas  $H_{0A}$  e  $H_{0B}$  são rejeitadas.

25. (-1,39; -1,05)

27. **a.**

Fonte	gl	SQ	QM	$f$	$F_{0,05}$
A	2	14.144,44	7.072,22	61,06	3,35
B	2	5.511,27	2.755,64	23,79	3,35
C	2	244.696,39	122.348,20	1.056,27	3,35
AB	4	1.069,62	267,41	2,31	2,73
AC	4	62,67	15,67	0,14	2,73
BC	4	331,67	82,92	0,72	2,73
ABC	8	1.080,77	135,10	1,17	2,31
Erros	27	3.127,50	115,83		
Total	53	270.024,33			

**d.**  $Q_{0,05;3,27} = 3,51$ ,  $w = 8,90$ , e todos os três níveis significativamente diferem uns dos outros.

29.

Fonte	gl	SQ	QM	$f$
A	2	12,896	6,448	1,04
B	1	100,041	100,041	16,10
C	3	393,416	131,139	21,10
AB	2	1,646	0,823	<1
AC	6	71,021	11,837	1,905
BC	3	1,542	0,514	<1
ABC	6	9,771	1,629	<1
Erros	72	447,500	6,215	
Total	95	1.037,833		

**a.** Nenhum efeito de interação é significativo.

**b.** Os principais efeitos do fator B e do fator C são significantes.

**c.**  $w = 1,89$ ; apenas as máquinas 2 e 4 não diferem significativamente uma da outra.

31. A coluna do valor P mostra que diversos efeitos de interação são significativos no nível 0,01.

33.

Fonte	gl	SQ	QM	$f$
A	6	67,32	11,02	
B	6	51,06	8,51	
C	6	5,43	0,91	0,61
Erros	30	44,26	1,48	
Total	48	168,07		

$F_{0,05;6,30} = 2,42$ ,  $f_C = 0,61$ , portanto  $H_{0C}$  não é rejeitada.

35. Fonte	gl	SQ	QM	f
A	4	28,88	7,22	10,7
B	4	23,70	5,93	8,79
C	4	0,62	0,155	<1
Erros	12	8,10	0,675	
Total	24	61,30		

Como  $F_{0,05;4;12} = 3,26$ , ambos A e B são significantes.

37. Fonte	gl	QM	f
A	2	2.207,329	2.259*
B	1	47,255	48,4*
C	2	491,783	503*
D	1	0,044	<1
AB	2	15,303	15,7*
AC	4	275,446	282*
AD	2	0,470	<1
BC	2	2,141	2,19
BD	1	0,273	<1
CD	2	0,247	<1
ABC	4	3,714	3,80
ABD	2	4,072	4,17*
ACD	4	0,767	<1
BCD	2	0,280	<1
ABCD	4	0,347	<1
Erros	36	0,977	
Total	71	93,621	

\*Indica razão F ratio significante.

39. a.  $\hat{\beta}_1 = 54,38$ ,  $\hat{\gamma}_{11}^{AC} = -2,21$ ,  $\hat{\gamma}_{21}^{AC} = 2,21$ .

b. Efeito			
Fonte	Contraste	QM	f
A	1.307	71.177,04	436,7
B	1.305	70.959,34	435,4
C	529	11.660,04	71,54
AB	199	1.650,04	10,12
AC	-53	117,04	<1
BC	57	135,38	<1
ABC	27	30,38	<1
Erros		162,98	

41. Fonte	SQ	f
A	136.640,02	1.007,6
B	139.644,19	1.029,8
C	24.616,02	181,5
D	20.377,52	150,3
AB	2.173,52	16,0
AC	2,52	<1
AD	58,52	<1
BC	165,02	1,2
BD	9,19	<1
CD	17,52	<1
ABC	42,19	<1
ABD	117,19	<1
ACD	188,02	1,4
BCD	13,02	<1
ABCD	204,19	1,5
Erros	4.339,33	
Total	328.607,98	

$F_{0,05;1;32} \approx 4,15$ , portanto apenas os quatro efeitos principais e a interação AB parecem significantes.

43. Fonte	gl	SQ	f
A	1	0,436	<1
B	1	0,099	<1
C	1	0,109	<1
D	1	414,12	851
AB	1	0,003	<1
AC	1	0,078	<1
AD	1	0,017	<1
BC	1	1,404	3,62
BD	1	0,456	<1
CD	1	2,190	4,50
Erros	5	2,434	

$F_{0,05;1;5} = 6,61$ , portanto somente o efeito principal do fator D é julgado significativo.

45. a. 1: (1), ab, cd, abcd; 2: a, b, acd, bcd; 3: c, d, abc, abd; 4: ac, bc, ad, bd.

b. Fonte	gl	SQ	f
A	1	14.028,13	53,89
B	1	92.235,13	345,33
C	1	3,13	<1
D	1	18,00	<1
AC	1	105,13	<1
AD	1	200,00	<1
BC	1	91,13	<1
BD	1	420,50	1,62
ABC	1	276,13	1,06
ABD	1	2,00	<1
ACD	1	450,00	1,73
BCD	1	2,00	<1
Blocos	7	898,88	<1
Erros	12	3.123,72	
Total	31	111.853,88	

$F_{0,01;1;12} = 9,33$ , portanto somente os efeitos principais de A e B são significantes.

47. a. ABFG; (1), ab, cd, ce, de, fg, acf, adf, adg, aef, acg, aeg, bcf, bdf, bdg, bef, beg, abcd, abce, abde, abfg, cdfg, cefg, defg, acdef, acdeg, bcdef, bcdeg, abcdfg, abcdefg, abdefg. {A, BCDE, ACDEFG, BFG}, {B, ACDE, BCDEFG, AFG}, {C, ABDE, DEFG, ABCFG}, {D, ABCE, CEF, ABDFG}, {E, ABCD, CDFG, ABEFG}, {F, ABCDEF, CDEG, ABG}, {G, ABCDEG, CDEF, ABF}.

b. 1: (1), aef, beg, abcd, abfg, cdfg, acdeg, bcdef; 2: ab, cd, fg, aeg, bef, acdef, bcdeg, abcdfg; 3: de, acg, adf, bcf, bdg, abce, cefg, abdefg; 4: ce, acf, adg, bcf, abde, defg, abcdefg.

49.  $SQA = 2,250$ ,  $SQB = 7,840$ ,  $SQC = 0,360$ ,  $SQAB = 52,563$ ,  $SQE = 10,240$ ,  $SQAC = 1,563$ ,  $SQAD = 7,563$ ,  $SQAD = 0,090$ ,  $SQAE = 4,203$ ,  $SQBC = 2,103$ ,  $SQBD = 0,010$ ,  $SQBE = 0,123$ ,  $SQCD = 0,010$ ,  $SQCE = 0,063$ ,  $SQDE = 4,840$ . Erro SQ = somas dos dois fatores SQs = 20,568. Erro QM = 2,057,  $F_{0,01;1;10} = 10,04$ , portanto somente o efeito principal D é significativo.

51. Fonte	gl	SQ	QM	f
Efeitos principais A	1	322,667	322,667	980,38
Efeitos principais B	3	35,623	11,874	36,08
Interação	3	8,557	2,852	8,67
Erros	16	5,266	0,329	
Total	23	372,113		

$F_{0,05;3;16} = 3,24$ , portanto as interações parecem presentes.

53. Fonte	gl	SQ	QM	f
A	1	30,25	30,25	6,72
B	1	144,00	144,00	32,00
C	1	12,25	12,25	2,72
AB	1	1.122,25	1.122,25	249,39
AC	1	1,00	1,00	0,22
BC	1	12,25	12,25	2,72
ABC	1	16,00	16,00	3,56
Erros	4	36,00	4,50	
Total	7			

Apenas o efeito principal para B e o efeito da interação AB são significantes em  $\alpha = 0,01$ .

55. a.  $\hat{\alpha}_1 = 9,00$ ,  $\hat{\beta}_1 = 2,25$ ,  $\hat{\delta}_1 = 17,00$ ,  $\hat{\gamma}_1 = 21,00$ ,  $(\hat{\alpha}\hat{\beta})_{11} = 0$ ,  $(\hat{\alpha}\hat{\delta})_{11} = 2,00$ ,  $(\hat{\alpha}\hat{\gamma})_{11} = 2,75$ ,  $(\hat{\beta}\hat{\delta})_{11} = 0,75$ ,  $(\hat{\beta}\hat{\gamma})_{11} = 0,50$ ,  $(\hat{\delta}\hat{\gamma})_{11} = 4,50$

b. Uma marcação de probabilidade normal sugere que os efeitos principais de A, C e D são importantes, e talvez a interação CD. Na realidade, agrupando a interação de 4 Sqs de fator três e a interação SQ de fator quatro para obter um SQE baseado em 5 gl e então construir

uma tabela ANOVA sugere que estes são os efeitos mais importantes.

57. Fonte	gl	SQ	QM	f	P
A	2	34,436	17,218	436,92	0,000
B	2	105,793	52,897	1.342,30	0,000
C	2	516,398	258,199	6.552,04	0,000
AB	4	6,868	1,717	43,57	0,000
AC	4	10,922	2,731	69,29	0,000
BC	4	10,178	2,545	64,57	0,000
ABC	8	6,713	839	21,30	0,000
Erros	27	1,064	39		
Total	53	692,372			

Todos os efeitos são significantes.

59. Baseado nos valores P da tabela ANOVA, os fatores estatisticamente importantes no nível  $\alpha = 0,01$  são do tipo adesivo e tempo de cura. O material condutor não tem um efeito estatisticamente significativo sobre a pressão de colagem. Não há interações significativas.

