Rayane Andrade

INTRODUÇÃO À ESTATÍSTICA

SÉTIMA EDIÇÃO

MARIO F. TRIOLA

Tradução

Alfredo Alves de Farias
Professor Adjunto/ UFMG

Revisão técnica

Eliana Farias e Soares, Ph.D.
Professora Adjunta / UFMG

com a colaboração de

Vera Regina L. F. Flores, M. Sc.
Professora Adjunta / UFMG

Prof.: Jorge
Disc.: Introdução a Estatística

LTC EDITORA

1 Introdução à Estatística

1-1 Aspectos Gerais

Define-se o termo estatística juntamente com os termos população, amostra, parâmetro e estatística (segundo significado).

1-2 Natureza dos Dados

Definem-se os dados quantitativos e os dados qualitativos juntamente com dados discretos e dados contínuos. Definem-se também os quatro níveis de mensuração (nominal, ordinal, intervalar e razão).

1-3 Usos e Abusos da Estatística

Apresentam-se exemplos de utilização benéfica da estatística e, ao mesmo tempo, algumas formas em que a estatística é usada para enganar. A utilização incorreta inclui pequenas amostras, números precisos, percentagens distorcidas, questões tendenciosas, gráficos enganosos e amostras mal extraídas.

1-4 Planejamento de Experimentos

Descrevem-se estudos observacionais e experimentos, juntamente com uma boa metodologia estatística. Dá-se ênfase à importância de uma boa amostragem. Definem-se e descrevem-se diversos métodos de amostragem, inclusive amostragem aleatória, amostragem estratificada, amostragem sistemática, amostragem por conglomerados e amostragem de conveniência.

1-5 Estatística com Calculadoras e Computadores

Discute-se a importância das calculadoras e dos computadores. A utilização das calculadoras é abordada em conjunto com pacotes STATDISK e Minitab.

Problema do Capítulo

Que podemos concluir desta pesquisa?

O programa de televisão ABC-Nightline realizou uma pesquisa em que solicitava a opinião dos espectadores sobre a permanência, ou não, da sede das Nações Unidas nos EUA. Para responder, os espectadores deviam pagar 50 centavos (americanos) para fazer uma chamada telefônica. Dos 186.000 que responderam, 67% disseram que a sede da ONU devia sair dos EUA. Com base nesses resultados amostrais, que podemos concluir sobre a opinião da população americana, sobre a permanência ou não da sede da ONU nos EUA?

1-1 Aspectos Gerais

Começamos nosso estudo de *estatística* observando que a palavra tem dois significados básicos. No primeiro sentido, o termo é usado em relação a números específicos obtidos de dados, conforme ilustrado nos exemplos seguintes:

- Em uma pesquisa, feita pela Bruskin-Goldring Research junto a 1.012 pessoas, a quem foi formulada pergunta sobre como utilizar um bolo de frutas, 13% responderam que deveria servir para calço de porta.
- Entre as pessoas com quem se fez um teste sobre o uso de drogas para admissão em novo emprego, 3,8% reagiram positivamente [de acordo com a American Management Association (Associação Americana de Gerenciamento)].
- O escore máximo de rebatidas de beisebol registrado até agora é de 0,442, obtido por James O'Neil em 1887.

A segunda acepção se refere à estatística como método de análise.

O Estado da Estatística

A palavra *estatística* provém do latim *status*, que significa *estado*. A primitiva utilização da estatística envolvia compilações de dados e gráficos que descreviam vários aspectos de um estado ou país. Em 1662, John Graunt publicou informes estatísticos sobre nascimentos e mortes. O trabalho de Graunt foi secundado por estudos de mortalidade e taxas de morbidade, tamanho de populações, rendas e taxas de desemprego. As famílias, os governos e as empresas se apóiam largamente em dados estatísticos. Assim é que as taxas de desemprego, de inflação, os índices do consumidor, as taxas de natalidade e mortalidade são calculadas cuidadosamente a intervalos regulares, e seus resultados são utilizados por empresários para tomarem decisões que afetam a futura contratação de empregados, níveis de produção e expansão para novos mercados.

DEFINIÇÃO

A **estatística** é uma coleção de métodos para planejar experimentos, obter dados e organizá-los, resumi-los, analisá-los, interpretá-los e deles extrair conclusões.

A estatística abrange muito mais do que o simples traçado de gráficos e o cálculo de médias. Neste livro veremos como tirar conclusões gerais e significativas que vão além dos dados originais. Em estatística, utilizamos extensamente os termos *população* e *amostra*. Esses termos, que passamos a definir, estão no próprio cerne da estatística.

DEFINIÇÕES

Uma **população** é uma coleção completa de todos os elementos (valores, pessoas, medidas etc.) a serem estudados.

Um **censo** é uma coleção de dados relativos a *todos* os elementos de uma população.

Uma **amostra** é uma subcoleção de elementos extraídos de uma população.

Por exemplo, uma pesquisa Nielsen típica de televisão utiliza uma *amostra* de 4000 lares e, com base nos resultados, formula conclusões acerca da *população* de todos os 97.855.392 lares nos EUA.

Estreitamente relacionados com os conceitos de população e amostra estão os conceitos de *parâmetro* e *estatística*. As definições seguintes são de fácil memorização.

DEFINIÇÕES

Um **parâmetro** é uma medida numérica que descreve uma característica de uma *população*.

Uma **estatística** é uma medida numérica que descreve uma característica de uma *amostra*.

Consideremos um exemplo. Em uma pesquisa, feita pela Bruskin-Goldring Research com 1015 pessoas escolhidas aleatoriamente, 269 (ou 26,5%) possuíam computador. Como a cifra de 26,5% se baseia em uma amostra, e não em toda a população, trata-se de uma *estatística* (e não um parâmetro). Já se uma pesquisa feita entre os 50 governadores estaduais dos EUA mostra que 42 (ou 84%) possuem computador, a cifra de 84% é um *parâmetro* porque se baseia em toda a população de governadores.

Um aspecto importante da estatística é sua aplicabilidade óbvia a situações reais e relevantes; em todo este livro encontraremos ampla diversidade dessas aplicações.

1-2 A Natureza dos Dados

Alguns conjuntos de dados (como alturas) consistem em números, enquanto outros são não-numéricos (como sexo). Aplicam-se as expressões *dados quantitativos* e *dados qualitativos* para distinguir esses dois tipos.

DEFINIÇÕES

Os **dados quantitativos** consistem em números que representam contagens ou medidas.

Os **dados qualitativos** (ou **dados categóricos**, ou **atributos**) podem ser separados em diferentes categorias que se distinguem por alguma característica não-numérica.

O Conjunto de Dados 4 do Apêndice B registra as quantidades de alcatrão em diferentes marcas de cigarros; esses valores representam dados quantitativos, mas as diversas marcas constituem dados qualitativos.

Podemos ainda descrever os dados quantitativos distinguindo entre os tipos discreto e contínuo.

DEFINIÇÕES

Os **dados discretos** resultam de um conjunto finito de valores possíveis, ou de um conjunto enumerável desses valores. (Ou seja, o número de valores possíveis é 0, ou 1, ou 2 etc.)

Os **dados contínuos (numéricos)** resultam de um número infinito de valores possíveis que podem ser associados a pontos em uma escala contínua de tal maneira que não haja lacunas ou interrupções.

Quando os dados representam contagens, são discretos; quando representam mensurações, são contínuos. O número de ovos que as galinhas põem constitui dados *discretos*, porque representa uma contagem; já a quantidade de leite que as vacas produzem constitui dados *contínuos*, porque representa mensurações que podem tomar qualquer valor em um intervalo contínuo.

Outra maneira comum de classificar dados consiste em utilizar quatro níveis de mensuração: nominal, ordinal, intervalar e razão.

DEFINIÇÃO

O **nível nominal de mensuração** é caracterizado por dados que consistem apenas em nomes, rótulos ou categorias. Os dados não podem ser dispostos segundo um esquema ordenado (como de baixo para cima).

Se associamos o termo *nominal* a "nome somente", o significado é fácil de memorizar. Um exemplo de dado nominal é o partido político a que cada senador dos EUA pertence.

EXEMPLO Seguem outros exemplos de dados amostrais ao nível nominal de mensuração.

1. Respostas do tipo "sim", "não" ou "indeciso".
2. O sexo dos estudantes em uma turma de estatística.

Como as categorias carecem de qualquer significado ordinal ou numérico, os dados precedentes não podem ser utilizados em cálculos. Assim é que não podemos tirar a "média" de 20 mulheres e 15 homens. *Cuidado*: Por vezes atribuem-se números a categorias (mormente quando os dados são computadorizados), mas tais números não têm qualquer significado para efeito de cálculo, e a média calculada com base neles em geral não tem sentido. Poderíamos citar o fato de que a Gallup Organization computou dados de uma pesquisa em que se atribui o "valor" 0 aos democratas, 1 aos republicanos e 2 aos independentes. Mesmo estando diante de rótulos numéricos, os dados permanecem no nível nominal e não podemos fazer cálculos com eles.

DEFINIÇÃO

O **nível ordinal de mensuração** envolve dados que podem ser dispostos em alguma ordem, mas as diferenças entre os valores dos dados não podem ser determinadas, ou não têm sentido.

EXEMPLO Dão-se a seguir exemplos de dados ao nível ordinal de mensuração.

1. Um editor classifica alguns originais como "excelentes", alguns como "bons" e alguns como "maus". (Não podemos determinar uma diferença quantitativa entre "bom" e "mau".)
2. Um comitê de preparação olímpica classifica Gail em 3.º, Diana em 7.º e Kim em 10.º. (Podemos determinar a diferença entre os 3.º e 7.º lugares mas a diferença de 4 não tem qualquer significado.)

Esse nível ordinal dá informações sobre comparações relativas, mas os graus de diferença não servem para cálculos. Os dados em nível ordinal não devem, pois, ser utilizados em cálculos.

Censo do Ano 2000

O censo nacional (dos EUA) do ano 2000 será mais rápido, menos dispendioso e mais preciso do que o censo de 1990. Ao contrário do censo de 1990, o Censo de 2000 utilizará métodos de amostragem para obter resultados mais precisos. Em 1990, os agenciadores voltaram até seis vezes às 35 milhões de casas que não remeteram os formulários preenchidos; mas, em 2000, essas casas omissas serão submetidas a uma amostragem. Espera-se que a amostragem produza resultados mais precisos do que as tentativas de atingir cada casa individualmente. O censo de 2000 custará cerca de $4 bilhões, o que significa $1 milhão menos do que o custo da repetição dos mesmos métodos de 1990. O censo de 2000 será mais eficiente — embora o censo de 1990 não tenha sido tão ineficiente como sugeriu o colunista Dave Barry: "O Departamento do Censo expede 100 milhões de formulários, 87 milhões dos quais chegam a um único destino em Albany."

DEFINIÇÃO

O **nível intervalar de mensuração** é análogo ao nível ordinal, com a propriedade adicional de que podemos determinar diferenças significativas entre os dados. Todavia, não existe um ponto de partida zero inerente, ou natural (onde *não* haja qualquer quantidade presente).

DEFINIÇÃO

O **nível de razão de mensuração** é o nível de intervalo modificado de modo a incluir o ponto de partida zero inerente (onde zero significa *nenhuma* quantidade presente). Para valores nesse nível, tanto as diferenças como as razões têm significado.

As temperaturas de 98,2°F e 98,6°F são exemplos de dados nesse nível intervalar de mensuração. Os valores se apresentam ordenados, e podemos determinar diferenças entre eles (em geral chamadas *distância* entre os dois valores). Todavia, não há ponto de partida natural. O valor 0°F pode parecer um ponto de partida, mas é inteiramente arbitrário, e não representa "ausência de calor". É um erro dizermos que 50°F é duas vezes mais quente do que 25°F. (Na escala Kelvin, as marcações de temperatura estão ao nível de razão de mensuração; essa escala tem um zero absoluto.)

EXEMPLO Seguem exemplos de dados ao nível intervalar de mensuração.

1. Os anos 1000, 2000, 1776 e 1944. (O tempo não começou no ano zero e, assim, 0 é arbitrário, e não um ponto de partida zero natural.)
2. As temperaturas anuais médias (em graus Celsius) das capitais dos 50 estados americanos.

Medida da Desobediência

Como coletar dados sobre algo que não se apresente mensurável, como o nível de desobediência do povo? O psicólogo Stanley Milgram planejou o seguinte experimento: Um pesquisador determinou que um voluntário acionasse um painel de controle que dava choques elétricos crescentemente dolorosos em uma terceira pessoa. Na realidade, não eram dados choques e a terceira pessoa era um ator. O voluntário começou com 15 volts e foi orientado a aumentar os choques de 15 em 15 volts. O nível de desobediência era o ponto em que a pessoa se recusava a aumentar a voltagem. Surpreendentemente, dois terços dos voluntários obedeceram às ordens mesmo que o ator gritasse e simulasse um ataque cardíaco.

EXEMPLO Dão-se a seguir exemplos de dados ao nível de razão de mensuração.

1. Pesos de artigos de material plástico descartados pelas residências (0 lb indica que nenhum plástico foi descartado, e 10 lb representam duas vezes 5 lb).
2. Duração (em minutos) de filmes.
3. Distâncias (em milhas) percorridas por carros em um teste de consumo de combustível.

Os valores de cada um desses conjuntos de dados podem ser dispostos em ordem, suas diferenças podem ser calculadas, e existe um ponto de partida zero inerente. *Este nível é chamado o nível de razão porque o ponto de partida torna as razões significativas.* Como um peso de 200 lb é *duas vezes* um peso de 100 lb, mas 50°F *não* é duas vezes mais quente do que 25°F, os pesos estão ao nível de razão, enquanto as temperaturas Fahrenheit estão em nível de intervalo. Para uma comparação e revisão concisas, deve-se estudar a Tabela 1-1 para ver as diferenças entre os quatro níveis de mensuração.

Ao aplicarmos a estatística a problemas reais, o nível de mensuração dos dados é um fator importante para determinarmos o processo a ser utilizado. Nossa compreensão dos quatro níveis de mensuração deve ser complementado pelo bom senso — uma ferramenta indispensável na estatística. Por exemplo, não tem sentido calcularmos a média dos números de inscrição de segurados no INSS, porque esses números *não* medem *nem* contam qualquer coisa; têm por função única e exclusiva identificar as pessoas. Tais números são, na verdade, *nomes* diferentes para as diversas pessoas e, como tais, não devem ser utilizados para cálculos. De modo geral, não devemos calcular médias de dados aos níveis nominal ou ordinal de mensuração.

TABELA 1-1 Níveis de Mensuração de Dados

Nível	Sumário	Exemplo	
Nominal	Categorias somente. Os dados não podem ser dispostos em um esquema ordenado.	Carros de alunos: 10 Corvettes 20 Ferraris 40 Porsches	Categorias ou nomes somente.
Ordinal	As categorias são ordenadas, mas não podemos estabelecer diferenças, ou estas não têm sentido.	Carros de alunos: 10 compactos 20 médios 40 grandes	Está determinada uma ordem: "compacto", "médio", "grande".
Intervalo	Podemos determinar diferenças entre valores, mas não há ponto de partida inerente. As razões não têm sentido.	Temperaturas no campus: 45°F 80°F 90°F	90°F não é duas vezes mais quente do que 45°F.
Razão	Como intervalo, mas com um ponto de partida inerente. As razões têm sentido.	Pesos de jogadores de *rugby* em uma faculdade: 150 lb 195 lb 300 lb	300 lb é duas vezes 150 lb.

1-2 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-8, identifique cada número como discreto ou contínuo.

1. Cada cigarro Camel tem 16,13 mg de alcatrão.
2. O altímetro de um avião da American Airlines indica uma altitude de 21.359 pés.
3. Uma pesquisa efetuada com 1015 pessoas indica que 40 delas são assinantes de um serviço de computador *on-line*.
4. O radar indica que Nolan Ryan rebateu a última bola a 82,3 mi/h.
5. De todos os escores SAT marcados no ano passado, 27 foram perfeitos.
6. De 1000 consumidores pesquisados, 930 reconheceram a marca de sopa Campbell.
7. O tempo total gasto anualmente por um motorista de táxi de Nova York ao dar passagem a pedestres é de 2,367 segundos.
8. Ao completar um programa de treinamento, Shaquille O'Neal pesava 12,44 lb menos do que no início do treinamento.

Nos Exercícios 9-18, determine o nível de mensuração mais adequado (nominal, ordinal, intervalo, razão).

9. Classificação como superior, acima da média, médio, abaixo da média ou pobre para encontros marcados com desconhecidos.
10. Conteúdo de nicotina (em miligramas) de cigarros Camel.
11. Números de inscrição do INSS.
12. Temperaturas (em graus Celsius) de uma amostra de contribuintes irritados por estarem sendo fiscalizados.
13. Anos em que os democratas ganharam as eleições presidenciais.
14. Graus finais (A, B, C, D, F) de estudantes de estatística.
15. Códigos de endereçamento postal.
16. Rendas anuais de enfermeiras.
17. Carros classificados como subcompacto, compacto, intermediário ou grande.
18. Cores de uma amostra de confeitos M&M.

1-2 Exercícios B: Acima do Básico

19. Presidentes americanos foram assassinados nos anos de 1865, 1881, 1901 e 1963. Qual é o nível de mensuração para esses anos? Explique sua resposta.
20. No quadrinho "Born Loser" (Perdedor nato) por Art Sansom, Brutus manifesta alegria por um aumento de temperatura de 1° para 2°. Ao lhe perguntarem a razão, respondeu: "Está agora duas vezes mais quente que hoje de manhã." Por que Brutus errou mais uma vez?

1-3 Usos e Abusos da Estatística

Usos da Estatística

As aplicações da estatística se desenvolveram de tal forma que, hoje, praticamente todo campo de estudo se beneficia da utilização de métodos estatísticos. Os fabricantes fornecem melhores produtos a custos menores através de técnicas de controle de qualidade. Controlam-se doenças com auxílio de análises que antecipam epidemias. Espécies ameaçadas são protegidas por regulamentos e leis que reagem a estimativas estatísticas de modificação do tamanho das populações. Visando reduzir as taxas de casos fatais, os legisladores têm melhor justificativa para leis como as que regem a poluição atmosférica, inspeções de automóveis, utilização do cinto de segurança e da bolsa de ar, e dirigir em estado de embriaguez. Citamos apenas esses exemplos, porque uma compilação completa das aplicações da estatística facilmente tomaria o resto deste livro.

Alguns estudantes escolhem um curso de estatística porque é exigido, mas um número cada vez maior o faz voluntariamente, porque reconhecem seu valor e aplicabilidade em qualquer campo em que pretendam trabalhar. Como os empregadores *gostam* de ver um curso de estatística no currículo de um candidato, o leitor que tiver estudado estatística levará vantagem ao procurar um emprego. Afora razões relacionadas com a obtenção de emprego e com a disciplina, o estudo da estatística pode tornar o leitor mais crítico em sua análise de informações, e menos sujeito a afirmações enganosas, como as que se acham comumente associadas a pesquisas, gráficos e médias. Como membro educado e responsável da sociedade, o leitor deve aguçar sua capacidade de reconhecer dados estatísticos distorcidos e de interpretar inteligentemente dados que se apresentem sem distorção.

Os Motoristas Mais Idosos São Mais Seguros do que os Mais Moços?

A American Association of Retired People — AARP (Associação Americana de Aposentados) alega que os motoristas mais idosos se envolvem em menor número de acidentes do que os mais jovens. Nos últimos anos, os motoristas com 16-19 anos de idade causaram cerca de 1,5 milhão de acidentes, em comparação com apenas 540.000 causados por motoristas com 70 anos ou mais, de forma que a alegação da AARP parece válida. Acontece, entretanto, que os motoristas mais idosos não dirigem tanto quanto os mais jovens. Em lugar de considerar apenas o *número* de acidentes, devemos examinar também as *taxas* de acidentes. Eis as taxas de acidentes por 100 milhões de milhas percorridas: 8,6 para os motoristas com idades de 16 a 19, 4,6 para os com idade de 75 a 79, 8,9 para os com idade de 80 a 84 e 20,3 para os motoristas com 85 anos de idade ou mais. Embora os motoristas mais jovens tenham de fato maior número de acidentes, os mais velhos apresentam as mais altas taxas de acidente.

Abusos da Estatística

Não é de hoje que ocorrem abusos com a estatística. Assim é que, há cerca de um século, o estadista Benjamin Disraeli disse: "Há três tipos de mentira: as mentiras, as mentiras sérias e a estatística." Já se disse também que "os números não mentem; mas os mentirosos forjam números" (*Figures don't lie; liars figure*) e que "se torturarmos os dados por bastante tempo, eles acabarão por admitir qualquer coisa". O historiador Andrew Lang disse que algumas pessoas usam a estatística "como um bêbado utiliza um poste de iluminação — para servir de apoio e não para iluminar". Todas essas afirmações se referem aos abusos da estatística, quando os dados são apresentados de forma enganosa. Alguns dos que abusam da estatística o fazem simplesmente por descuido ou ignorância; outros, porém, têm objetivos pessoais, pretendendo suprimir dados desfavoráveis enquanto dão ênfase aos dados que lhes são favoráveis. Passemos a alguns exemplos das diversas maneiras como os dados podem ser distorcidos.

Pequenas Amostras No Capítulo 6 veremos que as pequenas amostras não são necessariamente más; entretanto, os resultados obtidos com pequenas amostras podem por vezes ser usados como uma forma de "mentira" estatística. As preferências de apenas

10 dentistas por determinado dentifrício não devem servir de base para uma afirmação generalizada como "A pasta dentifrícia XYZ é recomendada por 7 em cada 10 dentistas." Mesmo que a amostra seja grande, ela deve ser não-tendenciosa e representativa da população de onde provém. Às vezes uma amostra pode parecer realmente grande (como em uma pesquisa com "2000 adultos americanos escolhidos aleatoriamente"), mas se se formulam conclusões acerca de subgrupos, como republicanos católicos do sexo masculino, tais conclusões podem estar baseadas em amostras assaz pequenas.

Números Precisos Às vezes os próprios números podem ser enganosos. Uma cifra, como um salário anual de \$37.735,29, pode parecer muito precisa, introduzindo alto grau de confiança em sua exatidão. Já a cifra \$37.700,00 não infunde o mesmo senso de precisão. Entretanto, uma estatística com muitas casas decimais não é necessariamente *precisa*.

Estimativas por Suposição Outra fonte de engano estatístico envolve estimativas que são, na verdade, suposições (ou, na linguagem popular, "palpites"), podendo apresentar erros substanciais. É preciso considerar a fonte da estimativa e a maneira como foi estabelecida. Quando o Papa visitou Miami, as fontes oficiais estimaram a multidão em 250.000 pessoas, mas, utilizando fotos aéreas e grades, o *Miami Herald* chegou a uma cifra mais precisa de apenas 150.000.