61. Fonte	gl	SQ	QM	f
A	4	285,76	71,44	0,594
B	4	227,76	56,94	0,473
C	4	2.867,76	716,94	5,958
D	4	5.536,56	1.384,14	11,502
Erros	8	962,72	120,34	$F_{0,05;4;8} = 3,84$
Total	24			

$H_{0A}$  e  $H_{0B}$  não podem ser rejeitadas enquanto  $H_{0C}$  e  $H_{0D}$  são rejeitadas.

## Capítulo 12

1. a. A tabela correspondente está baseada na repetição de cada valor de haste cinco vezes (a primeira para as folhas 0 e 1 e uma segunda para as folhas 2 e 3 etc.).

17	0	
17	2 3	
17	4 4 5	
17	6 7	
17		caule: centenas e dezenas
18	0 0 0 0 0 1 1	folha: unidades
18	2 2 2 2	
18	4 4 5	
18	6	
18	8	

Não há outliers nem lacunas significativas, e a distribuição é aproximadamente no formato de sino com um grau razoavelmente alto de concentração em torno de seu centro em aproximadamente 180.

0	8 8 9	
1	0 0 0 0	
1	3	
1	4 4 4 4	
1	6 6	
1	8 8 8 9	caule: unidades
2	1 1	folha: dezenas
2		
2	5	
2	6	
2		
3	0 0	

Um valor típico é em torno de 1,6, e há uma quantidade razoável de dispersão em torno deste valor. A distribuição é algo inclinada na direção dos valores grandes, e dois dos maiores podem ser candidatos a outliers.

b. Não, porque a observação com valores de x idênticos têm diferentes valores de y.

c. Não, porque os pontos não parecem cair todos próximos à reta ou curva simples.

3. Sim, Sim.

5. b. Sim.

c. Parece haver uma relação quadrática aproximada (os pontos caem próximos a uma parábola).

7. a. 5,050    b. 1,3    c. 130    d. -130

9. a. 0,095    b. -0,475    c. 0,830; 1,305  
d. 0,4207; 0,3446    e. 0,0036

11. a. -0,01; -0,10    b. 3,00; 2,50  
c. 0,3627    d. 0,4641

13. a. Sim, porque  $r^2 = 0,972$

15. a.

2	9
3	3 3 5 5 6 6 6 7 7 8 8 9
4	1 2 2 3 5 6 6 8 9
5	1
6	2 9
7	9
8	0

Valor típico é menor do que 40, uma quantidade razoável de variabilidade, inclinação positiva, dois potenciais outliers.

b. Não

c.  $y = 3,2925 + 0,10748x = 7,59$ . Não; perigo de extrapolação

d. 18,736; 71,605; 0,738, sim

17. a. Estimamos que 0,144 é a mudança esperada no conteúdo de cálcio associado com 1 mg/cm<sup>2</sup> de aumento na quantidade dissolvida de material. Aproximadamente, 86% da variação observada no conteúdo de cálcio pode ser atribuída à relação de regressão linear simples entre o conteúdo e a quantidade dissolvida de material.  
b. 10,88    c. 1,46

19. a.  $y = -45,5519 + 1,7114x$     b. 339,51  
c. -85,57    d. Os  $\hat{y}_s$  são 125,6; 168,4; 168,4; 211,1; 211,1; 296,7; 296,7; 382,3; 382,3; 467,9; 467,9; 553,4; 639,0; 639,0; uma reta de 45° através de (0, 0).

21. b.  $y = -2,182 + 0,660x$     c. 7,72    d. 7,72

23. a. 16.213,64; 16.205,45  
b. 414.235,71; sim, visto que  $r^2 = 0,961$ .

27.  $\hat{\beta}_1 = \frac{\sum x_i Y_i}{\sum x_i^2}$

29. Conjunto de dados

$r^2$	$s$	Mais efetivo: conjunto 3
1	0,43	4,03
2	0,99	4,03
3	0,99	1,90

Menos efetivo: conjunto 1

31. a. 0,001017    b. (-0,00956, -0,00516)

33. a. (0,081; 0,133)  
b.  $H_a: \beta_1 > 0,1$ , valor  $P = 0,277$ , não

35. a. (0,63; 2,44) é um IC de 95%.  
b. Sim.  $t \approx 3,6$ , valor  $P \approx 0,004$   
c. Não; extrapolação  
d. (0,54; 2,82), não

37. a. 0,110; 0,000262    b.  $t \approx 3,1$ , valor  $P \approx 0,014 < 0,10 = \alpha$ , portanto  $H_0$  deve ser rejeitada. Os dados parecem contradizer a hipótese principal.

39.  $f = 71,97$ ,  $s_{\beta_1} = 0,004837$ ,  $t = 8,48$ , valor  $P = 0,000$

43.  $d = 1,20$ ,  $gl = 13$ , e  $\beta \approx 0,1$ .

45. a. (77,80; 78,38)  
b. (76,90; 79,28), mesmo centro porém mais largo  
c. mais largo, visto que 115 é maior do que  $\bar{x}$   
d.  $t = -11$ , valor  $P = 0$

47. a. PI de 95% é (20,21; 43,69), não  
b. (28,53; 51,92), pelo menos 90%

49. (431,3; 628,5)

51. a. 0,40 está mais próximo de  $\bar{x} = 0,7495$   
b. (0,745; 0,875)  
c. (0,059; 0,523)

53. (a) mais estreito do que (b), (c) mais estreito do que (d), (a) mais estreito do que (c), (b) mais estreito do que (d)

55. a.  $x_2 = x_3 = 12$ , sendo  $y_2 \neq y_3$ .    b. Sim  
c.  $y = -19,670 + 3,2847x$     d. (24,93; 67,12)

57. Se, por exemplo, 18 é a idade mínima de elegibilidade, então para a maioria das pessoas  $y \approx x - 18$ .

59. a. 0,966

b. A porcentagem do peso da fibra seca para a primeira amostra tende a ser maior do que para a segunda.

c. Sem mudança    d. 93,3%

e.  $t = 14,9$ , valor  $P \approx 0$ , portanto parece haver tal relação.

61. a.  $r = 0,748$ ,  $t = 3,9$ , valor  $P = 0,001$ . Usando tanto  $\alpha = 0,05$  ou  $0,01$ , sim.  
b. 0,560 (56%), igual

63.  $r = 0,773$ , sendo  $t = 2,44 < 2,776$ ; portanto  $H_0: \rho = 0$  não pode ser rejeitada.

65. a. A marcação  $x$  é um pouco curvada, mas não influencia, portanto, em muito no tamanho pequeno da amostra. A marcação  $x$  é quase retilínea.    b.  $t = 6,3 \geq 3,355$ , portanto, parece haver uma relação linear.

67. a. Rejeita  $H_0$

b. Não. valor  $P = 0,00032 \Rightarrow z \approx 3,6 \Rightarrow r \approx 0,16$ , o que indica apenas uma relação fraca.

c. Sim, mas  $n \Rightarrow \rho \approx 0,022$ , muito grande, portanto sem significância prática.

69. a.  $t = -1,24 > -2,201$ , portanto  $H_0$  não pode ser rejeitada.

b. 0,970

71. a. 0,507    b. 0,712    c. valor  $P = 0,0013 < 0,01 = \alpha$ , portanto rejeita  $H_0: \beta_1 = 0$  e conclui que há uma relação linear útil.    d. Um IC de 95% é (1,056; 1,275).    e. 1,0143; 0,2143

73. a.  $y = 14,1904 - 0,14892x$     b.  $t = -1,43$ , portanto, não rejeita  $H_0: \beta_1 = -0,10$ .    c. Não;  $\sum(x - \bar{x})^2 = 143$ , neste caso, e 182 para os dados fornecidos.    d. Um IC de 95% para  $\mu_{y=28}$  é (9,599; 10,443).

75. a. Uma relação linear substancial  
 b.  $y = -0,08259 + 0,044649x$   
 c. 98,3%  
 d. 0,7702, -0,0902 e. Sim;  $t = 19,96$   
 f. (0,0394; 0,0499) g. (0,762; 0,858)  
 79. b. 0,573  
 85.  $t = -1,14$ , portanto é plausível que  $\beta_1 = \gamma_1$ .