Porcentagens Distorcidas Por vezes utilizam-se porcentagens confusas ou distorcidas. Em um anúncio de página inteira, a Continental Airlines anuncia melhores serviços. No tocante ao caso de bagagem extraviada, o anúncio afirmava que "se trata de uma área em que já melhoramos 100% nos últimos seis meses". Em um editorial criticando essa estatística, o *New York Times* interpretou corretamente a melhora de 100% como significando que agora não se extravia mais qualquer bagagem — o que ainda não foi conseguido pela Continental Airlines.

Cifras Parciais "Noventa por cento dos carros vendidos nos EUA nos últimos 10 anos ainda estão rodando." Milhões de consumidores ouviram esta mensagem e ficaram com a impressão de que esses carros devem ter sido muito bem construídos para durarem tanto. O que o fabricante não mencionou foi que 90% dos carros por ele vendidos, o foram nos últimos três anos. A alegação, embora tecnicamente correta, era enganosa, por não apresentar os resultados completos.

Distorções Deliberadas No livro *Tainted Truth*, Cynthia Crossen cita um exemplo da revista *Corporate Travel* que publicou dados mostrando que, entre as companhias locadoras de carros, a Avis foi a vencedora em uma pesquisa junto aos locatários. Quando a Hertz solicitou informações detalhadas sobre a pesquisa, as respostas desapareceram e o coordenador da pesquisa se demitiu. A Hertz processou a Avis (por falsa propaganda baseada na pesquisa) e a revista; chegou-se a um acordo.

Perguntas Tendenciosas As perguntas em uma pesquisa podem ser formuladas de modo a "sugerirem" uma resposta. Um caso famoso envolve o candidato à presidência dos EUA, Ross Perot, que formulou a seguinte pergunta em um questionário: "O presidente deve ter o poder de vetar decisões do Congresso?" Noventa e sete por cento das respostas foram "sim". Entretanto, o percentual de respostas "sim" caiu para 57% quando a pergunta foi "O presidente deve ter, ou não, o poder de *vetar decisões do Congresso*?" Às vezes as perguntas se apresentam involuntariamente tendenciosas em virtude de fatores como a ordem dos itens a serem considerados. Por exemplo, uma pesquisa alemã formulou estas duas perguntas:

- O leitor diria que o tráfego contribui em maior ou menor grau do que a indústria para a poluição atmosférica?
- O leitor diria que a indústria contribui em maior ou menor grau do que o tráfego para a poluição atmosférica?

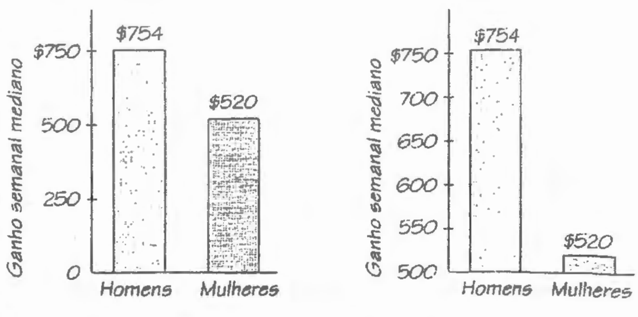
Quando o tráfego foi mencionado em primeiro lugar, 45% acusaram o tráfego e 32% acusaram a indústria; quando a indústria foi citada em primeiro lugar, as porcentagens se modificaram grandemente para 24% e 57%, respectivamente.

Pesquisa do *Literary Digest*

No campanha presidencial de 1936, a revista *Literary Digest* fez uma pesquisa e concluiu pela vitória de Alf Landon, mas Franklin D. Roosevelt venceu por larga margem. Maurice Bryson observa: "Foram enviados 10 milhões de cédulas — amostra a eleitores em potencial, mas apenas 2,3 milhões foram devolvidos. Como todos devem saber, tais amostras são quase sempre tendenciosas." Bryson afirma também: "As respostas voluntárias a questionários enviados pelo correio constituem talvez o método mais comum de coleta de dados sobre ciências sociais encontrado pelos estatísticos, e é também talvez o pior." (Ver Bryson, "The *Literary Digest* Poll: Making of a Statistical Myth", *The American Statistician*, Vol. 30, N.º4.)

Gráficos Enganosos Muitos dispositivos visuais — como gráficos em barras e gráficos em setores — podem ser utilizados para exagerar ou diminuir a verdadeira natureza de um conjunto de dados. (Tais recursos serão discutidos no Capítulo 2.) Os dois gráficos da Figura 1-1 representam *os mesmos dados* do Bureau of Labor Statistics (Departamento de Estatística do Trabalho), mas a parte (b) tem como objetivo exagerar a diferença entre os ganhos dos homens e os das mulheres. Não partindo do zero no eixo vertical, o gráfico (b) tende a produzir uma impressão subjetiva errônea. A Figura 1-1 nos dá uma lição importante. Devemos analisar as informações *numéricas* contidas em um gráfico, não nos deixando enganar por sua forma geral.

Pictográficos Os desenhos de objetos, chamados pictográficos, também podem levar-nos a erro. Os objetos comumente usados para ilustrar dados incluem sacos de dinheiro, pilhas de moedas, tan-



(a) (b)

Fig. 1-1 Ganhos de profissionais de tempo integral.

ques do exército (para despesas militares), vacas (para produção de laticínios), barris (para produção de petróleo) e casas (para construção). Ao desenhar tais objetos, o artista pode criar impressões falsas que distorcem as diferenças. Se duplicamos o lado de um quadrado, a área não é apenas duplicada, e sim quadruplicada; duplicando cada aresta de um cubo, seu volume não é apenas duplicado, e sim multiplicado por oito. Se os impostos dobram a cada década, um desenhista pode representar os aumentos de imposto por um saco de dinheiro para o primeiro ano e um segundo saco duas vezes mais fundo, duas vezes mais alto e duas vezes mais largo para o segundo ano. Ao invés de aparecerem duplicados, os impostos se apresentarão aumentados oito vezes; o desenho distorce, assim, a realidade.

Pressão do Pesquisador Quando se formulam perguntas a indivíduos pesquisados, esses freqüentemente dão respostas favoráveis à sua auto-imagem. Em uma pesquisa telefônica, 94% dos que responderam disseram que lavam suas mãos após usar um banheiro, mas a observação em lugares tais como a Estação Penn, em Nova York e Golden Gate Park em San Francisco mostraram que o percentual efetivo é de apenas 68%.

Más Amostras Outra fonte de estatística enganosa são os métodos inadequados de coleta de dados. É comum um pesquisador analisar dados e formular conclusões errôneas porque o método de coleta de dados foi deficiente.

Um exemplo típico é a pesquisa "Nightline" em que 186.000 espectadores de televisão pagaram 50 centavos para discar um número de telefone "900" dando sua opinião sobre se a sede das Nações Unidas deve permanecer nos EUA. Os resultados mostraram que 67% dos que foram consultados eram favoráveis a que a sede da ONU saísse dos EUA. No começo deste capítulo perguntamos o que se poderia concluir quanto à opinião geral da população sobre a permanência da ONU nos EUA. Como os próprios espectadores é que decidiram se seriam incluídos na pesquisa, temos um exemplo de pesquisa auto-selecionada, que se define como segue.

DEFINIÇÃO
Uma **pesquisa auto-selecionada** é uma pesquisa em que os próprios entrevistados decidem se serão incluídos.

Em tais pesquisas, o que freqüentemente ocorre é que participam apenas aqueles que têm uma opinião firmada, resultando daí que a amostra dos que respondem não é representativa da população como um todo. Como 67% dos 186.000 pesquisados eram favoráveis à mudança da sede da ONU dos EUA, *nada podemos concluir sobre a população em geral, dada a maneira como se obteve a amostra.* Na realidade, Ted Koppel reportou que uma pesquisa "científica" de 500 pessoas revelou que 72% delas desejavam que a sede da ONU *permanecesse* nos EUA. Nessa pesquisa de 500 pessoas, os que responderam foram selecionados aleatoriamente pelo pesquisador, de modo que o resultado tende muito mais a refletir a verdadeira opinião da população em geral.

Uma pesquisa auto-selecionada é apenas uma das maneiras como o método de coleta de dados pode ser seriamente prejudicado. Em vista de sua importância, dedicaremos a próxima seção ao método de amostragem ou coleta de dados.

Os exemplos precedentes constituem uma pequena amostra das maneiras como a estatística pode ser utilizada de forma enganosa. Livros inteiros têm sido dedicados a esse assunto, inclusive o clássico *How to Lie with Statistics*, de Darrell Huff, *The Figure Finaglers*, de Robert Reichard, e *Tainted Truth*, de Cynthia Crossen. O entendimento de tais práticas será de grande auxílio na avaliação dos dados estatísticos encontrados em situações cotidianas.

1-3 Exercícios A: Habilidades e Conceitos Básicos

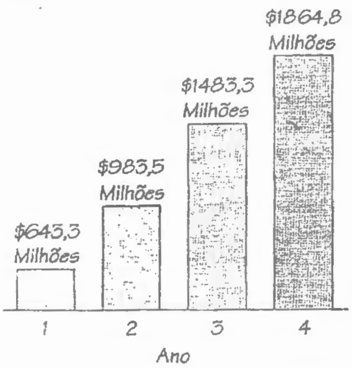
1. Uma pessoa foi encarregada de pesquisar o reconhecimento da marca Nike, devendo contactar por telefone 1500 consumidores nos EUA. Por que razão é incorreta a utilização de listas telefônicas como população para fornecer a amostra?
2. Setenta e dois por cento dos americanos espremem o tubo de dentifrício a partir da parte superior. Essa observação, assim como outras também não muito sérias, é apresentada em *The First Really Important Survey of American Habits* (a primeira pesquisa realmente importante dos hábitos dos americanos). Esses resultados se baseiam em 7000 respostas a 25.000 questionários enviados pelo correio. Qual o lado errado dessa pesquisa?
3. Um relatório patrocinado pela Florida Citrus Commission concluiu que os níveis de colesterol podem ser reduzidos mediante ingestão de produtos cítricos. Por que razão a conclusão poderia ser suspeita?
4. Uma funcionária tem um salário anual de $40.000, mas é informada de que terá uma redução de 10% no pagamento em virtude do declínio dos lucros da companhia. É informada também de que no próximo ano terá um aumento de 10%. A situação não se afigura tão má, porque a redução de 10% parece ser compensada pelo aumento de 10%.

   a. Qual a renda anual após o corte de 10%?
   b. Com base na renda anual da parte a, determine a renda anual após o aumento de 10%. O corte de 10% seguido do aumento de 10% restituem à funcionária o salário original de $40.000?

5. A revista *Glamour* publicou o seguinte resultado de uma pesquisa: "Setenta e nove por cento dos que responderam à nossa pesquisa de agosto afirmaram crer que os americanos se tornaram demasiadamente propensos a apelar para a justiça em casos corriqueiros." A questão foi publicada na revista e os leitores podiam responder pelo correio, fax ou *e-mail* (Tellus@Galamour.com). Até que ponto é válido o resultado de 79%?
6. ADT Security Systems advertiu que "quando você sai de férias, os ladrões começam a agir". O anúncio afirmava que "de acordo com estatísticas do FBI, mais de 26% dos assaltos a residências ocorriam entre o Memorial Day [feriado que homenageia os soldados mortos na guerra] e o Dia do Trabalho". Em que ponto essa afirmação é enganosa?
7. Em um estudo sobre crimes cometidos no campus de uma universidade por estudantes sob efeito do álcool ou das drogas, foram pesquisados 1.875 estudantes. Um artigo no *USA Today* notou: "Oito por cento dos estudantes que respondem anonimamente afirmam ter cometido um crime no campus. E 62% desse grupo dizem ter agido sob a influência do álcool ou das drogas." Supondo que o número de estudantes que responderam anonimamente seja de 1.875, quantos efetivamente cometeram um crime no campus sob a influência do álcool ou das drogas?
8. Um estudo realizado pelo Insurance Institute for Highway Safety (Instituto de Segurança nas Rodovias) constatou que o Chevrolet Corvette acusa o mais elevado índice de acidentes fatais — "5,2

mortes para cada 10.000". O carro com menor índice de acidentes fatais foi o Volvo, com apenas 0,6 morte por 10.000. Significa isto que o Corvette não é tão seguro quanto o Volvo?

9. O jornal *Newport Chronicle* afirma que as mães grávidas podem aumentar suas chances de ter um bebê sadio comendo lagostas. A alegação se baseia em um estudo mostrando que as crianças nascidas de mães que comem lagostas têm menos problemas de saúde do que as nascidas de mães que não comem lagostas. Qual é o erro nesta alegação?
10. Uma pesquisa inclui o seguinte item: "Registre sua altura em polegadas." Com isso pretende-se obter e analisar as alturas dos que respondem. Identifique os dois problemas neste item.
11. "De acordo com uma pesquisa de âmbito nacional feita por 250 agências de empregos, os sapatos gastos constituem o motivo mais comum para que um homem que procura emprego não cause boa impressão à primeira vista." Os jornais apresentavam essa alegação com base em uma pesquisa encomendada pela Kiwi Brands, produtores de graxa para sapatos. Faça um comentário sobre a razão por que os resultados de tal pesquisa podem ser questionados.
12. Em um suplemento de propaganda inserido no *Time*, os aumentos das despesas com o combate à poluição foram ilustrados em um gráfico como o que aparece a seguir. O que está errado com a figura?



1-3 Exercícios B: Além do Básico

13. Um artigo no *New York Times* afirmou que a duração média da vida de 35 regentes de orquestra do sexo masculino era de 73,4 anos, em contraste com a média de 69,5 anos para a população masculina em geral. A vida mais longa foi atribuída a fatores como satisfação pessoal e motivação. Há uma falha fundamental na conclusão de que os regentes de orquestra do sexo masculino vivem mais. Qual é?
14. Um pesquisador do Sloan-Kettering Cancer Research Center foi criticado certa vez por adulterar dados. Entre seus dados estavam cifras obtidas de seis grupos de ratos, com 20 ratos em cada grupo. Foram dados os seguintes valores como porcentagens de sucesso: 53%, 58%, 63%, 46%, 48%, 67%. O que está errado?
15. Procure identificar as quatro maiores falhas no seguinte. Um jornal realizou uma pesquisa solicitando a resposta dos leitores a esta pergunta: "Você apóia o desenvolvimento de armas atômicas que poderiam matar milhões de pessoas inocentes?" Relata-se que 20 leitores responderam, 87% com "não" e 13% com "sim".
16. Um editorial do *New York Times* criticou um anúncio que alegava que determinado anti-séptico bucal "reduzia em mais de 300% as placas nos dentes".

    a. Removendo-se 100% de uma quantidade, quanto resta?
    b. Que significa reduzir as placas em mais de 300%?

1-4 Planejamento de Experimentos

Os estudos que utilizam métodos estatísticos vão desde os que são bem concebidos e executados, dando resultados confiáveis, aos que são concebidos deficientemente e mal executados, levando a conclusões enganosas e sem qualquer valor real. Eis alguns pontos importantes para o planejamento de um estudo capaz de produzir resultados válidos:

1. Identificar com precisão a questão a ser respondida e definir com clareza a população de interesse.
2. Estabelecer um plano para coleta de dados. Esse plano deve descrever detalhadamente a realização de um *estudo observacional* ou de um *experimento* (ambos definidos a seguir), e deve ser elaborado cuidadosamente, de modo que os dados coletados representem efetivamente a população em questão.
3. Coletar os dados. Devemos ser extremamente cautelosos, para minimizar os erros que podem resultar de uma coleta tendenciosa de dados.
4. Analisar os dados e tirar conclusões. Identificar também possíveis fontes de erros.

Os estudos que requerem métodos estatísticos decorrem tipicamente de duas fontes comuns: estudos observacionais e experimentos.

DEFINIÇÕES

Em um **estudo observacional**, verificamos e medimos características específicas, mas não tentamos manipular ou modificar os elementos a serem estudados.

Em um **experimento**, aplicamos determinado *tratamento* e passamos então a observar seus efeitos sobre os elementos a serem pesquisados.

Por exemplo, um estudo observacional pode envolver uma pesquisa de cidadãos para determinar que porcentagem da população é a favor do registro de armas de fogo. Um experimento pode envolver o tratamento com um remédio ministrado a um grupo de pacientes a fim de determinar sua eficiência na cura. No caso da arma de fogo, coligimos dados sem modificar as pessoas a serem pesquisadas; já o tratamento por um remédio envolve a modificação das pessoas.

Os experimentos bem planejados costumam envolver um grupo a quem é dado um tratamento particular (chamado *grupo de tratamento*) e um segundo *grupo de controle* ao qual não se administra o tratamento. Por exemplo, o experimento sobre pólio realizado em 1954 envolveu um grupo de tratamento de crianças em quem foi injetada a vacina Salk, e um grupo de controle de crianças que recebeu um remédio neutro (*placebo*). Em experimentos deste tipo, ocorre um *efeito placebo* quando um indivíduo não tratado acredita estar recebendo o tratamento e alega uma melhora nos sintomas. O efeito placebo pode ser contrabalançado fazendo-se um experimento *cego*, uma técnica em que o indivíduo não sabe se está recebendo o tratamento ou um placebo. O experimento sobre pólio foi do tipo *duplo-cego*, em que as crianças que recebiam a injeção não sabiam se estavam recebendo a vacina Salk ou um placebo, e os médicos que davam a injeção e avaliavam os resultados também não sabiam.

As Pesquisas Políticas Crescem

Em "Consulting the Oracle", um artigo para o *U.S. News and World Report*, o autor Stephen Budiansky escreve que o Presidente Kennedy encomendou 16 pesquisas em seus três anos de mandato, Nixon encomendou 233 pesquisas em seus seis anos, e Clinton encomendou entre 100 e 150 pesquisas em seus primeiros 2,5 anos. As pesquisas de Clinton custaram entre $30.000 e $45.000 cada uma, o que dá um custo de $30 por pessoa. Budiansky relata que a pesquisa é complicada por máquinas de resposta e por pessoas que se recusam a cooperar, mas as boas pesquisas incluem tentativas repetidos para obter respostas das que não estão em casa ou se recusam a responder. Não levar em conta os que não respondem pode resultar em uma amostra que não represente adequadamente a população.

Ao planejar um experimento para testar a eficiência de um ou mais tratamentos, devemos ter o cuidado de atribuir as *unidades experimentais* (ou indivíduos) aos diferentes grupos de tal modo que esses grupos sejam bem semelhantes. (Tais grupos semelhantes de unidades experimentais são chamados *blocos*.) Uma abordagem eficiente consiste em utilizar um *planejamento experimental completamente aleatorizado*, que exige que as unidades experimentais sejam divididas em diferentes grupos mediante um processo de seleção *aleatória*. Assim é que tal planejamento pode envolver a atribuição aleatória de pessoas a um grupo tratado com aspirina e a um grupo de controle que não é tratado. Outro processo consiste em utilizar um *planejamento controlado rigorosamente*, com unidades experimentais escolhidas cuidadosamente, de modo que os diferentes grupos (ou blocos) sejam tão semelhantes quanto possível. Com um planejamento rigorosamente controlado, podemos tentar formar grupos de tratamento e de controle que incluam pessoas semelhantes em idade, peso, pressão sanguínea etc. É importante também considerar a *replicação*, que exige tamanhos de amostra suficientemente grandes que reduzam os efeitos da variação amostral aleatória. O experimento com a pólio foi um planejamento experimental completamente aleatorizado, porque os indivíduos em ambos os grupos, de tratamento e de controle, foram selecionados aleatoriamente. Incorporou a replicação incluindo números muito grandes (200.000) de indivíduos em cada grupo.

Na realização de experimentos, os resultados por vezes são comprometidos pelo confundimento.

DEFINIÇÃO

Ocorre o **confundimento** quando os efeitos de duas ou mais variáveis não podem distinguir-se uns dos outros.

Por exemplo, se estamos realizando um experimento para testar a eficiência de um novo retardante no incêndio em uma sarça, e repentinamente começa a chover, ocorre o confundimento porque é impossível distinguir entre o efeito do retardante e o efeito da chuva.

Um dos erros mais graves consiste em uma forma inadequada de coleta de dados. Nunca é demais enfatizarmos este importante ponto:

Dados coletados de forma descuidada podem ser tão inúteis que nenhum processamento estatístico consegue salvá-los.

Notamos na Seção 1-3 que uma pesquisa auto-selecionada é uma pesquisa em que as próprias pessoas decidem se vão responder ou não. As pesquisas auto-selecionadas são muito comuns, mas seus resultados em geral não têm utilidade para fazer inferências válidas sobre toda uma população.

Passamos agora a definir e descrever os cinco métodos mais comuns de amostragem.

DEFINIÇÃO

Em uma **amostra aleatória**, os elementos da população são escolhidos de tal forma que cada um deles tenha *igual chance* de figurar na amostra. (Escolhe-se uma **amostra aleatória simples** de *n* elementos, de maneira que toda a mostra de tamanho *n* possível tenha a mesma chance de ser escolhida.)

As amostras aleatórias podem ser escolhidas por diversos métodos, inclusive a utilização de tabelas de números aleatórios e de computadores para gerar números aleatórios. Com a amostragem aleatória, espera-se que todos os grupos da população sejam representados na amostra de forma aproximadamente proporcional. Uma amostragem descuidada pode facilmente resultar em uma amostra tendenciosa, com características assaz diferentes das da população que a originou. Em contrapartida, a amostragem aleatória é cuidadosamente planejada para evitar qualquer tendenciosidade. Por exemplo, a utilização de catálogos telefônicos elimina automaticamente todos aqueles cujos telefones não figurem no catálogo, e a exclusão desse segmento da população pode facilmente conduzir a resultados falsos. Em Los Angeles, por exemplo, 42,5% dos números de telefones não estão no catálogo (com base em dados da Survey Sampling, Inc.). Os pesquisadores costumam contornar esse problema utilizando computadores para gerar números de telefone, de modo que todos os números sejam possíveis. Eles devem também ter o cuidado de incluir os que inicialmente não foram encontrados ou se recusaram a responder. A Companhia de Pesquisas Harris constatou que a taxa de recusa para entrevistas telefônicas é em geral de 20%, no mínimo. O fato de ignorarmos os que inicialmente se recusam a responder pode concorrer para que nossa amostra seja tendenciosa.

DEFINIÇÃO

Com a **amostragem estratificada**, subdividimos a população em, no mínimo, duas subpopulações (ou estratos) que compartilham das mesmas características (como sexo) e, em seguida, extraímos uma amostra de cada estrato.

Em uma pesquisa sobre a Emenda Constitucional da Igualdade de Direitos, poderíamos utilizar o sexo como base para a criação de dois estratos. Após obter uma relação dos homens e uma relação das mulheres, aplicamos um método conveniente (como a amostragem aleatória) para escolher determinado número de elementos de cada relação. Quando os diversos estratos têm tamanhos amostrais que refletem a população global, temos o que se chama amostragem *proporcional*. No caso de alguns estratos não serem representados na proporção adequada, então os resultados poderão ser ajustados ou ponderados convenientemente.

Para um tamanho fixo de amostra, se escolhemos aleatoriamente elementos de diferentes estratos, temos chance de obter resultados mais consistentes (e menos variáveis) do que com a simples escolha de uma amostra aleatória de toda a população. Por essa razão, costuma-se usar a amostragem estratificada para reduzir a variação nos resultados.

DEFINIÇÃO

Na **amostragem sistemática**, escolhemos um ponto de partida, e selecionamos cada k-ésimo elemento (como por exemplo cada 50.° elemento) da população.

Por exemplo, se a Motorola quisesse fazer uma pesquisa sobre seus 107.000 empregados, poderia partir de uma relação completa dos mesmos e selecionar cada 100.° empregado, obtendo uma amostra de 1.070 elementos. Esse método é simples e utilizado com freqüência.

DEFINIÇÃO

Na **amostragem por conglomerados**, começamos dividindo a área da população em seções (ou conglomerados); em seguida escolhemos algumas dessas seções e, finalmente, tomamos *todos* os elementos das seções escolhidas.

Uma diferença importante entre a amostragem por conglomerados e a amostragem estratificada é que a amostragem por conglomerados utiliza *todos* os elementos dos conglomerados selecionados, enquanto a amostragem estratificada utiliza uma *amostra* de membros de cada estrato. Pode-se encontrar um exemplo de amostragem por conglomerado em uma pesquisa pré-eleitoral, onde escolhemos aleatoriamente 30 zonas eleitorais e pesquisamos todos os elementos de cada uma das zonas escolhidas. Esse método é muito mais rápido e menos dispendioso do que a escolha de um indivíduo de cada uma das inúmeras zonas da área populacional. Os resultados podem ser ajustados ou ponderados para corrigir qualquer representação desproporcionada de grupos. A amostragem por conglomerados é extensamente utilizada pelo governo e por organizações particulares de pesquisa.

Meta-análise

O termo *meta-análise* se refere a uma técnica de estudo que, essencialmente, combina os resultados de outros estudos. Tem a vantagem de permitir que amostras menores separadas sejam combinadas em uma única amostra grande, tornando mais significativos os resultados globais. Tem também a vantagem de utilizar um trabalho já feito. Por outro lado, tem a desvantagem de ser apenas tão boa quanto o tenham sido os estudos básicos. Se esses estudos apresentam falhas, pode ocorrer o fenômeno *"garbage in, garbage out"* (N. do T.: "O que sai é tão bom como o que entra.") A utilização da meta-análise é de uso corrente em pesquisas médicas e psicológicas. Um exemplo: "Reversal of Left Ventricular Hypertrophy in Essential Hypertension: A Meta-analysis of Randomized Double-blind Studies", por Schmieder, Martus e Klingbeil, *Journal of the American Medical Association*, Vol. 275, No.19.

DEFINIÇÃO

Na **amostragem de conveniência**, simplesmente utilizamos resultados que já estão disponíveis.

Em alguns casos, os resultados da amostragem de conveniência podem ser assaz bons, mas em outros casos podem apresentar séria tendenciosidade. Ao fazer uma pesquisa sobre pessoas canhotas, seria conveniente um estudante pesquisar seus próprios colegas de classe, porque estão ao seu alcance imediato. Mesmo que tal amostra não seja aleatória, os resultados devem ser bem satisfatórios. Em contrapartida, poderia ser muito conveniente (e talvez mesmo lucrativo) para a ABC News fazer uma pesquisa pedindo aos espectadores que liguem para um número de telefone "900" para registrar suas opiniões, mas essa pesquisa seria auto-selecionada e os resultados seriam provavelmente tendenciosos.

A Figura 1-2 ilustra os cinco métodos mais comuns de amostragem que acabamos de descrever. Essas descrições pretendem ser breves e gerais. O conhecimento aprofundado desses diversos métodos, que permita sua utilização com proveito, exige um estudo muito mais extenso, que ultrapassa o nível de um curso introdutório. Para manter esta seção em perspectiva, notemos que este texto fará referência freqüente a dados "selecionados aleatoriamente", o que significa que os dados foram selecionados de modo que todos os elementos da população têm a mesma chance de serem escolhidos. Conquanto não façamos referência freqüente aos outros métodos de amostragem, devemos ter consciência de que eles existem, e que o método de amostragem exige planejamento e execução cuidadosos. Os métodos apresentados em todo este texto dependem de amostras que tenham sido obtidas cuidadosamente. Além disso, o tamanho da amostra deve sempre ser suficientemente grande para os propósitos em vista. (Os problemas de tamanho da amostra são abordados mais adiante, especialmente no Capítulo 6.) Muitas pessoas acreditam que as grandes amostras são sempre boas, mas mesmo essas podem ser totalmente desprovidas de valor, se os dados tiverem sido coletados de maneira negligente. Finalmente, se estamos medindo uma característica (como altura) de um conjunto de indivíduos, podemos obter resultados mais precisos se fizermos nós mesmos as medidas, em vez de pedirmos aos indivíduos que indiquem os valores. Este último procedimento pode resultar em um número desproporcionado de resultados arredondados, assim como muitos resultados que refletem valores *desejados* em lugar de valores *efetivos*.

Não importa quão bem planejemos e executemos o processo de coleta de amostras, há sempre a possibilidade de um erro nos resultados. Como exemplo, escolha aleatoriamente 1000 adultos e pergunte a eles se têm o curso secundário completo, registrando a porcentagem de respostas "sim". Escolhido um outro grupo de 1000 indivíduos, é provável que se obtenha uma porcentagem amostral *diferente*.

DEFINIÇÕES

Um **erro amostral** é a diferença entre um resultado amostral e o verdadeiro resultado populacional; tais erros resultam de flutuações amostrais aleatórias.

Ocorre um **erro não-amostral** quando os dados amostrais são coletados, registrados ou analisados incorretamente. Tais erros resultam de um erro que não seja uma simples

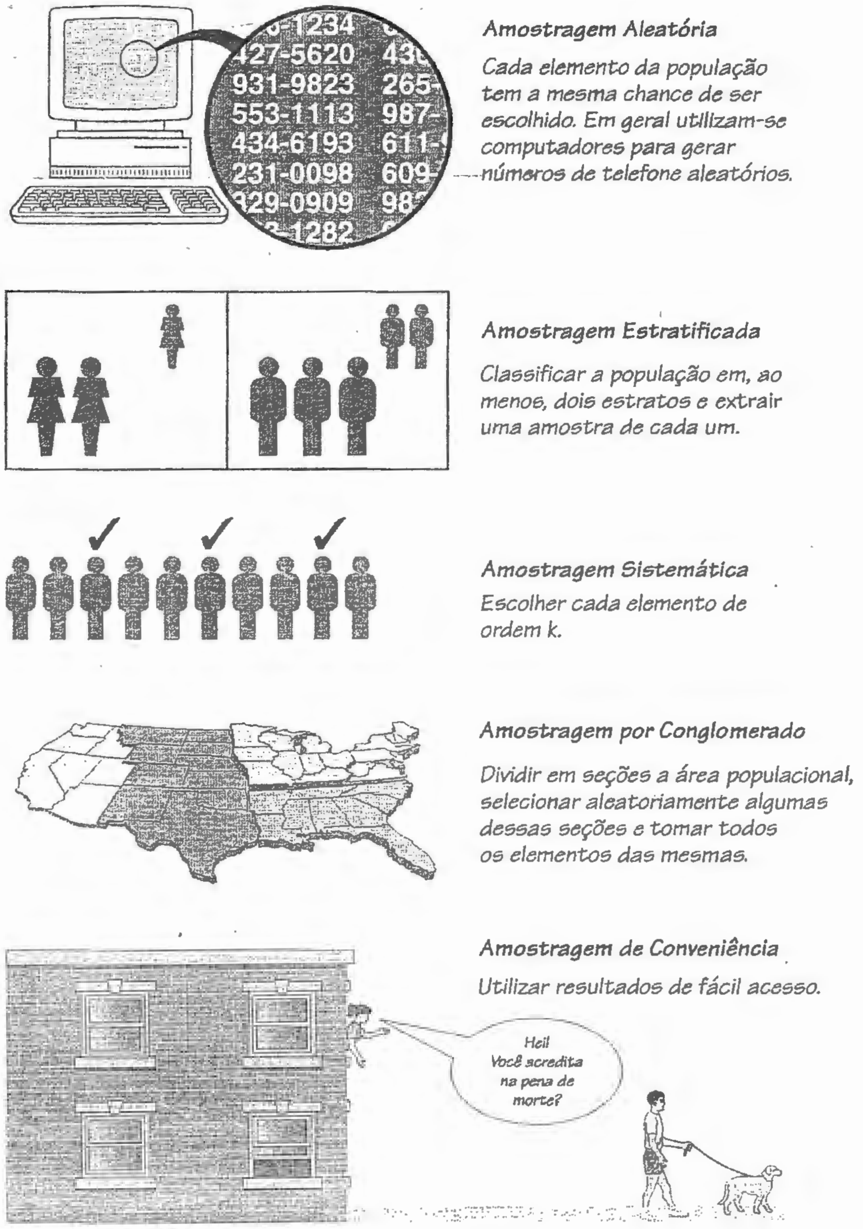


Fig. 1-2 Métodos comuns de amostragem.

flutuação amostral aleatória, como a escolha de uma amostra não-aleatória e tendenciosa, a utilização de um instrumento de mensuração defeituoso, uma questão formulada de modo tendencioso, um grande número de recusas de resposta ou a cópia incorreta dos dados amostrais.

Se extrairmos uma amostra cuidadosamente, de forma que ela represente realmente a população, podemos aplicar os métodos descritos neste livro para analisar o erro amostral, mas devemos ter o máximo cuidado em minimizar os erros não-amostrais.

Hawthorne e os Efeitos do Experimentador

O bem conhecido *efeito placebo* ocorre quando um indivíduo *não tratado* acredita incorretamente que está recebendo um tratamento real e reporta uma melhora dos sintomas. O *efeito Hawthorne* ocorre quando indivíduos *tratados* respondem de maneira um tanto diferente, simplesmente porque são partes de um experimento. (Esse fenômeno foi chamado "efeito Hawthorne" porque foi observado pela primeira vez em um estudo levado a efeito em operários da fábrica da Western Electric, em Hawthorne.) Ocorre um *efeito de experimentador* (às vezes chamado efeito Rosenthall) quando o pesquisador ou experimentador involuntariamente influencia o indivíduo pesquisado, através de fatores como expressão facial, tom de voz ou atitude.

1-4 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-4, determine se a descrição dada corresponde a um estudo observacional ou a um experimento.

1. Mede-se o conteúdo de alcatrão, nicotina e monóxido de carbono em diferentes marcas de cigarro (conforme Conjunto de Dados 4 no Apêndice B).
2. Pede-se a fumantes que reduzam à metade o número de cigarros consumidos diariamente, para que se possam medir os efeitos sobre a freqüência de pulsação.
3. Em uma turma de educação física, estuda-se o efeito dos exercícios físicos sobre a pressão sanguínea, determinando-se que metade dos estudantes *ande* uma milha cada dia, enquanto a outra metade *corra* uma milha diária.
4. Estuda-se a relação entre os pesos de ursos e seus comprimentos, tomando-se as medidas em ursos anestesiados.

Nos Exercícios 5-16, identifique o tipo de amostragem utilizado: aleatória, estratificada, sistemática, por conglomerado ou de conveniência.

5. Quando escreveu *Women and Love: A Cultural Revolution*, a autora Shere Hite baseou suas conclusões em 4.500 respostas a 100.000 questionários distribuídos a mulheres.
6. Um psicólogo da Universidade de Nova York faz uma pesquisa sobre todos os estudantes de cada uma de 20 turmas selecionadas aleatoriamente.
7. Um sociólogo na Universidade de Charleston seleciona 12 homens e 12 mulheres de cada uma de quatro turmas de inglês.
8. A empresa Sony seleciona cada 200.º CD de sua linha de produção e faz um teste de qualidade rigoroso.
9. Um cabo eleitoral escreve o nome de cada senador dos EUA. em cartões separados, mistura-os e extrai 10 nomes.
10. O gerente comercial da America Online testa uma nova estratégia de vendas selecionando aleatoriamente 250 consumidores com renda inferior a $50.000 e 250 consumidores com renda de ao menos $50.000.
11. O programa Planned Parenthood (Planejamento Familiar) pesquisa 500 homens e 500 mulheres sobre seus pontos de vista sobre o uso de anticoncepcionais.
12. Um pesquisador de mercado da American Airlines entrevista todos os passageiros de cada um de 10 vôos selecionados aleatoriamente.
13. Um pesquisador médico da Universidade Johns Hopkins entrevista todos os portadores de leucemia em cada um de 20 hospitais selecionados aleatoriamente.
14. Um repórter da revista *Business Week* entrevista todo 50.º gerente geral constante da relação das 1000 empresas com maior cotação de suas ações.
15. Um repórter da revista *Business Week* obtém uma relação numerada das 1000 empresas com maiores cotações de ações na bolsa, utiliza um computador para gerar 20 números aleatórios e então entrevista os gerentes gerais das empresas correspondentes aos números extraídos.
16. Ao fazer uma pesquisa para um noticiário vespertino de Boston, um repórter da NBC entrevista 15 pessoas que saem do auditório da IRS.

1-4 Exercícios B: Além do Básico

17. *Aberta* e *fechada* são dois tipos de questões de uma pesquisa. Uma questão aberta permite uma resposta livre, enquanto uma questão fechada comporta apenas uma resposta fixa. Alguns exemplos baseados em pesquisas Gallup.
Questão aberta: Na opinião do leitor, que se pode fazer para reduzir o crime?
Questão fechada: Qual das seguintes medidas mais contribuiria para a redução da criminalidade?
    - Contratar mais policiais.
    - Fazer com que os pais eduquem melhor os filhos.
    - Melhorar as condições sociais e econômicas nas favelas.
    - Ampliar os esforços para reabilitação nas cadeias.
    - Aplicar sentenças mais severas aos criminosos.
    - Reformar os tribunais.

    a. Quais são as vantagens e as desvantagens das questões abertas?
    b. Quais as vantagens e as desvantagens das questões fechadas?
    c. Que tipo é mais fácil de analisar com processos estatísticos formais? Por quê?
18. Descreva detalhadamente um método que poderia ser usado para obter uma amostra aleatória simples das alturas de cinco alunos de sua turma de estatística.

1-5 Estatística com Calculadoras e Computadores

Um subproduto importante do programa espacial dos EUA é a invenção do *chip* de microprocessador — uma invenção que teve profunda influência na aplicação da estatística. A instalação de *chips* de microprocessador em calculadoras e computadores eliminou a tremenda tarefa de cálculos monótonos, tornando o uso da estatística mais acessível a muitas pessoas. Descreveremos brevemente, nesta seção, o papel das calculadoras e dos computadores na estatística.

Calculadoras

Os estudantes de estatística cedo descobrem que uma calculadora é um de seus melhores auxiliares. Além de ter as operações básicas ($+$, $-$, $\times$, $\div$, $\sqrt{\ }$ etc.), muitas calculadoras apresentam hoje recursos estatísticos especiais, como média, desvio-padrão e resultados de correlação/regressão. (Esses tópicos serão abordados em capítulos posteriores.) Além de possibilitar o cálculo de expressões complicadas e de certas operações estatísticas, algumas calculadoras também permitem a introdução e armazenagem de programas especiais a serem utilizados durante todo o curso. A TI-83 da Texas Instruments é um excelente exemplo de calculadora perfeitamente adaptável a um curso introdutório de estatística. É programável, pode exibir gráficos e tem não poucas funções estatísticas especiais incluídas.

Existe um disco separado com programas escritos para a TI-82 e TI-83, e esses programas podem ser transferidos de um computador para a calculadora. Alguns professores de estatística exigem que todos os seus alunos utilizem uma calculadora TI-83, outros exigem qualquer calculadora que processe estatística bivariada e outros finalmente aceitam o uso de qualquer calculadora. Para o estudante que ainda não tem uma calculadora, recomenda-se uma que seja capaz de processar estatística de duas variáveis. Qualquer que seja a calculadora escolhida, o manual que a acompanha é um guia valioso. Em caso de dúvida, consulte o manual e procure fazer os exemplos apresentados. Se ainda assim tiver dificuldade, recorra ao seu professor.

Computadores

O computador desempenha hoje papel relevante em quase todos os aspectos da análise estatística. A ampla diversidade de computadores e pacotes de software possibilitou a utilização da estatística por pessoas com diferentes tipos de formação matemática, mas também criou maior oportunidade de uso indevido da estatística. É importante reconhecer que tanto os pacotes de software como os computadores têm uma limitação muito séria: eles seguem cegamente as instruções, ainda que inadequadas ou mesmo absurdas. O computador não raciocina, e não pode formular julgamentos. A compreensão dos princípios da estatística é pré-requisito importante para a correta interpretação de resultados obtidos por computador. Mesmo que o leitor não venha a usar efetivamente os computadores neste curso, deve procurar desenvolver habilidade em interpretar resultados de análise estatística obtidos em um computador, como os que ocorrem em todo este texto.

Faremos referência freqüentemente a dois pacotes em particular: O STATDISK e Minitab. O STATDISK apresenta uma vantagem importante: é um programa fácil de ser usado. O Minitab já é um pacote estatístico de nível mais elevado, mas também é de utilização relativamente fácil.

Com o STATDISK e o Minitab, os programas são escolhidos de uma barra de ferramentas no topo da tela, como segue:

STATDISK: `File Edit Analysis Data Help`

Minitab:

`File Edit Manip Calc Stat Graph Editor Window Help`

Utilizando STATDISK ou Minitab, podemos familiarizar-nos melhor com a operação geral de um computador. Os exemplos que seguem ilustram alguns aspectos básicos de STATDISK e Minitab:

Para introduzir um *novo* conjunto de dados:

STATDISK: Selecionar `Data` da barra de ferramentas e escolher então a opção `Sample Editor`.

Minitab: Selecionar `File` da barra principal e escolher então a opção `New Worksheet`.

Para *salvar* e nomear um conjunto de dados:

STATDISK Selecionar `File` da barra principal e escolher então a opção `Save As`.

Minitab: Selecionar `File` da barra principal e escolher então a opção `Save Worksheet As...`

Para *abrir* um arquivo de dados previamente armazenado:

STATDISK: Selecionar `File` da barra principal e escolher então a opção `Open`.

Minitab: Selecionar `File` da barra principal e escolher então a opção `Open Worksheet`.

Para *imprimir* resultados:

STATDISK: Selecionar `File` da barra principal e escolher a opção `Print`.

Minitab: Selecionar `File` da barra principal e escolher a opção `Print Window`.

Para *sair* do programa:

STATDISK: Selecionar `File` da barra principal e escolher então a opção `Quit`.

Minitab: Selecionar `File` da barra principal e escolher então a opção `Exit`.

STATDISK e Minitab são ambos capazes de realizar quase todas as operações importantes abordadas neste livro.

Apresentamos apenas algumas características de STATDISK e Minitab, mas a utilização desses programas é abordada com maior detalhe em *STATDISK Student Laboratory Manual and Workbook* (*7.ª edição*) e em *Minitab Student Laboratory Manual and Workbook* (*7.ª edição*). As características e a apresentação de alguns resultados dados por esses programas são também discutidos em todo este livro, sempre que adequado.

Alguns professores de estatística preferem outros pacotes como SPSS, SAS, BMDP, Execustat, Systat, Mystat ou Statgraphics. Qualquer que seja o pacote escolhido, o estudante sempre se beneficiará, melhorando seus conhecimentos em uma área que se tornou tão importante.

Deixe o Computador Ligado

Algumas pessoas costumam desligar o computador logo após o término de determinada tarefa, enquanto outras deixam-no ligado até que não precisem mais utilizá-lo naquele dia. O painel de circuitos e os *chips* do computador sofrem com esses ciclos de liga/desliga. Mas o monitor pode se danificar quando a mesma imagem é deixada na tela por períodos de tempo muito longos. O *tempo médio entre interrupções* (MTBF = Mean Time Between Failures) para o disco rígido já foi de 5000 horas, mas hoje é de cerca de 30.000 horas. Considerando os efeitos danosos dos ciclos *on/off* sobre o painel de circuitos e os *chips* do computador, e o grande MTBF para discos rígidos, faz sentido deixar o computador ligado até o fim do dia, desde que a tela do monitor possa ser protegida utilizando-se um programa para descansar a tela. Muitas pessoas utilizam essa estratégia, que se originou em parte de uma análise estatística de eventos passados.

1-5 Exercícios A: Habilidades e Conceitos Básicos

Exercícios iniciais com calculadora: Nos Exercícios 1-8, as expressões apresentadas são análogas às que se encontram em diferentes partes do livro. Utilize sua calculadora para obter os valores indicados.

1. $\dfrac{3{,}44 + 2{,}67 + 2{,}09 + 1{,}87 + 3{,}11}{5}$

2. $\sqrt{\dfrac{(2-5)^2 + (4-5)^2 + (9-5)^2}{3-1}}$

3. $\sqrt{\dfrac{3(101) - 15^2}{6}}$

4. $\dfrac{(12 - 8{,}5)^2}{8{,}5} + \dfrac{(22 - 25{,}3)^2}{25{,}3}$

5. $\dfrac{1{,}96^2 \cdot 0{,}25}{0{,}03^2}$

6. $\dfrac{102{,}7 - 100{,}0}{\dfrac{14{,}2}{\sqrt{50}}}$

7. $\dfrac{15!}{9!6!}$ (*Sugestão:* $6! = 6 \times 5 \times 4 \times 3 \times 2 \times 1$)

8. $\dfrac{8(56{,}80) - (14{,}60)(26)}{\sqrt{8(32{,}9632) - (14{,}60)^2}\ \sqrt{8(104) - (26)^2}}$

1-5 Exercícios B: Além do Básico

9. Carregar STATDISK ou Minitab e abrir o arquivo do conjunto de dados indicado a seguir. Esses dados já estão armazenados. Escreva os três primeiros valores relacionados.

STATDISK: `BLUE.SDD` (pesos de balas M&M azuis)

Minitab: `ALFALFA.MTW` (safras de alfafa de diversas variedades em diferentes plantações)

10. Carregar STATDISK ou Minitab e salvar as seguintes quantidades de alcatrão (em miligramas por cigarro) para 15 cigarros diferentes. Salve os dados com o nome de arquivo `CIGTAR`.

16 16 9 8 16 13 15 9 2 15 15 9 14 6 18

Abra o arquivo de dados para verificar que foram realmente salvos e obtenha uma apresentação impressa dos mesmos.