## Capítulo 13

1. a. 6,32; 8,37; 8,94; 8,37 e 6,32 b. 7,87; 8,49; 8,83; 8,94 e 2,83 c. O desvio é provavelmente muito menor para os valores de  $x$  da parte (b).
3. a. Sim. b. -0,31; -0,31; 0,48; 1,23; -1,15; 0,35; -0,10; -1,39; 0,82; -0,16; 0,62; 0,09; 1,17; -1,50; 0,96; 0,02; 0,65; -2,16; -0,79; 1,74. Neste caso,  $e/e^*$  varia entre 0,57 e 0,65, portanto  $e^*$  está próximo de  $e/s$ .  
 c. Não.
5. a. Cerca de 98% da variação observada na espessura é explicada pela relação.  
 b. Uma relação não-linear
7. a. Não. b.  $e/s$  são -16,60; 9,70; 19,00; -0,70; 11,40;  $e^*/s$  são -1,55; 0,68; 1,25; -0,05; -1,06; uma função quadrática.
9. Para o conjunto 1, uma regressão linear simples é apropriada. Uma regressão quadrática é razoável para o conjunto 2. No conjunto 3 (13; 12,74) parece muito inconsistente com os dados remanescentes. O coeficiente angular estimado para o conjunto 4 depende muito da observação simples (19; 12,5) e não há evidência de uma relação linear.
11. c.  $V(\hat{Y}_i)$  aumenta e  $V(Y_i - \hat{Y}_i)$  diminui.
13.  $t$  com  $n - 2$  gl; 0,02
15. a. Um padrão curvado b. Um padrão linear  
 c.  $Y = \alpha x^\beta \cdot \epsilon$  d. Um IP de 95% é (3,06; 6,50).  
 e. Um resíduo padronizado correspondendo à terceira observação é um pouco grande. Há apenas dois resíduos padronizados positivos, mas os outros dois são essencialmente 0. Os padrões em uma marcação de resíduo padronizada e uma marcação de probabilidade são marginalmente aceitáveis.
17. a.  $\sum x'_i = 15,501$ ,  $\sum y'_i = 13,352$ ,  $\sum (x'_i)^2 = 20,228$ ,  $\sum x'_i y'_i = 18,109$ ,  $\sum (y'_i)^2 = 16,572$ ,  $\hat{\beta}_1 = 1,254$ ,  $\hat{\beta}_0 = -0,468$ ,  $\hat{\alpha} = 0,626$ ,  $\hat{\beta} = 1,254$  c.  $t = -1,07$ , portanto, não rejeita  $H_0$ . d.  $H_0: \beta = 1$ ,  $t = -4,30$ , portanto, rejeita  $H_0$ .
19. a. Não b.  $Y' = \beta_0 + \beta_1 \cdot (1/t) + \epsilon'$ , onde  $Y' = \ln(Y)$ , portanto,  $Y = \alpha e^{\beta t} \cdot \epsilon$ . c.  $\hat{\beta} = \hat{\beta}_1 = 3,735,45$ ,  $\hat{\beta}_0 = -10,2045$ ,  $\hat{\alpha} = (3,70034) \cdot (10^{-5})$ ,  $\hat{y}' = 6,7748$ ,  $\hat{y} = 875,5$  d.  $SQE = 1,39587$ ,  $SQPE = 1,36594$  (usando valores transformados),  $f = 0,33 < 8,68 = F_{0,01;1;15}$ , portanto não rejeita  $H_0$ .
21. a.  $\hat{\mu}_{Y,x} = 18,14 - 1,485/x$  b.  $\hat{y} = 15,17$
23. Para o modelo exponencial,  $V(Y|x) = \alpha^2 e^{2\beta x} \sigma^2$ , que depende de  $x$ . Um resultado semelhante satisfaz o modelo de potência.
25. a. O ponto estimado de  $\beta_1$  é  $\hat{\beta}_1 = 0,1772$  e a razão ímpar estimada é 1,194.  $H_0: \beta_1 = 0$  é rejeitada em favor da conclusão de que a experiência parece afetar a probabilidade de desempenho com sucesso da tarefa.
27. b. 52,88; 0,12 c. 0,895 d. Não  
 e. (48,54; 57,22) f. (42,85; 62,91)
29. a.  $SQE = 103,37$ ,  $s^2 = 51,69$  b. 0,961  
 c.  $t = -3,83 > -4,303$ , portanto  $H_0: \beta_2 = 0$  não pode ser rejeitada. Os dados não pedem rigorosamente pela inclusão do termo quadrático (talvez uma consequência do  $n$  ser pequeno;  $H_0$  seria rejeitada no nível de significância 0,10). d.  $\beta_1: (-19,87; 35,99)$   $\beta_2: (-5,18; 1,50)$  e. (100,08; 129,34) f. 2,19
31. a.  $13,636 + 11,406x - 1,7155x^2$   
 b. Sim. Sim, (6, 20) c. 2,040; 0,947. O teste  $F$  do modelo utilitário via MINITAB resulta em  $f = 35,9$ , valor  $P = 0,003$ , sugerindo um modelo útil.  
 d. Sim, sim e. (28,35; 35,28)
33. a. 0,9671; 0,9407  
 b.  $0,0000492x^3 - 0,000446058x^2 + 0,007290688x + 0,96034944$  c.  $t = 2 < 3,182 = t_{0,025;3}$ , portanto o termo cúbico deveria ser deletado. d. Idêntico  
 e. 0,987; 0,994, sim
35.  $\hat{y} = 7,6883e^{0,1799x - 0,0022x^2}$
37. a. 4,9 b. Quando o número de entregas é mantido fixo, a mudança média no tempo de viagem associada ao aumento de uma milha na distância percorrida é 0,060 h. Quando a distância percorrida é mantida fixa, a mudança média no tempo de viagem associado com uma entrega extra é 0,900 h c. 0,9861
39. a. 77,3 b. 40,4
41.  $f = 24,4 > 5,12 = F_{0,001;6;30}$ , portanto o valor  $P \ll 0,001$ . O modelo mostrado parece ser útil.
43. a. 48,31; 3,69 b. Não. Se  $x_1$  aumenta tanto  $x_3$  quando  $x_2$  deverão mudar. c. Sim, visto que  $f = 18,924$ , o valor  $P = 0,001$ . d. Sim, usando  $\alpha = 0,01$ , visto que  $t = 3,496$  e valor  $P = 0,003$ .
45. a.  $f = 87,6$ , valor  $P = 0$ , parece haver uma relação linear útil entre  $y$  e pelo menos um dos previsores.  
 b. 0,935 c. (9,095; 11,087)

47. **b.** valor  $P = 0,000$ , portanto conclui-se que o modelo é útil. **c.** valor  $P = 0,034 \leq 0,05 = \alpha$ , portanto rejeita  $H_0$ ;  $\beta_3 = 0$ ; % de lixo parece fornecer informação adicional útil. **d.** (1.479,8; 1.531,1), precisão razoável **e.** Um IP de 95% é (1.435,7; 1.575,2).
49. **a.** 96,8303; -5,8303 **b.**  $f = 14,9 \geq 8,02 = F_{0,05;2,9}$ , portanto rejeita  $H_0$  e conclui que o modelo é útil. **c.** (78,28; 115,38) **d.** (38,50; 155,16) **e.** (46,91; 140,66) **f.** Não. valor  $P = 0,208$ , portanto  $H_0$ ;  $\beta_1 = 0$  não pode ser rejeitada.
51. **a.** Não **b.**  $f = 5,04 \geq 3,69 = F_{0,05;5,8}$ . Parece que há uma relação linear útil. **c.** 6,16; 3,304, (16,67; 31,91) **d.**  $f = 3,44 < 4,07 = F_{0,05;3,8}$ , portanto  $H_0$ ;  $\beta_3 = \beta_4 = \beta_5 = 0$  não pode ser rejeitada. Os termos quadráticos podem ser deletados.
55. **a.** A variável dependente é  $\ln(q)$ , e os preditores são  $x_1 = \ln(a)$  e  $x_2 = \ln(b)$ ;  $\hat{\beta} = \hat{\beta}_1 = 0,9450$ ,  $\hat{\gamma} = \hat{\beta}_2 = 0,1815$ ,  $\hat{\alpha} = 4,7836$ ,  $\hat{q} = 18,27$ .  
**b.** Agora regride  $\ln(q)$  contra  $x_1 = a$  e  $x_2 = b$ .  
**c.** (1,24; 5,78)
57. 

$k$	$R^2$	$R^2$ ajustado	$C_k$
1	0,676	0,647	138,2
2	0,979	0,975	2,7
3	0,9819	0,976	3,2
4	0,9824		4

**a.** O modelo com  $k = 2$  **b.** Não
59. O modelo com preditores  $x_1, x_3, e x_5$
61. Não. Todos valores de  $R^2$  são muito menores do que 0,9.
63. O impacto destas duas observações deve ser melhor investigado. Não inteiramente. A eliminação da observação de número 6 seguida pela re-regressão deve também ser considerada.
65. **a.** As duas distribuições têm quantidades semelhantes de variabilidade, são ambas razoavelmente simétricas e não contêm outliers. A principal diferença é que a mediana dos valores de rachadura é em torno de 840, enquanto ela vale cerca de 480 para valores de não-rachadura. Um IC de 95% para a diferença entre as médias é (132; 557).  
**b.**  $r^2 = 0,577$  para o modelo de regressão linear simples, valor  $P$  para o modelo utilitário = 0, mas um resíduo padronizado é -4,11! Incluindo um indicador para rachadura/não rachadura não fornece o ajuste nem inclui um indicador e preditor de interação.
67. **a.** Quando sexo, peso e taxa de batida cardíaca são mantidos fixos, estimamos que a mudança média em  $VO_2$  máximo associado a um aumento de 1-minuto no tempo de caminhada é -0,0996. **b.** Quando peso, tempo de caminhada e taxa de batida cardíaca são mantidos fixos, a estimativa da diferença da média entre  $VO_2$  máximo para homens e mulheres é 0,6566. **c.** 3,669, -0,519 **d.** 0,706 **e.**  $f = 9,0 \geq 4,89 = F_{0,01;4;15}$ , portanto parece haver uma relação útil.
69. **a.** Não. Há uma curvatura substancial na marcação de dispersão. **b.** A regressão cúbica produz  $R^2 = 0,998$  e um IP de 95% de (261,98; 295,62), e o preditor cúbico parece ser importante (valor  $P = 0,001$ ). Uma regressão de  $y$  versus  $\ln(x)$  tem  $r^2 = 0,991$ , mas há um resíduo padronizado muito grande e a marcação do resíduo padronizado não é satisfatória.
71. **a.**  $R^2 = 0,802$ ,  $f = 21,03$ , valor  $P = 0,000$ . pH é um candidato à deleção. Observe que há um resíduo padronizado extremamente grande. **b.**  $R^2 = 0,920$ ,  $R^2$  ajustado = 0,774,  $f = 6,29$ , valor  $P = 0,002$  **c.**  $f = 1,08$ , valor  $P > 0,10$ , não rejeita  $H_0$ ;  $\beta_6 = \dots = \beta_{20} = 0$ . O grupo de preditores de segunda ordem não parece ser útil. **d.**  $R^2 = 0,871$ ,  $f = 28,50$ , valor  $P = 0,000$ , e valor  $P = 0,000$  e agora todos os seis preditores são julgados importantes (o maior valor  $P$  para qualquer razão  $t$  é 0,016); a importância de  $pH^2$  foi mascarada no teste de (c). Observe que há dois resíduos padronizados muito grandes.
73. **a.**  $f = 1,783$ , portanto o modelo parece útil.  
**b.**  $t = -48,1 \leq -6,689$ , portanto mesmo no nível 0,001 o preditor quadrático deveria ser retido.  
**c.** Não **d.** (21,07; 21,65) **e.** (20,67; 22,05)
75. **a.**  $f = 30,8 \geq 9,55 = F_{0,01;2;7}$ , portanto o modelo parece útil. **b.**  $t = -7,69$  e valor  $P < 0,001$ , portanto retém o preditor quadrático. **c.** (44,01; 47,91)
77. **a.** 231,75 **b.** 0,903 **c.**  $f = 41,9$ , indicando uma relação útil. **d.** (220,9; 238,1)
79. Há várias escolhas razoáveis em cada caso.
81. **a.**  $f = 106$ , valor  $P \approx 0$  **b.** (0,014; 0,068)  
**c.**  $t = 5,9$ , rejeita  $H_0$ ;  $\beta_4 = 0$ , porcentual de pessoas não brancas parece ser importante.  
**d.** 99,514,  $y - \hat{y} = 3,486$