Vocabulário

estatística	nível de razão de mensuração
população	pesquisa auto-selecionada
censo	estudo observacional
amostra	experimento
parâmetro	confundimento
estatística	amostra aleatória
dados quantitativos	amostra aleatória simples
dados qualitativos	amostragem estratificada
dados discretos	amostragem sistemática
dados contínuos	amostragem por conglomerados
nível nominal de mensuração	amostragem de conveniência
nível ordinal de mensuração	erro amostral
nível intervalar de mensuração	erro não-amostral

Revisão

Iniciamos este capítulo com uma descrição geral da natureza da estatística e abordamos diferentes aspectos da natureza dos dados. Ilustramos com exemplos usos e abusos da estatística. Discutimos o planejamento de experimentos enfatizando a importância dos métodos de boa amostragem. Encerramos o capítulo com uma rápida discussão do papel das calculadoras e dos computadores. Ao completar o estudo deste capítulo, o estudante deve ser capaz de:

- Distinguir entre uma população e uma amostra
- Distinguir entre um parâmetro e uma estatística
- Identificar o nível de mensuração (nominal, ordinal, intervalar, razão) de um conjunto de dados
- Reconhecer a importância dos métodos de boa amostragem, bem como a séria deficiência dos métodos viciados de amostragem
- Reconhecer que as pesquisas auto-selecionadas não podem servir de base para formar conclusões válidas sobre uma população

Exercícios de Revisão

1. O Laboratório de Teste de Produtos para o Consumidor seleciona uma dúzia de pilhas (indicadas como de 9 volts) de cada um dos fabricantes, e testa a capacidade efetiva de cada uma.
   a. Os valores obtidos são discretos ou contínuos?
   b. Identifique o nível de mensuração (nominal, ordinal, intervalar, razão) para as voltagens.
   c. Que tipo de amostragem (aleatória, estratificada, sistemática, por conglomerado, de conveniência) está sendo utilizado?
   d. Trata-se de um estudo observacional ou de um experimento?
   e. Qual é o efeito relevante da utilização, pelo consumidor, de pilhas rotuladas como de 9 volts, quando, na realidade, seu nível de voltagem é muito diferente?
2. Os pesquisadores do Laboratório de Teste de Produtos para o Consumidor testam amostras de protetores eletrônicos contra variações de corrente para determinar os níveis de voltagem que podem danificar os computadores. Para cada um dos casos seguintes, determine qual dos quatro níveis de mensuração (nominal, ordinal, intervalar, razão) é apropriado.
   a. Os níveis de voltagem que causam dano.
   b. Postos (primeiro, segundo, terceiro etc.) por ordem de qualidade para uma amostra de protetores.
   c. Relacionar os protetores como "recomendado, aceitável, não-aceitável".
   d. As temperaturas das salas em que os protetores são testados.
   e. Os países em que os protetores foram fabricados.
3. A revista *Business Week* faz uma pesquisa, enviando pelo correio um questionário a 5000 pessoas que investem em títulos. Com base nos resultados, os editores das revista concluem que a maioria dos investidores nos EUA estão pessimistas quanto à economia. Qual o erro desta conclusão?
4. Identifique cada cifra como discreta ou contínua.
   a. A Nielsen Media Research Organization (Organização de Pesquisas Nielsen) pesquisou 2027 adultos que assistem ao programa *Monday Night Football* na ABC.
   b. O Professor Fisher registrou os tempos gastos por estudantes de estatística para completarem um exame final, e o primeiro resultado foi 87,25 minutos.
   c. Kathy Patel pesou seu livro de estatística e obteve o valor de 1,87 lb.
5. Identifique o tipo de amostragem (aleatória, estratificada, sistemática, por conglomerado, de conveniência) utilizada em cada um dos casos seguintes:
   a. Obtém-se uma amostra de um produto extraindo-se cada 100.ª unidade da linha de montagem.
   b. Geram-se números aleatórios em um computador para selecionar números de série de carros a serem escolhidos para uma amostra de teste.
   c. Um fornecedor de peças para automóvel obtém uma amostra de todos os itens de cada um de 12 fornecedores selecionados aleatoriamente.
   d. Um fabricante de automóveis faz um estudo de mercado compreendendo testes de direção feitos por uma amostra de 10 homens e 10 mulheres em cada uma de quatro diferentes faixas etárias.
   e. Um fabricante de automóveis faz um estudo de mercado entrevistando clientes em potencial que solicitam teste de direção a um revendedor local.
6. Agenciadores do censo constataram que ao perguntar a idade das pessoas encontram mais pessoas com 50 anos do que com 49 ou 51. Pode explicar por que isso ocorre?
7. O leitor pretende fazer uma pesquisa em seu campus. Onde está o erro ao selecionar cada 50.° estudante que sai da lanchonete?

8. O *Southport Chronicle* reportou que uma corrida preliminar foi assistida por 8725 pessoas. Comente.

Exercícios Cumulativos de Revisão

Os exercícios cumulativos de revisão deste livro destinam-se a incorporar algum material de capítulos anteriores, uma característica que será implementada nos capítulos seguintes. Os exercícios desta seção utilizam conceitos aprendidos antes do estudo deste livro.

1. A pergunta seguinte, feita em uma pesquisa, teve repercussão quando as respostas sugeriram que cerca de 22% dos americanos achavam que o holocausto pode não ter existido.

   "Acha possível ou impossível que a exterminação de judeus pelos nazistas nunca tenha existido?"

   Uma pesquisa subseqüente revelou que os que responderam provavelmente se sentiram confusos pela dupla negativa da frase. Eis uma formulação adequada em uma pesquisa Roper subseqüente:

   "Acha possível que a exterminação de judeus pelos nazistas jamais ocorreu, ou está certo de que realmente aconteceu?"

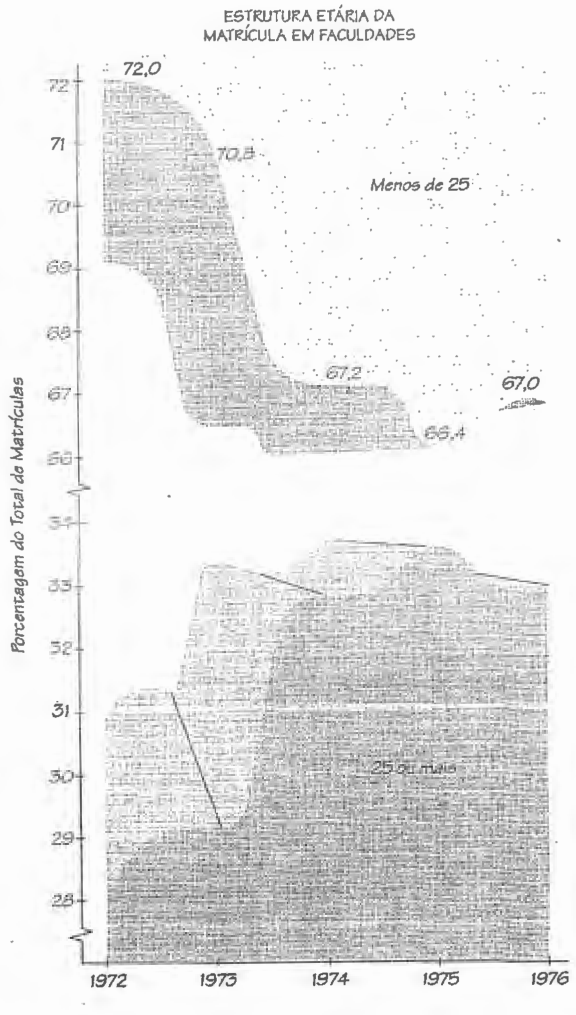
   Esta segunda versão parece substancialmente menos confusa? Pode formular a questão de modo que ela se apresente ainda mais clara do que nas duas versões?

2. Observe a figura a seguir. É semelhante a uma a que Edwin Tufte, autor de *The Visual Display of Quantitative Data*, se refere quando observa: "Este pode muito bem ser o pior gráfico jamais dado à imprensa". Observe que o gráfico relaciona "quase por acaso, apenas cinco conjuntos de dados (pois a divisão dentro de cada ano soma 100 por cento)". Examine primeiro o gráfico e identifique a informação que ele procura transmitir. Faça então um novo gráfico retratando a mesma situação.

Projeto para Computador

Recorra ao Conjunto de Dados 2 no Apêndice B e considere as 106 temperaturas (em graus Fahrenheit) encontradas na última coluna (Dia 2, 12 horas). Os pesquisadores da Universidade de Maryland coletaram dados sobre a temperatura do corpo humano e constataram que a média não era de 98,6°F, valor que quase todos nós supomos ser a média correta. Utilizando o STATDISK ou o Minitab, introduza as 106 temperaturas e as salve sob o nome BODYTEMP.

O objetivo deste projeto de computador é introduzir os dados e armazená-los em um disquete de computador. Isso permitirá termos os dados disponíveis para usá-los no Capítulo 2, contribuindo também para aumentar sua capacidade de introduzir e armazenar dados em um computador — uma técnica criticamente importante nos dias atuais.



DOS DADOS À DECISÃO

Dados Mal Representados

Obtenha um exemplo de um jornal ou uma revista em que os dados tenham sido apresentados de maneira enganosa. Identifique a fonte (incluindo data de publicação) de onde o exemplo foi tirado. Explique como a apresentação é enganosa e sugira uma forma mais honesta de apresentar os dados.

ATIVIDADES EM GRUPO

1. *Atividade extraclasse:* Divida em grupos de cinco e colete 50 valores utilizando amostragem aleatória, conforme descrita na Seção 1-4. Repita então a coleta de 50 valores para cada um dos outros quatro métodos de amostragem: estratificada, sistemática, por conglomerados e de conveniência. Em cada caso, calcule a "média". (A média é definida no Capítulo 2 e se obtém somando-se todos os valores e dividindo-se o total pelo número de valores.) Descreva inicialmente, com detalhe, o processo utilizado para cada método de amostragem; relacione então os valores e compare as cinco médias. Os diferentes métodos parecem dar os mesmos resultados? Os dados devem ser extraídos de uma população como idades dos livros em uma biblioteca, ou idades dos carros no estacionamento da faculdade.
2. *Atividades na classe:* Divida em grupos de três ou quatro e utilize os dados a seguir para construir um gráfico que exagere os aumentos nos pontos altos da Média Industrial Dow Jones. Construa outro gráfico que minimize a importância desses aumentos, e finalmente construa um terceiro gráfico que represente os dados sem qualquer tendenciosidade.

Década	1950	1960	1970	1980
Dow Alto	683	995	1.052	2.796

3. *Atividades na classe:* Divida em grupos de três ou quatro. Suponha que deve fazer uma pesquisa junto a estudantes de tempo integral de sua faculdade. Planeje e descreva detalhadamente um processo para obter uma amostra *aleatória* de 100 estudantes.

entrevista

Paul Mones

Paul Mones é advogado, autor e consultor. Escreveu *When a Child Kills: Abused Children Who Kill Their Parents*. Escreveu também *Stalking Justice*, a verdadeira história de um detetive que utilizou pela primeira vez impressões digitais DNA para apanhar um assassino contumaz. Foi entrevistado em muitos programas importantes nos EUA, na Europa e na Austrália, inclusive "60 Minutos", "20/20" e "Larry King Live". Seus comentários apareceram em jornais e revistas como o *New York Times* e *Time*; foi correspondente legal para "NBC News". Treina médicos, advogados e oficiais de justiça, e testemunha perante comissões legislativas.

O senhor utiliza a estatística em seu trabalho como advogado?

Utilizo extensamente a estatística em meu trabalho. Com a dactiloscopia DNA, por exemplo, consideramos vários fatores e determinamos a probabilidade de obter uma seqüência específica de genótipos nas mesmas pessoas. Costumávamos atentar para três *loci* (posições que os genes ocupam nos cromossomos); em seguida passamos para cinco, mas agora estamos em sete. Estudamos uma amostra de referência e uma amostra de comparação para ver a freqüência com que determinada seqüência ocorre. Se a seqüência de um suspeito coincide em sete *loci*, existe uma boa chance de o suspeito ser culpado; em seguida determinamos a freqüência daquela seqüência na população. Em *Stalking Justice*, as chances eram de 1 em 750.000.000 de outra pessoa ter o mesmo perfil DNA do acusado no caso. Aplicamos o teste de hipóteses e determinamos o nível de significância para o perfil DNA específico. O DNA é também muito importante em investigações de paternidade e casos de estupro. Com os testes convencionais de sangue-enzima, poderíamos chegar a cerca de 10% da população. Isso significa que há uma chance em 10 de o culpado não ser o acusado. Com a dactiloscopia DNA, temos uma chance em 300.000.000, e entramos assim no domínio da inevitabilidade estatística. O processo é usado não somente para condenar pessoas, mas também para excluir suspeitos. Há um caso famoso na Carolina do Norte, em que duas testemunhas oculares testificaram que o acusado era um estuprador. Ele ficou preso 11 anos, mas foi liberado quando o DNA mostrou que ele não era o culpado. Nesse caso, a estatística e o DNA se revelaram muito mais precisos do que as identificações por testemunhas oculares.

Tenho utilizado a estatística em casos de homicídio, de abusos de crianças, de mulheres espancadas e de paternidade. Nos casos de paternidade hoje, os resultados do DNA são tão precisos que todo o sistema de julgamento está sendo abreviado. Os condenados simplesmente não vão a julgamento quando os resultados do DNA são bastante claros. Uma dúvida razoável se transforma em nenhuma dúvida. A grande pergunta é: "A estatística é tão poderosa a ponto de tirar do júri sua responsabilidade de tomar decisões?" Há exceções, mas, na maioria dos casos, a presença de uma forte evidência estatística é um instrumento eficiente para a tomada de decisões.

Como advogado que faz extenso uso da estatística, o senhor acha que todos os advogados deveriam conhecer os princípios da estatística?

Eles necessitam de muito mais. Se queremos dominar efetivamente nossa evidência, devemos ter algum fundamento estatístico. O problema é que, hoje, os advogados costumam recorrer a outros peritos. Nos casos de morte acidental, por exemplo, costumam recorrer a estatísticos para obter dados atuariais sobre a vida provável de alguém. É raro o advogado que sequer entende o que a estatística está dizendo, de forma que seria recomendável que qualquer pessoa desejosa de ingressar na carreira de advogado, estudasse estatística.

O senhor utiliza a estatística em seu trabalho com abuso de crianças e violência?

Tenho grande interesse na relação entre abuso de crianças e violência, e uma das melhores formas de convencer os legisladores, ou jurados, ou audiências é utilizar a estatística. Entre os adolescentes que matam seus pais, sabemos que um dos maiores fatores de risco ocorre quando as crianças vêem seus pais espancarem suas mães. O estado do Texas fez uma pesquisa e constatou que, entre os meninos que cometeram homicídio, 66 por cento haviam matado alguém que, de alguma forma, fizera mal às suas mães. Quanto mais conhecemos acerca de uma população, mais sabemos que pesquisas estatísticas devemos fazer.

Recomenda a estatística para os alunos de universidades hoje?

A estatística não é somente para os que lidam com ciências exatas. Proporciona importantes recursos para os que desejam tornar-se advogados, médicos, enfermeiros ou policiais. Verifiquei que posso apreciar melhor notícias sobre eventos, notícias financeiras e demonstrações de lucros e perdas para ações. A estatística é mais importante do que boa parte da matemática básica ensinada. Utilizo muito mais a estatística do que a geometria ou a trigonometria.

Que outros conhecimentos são importantes para o universitário de hoje?

Em uma época em que tudo está sendo computadorizado, as pessoas estão dando cada vez menos atenção à comunicação, de forma que a arte da palavra falada está sendo um tanto negligenciada. Pode haver milhões de pessoas capazes de utilizarem um computador, mas muito poucas em condições de se dirigirem a uma assembléia. As pessoas pecisam também ter capacidade de comunicar suas idéias por escrito.

Rayane Andrade

Triola

2 Descrição, Exploração e Comparação de Dados

2-1 Aspectos Gerais

O capítulo apresenta tabelas, gráficos e medidas importantes que podem ser utilizados para descrever ou explorar um conjunto de dados, ou comparar dois ou mais conjuntos. Em capítulos posteriores serão utilizados muitos conceitos importantes ora introduzidos.

2-2 Resumo de Dados com Tabelas de Freqüência

Descreve-se a construção de tabelas de freqüência, tabelas de freqüência relativa e tabelas de freqüência acumulada. Essas tabelas são úteis para condensar grandes conjuntos de dados, facilitando o seu manuseio.

2-3 Representação Pictórica de Dados

Apresentamos métodos de construção de histogramas, histogramas de freqüências relativas, gráficos por pontos, gráficos tipo ramo-e-folha, gráfico em setores, diagramas de Pareto e diagramas de dispersão. Tais gráficos auxiliam grandemente a visualização de características dos dados que, de outra forma, permaneceriam encobertas.

2-4 Medidas de Tendência Central

As medidas de tendência central são tentativas de determinação de valores que representam conjuntos de dados. Definimos as seguintes medidas de tendência central: média, mediana, moda, ponto médio e média ponderada. Abordamos também o conceito de assimetria.

2-5 Medidas de Variação

As medidas de variação são números que refletem o grau de dispersão entre os valores de um conjunto de dados. Definem-se as seguintes medidas de variação: amplitude, desvio-padrão, desvio médio e variância. Tais medidas têm extrema importância em análise estatística.

2-6 Medidas de Posição

Define-se o escore padronizado (ou escore z), mostrando como identificar valores atípicos. Definem-se também percentis, quartis e decis, utilizados para comparar valores dentro do mesmo conjunto de dados.

2-7 Análise Exploratória de Dados (EDA – Exploratory Data Analysis)

Apresentamos técnicas para explorar dados com o resumo de cinco números e com diagramas de caixas (*boxplots*). Estes últimos são especialmente adaptados para comparar diferentes conjuntos de dados.

Problema do Capítulo

As latas de alumínio de 12 oz podem ter menor espessura para reduzir o custo?

O Conjunto de Dados 15 do Apêndice B inclui estas duas amostras:

1. Latas de alumínio de 12 oz com espessura de 0,0109 in. (0,0278 cm) (reproduzido como Tabela 2.1)
2. Latas de alumínio com espessura de 0,0111 in. (0,0282 cm)

Exploraremos os valores da Tabela 2.1, que relaciona as cargas axiais (em libras) da amostra de latas de alumínio de 0,0109 de espessura. Este conjunto de dados foi fornecido por um estudante que utilizou a edição anterior deste livro. Trata-se de uma funcionária da companhia que fabrica essas latas; ela utiliza métodos aprendidos em seu curso introdutório de estatística. O autor agradece essa contribuição.

A carga axial de uma lata é o peso máximo suportado por seus lados, e é medida utilizando-se uma placa para aplicar uma pressão crescente ao topo da lata, até que ela ceda. É importante termos uma carga axial suficientemente grande a fim de a lata não ceder quando se coloca a tampa sob pressão. Nesse processo de fabricação, os topos das latas são colocados no lugar com uma pressão que varia de 158 a 165 libras.

As latas menos espessas têm a vantagem óbvia de utilizar menos material, o que reduz o custo, mas não são provavelmente tão resistentes quanto as mais espessas. A empresa que fabrica essas latas costuma utilizar uma espessura de 0,0111 in. mas está testando latas de menor espessura. Com os métodos deste capítulo, exploraremos o conjunto de dados (reproduzido na Tabela 2-1) para essas latas menos espessas (0,0109 in. de espessura). E determinaremos, afinal, se essas latas menos espessas podem realmente ser usadas.

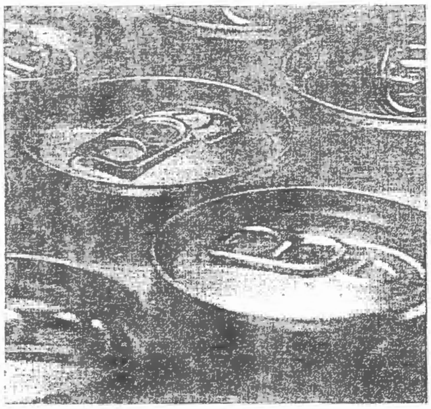


TABELA 2-1 Cargas Axiais de Latas de 0,0109 in. (0,0278 cm)

270	273	258	204	254	228	282
(122)*	(124)	(117)	(93)	(115)	(103)	(128)
278	201	264	265	223	274	280
(126)	(91)	(119)	(120)	(101)	(124)	(104)
250	275	281	271	263	277	275
(113)	(125)	(127)	(123)	(119)	(126)	(125)
278	260	262	273	274	286	236
(126)	(118)	(119)	(124)	(124)	(130)	(207)
290	286	278	283	262	277	295
(132)	(130)	(126)	(128)	(119)	(126)	(134)
274	272	265	275	263	251	289
(124)	(123)	(120)	(125)	(119)	(114)	(131)
242	284	241	276	200	278	283
(110)	(129)	(109)	(125)	(91)	(126)	(128)
269	282	267	282	272	277	261
(122)	(128)	(121)	(128)	(123)	(126)	(118)
257	278	295	270	268	286	262
(117)	(126)	(134)	(122)	(122)	(130)	(119)
272	268	283	256	206	277	252
(123)	(122)	(128)	(116)	(93)	(126)	(114)
265	263	281	268	280	289	283
(120)	(119)	(127)	(121)	(127)	(131)	(128)
263	273	209	259	287	269	277
(119)	(124)	(95)	(117)	(130)	(122)	(126)
234	282	276	272	257	267	204
(104)	(128)	(125)	(123)	(117)	(121)	(93)
270	285	273	269	284	276	286
(122)	(129)	(123)	(122)	(129)	(125)	(130)
273	289	263	270	279	206	270
(124)	(131)	(119)	(122)	(127)	(93)	(122)
270	268	218	251	252	284	278
(122)	(122)	(99)	(114)	(114)	(129)	(126)
277	208	271	208	280	269	270
(126)	(94)	(123)	(94)	(127)	(122)	(122)
294	292	289	290	215	284	283
(133)	(132)	(131)	(132)	(98)	(129)	(128)
279	275	223	220	281	268	272
(127)	(125)	(101)	(100)	(127)	(121)	(123)
268	279	217	259	291	291	281
(122)	(127)	(98)	(117)	(132)	(132)	(127)
230	276	225	282	276	289	288
(104)	(125)	(102)	(128)	(125)	(131)	(131)
268	242	283	277	285	293	248
(122)	(110)	(128)	(126)	(129)	(133)	(112)
278	285	292	282	287	277	266
(126)	(129)	(132)	(128)	(130)	(126)	(121)
268	273	270	256	297	280	256
(122)	(124)	(122)	(116)	(135)	(127)	(116)
262	268	262	293	290	274	292
(119)	(122)	(119)	(133)	(132)	(124)	(132)

*Os números entre parênteses são as cargas axiais em kg.

2-1 Aspectos Gerais

Às vezes coletamos dados visando a um fim específico. Por exemplo, um estudo sobre a segurança dos elevadores de um edifício exigiria dados relativos ao peso médio das pessoas que os utilizam. Em outros casos, coletamos ou obtemos dados não com uma finalidade específica, mas porque desejamos explorá-los para ver o que pode ser revelado. A um geólogo podem interessar os intervalos de tempo entre as erupções do gêiser Old Faithful — são elas igualmente espaçadas ao longo do tempo, ou alguns intervalos de tempo são mais freqüentes do que outros? Em ambas as circunstâncias, necessitamos de uma diversidade de recursos que contribuam para entendermos o conjunto de dados. Este capítulo apresenta tais recursos.

Ao analisarmos um conjunto de dados, devemos determinar em primeiro lugar se se trata de uma *amostra* ou de uma *população* completa. Essa determinação afetará não somente os métodos utilizados, mas também as conclusões a que chegarmos. Utilizamos métodos de **estatística descritiva** para resumir ou descrever as características importantes de um conjunto conhecido de dados populacionais, e recorremos a métodos de **inferência estatística** quando utilizamos dados amostrais para fazer inferências (ou generalizações) sobre uma população. Quando um professor calcula a média final de um exame para determinada turma, o resultado é um exemplo de uma estatística descritiva, se consideramos a população como toda a turma. Mas se afirmamos que o resultado é uma estimativa da média do exame final de todas as turmas, estamos fazendo uma inferência que ultrapassa o âmbito dos dados conhecidos.

A estatística descritiva e a inferência estatística são dois grandes ramos da estatística. Neste capítulo abordamos os conceitos básicos da estatística descritiva.

Características Importantes dos Dados

Com os recursos da estatística descritiva, podemos entender melhor um conjunto de dados através de suas características. As três características seguintes são extremamente importantes e proporcionam uma visão bastante satisfatória:

1. A natureza ou forma da distribuição dos dados, como forma de sino, uniforme ou assimétrica.
2. Um valor representativo, como uma média.
3. Uma medida de dispersão ou variação.

Podemos conhecer alguma coisa da natureza ou forma da distribuição organizando os dados e construindo gráficos, como nas Seções 2-2 e 2-3. Na Seção 2-4, veremos como obter valores representativos. Avaliaremos a extensão da dispersão, ou variação entre dados, com auxílio dos recursos da Seção 2-5. Na Seção 2-6 definiremos medidas de posição que nos permitem melhor analisar ou comparar diversos valores. E na Seção 2-7 estudaremos métodos de exploração de conjuntos de dados.

2-2 Resumo de Dados com Tabelas de Freqüência

Ao estudarmos grandes conjuntos de dados, é conveniente organizá-los e resumi-los, construindo uma tabela de freqüências.

DEFINIÇÃO

Uma **tabela de freqüências** relaciona categorias (ou classes) de valores, juntamente com contagens (ou freqüências) do número de valores que se enquadram em cada categoria.

A Tabela 2-2 é uma tabela de freqüências com 10 classes (ou categorias). A freqüência de determinada classe é o número de observações originais que se enquadram naquela classe. Por exemplo, a primeira classe na Tabela 2-2 tem uma freqüência de 9, indicando que há 9 valores entre 200 e 209 inclusive.

TABELA 2-2 Cargas Axiais de Latas de Alumínio

Carga Axial	Freqüência
200-209	9
210-219	3
220-229	5
230-239	4
240-249	4
250-259	14
260-269	32
270-279	52
280-289	38
290-299	14

Começaremos apresentando alguns termos-padrão no estudo de tabelas de freqüência e, em seguida, descreveremos um processo para construí-las. (Há vários pacotes estatísticos que constroem essas tabelas automaticamente.)

DEFINIÇÕES

Limites Inferiores de Classes são os menores números que podem efetivamente pertencer às diferentes classes. (Na Tabela 2-2 os limites inferiores de classe são 200, 210, ..., 290.)

Limites Superiores de Classes são os maiores números que podem efetivamente pertencer às diferentes classes. (A Tabela 2-2 tem os limites superiores de classe 209, 219, ..., 299.)

Fronteiras de Classes são os números usados para separar classes, mas sem as lacunas criadas pelos limites de classe. São obtidos como segue: Determinamos o tamanho da lacuna entre o limite superior de uma classe e o limite inferior da classe seguinte, adicionamos metade desse valor a cada limite superior de classe, obtendo as fronteiras superiores de classes; subtraímos metade daquele valor de cada limite inferior de classe, obtendo as fronteiras inferiores de classe. (Na Tabela 2-2 as fronteiras de classe são 199,5, 209,5, 219,5, ..., 299,5.)

Marcas de Classe são os pontos médios das classes. (Na Tabela 2-2 os pontos médios são 204,5, 214,5, ..., 294,5.) Cada marca de classe é obtida somando-se o limite inferior ao limite superior correspondente, e dividindo-se o resultado por 2.

Amplitude de Classe é a diferença entre dois limites de classe inferiores consecutivos ou entre duas fronteiras inferiores de classe consecutivas. (Na Tabela 2-2 a amplitude de classe é 10.)

As definições de marca de classe e fronteira de classe podem ser enganosas. Devemos ter o cuidado de evitar o erro de tomar como amplitude de classe a diferença entre o limite inferior de classe e o correspondente limite superior. Veja a Tabela 2-2 e note que a amplitude de classe é 10, e não 9. (Os estudantes costumam ter dificuldade com as fronteiras de classe. Veja a discussão na seção seguinte.) Observe os limites de classe na Tabela 2-2 e note que há uma lacuna entre 209 e 210, outra entre 219 e 220 e assim por diante. As fronteiras de classe basicamente dividem diferenças e preenchem as lacunas, facilitando a construção de certos gráficos. Examine cuidadosamente, durante algum tempo, a definição de fronteira de classe, até ter entendido perfeitamente.

O processo de construção de uma tabela de freqüência envolve os seguintes passos:

Passo 1: *Decidir o número de classes de sua tabela de freqüência.* A título de orientação, o número de classes deve ficar entre 5 e 20. O número efetivo de classes pode depender da conveniência de utilizar números arredondados ou de outros fatores subjetivos. Com notas de testes, por exemplo, pode ser conveniente utilizar 10 classes: 50-54, 55-59, 60-64, ..., 95-99.

Passo 2: *Determinar a amplitude de classe, dividindo a amplitude pelo número de classes.* (A amplitude é a diferença entre o maior e o menor valor.) Arredonde o resultado *para mais*, até um número conveniente. Esse arredondamento para mais não somente é conveniente como também garante que todos os valores sejam incluídos na tabela de freqüências. (Se o número de classes divide *exatamente* a amplitude, é preciso acrescentar mais uma classe para que todos os dados sejam incluídos.)

$$\text{Amplitude de classe} = \frac{\text{amplitude}}{\text{número de classes}} \text{ arredondado } \textit{para mais}$$

Passo 3: *Escolher como limite inferior da primeira classe o menor valor observado ou um valor ligeiramente inferior a ele.* Esse valor serve como ponto de partida.

Passo 4: *Some a amplitude de classe ao ponto de partida, obtendo o segundo limite inferior de classe.* Adicione a amplitude de classe ao segundo limite inferior para obter o terceiro; e assim por diante.

Passo 5: *Relacione os limites inferiores de classe em uma coluna e introduza os limites superiores, que podem ser facilmente determinados a esta altura.*

Passo 6: *Represente cada observação por um pequeno traço na classe apropriada e, com auxílio desses traços, determine a freqüência total de cada classe.*

Como a determinação do número de classes ainda não é uma imposição legal, podemos tomar um número diferente de classes que resulte em uma tabela de freqüências diferente e igualmente correta. Novamente frisamos que a prioridade deve ser a obtenção de uma tabela com valores convenientes e compreensíveis.

Autores Identificados

Em 1787-88, Alexander Hamilton, John Jay e James Madison publicaram anonimamente os famosos panfletos *Federalist*, como uma tentativa de convencer os nova-iórquinos a ratificarem a nascente Constituição. A identidade da maioria dos autores dos panfletos tornou-se conhecida, mas a autoria de doze deles foi contestada. Através da análise estatística das freqüências de diversas palavras, podemos agora concluir que James Madison foi o autor *provável* desses 12 panfletos. Em muitos deles, a evidência da autoria de Madison é esmagadora, a ponto de podermos considerá-la praticamente certa.

EXEMPLO Construa uma tabela de freqüências para as 175 cargas axiais de latas de alumínio da Tabela 2-1.

SOLUÇÃO Indicaremos os passos que conduzem à tabela de freqüências mostrada na Tabela 2-2.

Passo 1: Começamos escolhendo 10 como o número de classes. (Muitos estatísticos recomendam de modo geral o uso de 10 classes, mas utilizam um número menor de classes para conjuntos menores de dados, e um número maior para conjuntos maiores.)

Passo 2: Com um mínimo de 200 e um máximo de 297, a amplitude total é $297 - 200 = 97$.

$$\begin{aligned}\text{intervalo de classe} &= \text{arredondamento de } \frac{97}{10} \text{ para cima}\\ &= \text{arredondamento de } 9{,}7 \text{ para cima}\\ &= 10 \textbf{ (arredondamento para cima pela conveniência de termos um número inteiro)}\end{aligned}$$

Passo 3: O menor valor é 200. Como é um valor conveniente, tomamo-lo como ponto de partida e limite inferior da primeira classe.

Passo 4: Adicionando a amplitude de classe 10 ao limite inferior 200, obtemos o próximo limite inferior 210. Prosseguindo, obtemos os outros limites 220, 230 etc.

Passo 5: Esses limites inferiores sugerem os seguintes limites superiores de classe:

200 209
210 219
etc .

Passo 6: A coluna direita da Tabela 2-2 apresenta as contagens, ou freqüências.

A Tabela 2-2 nos dá informações úteis tornando a lista de cargas axiais mais inteligível, mas perdemos a precisão dos dados originais. Por exemplo, a primeira classe 200-209 indica 9 observações, mas não há maneira de sabermos, pela tabela, quais são precisamente esses valores. Não podemos reconstruir os 175 valores iniciais das cargas axiais com base na tabela de freqüências; sacrificamos a exatidão dos dados originais para termos dados mais compreensíveis.

Na construção de tabelas de freqüência, devemos observar as seguintes diretrizes:

1. *As classes devem ser mutuamente excludentes.* Ou seja, cada valor original deve pertencer exatamente a uma, e uma só classe.
2. *Todas as classes devem ser incluídas,* mesmo as de freqüência zero.
3. *Procurar utilizar a mesma amplitude para todas as classes,* embora eventualmente seja impossível evitar intervalos com extremidade aberta, como "65 anos ou mais".
4. *Escolher números convenientes para limites de classe.* Arredondar para cima a fim de ter menos casas decimais, ou utilizar números adequados à situação.
5. *Utilizar entre 5 e 20 classes.*
6. *A soma das freqüências das diversas classes deve ser igual ao número de observações originais.*

Tabela de Freqüências Relativas

Uma modalidade importante da tabela básica de freqüência utiliza **freqüências relativas**, que se obtêm dividindo a freqüência de cada classe pela freqüência total. A **tabela de freqüências relativas** tem os mesmos limites de classe que a tabela de freqüências; apenas, apresenta freqüências relativas em lugar das freqüências absolutas.

$$\text{freqüência relativa} = \frac{\text{freqüência da classe}}{\text{freqüência total}}$$

A Tabela 2-3 apresenta as freqüências relativas das 175 cargas axiais resumidas na Tabela 2-2. A primeira classe tem uma freqüência relativa de 9/175 = 0,051. (As freqüências relativas também podem ser apresentadas como porcentagens; isto é, 0,051 pode expressar-se como 5,1%.) A segunda classe tem uma freqüência relativa de 3/175 = 0,017 etc. Quando calculadas corretamente, a soma das freqüências relativas deve ser 1 (ou 100%), admitindo-se pequenas discrepâncias como conseqüência de arredondamentos.

TABELA 2-3 Freqüência Relativa das Cargas Axiais de Latas de Alumínio

Carga Axial	Freqüência Relativa
200-209	0,051
210-219	0,017
220-229	0,029
230-239	0,023
240-249	0,023
250-259	0,080
260-269	0,183
270-279	0,297
280-289	0,217
290-299	0,080

TABELA 2-4 Freqüência Acumulada das Cargas Axiais

Carga Axial	Freqüência Acumulada
Menos de 210	9
Menos de 220	12
Menos de 230	17
Menos de 240	21
Menos de 250	25
Menos de 260	39
Menos de 270	71
Menos de 280	123
Menos de 290	161
Menos de 300	175

As tabelas de freqüência relativa facilitam a compreensão da distribuição e a comparação de diferentes conjuntos de dados. Assim, é mais fácil dizer que 5,1% das latas têm carga axial entre 200 e 209 lb do que dizer que 9 das 175 latas têm carga axial entre aqueles valores. Veja também o Exercício 21, para exemplo de uma situação em que a comparação é facilitada pelo uso de tabelas de freqüência relativa.

Tabela de Freqüências Acumuladas

Obtemos outra variante da tabela de freqüências quando desejamos as freqüências acumuladas. A **freqüência acumulada** de uma classe é a soma das freqüências daquela classe e de todas as classes que a antecedem. A Tabela 2-4, que representa as mesmas 175 latas de alumínio da Tabela 2-2, é um exemplo de **tabela de freqüência acumulada**, onde se registram as freqüências acumuladas em lugar das freqüências das classes individuais. A comparação da coluna de freqüências da Tabela 2-2 com a coluna de freqüências acumuladas da Tabela 2-4 mostra que os valores das freqüências acumuladas se obtêm partindo da freqüência da primeira classe e somando sucessivamente as freqüências de cada classe subseqüente. Por exemplo, há 9 valores inferiores a 210, 9 + 3 = 12 valores inferiores a 220 e assim por diante. Construída corretamente, a última freqüência acumulada deve ser igual ao total de observações no conjunto.

Com as tabelas de freqüência, podemos identificar a natureza geral da distribuição dos dados, bem como construir gráficos que facilitem a visualização dessa distribuição. Na próxima seção estudaremos esses gráficos.

2-2 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-4, identifique, para cada tabela de freqüências, a amplitude da classe, os pontos médios das classes e as fronteiras de classe.

1.

Ausências	Freqüência
0–5	39
6–11	41
12–17	38
18–23	40
24–29	42

2.

Ausências	Freqüência
0–9	22
10–19	40
20–29	71
30–39	44
40–49	23

3.

Peso (kg)	Freqüência
0,0–1,9	20
2,0–3,9	32
4,0–5,9	49
6,0–7,9	31
8,0–9,9	18

4.

Peso (kg)	Freqüência
0,0–4,9	60
5,0–9,9	58
10,0–14,9	61
15,0–19,9	62
20,0–24,9	59

Nos Exercícios 5-8, construa a tabela de freqüências relativas correspondente à tabela de freqüências do exercício indicado.

5. Exercício 1
6. Exercício 2
7. Exercício 3
8. Exercício 4

Nos Exercícios 9-12, construa a tabela de freqüências acumuladas correspondente à tabela de freqüências do exercício indicado.

9. Exercício 1
10. Exercício 2
11. Exercício 3
12. Exercício 4
13. Compare a *distribuição* de dados do Exercício 1 com a *distribuição* de dados do Exercício 2. Qual é a diferença básica?
14. Compare a *distribuição* dos dados do Exercício 3 com a *distribuição* de dados do Exercício 4. Qual é a diferença básica?

Nos Exercícios 15-16, use a informação dada para determinar limites superior e inferior da primeira classe. (Os dados constam do Apêndice B, mas não é preciso recorrer ao apêndice para esses exercícios.)

15. Um conjunto de dados consiste em pesos de metal coletados de lares em uma semana; esses pesos variam de 0,26 lb a 4,95 lb. Desejamos construir uma tabela de freqüências com 10 classes.
16. Uma amostra de bombons M&M tem pesos que vão de 0,838 g a 1,033 g. Desejamos construir uma tabela de freqüências com 12 classes.

Nos Exercícios 17-20, construa uma tabela de freqüências com os valores indicados.

17. Para o Conjunto de Dados 3 do Apêndice B, construa uma tabela de freqüências dos pesos de ursos. Tome 11 classes, começando com 0 como limite inferior de classe.
18. Para o Conjunto de Dados 2 do Apêndice B, construa uma tabela de freqüências das temperaturas à meia-noite do segundo dia. Tome 8 classes, começando com 96,5 como limite inferior.
19. Para o Conjunto de Dados 16 do Apêndice B, construa uma tabela de freqüências para os intervalos de tempo entre erupções do gêiser Old Faithful no Parque Nacional de Yellowstone. Tome 7 classes, começando com um limite inferior de 56 min.; adote uma amplitude de classe de 8 min.
20. Para o Conjunto de Dados 11 do Apêndice B, construa uma tabela de freqüências com 12 classes para os pesos de todos os 100 bombons M&M.

2-2 Exercícios B: Além do Básico

21. A seguir é dada uma tabela de freqüências de consumo de álcool antes da prisão, para prisioneiros do sexo masculino cumprindo penas por dirigirem embriagados e a tabela correspondente para mulheres (com base em dados do Ministério da Justiça dos EUA). Construa primeiro as tabelas de freqüência relativa e use em seguida os resultados para comparar as duas amostras. Note que é difícil comparar as freqüências originais, mas é muito mais fácil comparar as freqüências relativas.

Etanol Consumido por Homens (oz)	Freqüência
0,0–0,9	249
1,0–1,9	929
2,0–2,9	1545
3,0–3,9	2238
4,0–4,9	1139
5,0–9,9	3560
10,0–14,9	1849
15,0 ou mais	1546

Etanol Consumido por Mulheres (oz)	Freqüência
0,0–0,9	7
1,0–1,9	52
2,0–2,9	125
3,0–3,9	191
4,0–4,9	30
5,0–9,9	201
10,0–14,9	43
15,0 ou mais	72

22. A seguir são listados dois conjuntos de dados que se supõe serem as alturas (em polegadas) de homens adultos escolhidos aleatoriamente. Um conjunto consiste em alturas obtidas efetivamente de um conjunto aleatório de homens adultos, mas o outro conjunto consiste em números "fabricados". Construa uma tabela de freqüências para cada conjunto de dados. Examinando as duas tabelas de freqüências, identifique o conjunto que lhe parece ser falso, e justifique sua conclusão.

a. 70 73 70 72 71 73 71 67 68 72 67 72 71 73
72 70 72 68 71 71 71 73 69 73 71 66 77 67

b. 70 73 70 72 71 66 74 76 68 75 67 68 71 77
66 69 72 67 77 75 66 76 76 77 73 74 69 67

23. A tabela de freqüências a seguir resume dados do Departamento do Censo dos EUA. Recorde as 5 diretrizes para construção de tabelas de freqüências e identifique as diretrizes que não foram seguidas.

Idade	População dos EUA(milhões)
Menos de 15	55
15–24	37
25–44	82
45 ou mais	79

24. Ao construir uma tabela de freqüências, Sturges sugere que o número ideal de classes pode ser aproximado por $1 + (\log n)/(\log 2)$, onde n é o número de observações. Com esta orientação, determine o número ideal de classes (arredondado para menos, e não para mais) para um conjunto de dados com número de elementos igual a

a. 50 **b.** 100 **c.** 150
d. 500 **e.** 1000 **f.** 50.000

2-3 Representação Pictórica de Dados

Na Seção 2-2, utilizamos tabelas de freqüências para transformar coleções de dados brutos em sumários organizados e compreensíveis. O objetivo precípuo desta seção é apresentar métodos de representação de dados em uma forma pictórica que nos permita visualizar facilmente a natureza da distribuição.

Histogramas e a Forma dos Dados

Um recurso gráfico, comum e importante, para apresentação de dados é o histograma, do qual temos um exemplo na Figura 2-1. Um **histograma** consiste em uma escala horizontal para os valores dos dados a serem representados, uma escala vertical para as freqüências e barras para representar os valores das freqüências das diversas classes. Em geral, a construção de um histograma para representar um conjunto de valores é precedida de uma tabela completa de freqüências daqueles valores. Cada barra é delimitada pela fronteira inferior de classe à esquerda e pela fronteira superior de classe à direita. Obtém-se, entretanto, melhor legibilidade tomando-se os pontos médios das classes em lugar das fronteiras das classes. O histograma da Figura 2-1 corresponde diretamente à tabela de freqüências (Tabela 2-2 da seção anterior).

Antes de construir um histograma com base em uma tabela de freqüências, devemos atentar para as escalas usadas nos eixos vertical e horizontal. A freqüência máxima (ou maior número mais próximo conveniente) deve sugerir o maior valor para a escala vertical; 0 deve ser a base. Na Figura 2-1, a escala vertical vai de 0 a 60. A escala horizontal deve ser construída de modo a abranger todas as classes da tabela de freqüências. Idealmente, devemos procurar seguir a regra empírica, segundo a qual a altura vertical do histograma deve ser cerca de três quartos da largura total. Ambos os eixos devem ser demarcados sem qualquer ambigüidade.

Um **histograma de freqüências relativas** tem a mesma forma e a mesma escala horizontal que um histograma, mas a escala vertical apresenta *freqüências relativas* em lugar de freqüências absolutas, como na Figura 2-2. A Figura 2-1 pode ser modificada para um histograma de freqüências relativas simplesmente

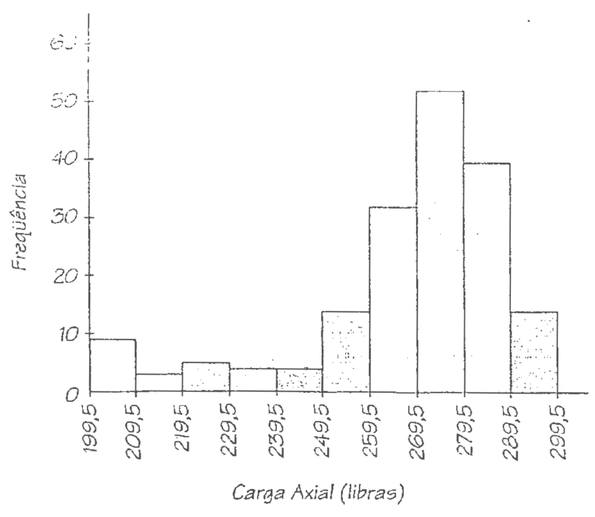


Fig. 2-1 Histograma das cargas axiais de latas de alumínio.

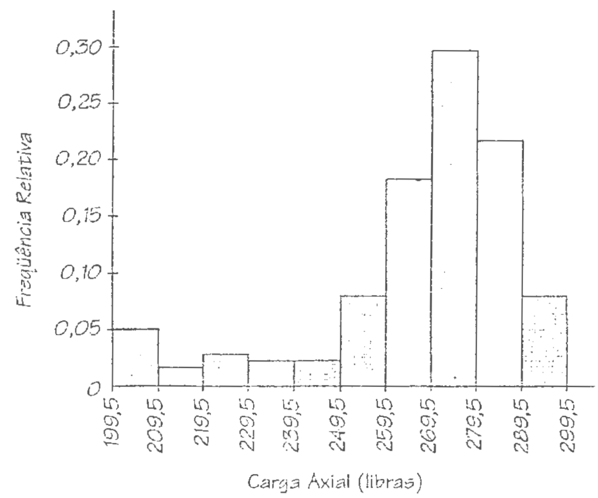
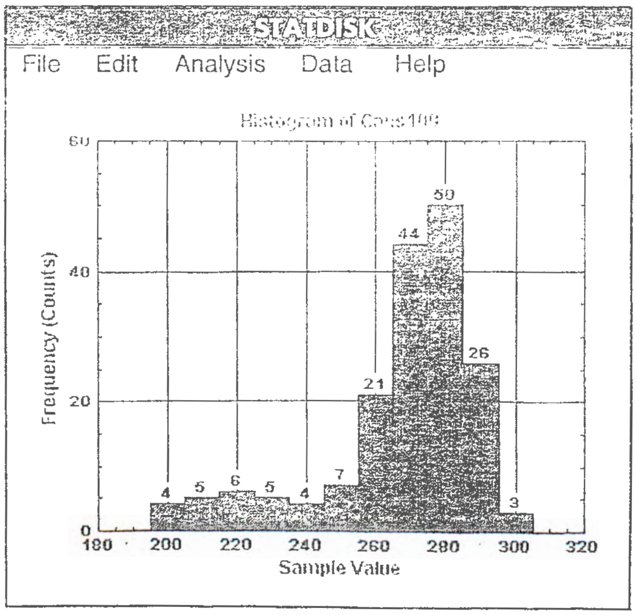


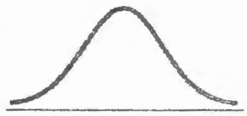
Fig. 2-2 Histograma das freqüências relativas das cargas axiais de latas de alumínio.

designando a escala vertical como "freqüência relativa" e modificando os valores respectivos para a escala de 0 a 0,300, conforme a Figura 2-2. (A maior freqüência relativa para esse conjunto de dados é 0,297, de forma que tem sentido tomar 0,300 como valor máximo na escala vertical; o fato de a maior freqüência relativa ser 0,297 e o maior valor ser 297 é mera coincidência.) Assim como o histograma da Figura 2-1 representa a tabela de freqüências da Tabela 2-2, o histograma de freqüências relativas da Figura 2-2 representa a tabela de freqüências relativas da Tabela 2-3.