## Capítulo 14

1. **a.** Rejeita  $H_0$ . **b.** Não rejeita  $H_0$ .  
**c.** Não rejeita  $H_0$ . **d.** Não rejeita  $H_0$ .
3. Não.  $\chi^2 = 1,57$  e valor  $P > 0,10$ , portanto  $H_0$  não pode ser rejeitada.
5.  $\chi^2 = 6,61 < 14,684 = \chi_{0,10;9}^2$ , portanto não rejeita  $H_0$ .
7.  $\chi^2 = 4,03$  e valor  $P > 0,10$ , portanto não rejeita  $H_0$ .
9. **a.** [0; 0,2231), [0,2231; 0,5108), [0,5108; 0,9163), [0,9163; 1,6094) e [1,6094;  $\infty$ ) **b.**  $\chi^2 = 1,25 < \chi_{\alpha;4}^2$  para qualquer  $\alpha$  razoável, portanto a distribuição exponencial especificada é bem plausível.



11. a.  $(-\infty; -0,97)$ ,  $[-0,97; -0,43]$ ,  $[-0,43; 0)$ ,  $[0; 0,43)$ ,  $[0,43; 0,97)$ , e  $[0,97; \infty)$  b.  $(-\infty; 0,49806)$ ,  $[0,49806; 0,49914)$ ,  $[0,49914; 0,5)$ ,  $[0,5; 0,50086)$ ,  $[0,50086; 0,50194)$  e  $[0,50194; \infty)$  c.  $\chi^2 = 5,53$ ,  $\chi^2_{0,10;5} = 9,236$ , portanto o valor  $P > 0,10$  e a distribuição normal especificada é plausível.
13.  $\hat{p} = 0,0843$ ,  $\chi^2 = 280,3 > \chi^2_{\alpha,1}$  para qualquer  $\alpha$  tabulado, portanto o modelo fornece um ajuste insatisfatório.
15. A verossimilhança é proporcional a  $\theta^{233}(1 - \theta)^{367}$ , desde que  $\hat{\theta} = 0,3883$ . As contagens estimadas esperadas são 21,00; 53,33; 50,78; 21,50 e 3,41. Combinando as células 4 e 5,  $\chi^2 = 1,62$ , portanto não rejeita  $H_0$ .
17.  $\hat{\lambda} = 3,167$ , a partir do qual  $\chi^2 = 103,98 \gg \chi^2_{\alpha,k-1} = \chi^2_{\alpha,7}$  para qualquer  $\alpha$  tabulado, portanto a distribuição de Poisson fornece um ajuste muito insatisfatório.
19.  $\hat{\theta}_1 = (2n_1 + n_3 + n_5)/2n = 0,4275$ ,  $\hat{\theta}_2 = 0,2750$ ,  $\chi^2 = 29,1$ ,  $\chi^2_{0,01;3} = 11,344$ , portanto rejeita  $H_0$ .
21. Sim. A hipótese nula de uma distribuição de população normal não pode ser rejeitada.
23. O MINITAB fornece  $r = 0,967$ , e como  $c_{0,10} = 0,9707$  e  $c_{0,05} = 0,9639$ ,  $0,05 < \text{valor } P < 0,10$ . Utilizando  $\alpha = 0,05$ , a normalidade é julgada plausível.
25.  $\chi^2 = 23,18 \geq 13,277 = \chi^2_{0,01;4}$ , portanto  $H_0$  é rejeitada. As proporções parecem ser diferentes.
27. Sim.  $\chi^2 = 44,98$  e valor  $P < 0,001$ .
29.  $p_{ij}$  = proporção da  $j$ -ésima combinação de sexo resultante do  $i$ -ésimo genótipo.  $\chi^2 = 6,46$ , portanto o valor  $P > 0,10$  e a hipótese nula de homogeneidade não pode ser rejeitada.
31. Sim.  $\chi^2 = 14,15$ , portanto  $0,025 < \text{valor } P < 0,03$  e  $H_0$  deve ser rejeitada no nível de significância 0,05.
35.  $N_{ij}/n, n_k N_{ij}/n, 24$
37.  $\chi^2 = 3,65 < 5,992 = \chi^2_{0,05;2}$ , portanto  $H_0$  não pode ser rejeitada.
39. Sim.  $\chi^2 = 131$  e valor  $P < 0,001$ .
41.  $\chi^2 = 22,4$  e valor  $P < 0,001$ , portanto a hipótese nula de independência é rejeitada.
43. Valor  $P = 0$ , portanto a hipótese nula de homogeneidade é rejeitada.
47. a. Valor do teste estatístico = 19,2, valor  $P = 0$ .  
 b. Evidência, na melhor hipótese relação de uma semana; valor do teste estatístico = -2,13  
 c. Valor do teste estatístico = -0,98, valor  $P > 0,10$   
 d. Valor do teste estatístico = 3,3;  $0,01 < \text{valor } P < 0,05$

## Capítulo 15

1.  $s_+ = 35$  e  $14 < 35 < 64$ , portanto  $H_0$  não pode ser rejeitada.
3.  $s_+ = 18 \leq 21$ , portanto  $H_0$  é rejeitada.
5. Rejeita  $H_0$  se tanto  $s_+ \geq 64$  ou  $s_+ \leq 14$ . Porque  $s_+ = 72$ ,  $H_0$  é rejeitada.
7.  $s_+ = 442,5$ ,  $z = 2,89 \geq 1,645$ , de modo que rejeita  $H_0$ .
9. 

$d$	0	2	4	6	8	10	12	14	16	18	20
$p(d)$	$\frac{1}{24}$	$\frac{3}{24}$	$\frac{1}{24}$	$\frac{4}{24}$	$\frac{2}{24}$	$\frac{2}{24}$	$\frac{2}{24}$	$\frac{4}{24}$	$\frac{1}{24}$	$\frac{3}{24}$	$\frac{1}{24}$
11.  $w = 37$  e  $29 < 37 < 61$ , portanto  $H_0$  não pode ser rejeitada.
13.  $z = 2,27 < 2,58$ , portanto  $H_0$  não pode ser rejeitada. Valor  $P \approx 0,023$
15.  $w = 39 \leq 41$ , portanto  $H_0$  é rejeitada.
17.  $(\bar{x}_{(5)}, \bar{x}_{(32)}) = (11,15; 23,80)$
19.  $(-13,0; -6,0)$
21.  $(d_{ij(5)}; d_{ij(21)}) = (16, 87)$
23.  $k = 14,06 \geq 6,251$ , portanto rejeita  $H_0$ .
25.  $k = 9,23 \geq 5,992$ , portanto rejeita  $H_0$ .
27.  $f_r = 2,60 < 5,992$ , portanto não rejeita  $H_0$ .
29.  $f_r = 9,62 > 7,815 = \chi^2_{0,05;3}$ , portanto rejeita  $H_0$ .
31.  $(-5,9, -3,8)$
33. a. 0,021      b.  $c = 14$  igual a  $\alpha = 0,058$ ;  $y = 12$ , portanto  $H_0$  não pode ser rejeitada.
35.  $w' = 26 < 27$ , portanto não rejeita  $H_0$ .

## Capítulo 16

1. Todos os pontos do gráfico caem entre os limites de controle.
3. 0,9802; 0,9512; 53
5. a. 0,0301      b. 0,2236      c. 0,6808
7. LCI = 12,20, LCS = 13,70. Não.
9. LCI = 94,91, LCS = 98,17. Parece haver um problema no 22º dia.
11. a. 200      b. 4,78      c. 384,62 (maior), 6,30 (menor)

13. LCI = 12,37, LCS = 13,53
15. a. LCI = 0, LCS = 6,48  
b. LCI = 0,48, LCS = 6,60
17. LCI = 0,045, LCS = 2,484. Sim, desde que todos os pontos estejam dentro dos limites de controle.
19. a. LCI = 0,105, LCS = 0,357  
b. Sim, desde que  $0,39 > LCS$ .
21.  $\bar{p} > 3/53$
23. LCI = 0, LCS = 10,1
25. Quando a área = 0,6, LCI = 0 e LCS = 14,6; quando a área = 0,8, LCI = 0 e LCS = 13,4; quando a área = 1,0, LCI = 0 e LCS = 12,6.
27. 