Geração de Histogramas com o Uso de Calculadoras e Computadores

Apresentamos a seguir um histograma, feito por STATDISK, das cargas axiais de latas de alumínio com que estamos trabalhando





Em forma de sino

neste capítulo. A apresentação STATDISK é obtida utilizando-se `Data` da barra principal de ferramentas e introduzindo-se os dados com auxílio da opção `Sample Editor`. Utilizam-se então os comandos `Copy` e `Paste` para usar os dados no programa `Histogram`, que também se encontra sob `Data`. (Os comandos copy e paste são comuns a muitos programas Windows.) A apresentação do histograma pode ser obtida da versão Windows de Minitab, introduzindo primeiro os dados sob a coluna C1 na grade de dados. Utilizam-se então as opções `Graph` e `Histogram`. Pode-se gerar um histograma também em algumas calculadoras gráficas, como a TI-82 e a TI-83.

As tabelas de freqüências e os gráficos tais como histogramas permitem-nos ver como se distribuem nossos dados; a distribuição dos dados é uma característica extremamente importante. As Figuras 2-3 e 2-4 são histogramas de dados reais (ver Conjuntos de Dados 12 e 13 no Apêndice B) com distribuições fundamentalmente diferentes.

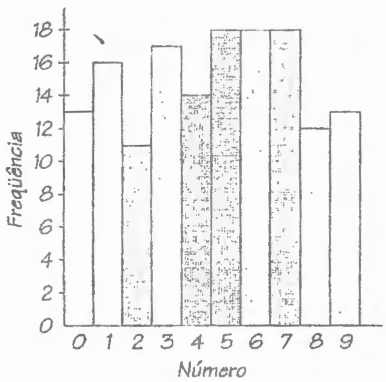


Fig. 2-3 Histograma dos resultados de uma loteria.

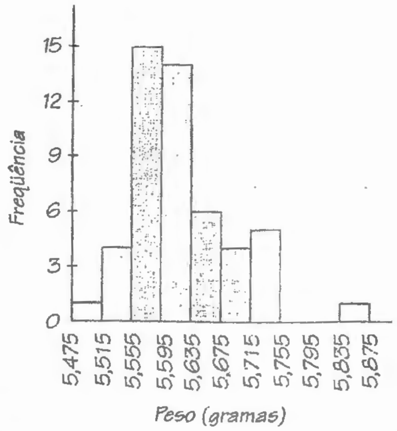


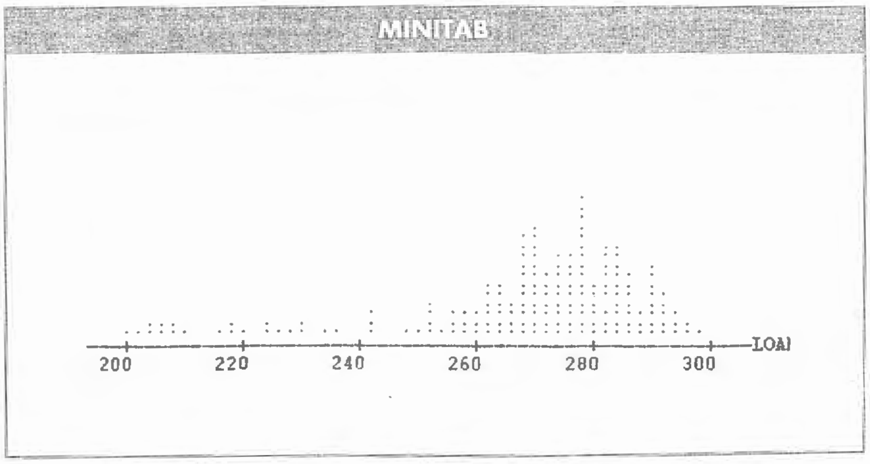
Fig. 2-4 Histograma dos pesos de moedas de 25 cents.

A Figura 2-3 é basicamente plana, ou uniforme, enquanto a Figura 2-4 tem aproximadamente a forma de um sino, no sentido de que se assemelha à segunda figura anterior sem número ilustrada aqui. Como a Figura 2-3 mostra algarismos selecionados da Loteria Pick Three de Maryland, é de se esperar que todos os algarismos sejam igualmente prováveis e que o histograma seja basicamente plano, como na Figura 2-3. Qualquer discrepância sensível da forma plana, ou uniforme, sugere que há algo errado com a loteria.

A forma de sino dos pesos das moedas de 25 centavos de dólar da Figura 2-4 é típica de uma ampla diversidade de circunstâncias, especialmente em processos de fabricação. Muitos processos estatísticos exigem que um conjunto de dados tenha uma distribuição em forma de sino análoga à apresentada na Figura 2-4, e uma maneira de verificar esse comportamento consiste em construir um histograma.

Gráficos por Pontos

A figura a seguir é um gráfico por pontos dos mesmos dados relativos a latas de alumínio relacionados na Tabela 2-1, obtido com o programa Minitab. (Com Minitab, introduzimos os dados



e selecionamos as opções `Graph`, `Character Graphs` e `Dotplot`.) Por esta ilustração, é muito fácil ver que um **gráfico por pontos** consiste em um gráfico em que cada observação é representada por um ponto ao longo da escala de valores. O ponto mais à esquerda, por exemplo, representa a carga axial de 200 lb. Quando os valores ocorrem mais de uma vez, são marcados como pontos em colunas verticais acima do valor correspondente na escala. Assim é que, nos dados da Tabela 2-1, a carga de 204 ocorre duas vezes, e esses valores são representados pelos dois pontos situados acima da locação correspondente a 204. (Este gráfico por pontos utiliza 10 intervalos para representar uma amplitude de 20 libras, de forma que cada traço na escala horizontal tracejada representa dois valores. O traço logo antes de 200 representa os valores de 199 e 200.)

O gráfico por pontos é análogo ao histograma pelo fato de permitir que vejamos a *distribuição* dos dados.

Gráficos Ramo-e-folhas

Já vimos que a construção de uma tabela de freqüências e do histograma correspondente nos dá informações valiosas sobre a natureza da distribuição dos dados, mas há a desvantagem de perdermos alguns detalhes sobre os mesmos. Em geral, não podemos recompor os dados originais a partir da tabela de freqüências ou do histograma. Vamos introduzir agora os gráficos do tipo ramo-e-folhas, que permitem vermos a distribuição dos dados *sem* perda de informação no processo.

Em um **gráfico ramo-e-folhas**, classificamos os dados segundo um padrão que revela a distribuição subjacente. O padrão consiste em separar um número (como 257) em duas partes — em geral, o primeiro ou os dois primeiros algarismos (25) e o outro algarismo (7). O ramo consiste nos algarismos mais à esquerda (25 neste caso), e as folhas consistem nos algarismos mais à direita (7, no caso). O método é ilustrado no exemplo seguinte.

EXEMPLO Construa um gráfico ramo-e-folhas com as cargas axiais de latas de alumínio da Tabela 2-1.

SOLUÇÃO Tomando os dois algarismos mais à esquerda como ramos, estes serão 20, 21, ..., 29. Traçamos então uma reta vertical e relacionamos as folhas conforme mostrado a seguir. O primeiro valor na Tabela 2-1 é 270; incluímos este valor registrando um 0 na linha (ramo) para 27. Continuamos a incluir todos os 175 valores, e compomos as folhas (os algarismos localizados à direita) de forma que os números se disponham em ordem crescente. A primeira linha representa os números 200, 201, 204, 204, 206 etc.

Ramo	Folhas
20	014466889
21	578
22	03358
23	0046
24	1228
25	01122466677899
26	01222223333345556778888888889999
27	00000001122222333333444455556666677777778888889999
28	0001111222222333333444455566667789999 9
29	00011222334557

Deitando a página, podemos ver a distribuição desses dados. Eis a grande vantagem do gráfico ramo-e-folhas: Podemos visualizar a distribuição dos dados e, ainda assim, conservar toda a informação da lista original; se necessário, podemos recompor a relação original de valores.

O leitor notará que as linhas de algarismos em um gráfico ramo-e-folhas são análogas, em natureza, às barras de um histograma. Uma das diretrizes para a construção de histogramas é que o número de classes esteja entre 5 e 20; essa mesma orientação se aplica aos gráficos ramo-e-folhas, pelas mesmas razões. Tais gráficos podem ser *ampliados* de modo a incluírem mais linhas, como podem também ser *condensados*, para reduzir o número de linhas. O gráfico ramo-e-folhas do exemplo precedente pode ser ampliado subdividindo-se as linhas entre aquelas com os algarismos 0 a 4 e as que contêm os algarismos 5 a 9. Mostramos aqui esse ramo-e-folhas ampliado. Quando se torna necessário *reduzir* o número de linhas, podemos *condensar* um gráfico ramo-e-folhas combinando linhas adjacentes, conforme ilustrado a seguir. Note que separamos por um asterisco os algarismos nas folhas associadas a cada ramo. Cada linha no gráfico condensado deve conter precisamente *um* asterisco, de modo que a forma do gráfico não sofra distorção.

Ramo	Folhas
20	0144
20	66889
21	
21	578
22	033
22	58
23	004
23	6
24	122
24	8
25	011224
25	66677899
26	0122222333334
26	5556778888888889999
27	0000000112222223333334444
27	5555666667777777888888999
28	0001111222222333333444
28	555666677899999
29	00011222334
29	557

78-79	07*4	←Esta linha representa 780, 787, 794.
80-81	*55	←Esta linha representa 815, 815.
82-83	9*	←Esta linha representa 829.
84-85	*	←Esta linha não tem dados.
86-87	79*0	←Esta linha representa 867, 869, 870.

Outra vantagem dos gráficos ramo-e-folhas é que sua construção constitui um processo rápido e fácil para *ordenar* os dados. A ordenação dos dados é necessária em vários processos estatísticos, como o cálculo da mediana (abordado na Seção 2-4) e a determinação de percentis ou quartis (Seção 2-6).

A Utilização de Computadores para Gráficos Ramo-e-folhas

O STATDISK não faz gráficos ramo-e-folhas, mas o Minitab os faz. Com o Minitab, introduza os dados na coluna C1 e utilize as opções `Graph, Character Graphs` e `Stem-and-Leaf`. A apresentação Minitab inclui uma coluna adicional de totais acumulados.

Diagramas de Pareto

Consideremos a afirmação: De 75.200 mortes por acidente nos EUA, em um ano recente, 43.500 foram causadas por veículos motorizados, 12.200 por quedas, 6.400 por envenenamento, 4.600 por afogamento, 4.200 por incêndios, 2.900 por ingestão de alimentos ou de um objeto, e 1.400 por armas de fogo (com base em dados do Conselho de Segurança Nacional). O ponto fraco

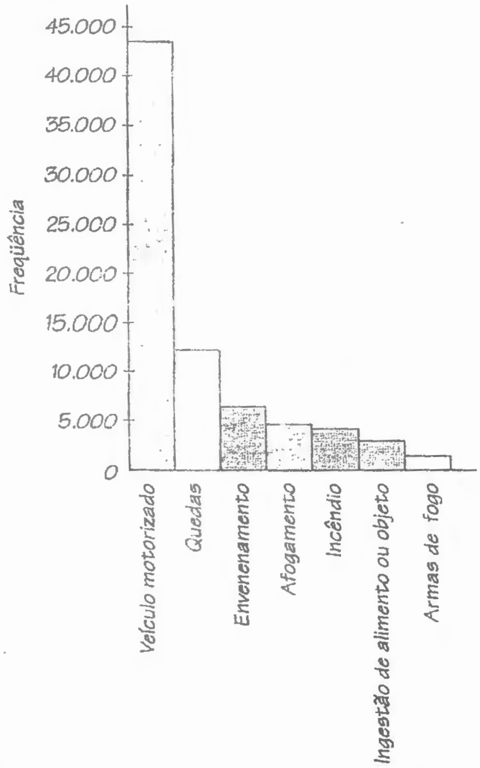


Fig. 2-5 Diagrama de Pareto: causas de mortes acidentais.

dessa afirmação escrita é não caracterizar bem um relacionamento entre categorias diferentes de dados qualitativos. Uma forma mais conveniente de indicar relações entre dados qualitativos é a construção de um diagrama de Pareto. (Recorde, da Seção 1-2, que os dados qualitativos representam uma característica não-numérica, como os tipos de morte acidental relacionados aqui.) Um **diagrama de Pareto** é um gráfico em barras para dados qualitativos, com as barras ordenadas de acordo com a freqüência. Tal como no caso dos histogramas, as escalas verticais em um diagrama de Pareto podem representar freqüências absolutas ou freqüências relativas. A barra mais alta fica à esquerda, e as barras menores na extrema direita, conforme a Figura 2-5. Dispondo as barras por ordem de freqüência, o diagrama de Pareto focaliza a atenção sobre as categorias mais importantes. Pela Figura 2-5, podemos ver que as mortes acidentais causadas por veículos motorizados representam um problema muito mais sério do que as outras categorias. Embora as mortes acidentais causadas por armas de fogo mereçam considerável atenção dos jornais, elas constituem um problema relativamente pequeno quando comparadas com as outras categorias.

Florence Nightingale

Florence Nightingale (1820-1910) é conhecida por muitas como a fundadora da profissão de enfermeira, mas ela também salvou milhares de vidas utilizando a estatística. Ao encontrar um hospital em más condições sanitárias e sem suprimentos, tratou de melhorar essas condições e passou a utilizar a estatística para convencer as autoridades da necessidade de uma reforma médica mais ampla. Elaborou gráficos originais para mostrar que, durante a guerra da Criméia, morreram mais soldados em conseqüência de más condições sanitárias do que em combate. Florence Nightingale foi a pioneira na utilização não só da estatística social como das técnicas de gráficos.

Gráficos em Setores

Tal como os diagramas de Pareto, os gráficos em setores são utilizados para ilustrar dados qualitativos de modo mais compreensível. A Figura 2-6 é um exemplo de **gráfico em setor**, que ilustra graficamente dados qualitativos como fatias de uma torta. A construção de tal gráfico exige a divisão da torta em pedaços com as devidas proporções. Se a categoria de veículos motorizados responde por 57,8% do total de acidentes, então o setor que representa veículos motorizados deve ser 57,8% do total. (O ângulo central deve ser $0{,}578 \times 360° = 208°$.)

O diagrama de Pareto da Figura 2-5 e o gráfico em setores da Figura 2-6 representam os mesmos dados, mas uma comparação

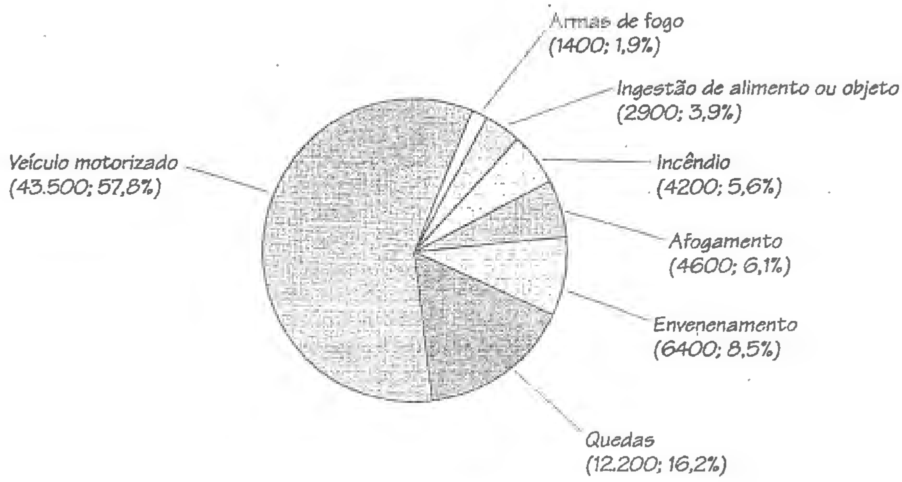
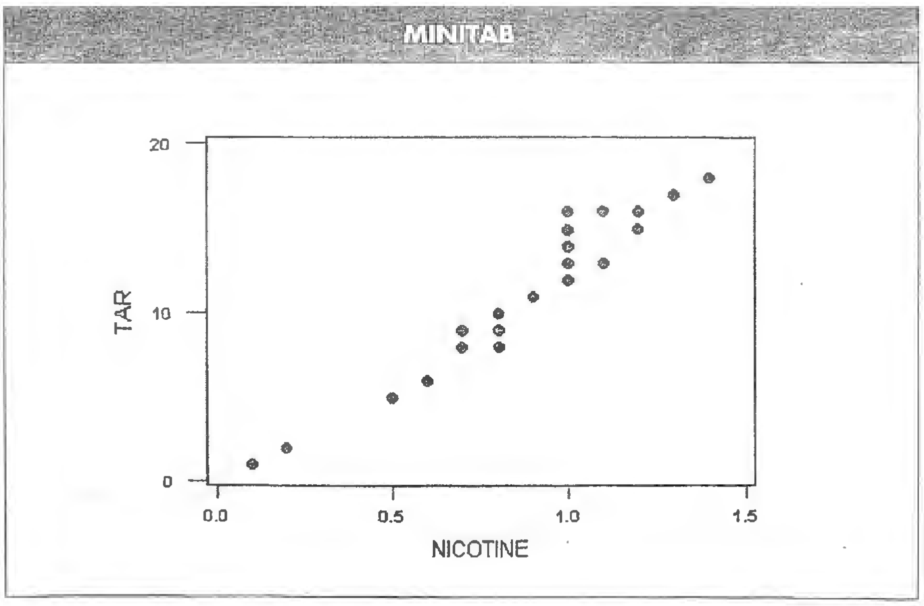


Fig. 2-6 Gráfico em setores: causas de mortes acidentais.

dos dois mostrará provavelmente um melhor desempenho do diagrama de Pareto para evidenciar os tamanhos relativos das diversas componentes.

Diagramas de Dispersão

Às vezes temos dados emparelhados de uma forma que associa cada valor de um conjunto a um determinado valor de um segundo conjunto. Um **diagrama de dispersão** é um gráfico dos dados emparelhados (x, y), com um eixo x horizontal e um eixo y vertical. Para construir manualmente um diagrama de dispersão, traçamos um eixo horizontal para os valores da primeira variável e um eixo vertical para os valores da segunda variável e marcamos os pontos. O padrão dos pontos assim marcados costuma ajudar a determinar se existe algum relacionamento entre as duas variáveis. (Esse tópico será abordado extensamente quando tratarmos da correlação na Seção 9-2.) Utilizando os dados referentes à nicotina e alcatrão presentes em cigarros (Conjunto de Dados 4 do Apêndice B), geramos, com o Minitab, o diagrama de dispersão mostrado na figura. (Para obter esse gráfico, começamos introduzindo ou recuperando os dois conjuntos de dados emparelhados, de forma que eles apareçam nas colunas C1 e C2. Recorremos então às opções `Graph` e `Plot`. O STATDISK e a calculadora TI-83 também são planejados para gerar diagramas de dispersão.) Com base nesse gráfico, parece haver uma relação entre os conteúdos de alcatrão e nicotina nos cigarros, evidenciada pelo padrão dos pontos.



Tinta Invisível

O *National Observer* certa vez contratou uma firma para fazer uma pesquisa confidencial através do correio. O editor Henry Gemmill assegurou em uma circular que "cada resposta individual seria considerada confidencial, mas que, combinada a sua resposta com as outras em todo o país, teríamos um perfil de nossos assinantes". Um assinante sagaz utilizou um raio ultravioleta para detectar um código escrito na pesquisa com tinta invisível. Esse código poderia ser utilizado para identificar o autor da resposta. Gemmill não sabia que esse processo estava sendo usado, e desculpou-se publicamente. O caráter confidencial foi mantido, conforme prometido, mas a anonimidade não havia sido prometida diretamente, de forma que não foi mantida.

Outros Gráficos

Há inúmeros outros recursos pictóricos, além dos que acabamos de indicar, para representar dados de forma interessante e eficiente. O Exercício 27 se refere a um polígono de freqüência, que é uma variante do histograma. Na Seção 2-7 são apresentados diagramas em caixas (*boxplots*), muito úteis para visualizar uma distribüição de dados. Os pictogramas ilustram dados por meio de figuras de objetos ou pessoas, como soldados, tanques, aviões, pilhas de moedas ou sacos de dinheiro. No Capítulo 12, diversos gráficos ilustram padrões de dados ao longo do tempo.

Considere a figura no encarte, tida talvez como "o melhor gráfico estatístico jamais traçado". A figura inclui seis variáveis diferentes relativas à marcha do exército de Napoleão sobre Moscou em 1812. A faixa grossa à esquerda ilustra o tamanho do exército quando começou a invasão da Rússia a partir da Polônia. A faixa inferior descreve a retirada de Napoleão, com as correspondentes temperaturas e datas. Embora elaborado em 1861 por Charles Joseph Minard, esse gráfico é considerado engenhoso mesmo pelos padrões atuais.

Nesta seção focalizamos a natureza ou a forma da distribuição de dados e os métodos de representá-los graficamente. Nas seções seguintes abordaremos outras maneiras de avaliar características de dados.

2-3 Exercícios A: Habilidades e Conceitos Básicos

1. Os visitantes do Parque Nacional de Yellowstone consideram uma erupção do gêiser Old Faithful uma atração que não pode ser perdida. A tabela de freqüências a seguir resume uma amostra de tempos (em minutos) entre erupções. Construa um histograma para a tabela de freqüências dada. Se um guia turístico deseja garantir que seus turistas presenciem uma erupção, qual o tempo mínimo que devem permanecer no parque?

Tempo	Freqüência
40–49	8
50–59	44
60–69	23
70–79	6
80–89	107
90–99	11
100–109	1

2. Obtiveram-se na faculdade do autor os dados ao lado referentes aos carros de estudantes e aos de professores e funcionários. Construa um histograma de freqüências relativas para cada conjunto de dados. Com base nos resultados, quais são as diferenças perceptíveis entre as duas amostras?