$l:$	1	2	3	4	5	6	7	8
$d_i:$	0	0,001	0,017	0	0	0,010	0	0
$e_i:$	0	0	0	0,038	0	0	0	0,054
$l:$	9	10	11	12	13	14	15	
$d_i:$	0	0,024	0,003	0	0	0	0,005	
$e_i:$	0	0	0	0,015	0	0	0	

  
Não há sinais fora de controle.
29.  $n = 5, h = 0,00626$
31. Probabilidades hipergeométricas (calculadas em uma calculadora HP21S) são 0,9919; 0,9317; 0,8182; 0,6775; 0,5343; 0,4047; 0,2964; 0,2110; 0,1464 e 0,0994, enquanto as probabilidades binomiais correspondentes são 0,9862; 0,9216; 0,8108; 0,6767; 0,5405; 0,4162; 0,3108; 0,2260; 0,1605 e 0,1117. A aproximação é satisfatória.
33. 0,9206; 0,6767; 0,4198; 0,2321; 0,1183; o plano com  $n = 100, c = 2$  é preferível.
35. 0,9981; 0,5968 e 0,0688
37. a. 0,010; 0,018; 0,024; 0,027; 0,027; 0,025; 0,022; 0,018; 0,014; 0,011    b. 0,0477; 0,0274    c. 77,3; 202,1; 418,6; 679,9; 945,1; 1.188,8; 1.393,6; 1.559,3; 1.686,1; 1.781,6
39. O gráfico  $\bar{x}$  é baseado nos desvios padrão da amostra: LCI = 402,42, LCS = 442,20. O gráfico  $\bar{x}$  é baseado nas variações de amostra: LCI = 402,36, LCS = 442,26 gráfico S: LCI = 0,55, LCS = 30,37. O gráfico R: LCI = 0, LCS = 82,75.
41. O gráfico S: LCI = 0, LCS = 2,3020; como  $s_{21} = 2,931 > LCS$ , o processo parece estar fora de controle neste momento. Como uma causa assinalável é identificada, recalculamos os limites depois da deleção: para um gráfico S, LCI = 0 e LCS = 2,0529; para um gráfico  $\bar{x}$ , LCI = 48,583 e LCS = 51,707. Todos os pontos em ambos os gráficos caem entre os limites de controle.
43.  $\bar{\bar{x}} = 430,65, s = 24,2905$ ; para um gráfico S, LCS = 62,43 quando  $n = 3$  e LCS = 55,11 quando  $n = 4$ ; para um gráfico  $\bar{x}$ , LCI = 383,16 e LCS = 478,14 quando  $n = 3$  e LCI = 391,09 e LCS = 470,21 quando  $n = 4$ .



# Índice

- Adequação do modelo à regressão, 484-488, 524-525  
ajustado, 504, 516
- Ajustamento exponencial, 44
- Amostra, 2  
aleatória simples, 7, 200  
conveniência, 7  
estratificada, 7
- Amostra aleatória, 200
- Amostra aleatória simples, 7
- Amostragem de aceitação, 632-637
- Amostragem estratificada, 7
- Amostragem reduzida, 636
- Amplitude, 32
- Amplitude do meio, 43
- Análise de dados qualitativos, 550-579
- Análise de resíduos, 484-488
- Análise de variância, 357  
em regressão, 459  
fator único (unidirecional), 357-384  
fatores múltiplos, 385-431  
livre de distribuição, 598-601
- ANOVA de dois fatores, 386-406
- ANOVA fator único, 357
- ANOVA multifatores, 385-431
- ANOVA três fatores, 406-415
- ANOVA unidirecional, 357
- Associação, 196
- Atribuindo análogos, 423
- Axiomas de probabilidade, 50
- Bloco principal, 422
- Blocos, 392, 394, 420
- Boxplot, 35  
comparativo, 37-38
- Caudas leves, 168
- Caudas pesadas, 168
- Caules repetidos, 12
- Causa atribuível, 607
- Causalidade, 196
- Censo, 2, 32, 228
- Centralização de valores  $x$ , 507
- Coefficiente de correlação  
da amostra, 468  
de estimativa por pontos. *Veja* estimativa por pontos  
de população, 471  
de teste de hipóteses. *Veja* Teste de hipóteses  
de variáveis aleatórias, 194  
do intervalo de confiança. *Veja* intervalo de confiança múltiplo, 516
- Coefficiente de correlação múltipla, 516
- Coefficiente de determinação múltipla, 504, 517  
ajustado, 504, 511, 536
- Coefficiente de determinação, 448
- Coefficiente de variação, 41
- Coefficientes de autocorrelação, 477
- Coincidências, 82
- Comando de carga  
estimativa do erro padrão, 230  
intervalo de confiança, 252-253, 267
- Combinação, 62
- Combinação linear, 213  
de variáveis normais, 214  
valor esperado da, 213  
variância de, 213
- Comparações múltiplas, 367, 377-378, 391, 402
- Complemento de um evento, 47
- Componente da variância, 395
- Computadores, 3
- Confusão, 420-422
- Contagens de células  
esperada estimada, 559, 570, 572  
esperada, 551  
observada, 551, 569
- Contraste, 384, 415
- Correção de continuidade, 148
- Covariância, 192
- Curva característica da operação, 633
- Curva de densidade, 129
- Curva de Gompertz, 501
- Dados, 2  
bivariados, 2  
de censos, 32, 228  
multivariáveis, 3, 19  
qualitativos (categorizados), 19, 29  
univariáveis, 2
- Dados atributos, 6
- Dados em pares, 332, 337-338, 584
- Deming, W. E., 6
- Desigualdade de Chebyshev, 104
- Desigualdade de Jensen, 177
- Desvio, 224
- Desvio aleatório, 436
- Desvio padrão  
amostra, 33  
estimativa por pontos. *Veja* Estimativa por pontos

- intervalo de confiança. *Veja* Intervalo de confiança
- população, 34
- teste de hipóteses. *Veja* Teste de hipóteses
- variável aleatória contínua, 138
- variável aleatória discreta, 102
- Desvios em relação à média, 32
- Determinação do tamanho da amostra
  - intervalo de confiança, 250, 257, 321
  - teste de hipóteses, 288, 292, 298, 318, 344
- Diagrama de caule e folha, 10
  - comparativo, 21
- Diagrama de Pareto, 24
- Diagrama de Venn, 48
- Diagramas de árvore, 59-60
- Dispersão entre os quartos, 35
- Distribuição beta, 162
- Distribuição binomial negativa, 116
  - Estimativa por pontos, 236
- Distribuição binomial, 108-111
  - aproximação de Poisson para, 119
  - aproximação normal para, 148
  - aproximação para hipergeométrica, 115
- tabelas, 109, 640-641
- Distribuição com intervalo para student, 367
  - tabela, 367, 656
- Distribuição da amostragem, 197
  - da média da amostra, 207-209
- Distribuição de Cauchy, 228
- Distribuição de Erlang, 157
- Distribuição de freqüência, 13
- Distribuição de Pareto, 132
- Distribuição de Poisson, 119
  - aproximação normal para, 212
  - aproximação para binomial, 120
  - em gráficos de controle, 622
  - estimativa por pontos, 233, 238
  - intervalo de confiança, 260
  - relação com a distribuição exponencial, 155
  - tabela, 120, 642-643
  - teste de hipóteses, 311, 561
- Distribuição de probabilidade conjunta, 179, 181, 185
- Distribuição de probabilidades (*Veja também listagens de distribuições individuais*)
  - Bernoulli, 92
  - beta, 162
  - binomial, 108-111
  - binomial negativa, 116
  - Cauchy, 228
  - condicional, 188
  - conjunta contínua, 181, 185
  - conjunta discreta, 179, 185
  - contínua, 129
  - da média da amostra, 207-209
  - de combinação linear, 213-214
  - de uma função, 176
  - discreta, 88-89
  - do total da amostra, 207-209
  - Erlang, 157
  - exponencial deslocada, 243
  - exponencial, 154
  - $F$ , 296, 362
  - família, 91
  - gama, 152
  - geométrica, 117
  - hiperexponencial, 175
  - hipergeométrica, 113-115, 190
  - Intervalo da distribuição de Student, 367
  - lognormal, 161
  - multinomial, 186
  - normal, 141
  - normal bivariada, 472
  - normal padrão, 142
  - Pareto, 132
  - Poisson, 119
  - qui-quadrado, 156
  - Rayleigh, 132
  - simétrica, 136
  - $t$ , 261
  - uniforme, 130
  - valor extremo, 170, 175
  - Weibull, 158
- Distribuição de Rayleigh, 132, 234
- Distribuição de valor extremo, 170, 175
- Distribuição de Weibull, 158
  - estimativa por pontos, 239
  - gráfico de probabilidade, 171
- Distribuição exponencial, 154, 243
  - estimativa por pontos, 228, 231, 237, 243
  - gráfico de probabilidade, 173
  - intervalo de confiança, 251-252
  - relação com a distribuição de Poisson, 155
  - teste de hipóteses, 312
- Distribuição  $F$ , 348, 361
  - tabela, 348, 361, 650, 655-658
- Distribuição gama, 153
  - estimativa pontual, 235
  - padrão, 153
- Distribuição geométrica, 117
- Distribuição hiperexponencial, 175
- Distribuição hipergeométrica, 113-115
  - aproximação binomial para, 116
  - multivariada, 190
- Distribuição inclinada, 168-169
- Distribuição lognormal, 161
  - de um produto, 211
- Distribuição multinomial, 186
- Distribuição normal, 141
  - aproximação para binomial, 148
  - aproximação para distribuições discretas, 147
  - aproximação para gama, 212
  - aproximação para Poisson, 212
  - bivariada, 472
  - estimativa por pontos, 227, 238
  - gráfico de probabilidade, 167, 425
  - intervalo de confiança, 249, 263, 271
  - intervalo de previsão, 265
  - intervalo de tolerância, 266
  - padrão, 142
  - percentil, 143, 146
  - tabela de padrão, 142, 643-644
  - teste baseado no gráfico de probabilidade, 565
  - teste de hipóteses, 286, 291, 312, 316-317
- Distribuição normal bivariada, 472
- Distribuição normal padrão, 142
  - tabela, 142
- Distribuição qui-quadrado, 156, 269, 552
  - tabela, 270, 552, 647, 657-658
  - valor crítico, 270, 552
- Distribuição simétrica, 136
- distribuição  $t$ , 261
  - tabela, 262, 645, 648-649
- Distribuição uniforme, 130
  - estimativa por pontos, 241
  - intervalo de confiança, 273
- distribuição  $z$ , 142
- Distribuição, de probabilidade. *Veja* Distribuição de probabilidade
- Distribuições de Bernoulli, 92
- Dotplot, 12