Idade	Estudantes	Funcionários e Professores
0–2	23	30
3–5	33	47
6–8	63	36
9–11	68	30
12–14	19	8
15–17	10	0
18–20	1	0
21–23	0	1

3. A tabela de freqüências a seguir dá as velocidades de motoristas multados pela polícia da cidade de Poughkeepsie. Esses motoristas estavam dirigindo em um trecho da zona de 30 mi/h, em Creek Road, que passa pela faculdade do autor. Construa um histograma para essa tabela de freqüências. O que essa distribuição sugere sobre o limite fixado comparado com o limite de velocidade constatado?

Velocidade	Freqüência
42–43	14
44–45	11
46–47	8
48–49	6
50–51	4
52–53	3
54–55	1
56–57	2
58–59	0
60–61	1

4. As companhias de seguro pesquisam continuamente as idades e as causas de morte. Construa um histograma de freqüências relativas correspondente à tabela de freqüências ao lado. Os dados se baseiam em um estudo da revista *Time* sobre vítimas fatais de armas de fogo na América durante uma semana. O que o histograma sugere quanto às idades dessas vítimas fatais?

Idade na Morte	Freqüência
16–25	22
26–35	10
36–45	6
46–55	2
56–65	4
66–75	5
76–85	1

Nos Exercícios 5 e 6, relacione os valores originais nos conjuntos de dados representados pelos dois gráficos ramo-e-folhas.

5.

Ramos	Folhas
57	017
58	13349
59	456678
60	23

6.

Ramos	Folhas
10	21 45 91
11	11 32 [illegible] 83
12	04 22 [illegible]
13	69

Nos Exercícios 7 e 8, construa o gráfico por pontos para os dados representados pelo ramo-e-folhas dos exercícios indicados.

7. Exercício 5
8. Exercício 6

Nos Exercícios 9-12, construa os gráficos ramo-e-folhas para os conjuntos de dados constantes do Apêndice B.

9. Os comprimentos (em polegadas) de ursos do Conjunto de Dados 3. (*Sugestão:* Inicialmente, arredonde os comprimentos para a polegada mais próxima.)
10. As taxas de pulsação das alunas de estatística do Conjunto 8.
11. Pesos (em gramas) das 50 moedas de 25 centavos de dólar relacionados no Conjunto de Dados 13. (Utilize um gráfico ramo-e-folhas ampliado com cerca de 8 linhas.)
12. Pesos (em libras) de artigos de plástico descartados por 62 residências: Recorra aos Dados 1 e arredonde inicialmente os pesos relacionados para o próximo décimo de libra (uma casa decimal). (Use um gráfico ramo-e-folhas ampliado com cerca de 11 linhas.)
13. Foi feito um estudo para determinar como as pessoas obtêm empregos. A tabela que segue relaciona dados de 400 pessoas escolhidas aleatoriamente. Os dados se baseiam em resultados do National Center for Career Strategies (Centro Nacional de Estratégias de Carreiras). Construa um diagrama de Pareto que corresponda aos dados em questão. Qual seria a abordagem mais eficiente para uma pessoa que deseje um emprego?

Fontes de Trabalho dos que Responderam à Pesquisa	Freqüência
Anúncios tipo "Procura-se"	56
Firmas de pesquisas	44
Rádio e televisão	280
Envio de correspondência em massa	20

14. Construa um gráfico em setores para os dados do Exercício 13. Compare o gráfico em setores com o diagrama de Pareto e indique qual deles melhor apresenta a importância relativa das fontes de trabalho.
15. Uma análise de descarrilamentos de trens mostrou que 23 descarrilamentos foram causados por más condições da linha, 9 foram devidos a falhas no equipamento, 12 foram atribuídos a erro humano e 6 tiveram outras causas. [*Fonte*: Dados da Federal Railroad Administration (Departamento Federal de Administração de Ferrovias).] Construa um gráfico em setores para representar os dados em questão.
16. Construa um diagrama de Pareto para os dados do Exercício 15. Compare o diagrama de Pareto com o gráfico em setores, e determine qual dos gráficos mostra com maior eficiência a importância relativa das causas de descarrilamentos de trens.

Nos Exercícios 17-18, use os dados emparelhados do Apêndice B para construir um diagrama de dispersão.

17. No Conjunto de Dados 4, utilize a escala horizontal para o alcatrão e a escala vertical para o monóxido de carbono. Com base no resultado, parece haver uma relação entre o alcatrão e o monóxido de carbono nos cigarros? Em caso afirmativo, descreva esse relacionamento.
18. No Conjunto de Dados 3, use a escala horizontal para os perímetros dos pescoços dos ursos e a escala vertical para os pesos dos animais. Com base no resultado, qual é a relação entre o tamanho do pescoço de um urso e o seu peso?

Nos Exercícios 19-22, recorra aos conjuntos de dados do Apêndice B.

a. Construa um diagrama.
b. Descreva a forma geral da distribuição, como forma de sino, uniforme ou assimétrica.

19. Conjunto de Dados 3 do Apêndice B: pesos de ursos. (Tome 11 classes com amplitude de 50 e comece com −0,5 como limite inferior de classe.)
20. Conjunto de Dados 11 do Apêndice B: pesos de 100 M&Ms. (Utilize 12 classes com amplitude de 0,017 e tome 0,8375 como limite inferior de classe.)
21. Conjunto de Dados 1 do Apêndice B: pesos de papel descartado por 62 residências em uma semana. (Tome 10 classes.)
22. Conjunto de Dados 12 do Apêndice B: os 300 números sorteados na loteria de Maryland (não é a loteria Pick Three).

Nos Exercícios 23-26, recorra à figura do encarte, que descreve a campanha de Napoleão na Rússia em 1812. A faixa grossa à esquerda ilustra o tamanho do exército quando ele começou a invadir a Rússia a partir da Polônia, e a faixa inferior descreve a retirada de Napoleão.

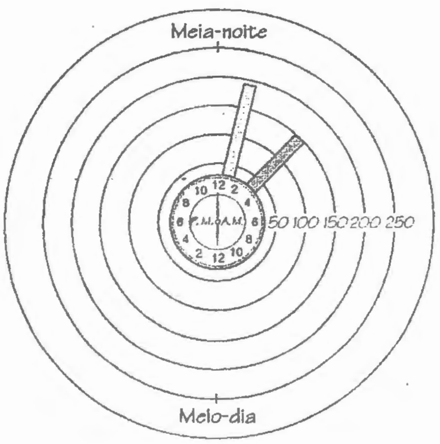
23. Determine a porcentagem dos combatentes que sobreviveram a toda a campanha.
24. Determine o número e a porcentagem dos que morreram cruzando o rio Berezina.
25. Quantos morreram, no retorno de Moscou, no intervalo de tempo em que a temperatura caiu de 16°F para −6°F?
26. Dos que chegaram a Moscou, quantos morreram no percurso de volta entre Moscou e Botr? (Observe que 33.000 homens não foram a Moscou, mas se juntaram aos que voltavam de lá.)

2-3 Exercícios B: Além do Básico

27. Um **polígono de freqüência** é uma variante de um histograma que utiliza segmentos de retas ligando pontos em lugar de barras. Construa um polígono de freqüências modificando o histograma da Figura 2-1 como segue: Inicialmente, substitua as fronteiras de classe na escala horizontal pelos pontos médios das classes. Em seguida, substitua as barras por pontos localizados acima de cada ponto médio a uma altura igual à freqüência da classe. Terceiro, ligue os pontos e prolongue o gráfico à direita e à esquerda, de modo que comece e termine com uma freqüência 0.
28. São fornecidas tabelas de freqüência dos 100 primeiros algarismos na representação decimal do número π e dos 100 primeiros algarismos da representação decimal de 22/7.

a. Construa histogramas que representem as tabelas de freqüências, e assinale quaisquer diferenças.
b. Os números π e 22/7 são ambos reais, mas diferem fundamentalmente um do outro; como?

π		22/7	
x	f	x	f
0	8	0	0
1	8	1	17
2	12	2	17
3	11	3	1
4	10	4	17
5	8	5	16
6	9	6	0
7	8	7	16
8	12	8	16
9	14	9	0



29. Com uma coleção de dados amostrais, construímos uma tabela de freqüências com 10 classes e, em seguida, construímos o histograma correspondente. Indique como o histograma é afetado se se duplica o número de classes mas se mantém a mesma escala vertical.

30. Em um estudo de seguro de acidentes com veículos motorizados no estado de Nova York, classificam-se as colisões fatais de acordo com a hora do dia, com os resultados constantes da tabela a seguir. [*Fonte:* Dados do New York State Department of Motor Vehicles (Departamento de Veículos Motorizados do Estado de Nova York).]

a. Complete o gráfico circular e construa um histograma.
b. Qual dos dois ilustra melhor os dados? Por quê?
c. Como o período de 4 às 6 horas da manhã é o que acusa menor número de colisões fatais, podemos concluir que esse período é o mais seguro para dirigir? Por que sim ou por que não?

Hora	Número de Acidentes Fatais
Manhã 12–2	194
2–4	149
4–6	100
6–8	131
8–10	119
10–12	160
Tarde 12–2	152
2–4	221
4–6	230
6–8	211
8–10	223
10–12	178

31. No artigo "Idades dos Atores e Atrizes Ganhadores do Oscar" (revista *Mathematics Teacher*), de Richard Brown e Gretchen Davis, utilizam-se gráficos ramo-e-folha para comparar as idades de atores e de atrizes no momento da premiação. Eis os resultados para os 34 últimos vencedores recentes de cada categoria.

Atores:	32	37	36	32	51	53	33	61	35	45	55	39
	76	37	42	40	32	60	38	56	48	48	40	
	43	62	43	42	44	41	56	39	46	31	47	
Atrizes:	50	44	35	80	26	28	41	21	61	38	49	33
	74	30	33	41	31	35	41	42	37	26	34	
	34	35	26	61	60	34	24	30	37	31	27	

a. Construa um ramo-e-folhas conjugado para esses dados. Os dois primeiros valores de cada grupo foram registrados a seguir.

Idade dos Atores	Ramo	Idade das Atrizes
	2	
72	3	
	4	4
	5	0
	6	
	7	
	8	

b. Utilizando os resultados da parte a, compare os dois conjuntos distintos de dados e explique quaisquer diferenças.

2-4 Medidas de Tendência Central

O objetivo fundamental desta seção é apresentar as medidas de tendência central importantes.

DEFINIÇÃO

Uma **medida de tendência central** é um valor no centro ou no meio de um conjunto de dados.

Enquanto as Seções 2-2 e 2-3 trataram de tabelas de freqüência e gráficos que revelam a natureza ou a forma da *distribuição* de um conjunto de dados, esta seção focaliza a determinação de valores típicos ou representativos de um conjunto de dados. Há diferentes maneiras de definir o centro e, assim, há diferentes definições de medidas de tendência central, inclusive a média, a mediana, a moda e o ponto médio. Comecemos com a média.

O Paradoxo do Tamanho de uma Turma

Há ao menos duas maneiras de obter o tamanho médio de uma turma, que podem ter resultados muito diferentes. Em uma faculdade, se tomarmos o número de alunos em 737 turmas, obtemos uma média de 40 alunos. Mas se formos compilar uma lista dos tamanhos de turma para cada estudante e utilizar essa lista, obteremos um tamanho médio de turma de 147. Essa grande discrepância é devida ao fato de que há muitos alunos em turmas grandes, mas poucos alunos em turmas pequenas. Sem alterar o número de turmas ou a faculdade, poderíamos reduzir o tamanho médio de turma, formando turmas com aproximadamente o mesmo tamanho. Isso melhoraria também o acompanhamento das aulas, que é melhor em turmas menores.

A Média

A média (aritmética) é, de modo geral, a mais importante de todas as mensurações numéricas descritivas. Na Figura 2-7 ilustramos a propriedade da média como centro do conjunto de dados, no sentido de que é um ponto de equilíbrio dos mesmos.

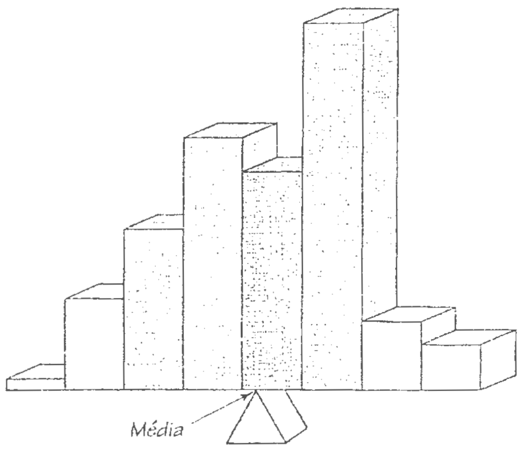


Fig. 2-7 A média como ponto de equilíbrio.
Um fulcro, colocado na posição da média, equilibrará o histograma.

DEFINIÇÃO

A **média aritmética** de um conjunto de valores é o valor obtido somando-se todos eles e dividindo-se o total pelo número de valores. Essa medida particular de tendência central será utilizada freqüentemente em todo o resto deste texto, e será designada simplesmente como **média**.

Esta definição pode expressar-se como na Fórmula 2-1, onde a letra grega Σ (sigma maiúsculo) indica um *somatório* de valores, de forma que Σx representa a soma de todos os valores. O símbolo n denota o **tamanho da amostra**, que é o número de valores em consideração.

Fórmula 2-1 $$\text{média} = \frac{\Sigma x}{n}$$

A média pode denotar-se por $\bar{x}$ (leia-se "x barra") se o conjunto de valores de que dispomos é uma amostra extraída de uma população maior; se todos os valores da população foram considerados, denotamos por μ (minúscula grega mu) a média calculada. (As estatísticas amostrais são em geral representadas por letras do alfabeto latino, como $\bar{x}$, ao passo que os parâmetros populacionais costumam representar-se por letras gregas, como μ.) Muitas calculadoras podem calcular a média de um conjunto de dados: introduzem-se os dados e aciona-se uma tecla $\bar{x}$. A introdução dos dados varia de uma calculadora para outra, de forma que é necessário consultar o respectivo manual.

EXEMPLO Relacionam-se a seguir os tempos (em anos) que os 10 primeiros presidentes americanos sobreviveram à posse. Calcule a média desta amostra.

10 29 26 28 15 23 17 25 0 20

SOLUÇÃO Aplica-se a Fórmula 2-1 para calcular a média. Primeiro somamos os valores.

$\Sigma x = 10 + 29 + 26 + 28 + 15 + 23 + 17 + 25 + 0 + 20 = 193$

Dividimos em seguida o total pelo número de valores. Como há 10 valores, temos $n = 10$ e

$$\bar{x} = \frac{193}{10} = 19{,}3$$

A média é, pois, 19,3 anos.

Para os 10 valores do exemplo precedente, 19,3 está no centro, de acordo com a definição de média. Outras definições de uma medida de tendência central envolvem diferentes percepções de como se determina o centro.

Seis Graus de Separação

Os psicólogos sociais, os historiadores, os cientistas políticos e os especialistas em comunicações estão entre os que se interessam pelo "Problema do Pequeno Mundo": Dadas duas pessoas quaisquer no mundo, quantas ligações intermediárias são necessárias para ligar as duas pessoas originais? O psicólogo social Stanley Milgram fez um experimento utilizando o sistema postal dos EUA. As pessoas foram instruídas a procurar contactar outras pessoas-alvo enviando um formulário a uma pessoa conhecida que julgassem estar próximo do alvo. Das 160 cadeias iniciais, apenas 44 foram completadas. O número de relacionamentos intermediários variou de 2 a 10, com uma mediana de 5. Utilizou-se um modelo matemático para mostrar que, se essas cadeias que faltavam fossem completadas, a mediana seria ligeiramente superior a 5. (Ver "The Small World Problem", de Stanley Milgram, *Psychology Today*, maio de 1967.)

A Mediana

DEFINIÇÃO

A **mediana** de um conjunto de valores é o valor do meio desse conjunto, quando os valores estão dispostos em ordem crescente (ou decrescente). A mediana é representada geralmente por $\tilde{x}$ (lê-se: "x *til*").

Notação

Σ	denota *somatório* de um conjunto de valores.
x	é a *variável* usada para representar valores individuais dos dados.
n	representa o *número de valores em uma amostra.*
N	representa o *número de valores em uma população.*
$\bar{x} = \frac{\Sigma x}{n}$	é a *média de um conjunto de valores amostrais.*
$\mu = \frac{\Sigma x}{N}$	denota a *média de todos os valores de uma população.*

Para calcular a mediana, disponha primeiro os valores em ordem (crescente ou decrescente); em seguida aplique um dos dois processos a seguir:

1. Se o número de valores é ímpar, a mediana é o número localizado exatamente no meio da lista.
2. Se o número de valores é par, a mediana é a média dos dois valores do meio.

EXEMPLO Calcule a mediana dos tempos de sobrevivência (em anos após a posse) dos cinco primeiros presidentes americanos.

10 29 26 28 15

SOLUÇÃO Inicialmente, ordenemos os valores:

10 15 26 28 29

O número de valores é 5, que é ímpar; assim, a mediana é precisamente o número do meio. Logo, a mediana deste conjunto de dados é 26.

EXEMPLO Os valores a seguir são os pagamentos (em dólares) feitos aos executantes de um concerto de rock. A média é \$8900. Calcule a mediana.

500 600 800 50.000 1000 500

SOLUÇÃO Ordenemos inicialmente os valores:

500 500 600 800 1000 50.000

O número de valores é 6, um número par; procuramos, pois, os dois valores do meio e obtemos a sua média. Os dois valores centrais são 600 e 800; a mediana é, pois, a soma desses valores dividida por 2, ou seja, \$700.

Neste conjunto, a média de \$8900 é fortemente afetada pelo valor atípico de \$50.000, o que não ocorre com a mediana de \$700.

Moda

DEFINIÇÃO

A **moda** de um conjunto de dados é o valor que ocorre com maior freqüência. Quando dois valores ocorrem com a mesma freqüência máxima, cada um deles é uma moda, e o conjunto se diz **bimodal**. Se mais de dois valores ocorrem com a mesma freqüência máxima, cada um deles é uma moda, e o conjunto é **multimodal**. Quando nenhum valor é repetido, o conjunto não tem moda. Costuma-se denotar a moda por M.

Um Cidadão Médio

O homem americano "médio" se chama Robert. Tem 31 anos, altura de 1,75 m, pesa 78 kg, seu manequim é 48, calça sapatos tamanho 43 e tem 85 cm de cintura. Consome anualmente 5,6 kg de massa, 11,8 kg de bananas, 1,8 kg de batatas fritas, 8,15 kg de sorvete e 35,8 kg de carne. Em cada ano, vê televisão durante 2567 horas e recebe 585 cartas ou assemelhados pelo correio. Após comer sua porção de batatas fritas, ler a correspondência e ver televisão, ele termina o dia com 7,7 horas de sono. O dia seguinte começa com 21 minutos de transporte para um emprego, onde trabalha 6,1 horas.

EXEMPLO Determine a moda dos seguintes conjuntos de dados.

a. 5 5 5 3 1 5 1 4 3 5
b. 1 2 2 2 3 4 5 6 6 6 7 9
c. 1 2 3 6 7 8 9 10

SOLUÇÃO

a. O número 5 é a moda, porque é o valor que ocorre com maior freqüência.
b. Os números 2 e 6 são ambos modas, porque ocorrem com a mesma freqüência máxima. O conjunto de dados é bimodal.
c. Não há moda, porque não há valor repetido.

Das diferentes medidas de tendência central que estamos considerando, a moda é a única que pode ser usada com dados em nível nominal de mensuração, conforme ilustrado no próximo exemplo.

EXEMPLO Um estudo sobre tempos de reação abrangeu 30 canhotos, 50 destros e 20 ambidestros. Embora não possamos tomar a média numérica dessas características, podemos afirmar que a moda é destra, que é a característica que ocorre com maior freqüência.

Ponto Médio

DEFINIÇÃO

O **ponto médio** é o valor que está a meio caminho entre o maior e o menor valor. Para obtê-lo, somamos esses valores extremos e dividimos o resultado por 2, como na fórmula a seguir:

$$\text{ponto médio} = \frac{\text{maior valor} + \text{menor valor}}{2}$$

EXEMPLO Determine o ponto médio dos tempos de sobrevivência (após a posse) dos 10 primeiros presidentes americanos:

10 29 26 28 15 23 17 25 0 20

SOLUÇÃO Obtém-se como segue o ponto médio:

$$\frac{\text{maior valor} + \text{menor valor}}{2} = \frac{29 + 0}{2} = 14{,}5 \text{ anos}$$

Embora o ponto médio não seja muito usado, incluímo-lo aqui para enfatizar o fato de que há diferentes maneiras de definir o centro de um conjunto de dados. (Veja também Exercícios 20-22.)

Ao nos referirmos ao valor médio de um conjunto de dados, devemos ser precisos, mencionando o termo exato, como média, mediana, moda ou ponto médio.

EXEMPLO Para os 175 valores de cargas axiais de latas de alumínio, relacionados na Tabela 2-1, determine (a) a média, (b) a mediana, (c) a moda e (d) o ponto médio.

SOLUÇÃO

a. Média: A soma dos 175 valores é 46.745; assim,

$$\bar{x} = \frac{46.745}{175} = 267{,}1 \text{ lb}$$

b. Mediana: Dispostos os valores em ordem crescente, verificamos que o 88.º valor, 273, está no meio exato, de modo que a mediana é 273,0. (Os valores podem facilmente ser dispostos em ordem crescente construindo-se um gráfico ramo-e-folhas, conforme vimos na Seção 2-3, ou utilizando-se um programa de computador como STATDISK ou Minitab.) Expressamos o resultado com mais uma casa decimal utilizando a regra do arredondamento que segue este exemplo.

c. Moda: A carga axial mais freqüente é 268 lb, que ocorre 9 vezes. É, pois, a moda.

d. Ponto médio: Obtemo-lo aplicando a fórmula

$$\text{intervalo médio} = \frac{\text{maior valor} + \text{menor valor}}{2} = \frac{297 + 200}{2} = 248{,}5 \text{lb}$$

Passamos a resumir os resultados acima.

média:	267,1 lb
mediana:	273,0 lb
moda:	268 lb
ponto médio:	248,5 lb

Já construímos uma tabela de freqüências e um histograma para os dados da Tabela 2-1, e vimos a distribuição dos dados. Temos agora informações importantes sobre o centro dos dados.

Regra do Arredondamento

Eis uma regra simples para arredondamento de respostas:

Tome uma decimal a mais, além das que aparecem nos dados.

Devemos arredondar apenas a resposta final, e não os valores intermediários. Por exemplo, a média de 2, 3, 5 é 3,333333333..., que pode ser arredondada para 3,3. Como os dados originais são expressos em números inteiros, arredondamos a resposta para o décimo mais próximo. Outro exemplo: a média de 2,1, 3,4 e 5,7 é arredondada para 3,73 com duas decimais (uma a mais em relação às que figuram nos valores originais).