- Duração média da execução, 613, 629  
 Duração média da operação, 613, 629
- Efeitos
- aleatório, 379
    - fixo, 379
    - misturado, 394-395
    - principal, 399
  - Efeitos aleatórios, 379, 394
  - Efeitos fixos, 379
  - Efeitos misturados, 394-295
  - Efeitos principais, 399
  - Eficiência assintótica relativa, 587, 593
  - Eliminação retroativa, 538
  - Empates, 585, 591
  - Equação do modelo
    - ANOVA, 373,-374, 379, 389, 394-395, 399, 403, 407, 411, 415
    - regressão, 487, 502, 511
  - Equações normais, 441, 503, 514
  - Erro aleatório, 436
  - Erro de soma dos quadrados, 363, 376, 445, 504
  - Erro do quadrado médio, 244, 535
  - Erro tipo I, 279
  - Erro tipo II, 279
    - teste de livre distribuição, 587, 593
    - Teste  $F$  ANOVA, 374
    - teste  $t$ , 292-293, 328-329, 460, 664
    - teste  $z$ , 288, 290, 298, 317, 343
  - Erro total esperado normalizado da estimativa, 536
  - Erro-padrão, 229
  - Erros em testes de hipóteses
    - tipo I, 278
    - tipo II, 278
  - Escala de densidade, 16
  - Espaço amostral, 44
  - Estatística, 3-4
    - descritiva, 3
    - inferência, 4
  - Estatística descritiva, 3
  - Estatística do teste, 278
  - Estatístico, 179
    - teste, 278
  - Estimador combinado, 328
  - Estimador consistente, 243
  - Estimador de Hodges-Lehmann, 244
  - Estimador de momentos, 236
  - Estimador por pontos, 221
    - carga de entrada, 230
    - com desvio, 224
    - consistente, 243
    - erro-padrão, 229
    - erro-padrão estimado, 229
    - máxima verossimilhança, 237
    - método dos momentos, 235
    - robusto, 228
    - sem desvio, 224
    - sem desvio de mínima variância, 227
  - Estimador por pontos. *Veja* Estimador por pontos
  - Estimador robusto, 228
  - Estimador sem desvio, 224
  - Estimador sem desvio de mínima variância, 226
  - Estimativa com desvio, 224, 228
  - Estimativa de máxima verossimilhança, 237-242
    - e testes de qualidade do ajustamento, 558
  - Estimativa  $M$ , 242
  - Estimativa por pontos
    - coeficiente angular da reta de regressão, 441
    - coeficiente de correlação, 471
    - combinação linear das médias, 371
    - desvio padrão, 223, 226, 240
    - diferença entre médias, 315
    - diferença entre proporções, 340
    - distribuição binomial negativa, 236
    - Distribuição de Cauchy, 228
    - distribuição de Poisson, 233, 238
    - Distribuição de Rayleigh, 234
    - distribuição de Weibull, 239
    - distribuição exponencial, 228, 231, 237
    - distribuição gama, 235
    - distribuição lognormal, 233
    - distribuição normal, 222, 227, 238
    - distribuição simétrica, 226
    - distribuição uniforme, 241
    - efeitos de ANOVA, 390, 406-407
    - em regressão, 446
    - Estimador de Hodges-Lehmann, 244
    - estimador  $M$ , 242
    - estimador combinado, 328
    - média, 222, 226-227
    - média da variável dependente em regressão, 443
    - método da máxima verossimilhança, 237
    - método dos mínimos quadrados, 440-441
    - método dos momentos, 235
    - não-tendenciosa, 224
    - não-tendencioso de mínima variância, 227
    - parâmetros de regressão, 441, 514
    - princípio de invariância, 240
    - proporção, 222, 224
    - robusta, 228
    - tamanho da população, 241
    - técnica de resposta aleatorizada, 235
    - Termo constante da reta de regressão, 441
    - variância, 223, 226, 328
  - Estimativa por pontos, 221
  - Estimativa por pontos. *Veja* Estimativa por pontos
  - Estrutura de amostragem, 6
  - Estudo analítico, 6
  - Estudo enumerativo, 6
  - Estudo observacional, 316
  - Estudo retrospectivo, 316
  - Evento composto, 47
  - Evento simples, 45
  - Evento(s), 46
    - complemento, 47
    - composto, 46
    - dependente, 74
    - disjunto, 48
    - exaustivo, 70
    - independente, 74, 77
    - intersecção, 47, 48
    - mutuamente exclusivo, 48
    - simples, 46
    - união, 47, 48
  - Eventos dependentes, 75
  - Eventos disjuntos, 48
  - Eventos exaustivos, 70
  - Eventos independentes, 75
  - Eventos mutuamente exclusivos, 48
  - Experimento, 8, 45
    - “double-blind”, 344
    - binomial, 105
    - bloco aleatorizado, 392
    - controlado aleatoriamente, 317
    - Experimento “double-blind”, 344
    - Experimento binomial, 105
    - Experimento controlado aleatoriamente, 317
    - Experimento de bloco aleatório, 392
    - Experimento de captura-recaptura, 241

- Experimento fatorial, 415  
 Experimento multinomial, 186, 551  
 Experimentos de simulação, 203  
 Extrapolação, risco de, 443
- Família de distribuições de probabilidade, 91  
 Família de taxa de erro, 371  
 Fator de correção de população finita, 115  
 Fator, 357  
 Fator de correção, 363  
 Freqüência, 13  
   acumulada, 23  
   relativa, 13  
 Freqüência acumulada, 24  
 Freqüência relativa, 13  
 Freqüência relativa e probabilidade, 51-52  
 Freqüência relativa limitante, 1-52  
 Função convexa, 177  
 Função de densidade de probabilidade marginal, 183, 191  
 Função de distribuição acumulada, 93, 132, 134  
   distribuição de probabilidades, 129  
   valor esperado, 137  
   variância, 138  
 Função de distribuição. *Veja* Função de distribuição cumulativa  
 Função de massa de probabilidade marginal, 181  
 Função de taxa de falha, 177  
 Função de variáveis aleatórias  
   distribuição de, 176  
   valor esperado de, 100-101, 137, 192, 219  
   variância de, 103-104, 138  
 Função de verossimilhança, 237  
 Função degrau, 93  
 Função densidade de probabilidade, 129  
   condicional, 188  
   conjunta, 181, 185  
   marginal, 182  
 Função densidade de probabilidade conjunta, 181, 185, 191  
 Função densidade, probabilidade da, 129  
 Função gama, 152  
   incompleta, 153, 154, 644  
   tabela da incompleta, 154  
 Função intrinsecamente linear, 490-491  
 Função logit, 497  
 Função massa de probabilidade conjunta, 180, 185
- Galton, Francis, 449  
 Gráfico da série temporal, 12  
 Gráfico de controle de média móvel ponderada exponencialmente, 638  
 Gráfico de dispersão, 434  
   ajustador, 496  
 Gráfico de probabilidade, 165, 166  
 Gráfico de reta, 90  
 Gráfico digidot, 12  
 Gráfico dos resíduos ajustado, 517  
 Gráfico seminormal, 173  
 Gráficos de controle, 580  
   características de desempenho, 612-614  
   dados dos atributos, 620, 622, 624, 638  
   localização, 608, 610, 611  
   média móvel ponderada exponencialmente, 638  
   regras suplementares, 614  
   robustos, 614-615  
   soma acumulada, 624-631  
   variação, 617, 618, 619  
 Gráficos de resíduos, 486, 517  
 Graus de liberdade  
   amostra única, 35, 261  
   ANOVA fator único, 364  
   ANOVA multifator, 390, 399, 407, 411
- Distribuição da amplitude studentizada, 367  
 distribuição  $F$ , 347, 361  
 distribuição qui-quadrado, 156  
 distribuição  $t$ , 261  
 duas amostras, 324  
 regressão, 445, 504, 516  
 teste de homogeneidade, 570  
 teste de independência, 573  
 testes de qualidade do ajustamento, 552, 559, 561
- Hipótese, 276  
 alternativa, 276  
 composta, 558  
 nula, 276  
 simples, 557  
 Hipótese alternativa, 276  
 Hipótese composta, 558  
 Hipótese estatística, 275  
 Hipótese nula, 276  
 Hipótese simples, 557  
 Histograma, 14  
   bimodal, 18  
   com inclinação negativa, 18  
   com inclinação positiva, 18  
   de dados contínuos, 15-16  
   de dados discretos, 14  
   de probabilidade, 91  
   multimodal, 18  
   simétrico, 18  
   unimodal, 18  
 Histograma bimodal, 18  
 Histograma de probabilidade, 91  
 Histograma inclinado, 18  
 Histograma multimodal, 18  
 Histograma simétrico, 18  
 Histograma unimodal, 18  
 Homogeneidade, teste de, 571
- Inclinação negativa, 18  
 Inclinação positiva, 18  
 Independência de pares, 79  
 Independência mútua, 78  
 Independência, teste de, 573  
 Inferência e estatística, 4  
 Inspeção de retificação, 636  
 Interação, 513  
   de dois fatores, 398-399, 407  
   de três fatores, 407  
   generalizada, 421  
 Interação da soma dos quadrados, 400, 407  
 Interação generalizada, 421  
 Interpretação objetiva da probabilidade, 52  
 Intersecção de eventos, 47, 48  
 Intervalo  
   aleatório, 247  
   confiança. *Veja* Intervalo de confiança  
   previsão. *Veja* Intervalo de previsão  
 Intervalo aleatório, 247  
 Intervalo de classe (classe), 15  
 Intervalo de confiança, 246  
   da combinação linear de médias, 371  
   da diferença das médias com dados em pares, 335  
   da diferença entre proporções, 345  
   da Distribuição de Poisson, 260  
   da distribuição exponencial, 251-252  
   da distribuição livre, 594-598  
   da distribuição uniforme, 273  
   da média de variável dependente na regressão, 462, 506, 520  
   da média, 250, 255, 264, 275

- da mediana, 273, 603
- da razão das proporções, 347
- da razão das variâncias, 350
- da razão dos desvios padrão, 350
- da soma de postos de Wilcoxon, 596
- da variância, 271
- de amostra grande, 255, 256, 320
- de contraste, 384
- de diferença entre médias, 320, 325, 328, 335
- de postos com sinais de Wilcoxon, 595
- de proporção, 256
- dedução do, 251
- determinação do tamanho da amostra, 250, 257, 321
- do coeficiente angular da reta de regressão, 454
- do coeficiente de correlação, 475
- do coeficiente de regressão, 505, 520
- do comando de carga, 252-253
- do desvio padrão, 271
- do termo constante da reta de regressão, 481
- limite do, 251
- monocaudal, 258
- simultâneo, 367, 375, 384
- sinal do, 603
- Intervalo de confiança com amostra grande, 255, 257, 320
- Intervalo de confiança da soma de postos de Wilcoxon, 596, 663
- Intervalo de confiança de sinal, 603
- Intervalo de confiança  $t$ , 263, 325, 328, 335, 371
- Intervalo de confiança  $z$ , 249, 254, 256, 320, 345
- Intervalo de estimativa. *Veja* Intervalo de confiança
- Intervalo de postos com sinais de Wilcoxon, 595, 662
- Intervalo de previsão, 265, 465, 506, 520
- Intervalo de tolerância, 266
- Intervalo interquartil, 614
- Intervalos de confiança de Scheffé, 384
- Intervalos de confiança simultâneos de Bonferroni, 464
- Intervalos de confiança simultâneos
  - ANOVA, 367, 378, 384
  - regressão, 464
- $K$ -tupla, 60
- Layout completo, 410
- Layout incompleto, 410
- Lei da probabilidade total, 70
- Leis de DeMorgan, 50
- Limite de confiança, 258, 263
- Limite do erro de estimativa, 251
- Limite inferior de confiança, 258, 263
- Limite superior de confiança, 258, 263
- Limites de controle, 606
  - recalculando, 612
- LOWESS (ajustador de gráfico de dispersão ponderado no local), 496
- Máscara em  $V$ , 624
- Média ( $s$ )
  - da amostra, 25
  - estimativa por pontos. *Veja* Estimativa por pontos da população, 25, 98
  - geral, 360
  - intervalo de confiança. *Veja* Intervalo de confiança de variável aleatória contínua. *Veja* Valor esperado de variável aleatória discreta.
  - Veja* Valor esperado da distribuição da amostra, 207-209
  - teste de hipóteses da. *Veja* Teste de hipóteses truncada, 328
- Média geral, 360
- Média truncada, 28
- Mediana
  - amostra, 25
  - distribuição contínua, 136
  - estimativa por pontos. *Veja* Estimativa por pontos
  - intervalo de confiança da. *Veja* Intervalo de confiança
  - população, 26
  - teste de hipóteses. *Veja* Teste de hipóteses
- Medidas resistentes, 34
- Meia-réplica, 422
- Método de comparações múltiplas de Tukey, 368, 378
- Método de Yates, 591
- Método  $T$ . *Veja* Método de Comparações múltiplas de Tukey
- Métodos de Taguchi, 606
- Métodos livres de distribuição, 580
- Métodos não-paramétricos. *Veja* Métodos de distribuição livre
- Mínimos quadrados ponderados, 487-488
- Mínimos quadrados, princípio dos, 440, 502
- MINITAB, 3
- Moda, 42, 125, 175
- Modelo aditivo, 387
- Modelo completo de segunda ordem, 512
- Modelo de regressão intrinsecamente linear, 492
- Modelo de regressão múltipla de primeira ordem, 512
- Modelo de regressão múltipla de segunda ordem, 512
- Modelo probabilístico, 433
- Modelo probabilístico linear, 444
- Modelo quadrático completo, 512
- Momento
  - da amostra, 235
  - da população, 235
- Momentos, método dos, 236
- Multicolinearidade, 541-542
- Níveis, 358
- Nível de confiança simultâneo, 371
- Nível de confiança, 246, 249, 250, 371
- Nível de previsão, 265
- Nível de qualidade aceitável, 634
- Nível de significância, 282
- Nível de significância observado, 302
- Nomograma de Kemp, 629-630
- Notação fatorial, 62
- Número médio do total inspecionado, 636
- Observação influente, 539-541
- Ordem-padrão, 416
- Padronização, 144
  - variáveis, 510, 534
- Par análogo, 423
- Parâmetro
  - distribuição de probabilidade, 91
  - escala, 170
  - formato, 171
  - localização, 170
  - não-centralidade, 375, 382
- Parâmetro de escala, 170
- Parâmetro de forma, 171
- Parâmetro de localização, 170
- Parâmetro de não-centralidade, 375, 382
- Parâmetro do declive na regressão, 437
- Pares casados, 337
- Pares disjuntos, 48
- Pares ordenados, 58
- Percentil
  - amostra, 28, 166
  - distribuição contínua, 135
  - distribuição normal, 143, 146
- Permutação, 61
- Projeto com medidas repetitivas, 393
- planos de amostragem dupla, 635-636
- planos de amostragem simples, 632-635
- Ponto fora da faixa, 26, 36
  - extremo, 36
  - moderado, 36



- Ponto fora da faixa extremo, 36
- População, 3
- conceitual, 6
  - hipotética, 6
- População conceitual, 6
- População hipotética, 6
- Porcentagem de truncamento, 28
- Postos, 582
- Postos com sinais, 582
- Potência, 293
- Princípio da razão de verossimilhança, 309-310
- Princípio de invariância, 240
- Probabilidade, 5, 43, 50
- anterior, 70
  - axiomas, 50
  - condicional, 66-67
  - histograma, 91
  - interpretação, 51
  - lei do total, 70
  - posterior, 70
  - propriedade da intersecção, 68-69
  - propriedade da união, 54-55
  - propriedade de multiplicação, 68-69, 76, 78
  - propriedade do complemento, 53
- Probabilidade anterior, 70
- Probabilidade condicional, 66-67
- distribuição de, 188
- Probabilidade função massa, 89
- condicional, 188
  - conjunta, 181, 185
  - marginal, 180
- Probabilidade posterior, 70
- Probabilidade subjetiva, 53
- Probabilidade total, lei da, 70
- Problema do aniversário, 82
- Procedimento de teste, 278
- Procedimentos combinados  $t$ , 328
- Procedimentos de soma acumulada, 624-631
- Processo, 6
- Processo de nascimento, 243
- Processo de Poisson, 120-121
- não-homogêneo, 125
- Proporção(ões)
- amostra, 29
  - Estimativa por pontos. *Veja* estimativa por pontos
  - intervalo de confiança. *Veja* intervalo de confiança
  - população, 30
  - teste de hipóteses. *Veja* Teste de hipóteses
- Propriedade de ausência de memória, 155
- Quadrado greco-latino, 430
- Quadrado latino, 410
- Quadrado médio, 361, 364
- Quadrados médios esperado, 366, 374, 381, 395, 401, 404, 406, 408, 413
- Quadro de amostragem, 6
- Qualidade média resultante, 636
- Quartil, 28
- Quarto do meio, 242
- Quarto inferior, 35
- Quarto superior, 35
- Razão das chances, 498
- razão  $t$ , 457
- Região de rejeição, 278
- bicaudal, 285
  - de cauda inferior, 279
  - de cauda superior, 279
- Regra da multiplicação para probabilidades, 68-69, 76, 78
- Regra do produto
- $K$ -tuplas, 59
  - pares ordenados, 58-59
- Regra empírica, 146
- Regressão, 433
- adequação do modelo, 484-488
  - análise residual, 484-488
  - coeficientes, 511, 517
  - cúbica, 502
  - e ANOVA, 458
  - efeito, 449
  - exponencial, 492
  - função, 511
  - intrinsecamente linear, 492
  - linear simples, 436-466
  - logística, 497
  - LOWESS (ajustador de gráfico de dispersão ponderado no local), 496
  - multicolinearidade, 541-542
  - múltipla, 511-549
  - não-linear, 490-499
  - observação influente, 539-541
  - origem da terminologia, 449
  - pela origem, 452
  - polinomial, 501-507
  - potência, 492
  - quadrática, 502
  - reta, 436, 440
  - robusta, 488
  - seleção de variáveis, 535-539
  - soma dos quadrados, 449
  - transformações, 490-496, 532-534
- Regressão cúbica, 502
- Regressão da função de potência, 492
- Regressão exponencial, 492
- Regressão linear simples, 436-466
- Regressão logística, 498
- Regressão MSDA (Mínima Soma dos Desvios Absolutos) MAD, 488
- Regressão múltipla, 511-549
- adequação do modelo, 524-525
  - equação do modelo, 511
  - estimativa por pontos, 514
  - intervalo de confiança, 520
  - intervalo de previsão, 520
  - modelos, 512-516
  - multicolinearidade, 541-542
  - observação influente, 539-541
  - padronização de variáveis, 533
  - seleção de variáveis, 535-539
  - teste de hipóteses, 518, 520, 523
  - teste de utilidade do modelo, 518
  - transformações, 532-534
- Regressão não-linear, 490-499
- Regressão passo a passo, 538-539
- Regressão polinomial, 501-507
- centralização de valores  $x$ , 506-507
  - estimativa por pontos, 502-503
  - intervalo de confiança, 505-506
  - intervalo de previsão, 506
  - modelo, 502
  - teste de hipóteses, 505
  - variável independente padrão, 510
- Regressão quadrática, 502
- Regressão robusta, 488
- Relação determinística, 483
- Relação linear, medida de, 195
- Réplica pelo quarto, 509
- Replicação fracionária, 422-426
- Resíduo, 444, 504, 516
- padrão, 485

- Resposta aleatória, 235
- Resultados igualmente prováveis, 55
- Resumo de cinco dígitos, 35
- Reta de regressão real, 436
- Reta dos mínimos quadrados, 440
- SAS, 3
- Seleção de variáveis, 535-539
- Seleção progressiva, 538
- Série temporal, 41, 477
- Significado prático, 308
- Significância
  - estatística, 307
  - prática, 308
- Significância estatística, 307
- Significativo, 303
- Simetria em conjunto de dados, 42
- Sistema *k-de-n*, 123
- Soma dos quadrados
  - erro. *Veja* Erro de Soma dos quadrados
  - interação. *Veja* Interação da soma dos quadrados
  - total. *Veja* soma dos quadrados total
  - tratamento. *Veja* soma dos quadrados do tratamento
- Soma dos quadrados dos resíduos, 449
- S-PLUS, 3
- Subestimativa em comparações múltiplas, 367-369
- Tabela ANOVA, 364
- Tabela de contingência, 569
- Tabela de probabilidade conjunta, 180
- Tabela *t*, 362
- Tabela *z*, 142
- Taxa de erro relativa ao experimento, 371
- Técnicas de contagem, 58-63
- Tentativa, 105
- Teorema binomial, 190
- Teorema de Bayes, 70-72
- Teorema do limite central, 209-211
- Teste bicaudal, 286
- Teste de amostras grandes, 289, 296, 319
- Teste de cauda inferior, 279
- Teste de cauda superior, 279
- Teste de falta de ajustamento, 490
- Teste de Fisher-Irwin, 346
- Teste de Friedman, 600
- Teste de grupo, 124
- Teste de hipóteses, 276
  - amostra grande, 289, 296, 319
  - ANOVA de teste *F*, 362, 376
  - ANOVA dois fatores, 397, 401, 403, 406
  - ANOVA fator único, 362, 376
  - ANOVA três fatores, 408
  - Ansari-Bradley, 604
  - bicaudal, 285
  - coeficiente angular da reta de regressão, 457
  - coeficiente de correlação, 473
  - coeficiente de regressão, 505, 520
  - da qualidade do ajustamento, 552, 559, 562
  - de cauda inferior, 280
  - de cauda superior, 279
  - desvio padrão, 312
  - determinação do tamanho da amostra, 288, 293, 298, 318, 344
  - diferença da média de dados em pares, 333
  - diferença entre médias, 316-317, 319, 325, 328, 333
  - diferença entre proporções, 342
  - Distribuição de Poisson, 311, 561
  - distribuição exponencial, 312
  - distribuição livre, 309, 581-593, 598-601
  - distribuição normal, 286, 291, 312, 316-317, 563, 565
  - e intervalos de confiança, 595
  - erros, 279
  - estatística de teste, 278
  - etapas, 286
  - falta de ajustamento na regressão, 490
  - homogeneidade de várias populações, 571
  - igualdade de coeficientes angulares, 481
  - igualdade de variâncias, 348
  - independência de fatores qualitativos, 573
  - média, 286, 289, 291
  - média da variável dependente na regressão, 464
  - mediana, 603
  - modelo reduzido x completo, 523
  - nível de significância, 282
  - percentil normal, 312
  - potência do teste, 293
  - probabilidades de erro tipo II, 288, 291-293, 298, 317, 328, 343, 374-376
  - procedimento de teste, 278
  - proporção, 297, 299
  - Quadrado latino, 411-412
  - região de rejeição, 278
  - Siegel-Tukey, 604
  - tendência, 589
  - teste da razão de verossimilhança, 308-309
  - Teste da soma de postos de Wilcoxon, 589, 591
  - Teste de Fisher-Irwin, 346
  - Teste de Friedman, 600
  - Teste de Kruskal-Wallis, 598-599
  - teste de McNemar, 355
  - Teste de postos com sinais de Wilcoxon, 583, 586
  - teste de sinal, 603
  - teste do gráfico de probabilidades para normalidade, 565
  - utilidade do modelo, 457, 518
  - valor *P*, 302, 303
  - variância, 312
- Teste de Kruskal-Wallis, 598-599
- Teste de Mann-Whitney, 589
- Teste de McNemar, 355
- Teste de normalidade de Ryan-Joiner, 566, 659
- Teste de postos com sinais de Wilcoxon, 583, 586, 660
- Teste de sinal, 603
- Teste de soma de postos de Wilcoxon, 590-591, 661
- Teste de utilidade do modelo
  - regressão linear simples, 457
  - regressão múltipla, 518
- Teste *F*
  - ANOVA, 362, 390, 401, 403, 408
  - igualdade de duas variâncias, 348
  - regressão, 458, 518, 523
- teste *F*, 349
- Teste monocaudal. *Veja* Teste de causa superior: teste de cauda inferior
- Teste nível  $\alpha$ , 282
- Teste qui-quadrado de ajustamento, 552, 559, 563
- teste qui-quadrado, 554
- Teste *t*, 290, 305-306, 325, 328, 333
- teste *z*, 287, 289, 297, 304, 316, 319, 342
- Testes de aderência, 550-565
- Tolerância percentual de defeitos em lote, 634
- Total da soma dos quadrados, 363, 376, 447
- Total geral, 363
- Total, amostra, 207
- Transformação de estabilização da variância, 378
- Transformação de Fisher, 474
- Transformações
  - de dados, 22
  - em gráficos de controle, 623
  - na ANOVA, 378
  - na regressão, 490-496, 532-534

- Tratamento, 385  
 Tratamento da soma dos quadrados, 363, 376
- União de eventos, 46, 47
- Valor ajustado, 444, 516
- Valor crítico  
 Amplitude Studentizada, 367  
*F*, 347  
 qui-quadrado, 270  
*t*, 262  
 tolerância, 266  
*z*, 143-144
- Valor crítico de tolerância, 266, 646  
 valor crítico *t*, 262  
 valor crítico *z*, 144
- Valor esperado  
 combinação linear de, 213  
 de função de variável aleatória, 100-101, 137, 192, 220  
 de uma diferença, 214  
 de variável aleatória contínua, 137  
 de variável aleatória discreta, 98  
 média da amostra, 207  
 regras de, 101-102  
 total da amostra, 207
- Valor fora da faixa moderado, 36  
 Valor médio, 98, 137, 192  
 Valor nulo, 277  
 Valor *P*, 302-303  
 Valor previsto, 444, 516
- Varição, coeficiente de, 39
- Variância(s)  
 amostra, 32  
 combinação linear, 213  
 de uma diferença, 214  
 estimador combinado, 328  
 estimativa por pontos. *Veja* Estimativa por pontos  
 função de variáveis aleatórias, 103-104, 138  
 intervalo de confiança. *Veja* Intervalo de confiança  
 média da amostra, 207  
 população, 33  
 regras, 35, 103  
 teste de hipóteses. *Veja* Teste de hipóteses  
 total da amostra, 207  
 variável aleatória contínua, 138  
 variável aleatória discreta, 102
- Variáveis aleatórias dependentes, 183  
 Variáveis aleatórias independentes, 184, 187  
 Variáveis aleatórias não-correlacionadas, 195  
 Variável aleatória binomial, 107  
 Variável aleatória de Bernoulli, 86  
 Variável aleatória padrão, 144  
 Variável contínua, 13  
 Variável de previsão, 434  
 Variável de resposta, 434  
 Variável dependente, 485  
 Variável discreta, 13  
 Variável explicativa, 434  
 Variável fictícia, 514  
 Variável independente, 434  
 Variável indicadora, 514  
 Variável(is) aleatória(s) contínua(s), 87, 128  
 Variável(is) aleatória(s) discreta(s), 87, 128  
 distribuição de probabilidades, 89-90  
 função de distribuição acumulada, 93  
 valor esperado, 98  
 variância, 102
- Variável(is) aleatória(s), 173  
 Bernoulli, 174  
 binomial, 107  
 binomial negativa, 116  
 contínua, 87, 128  
 dependente, 184  
 desvio padrão, 102, 138  
 discreta, 87  
 geométrica, 117  
 hipergeométrica, 113  
 independente, 184, 187  
 não-correlacionada, 195  
 valor esperado, 98  
 valor médio, 98, 137  
 variância, 102, 138
- Variável, 2  
 aleatória, 86  
 contínua, 13  
 de resposta, 433  
 dependente, 433  
 discreta, 13  
 explicativa, 433  
 independente, 433  
 preditiva, 433

# Probabilidade e Estatística para Engenharia e Ciências

Este livro visa fornecer uma introdução abrangente dos modelos e métodos estatísticos mais passíveis de serem encontrados e usados por estudantes em cursos de Engenharia e Ciências Naturais.

Apesar de os exemplos e exercícios terem sido desenvolvidos para cientistas e engenheiros, a maior parte dos métodos abrangidos é fundamental para análises estatísticas em várias outras disciplinas, de forma que os estudantes de Administração e Ciências Sociais também se beneficiarão desta leitura.

A obra introduz uma grande variedade de exercícios tanto após a apresentação de cada tópico quanto no final de cada capítulo, a fim de reforçar o aprendizado, e a didática do texto é enfatizada por inúmeros exemplos.

## Aplicações

Destina-se a estudantes dos cursos de Engenharia, Ciências Naturais e Estatística.

Leitura complementar para os cursos de Administração e de Ciências Sociais.



Para suas soluções de curso e aprendizado,  
visite [www.cengage.com.br](http://www.cengage.com.br)

ISBN 13 978-85-221-0924-1

ISBN 10 85-221-0924-9



9 788522 109241