A Média de uma Tabela de Freqüências. A Média Ponderada

Quando os dados estão resumidos em uma tabela de freqüências, podemos aproximar a média substituindo os limites de classe pelos pontos médios das classes e supondo que todos os elementos da classe se concentrem no respectivo ponto médio. Na Tabela 2-2, por exemplo, a primeira classe de 200-209 contém 9 valores que se situam em algum ponto entre os limites de classe, mas não sabemos os valores específicos desses 9 números. A fim de possibilitar os cálculos, supomos que todos os 9 valores se concentrem no ponto médio 204,5. Com 9 valores de 204,5, temos um total de $9 \times 204{,}5 = 1840{,}5$ que contribui para o total geral de todos os valores. O número de valores é igual à soma das freqüências, e assim podemos aplicar a Fórmula 2-2 para achar a média de uma tabela de freqüências. Na realidade, a Fórmula 2-2 não envolve um conceito fundamentalmente diferente: é apenas uma variante da Fórmula 2-1.

Fórmula 2-2 $\bar{x} = \dfrac{\Sigma(f \cdot x)}{\Sigma f}$ **média de uma tabela de freqüências**

onde x = ponto médio da classe
f = freqüência
$\Sigma f = n$

As cargas axiais das latas de alumínio da Tabela de Freqüências 2-2 foram introduzidas na Tabela 2-5, onde aplicamos a Fórmula 2-2. (Podemos também calcular a média de uma tabela de freqüências com uma calculadora TI-83: Introduzimos os pontos médios em L1, introduzimos as freqüências em L2 e utilizamos STAT, CALC, e 1 = Var Stats e introduzimos então L1, L2.) Quando utilizamos a coleção original de dados para calcular a média diretamente, obtivemos o valor 267,1, de modo que o valor da média ponderada baseada na tabela de freqüências é apenas ligeiramente diferente.

Em certas situações, os valores têm graus de importância diferentes, o que nos leva a calcular uma **média ponderada**, que é uma média dos valores afetados de pesos diferentes. Em tais casos, calculamos a média ponderada atribuindo pesos diferentes aos diversos valores, como se vê na Fórmula 2-3.

Fórmula 2-3 média ponderada : $\bar{x} = \dfrac{\Sigma(w \cdot x)}{\Sigma w}$

Suponha, por exemplo, que queiramos a média de 5 notas de teste (85, 90, 75, 80, 95), com os quatro primeiros testes valendo 15% cada um, e o último valendo 40%. Basta atribuirmos o peso 15 a cada uma das quatro primeiras notas, o peso 40 à última nota e calcularmos a média pela Fórmula 2-3, como segue:

$$\bar{x} = \frac{\Sigma(w \cdot x)}{\Sigma w}$$

$$= \frac{(15 \times 85) + (15 \times 90) + (15 \times 75) + (15 \times 80) + (40 \times 95)}{15 + 15 + 15 + 15 + 40}$$

$$= \frac{8750}{100} = 87{,}5$$

TABELA 2-5 Determinação de Σf e $\Sigma(f \cdot x)$

Carga Axial	Freqüência f	Ponto Médio da Classe x	$f \cdot x$
200-209	9	204,5	1.840,5
210-219	3	214,5	643,5
220-229	5	224,5	1.122,5
230-239	4	234,5	938,0
240-249	4	244,5	978,0
250-259	14	254,5	3.563,0
260-269	32	264,5	8.464,0
270-279	52	274,5	14.274,0
280-289	38	284,5	10.811,0
290-299	14	294,5	4.123,0
Total	$\Sigma f = 175$		$\Sigma(f \cdot x) = 46.757{,}5$

$$\bar{x} = \frac{\Sigma(f \cdot x)}{\Sigma f} = \frac{46.757{,}5}{175} = 267{,}2$$

Outro exemplo: As notas de provas podem ser calculadas atribuindo-se a cada conceito (literal) um certo número de pontos (A = 4, B = 3 etc.) e atribuindo-se então a cada número uma freqüência igual ao número de horas de crédito. Um conceito C em um curso de 3 créditos seria equivalente a um ponto médio de classe 2 com freqüência 3. Novamente aqui, podemos aplicar a Fórmula 2-3 para calcular esse tipo de média.

A Melhor Medida de Tendência Central

Vimos que, para os dados da Tabela 2-1, a média, a mediana, a moda e o ponto médio tinham os valores 267,1, 273,0, 268 e 248,5, respectivamente. Qual dessas medidas de tendência central é a melhor? Infelizmente, não há uma resposta única, porque não há critérios objetivos para determinar a medida mais representativa para todos os conjuntos de dados. As diversas medidas de tendência central têm diferentes vantagens e desvantagens, algumas das quais estão resumidas na Tabela 2-6. Uma vantagem importante da média é que leva em conta todos os valores, mas uma grande desvantagem é que às vezes pode ser seriamente afetada por alguns valores extremos. Essa desvantagem pode ser superada com o uso da média aparada, descrita no Exercício 25.

Assimetria

A comparação da média, mediana e moda pode nos dizer algo sobre a característica da assimetria, definida a seguir e ilustrada na Figura 2-8.

DEFINIÇÃO

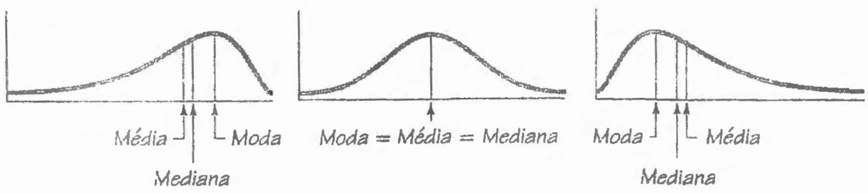
Uma distribuição de dados é **assimétrica** quando não é simétrica, estendendo-se mais para um lado do que para o outro. (Uma distribuição de dados é **simétrica** quando a metade esquerda do seu histograma é aproximadamente a imagem-espelho da metade direita.)

Os dados assimétricos *para a esquerda* dizem-se **negativamente assimétricos**; a média e a mediana estão à esquerda da moda. Embora nem sempre previsíveis, os dados negativamente assimétricos têm em geral a média à esquerda da mediana. (Veja Figura 2-8(a).) Os dados assimétricos *para a direita* dizem-se **positivamente assimétricos**; a média e a mediana estão à direita

TABELA 2-6 Comparação entre Média, Mediana, Moda e Ponto Médio

Medida	Definição	Quão Freqüente?	Existência	Leva em Conta todos os Valores?	Afetada pelos Valores Extremos?	Vantagens e Desvantagens
Média	$\bar{x} = \frac{\Sigma x}{n}$	"média" mais familiar	existe sempre	sim	sim	usada em todo este livro; funciona bem com muitos métodos estatísticos
Mediana	valor do meio	usada comumente	existe sempre	não	não	costuma ser uma boa escolha se há alguns valores extremos
Moda	valor mais freqüente	usada às vezes	pode não existir; pode haver mais de uma moda	não	não	apropriada para dados ao nível nominal
Ponto médio	$\frac{\text{alto} + \text{baixo}}{2}$	raramente usada	existe sempre	não	sim	muito sensível a valores extremos

Comentários gerais:
- Para um conjunto de dados aproximadamente simétrico com uma moda, a média, a mediana, a moda e o ponto médio tendem a coincidir.
- Para um conjunto de dados obviamente assimétrico, convém levar em conta a média e a mediana.
- A média é relativamente *confiável*; ou seja, quando as amostras são extraídas da mesma população, as médias tendem a ser mais constantes do que outras medidas (constantes no sentido de que as médias amostrais extraídas da mesma população não variam tanto quanto as outras medidas).



(a) Assimétrica para a esquerda (negativamente assimétrica): A média e a mediana estão *à esquerda* da moda.

(b) Simétrica (assimetria zero): A média, a mediana e a moda coincidem.

(c) Simétrica para a direita (positivamente assimétrica): A média e a mediana estão *à direita* da moda.

Fig. 2-8 Assimetria.

da moda. Novamente aqui, a maioria dos dados positivamente assimétricos tem a média à direita da mediana. (Veja Figura 2-8(c).)

Se examinarmos o histograma da Figura 2-1 para as cargas axiais de latas de alumínio que estamos considerando neste capítulo, veremos um gráfico que se apresenta assimétrico para a esquerda. Na prática, muitas distribuições de dados são simétricas. As distribuições assimétricas para a direita são mais comuns do que as assimétricas para a esquerda, porque em geral é mais fácil obter valores excepcionalmente grandes do que valores excepcionalmente pequenos. Com as rendas anuais, por exemplo, é impossível termos valores abaixo do limite inferior zero, mas há algumas pessoas que ganham milhões de dólares (ou reais) em um ano. As rendas anuais tendem, pois, a ser assimétricas para a direita, conforme a Figura 2-8(c).

2-4 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-4, determine (a) a média, (b) a mediana, (c) a moda e (d) o ponto médio.

1. Os valores a seguir são os pesos (em onças) de bifes constantes do cardápio de um restaurante como "Bifes Porterhouse de 20 onças" (dados coletados por um aluno do autor). Supõe-se que o peso seja de 21 oz porque os filés perdem cerca de uma onça ao serem cozidos. Os pesos a seguir parecem razoáveis?

17 20 21 18 20 20 20 18 19 19
20 19 21 20 18 20 20 19 18 19

2. Algarismos selecionados na Loteria Pick Three de Maryland:

0 7 3 6 2 7 6 6 6 3 8 1 7 8 7
1 6 8 6 9 5 2 1 5 0 3 9 9 0 7

3. Depósitos de nitrato (em kg por hectare) como parte da chuva ácida no estado de Massachusetts de julho a setembro dos últimos anos (com base em dados do Ministério da Agricultura dos EUA):

6,40 5,21 4,66 5,24 6,96 5,53 8,23 6,80 5,78 6,00 5,41

4. Concentrações sangue-álcool de 15 motoristas envolvidos em acidentes fatais e condenados à prisão (com base em dados do Ministério da Justiça dos EUA):

0,27 0,17 0,17 0,16 0,13 0,24 0,29 0,24
0,14 0,16 0,12 0,16 0,21 0,17 0,18

Nos Exercícios 5-8, determine a média, a mediana, a moda e o ponto médio de cada uma das duas amostras e compare os dois conjuntos de resultados.

5. Tempos de espera de clientes no Banco Jefferson Valley (onde todos os clientes formam uma fila única) e no Banco de Providence (onde os clientes entram em três filas de guichês diferentes):

Jefferson Valley: 6,5 6,6 6,7 6,8 7,1 7,3 7,4 7,7 7,7 7,7
Providence: 4,2 5,4 5,8 6,2 6,7 7,7 7,7 8,5 9,3 10,0

6. Amostras das idades (em anos) de carros de alunos e carros de professores e funcionários da faculdade, obtidas na faculdade do autor:

Alunos: 10 4 5 2 9 7 8 8 16 4 13 12
Profs. e Funcs.: 7 10 4 13 23 2 7 6 6 3 9 4

7. Largura máxima de amostras de crânios de egípcios do sexo masculino, de 4000 aC a 150 aD (com base em dados de *Ancient Races of the Thebaid* por Thomson e Randall-Maciver):

4000 aC: 131 119 138 125 129 126 131 132 126 128 128 131
150 aD: 136 130 126 126 139 141 137 138 133 131 134 129

8. Pesos (em libras) de papel e plástico descartado em residências durante uma semana (dados coletados para o Projeto do Lixo na Universidade do Arizona):

Papel: 9,55 6.38 2,80 6,98 6,33 6,16 10,00 12,29
Plástico: 2,19 2,10 1,41 0,63 0,92 1,40 1,74 2,87

Nos Exercícios 9-12, recorra ao conjunto de dados do Apêndice B e determine (a) a média, (b) a mediana, (c) a moda e (d) o ponto médio.

9. Conjunto de Dados 2 do Apêndice B: Temperaturas do corpo às 8 horas da manhã no dia 1
10. Conjunto de Dados 4 do Apêndice B: Conteúdo de nicotina de todos os cigarros relacionados
11. Conjunto de Dados 3 do Apêndice B: Pesos dos ursos
12. Conjunto de Dados 11 do Apêndice B: Pesos dos bombons M&M vermelhos.

Nos Exercícios 13-16, ache a média dos dados resumidos na tabela de freqüências dada.

13. Os visitantes do Parque Nacional de Yellowstone consideram uma erupção do Old Faithful uma atração que não deve ser perdida. A tabela de freqüências resume uma amostra de tempos (em minutos) decorridos entre as erupções.

Tempo	Freqüência
40–49	8
50–59	44
60–69	23
70–79	6
80–89	107
90–99	11
100–109	1

14. Na faculdade do autor obtiveram-se amostras de carros de estudantes e carros dos professores e funcionários da faculdade, com as respectivas idades (em anos). Essas idades estão resumidas na tabela de freqüência a seguir. Ache a idade média de ambos os grupos de carros. Com base nos resultados, percebe-se alguma diferença significativa entre as duas amostras? Em caso afirmativo, quais são elas?

Idade	Estudantes	Profs. e Funcs.
0–2	23	30
3–5	33	47
6–8	63	36
9–11	68	30
12–14	19	8
15–17	10	0
18–20	1	0
21–23	0	1

15. A tabela de freqüência a seguir dá as velocidades desenvolvidas por motoristas multados pela polícia da cidade de Poughkeepsie. Esses motoristas estavam dirigindo em uma zona de Creek Road com limite de velocidade de 30 mi/h. Compare a velocidade média observada com o limite de 30 mi/h.

Velocidade	Freqüência
42–43	14
44–45	11
46–47	8
48–49	6
50–51	4
52–53	3
54–55	1
56–57	2
58–59	0
60–61	1

16. As companhias de seguro pesquisam continuamente as idades na morte e as respectivas causas. Os dados se baseiam em um estudo da revista *Time* sobre as mortes causadas por armas de fogo na América durante uma semana. Que podemos concluir do resultado?

Idade na morte	Freqüência
16–25	22
26–35	10
36–45	6
46–55	2
56–65	4
66–75	5
76–85	1

2-4 Exercícios B: Além do Básico

17. Um estudante obtém as notas 60, 84 e 90 em testes, e 88 no exame final. Calcule a média ponderada das notas se cada teste corresponde a 20% e o exame final corresponde a 40% da nota final.
18. O boletim de um estudante acusa A em um curso de 4 créditos, A em um curso de 3 créditos, C em um curso de 3 créditos e D em um curso de 2 créditos. Atribuem-se pontos aos conceitos como segue: A = 4, B = 3, C = 2, D = 1, F = 0. Se as notas são ponderadas de acordo com as horas de crédito, determine a média ponderada arredondada para três decimais.
19. a. Calcule a média, a mediana, a moda e o ponto médio das seguintes rendas anuais (em dólares) de médicos autônomos (com base em dados da American Medical Association):

    108.000 236.000 179.000 206.000 236.000

    b. Se se adiciona um valor constante k a cada renda, como são afetados os resultados da parte (a)?
    c. Se os valores das rendas na parte (a) são multiplicados por uma constante k, como são afetados os resultados da parte (a)?
    d. Às vezes os dados são transformados, substituindo-se cada valor x por $\log x$. Para os valores dados de x, determine se a média dos valores de $\log x$ é igual a $\log \bar{x}$.
20. A **média harmônica** costuma ser usada como medida de tendência central para conjuntos de dados que consistem em taxas de variação, como por exemplo velocidades. Obtém-se a média harmônica dividindo-se o número n de valores pela soma dos *inversos* de todos os valores. Expressa-se como:

$$\frac{n}{\sum \frac{1}{x}}$$

(Nenhum valor pode ser zero.) Por exemplo, a média harmônica de 2, 4, 10 é

$$\frac{n}{\sum \frac{1}{x}} = \frac{3}{\frac{1}{2}+\frac{1}{4}+\frac{1}{10}} = \frac{3}{0,85} = 3,5$$

a. Quatro estudantes dirigem de Nova York à Flórida (1200 milhas) a uma velocidade de 40 mi/h (sim, é verdade!) e voltam à velocidade de 60 mi/h. Qual é sua velocidade média para a viagem de ida e volta? (Usa-se a média harmônica para calcular médias de velocidades.)

b. Um despachante da Kramden Bus Company calcula a velocidade média, em mi/h, do percurso de ida e volta de Boston a Providence. Dão-se a seguir os resultados obtidos em 14 viagens diferentes. Com base nesses dados, qual é a velocidade média de um ônibus nesse percurso?

42,6	41,3	38,2	42,9	43,4	43,7	40,8
34,2	40,1	41,2	40,5	41,7	39,8	39,6

21. A **média geométrica** é usada em administração e economia para achar taxas médias de variação, de crescimento, ou razões médias. Dados n valores (todos positivos), a média geométrica é a raiz n^{ma} do seu produto. Por exemplo, determina-se a média geométrica de 2, 4, 10 multiplicando-se os três valores — o que dá 80, e tomando-se a raiz cúbica do resultado (porque há três valores). O resultado é 4,3. O *fator de crescimento médio* para o dinheiro, composto às taxas anuais de juro de 10%, 8%, 9%, 12% e 7% pode ser determinado calculando-se a média geométrica de 1,10, 1,08, 1,09, 1,12 e 1,07. Calcule esse fator médio de crescimento.
22. A **média quadrática** é utilizada em geral em experimentos físicos. Em sistemas de distribuição de energia, por exemplo, as tensões e correntes são em geral dadas em termos de sua média quadrática. Obtém-se a média quadrática de um conjunto de valores elevando-se cada um ao quadrado, somando-se os resultados, dividindo-se o total pelo número n de valores e tomando-se a raiz quadrada do resultado. Por exemplo, a média quadrática de 2, 4, 10 é

$$\sqrt{\frac{\Sigma x^2}{n}} = \sqrt{\frac{4+16+100}{3}} = \sqrt{\frac{120}{3}} = \sqrt{40} = 6.3$$

Calcule a média quadrática dos seguintes valores de fornecimento de energia (em volts): 151, 162, 0, 81, −68.
23. As tabelas de freqüência costumam apresentar classes com intervalo aberto, como a tabela a seguir, que resume os tempos gastos em estudo por calouros (com base em dados de *The American Freshman*, em *USA Today*). Não se pode aplicar diretamente a Fórmula 2-2, porque o ponto médio da classe "mais de 20" não está definido. Calcule a média supondo que esta última classe seja realmente (a) 21-25, (b) 21-30, (c) 21-40. O que se pode concluir?

Horas de estudo por semana	Freqüência
0	5
1–5	96
6–10	57
11–15	25
16–20	11
Mais de 20	6

24. Quando os dados são resumidos em uma tabela de freqüências, pode-se achar a mediana identificando primeiro a *classe mediana* (a classe que contém a mediana). Supomos então que os valores se

distribuam uniformemente nessa classe, e interpolamos. Esse processo é descrito por

$$(\text{limite inferior da classe mediana}) + (\text{amplitude da classe})\left(\frac{\left(\frac{n+1}{2}\right) - (m+1)}{\text{freqüência da classe mediana}}\right)$$

onde n é a soma de todas as freqüências de classe e m é a soma das freqüências das classes que *precedem* a classe mediana. Utilize este processo e os dados da Tabela de Freqüências 2-2 para achar a carga axial mediana.

25. Como a média é muito sensível a valores extremos, é acusada de não ser uma medida *robusta* de tendência central. A **média aparada** é mais robusta. Para achar a média aparada em 10% de um conjunto de dados, primeiro ordenamos os dados, em seguida eliminamos 10% dos valores superiores e 10% dos valores inferiores, e calculamos a média dos valores restantes. Para os pesos de ursos do Conjunto de Dados 3 do Apêndice B, determine (a) a média; (b) a média aparada em 10%; (c) a média aparada em 20%. Compare os resultados.
26. Consultando um almanaque, um pesquisador determina o salário médio dos professores para cada estado americano. Soma esses 50 valores e divide o total por 50, para obter a média. O resultado é igual ao salário médio nacional dos professores? Por quê?

2-5 Medidas de Variação

Esta seção aborda a característica da variação, de grande importância para a estatística, sendo, por isso, uma das principais de todo o livro. O leitor deve dominar os seguintes conceitos-chave: (1) a variação se refere a quanto os valores podem diferir entre si e pode ser medida por números específicos; (2) os números relativamente próximos uns dos outros têm baixas medidas de variação, enquanto os valores mais dispersos têm maior medida de variação; (3) o desvio-padrão é uma medida de variação particularmente importante, e devemos saber calculá-lo para um conjunto de valores; (4) os valores dos desvios-padrão devem ser *interpretados* corretamente.

Muitos bancos costumavam exigir que os clientes formassem filas separadas para os diversos guichês, mas recentemente passaram a adotar fila única. Qual o motivo dessa modificação? O tempo médio de espera não se modifica, porque a fila de espera não afeta a eficiência dos caixas. A adoção de fila única se deveu ao fato de os clientes preferirem tempos de espera mais *consistentes* com menor variação. Assim é que milhares de bancos efetuaram uma modificação que resultou em uma variação menor (e clientes mais satisfeitos), mesmo que a média não tenha sido afetada. Consideremos agora a mesma amostra de dados bancários usada no Exercício 5 da seção precedente. Os valores relacionados são tempos de espera (em minutos) de clientes.

Um Bom Conselho aos Jornalistas

O colunista Max Frankel escreveu no *The New York Times*: "As escolas de jornalismo não dão a devida importância à estatística, e algumas permitem que seus estudantes se formem sem qualquer treinamento com números. Como podem tais repórteres escrever conscientemente sobre comércio, bem-estar social, crime, ou tarifas aéreas, saúde e nutrição? O uso descuidado pela mídia de números sobre a incidência de acidentes ou de doenças assusta o povo, deixando-o vulnerável aos truques jornalísticos, à demagogia política, e à fraude comercial." O colunista cita diversos casos, inclusive o exemplo de um artigo de página inteira sobre o déficit da cidade de Nova York, com uma promessa do prefeito daquela cidade de cobrir um déficit orçamentário de $2,7 bilhões; mas em todo o artigo não se menciona uma vez sequer o *total* do orçamento, de modo que a cifra de $2,7 bilhões por si só pouco significa.

Banco Jefferson Valley (Fila única)	6,5 6,6 6,7 6,8 7,1 7,3 7,4 7,7 7,7 7,7
Banco da Providência (Fila múltipla)	4,2 5,4 5,8 6,2 6,7 7,7 7,7 8,5 9,3 10,0

Os clientes do Jefferson Valley Bank entram em uma fila única que é atendida por três caixas. Os clientes do Bank of Providence podem entrar em qualquer uma de três filas que conduzem a três guichês. Se fizermos o Exercício 5 da Seção 2-4, veremos que ambos os bancos têm a mesma média de 7,15, a mesma mediana de 7,20, a mesma moda de 7,7 e o mesmo ponto médio de 7,10. Com base apenas nestas medidas de tendência central, poderíamos admitir que os tempos de espera nos dois bancos fossem praticamente os mesmos. Todavia, esquadrinhando os tempos de espera originais, constataríamos uma diferença fundamental: O Jefferson Valley Bank tem tempos de espera com muito menos *variação* do que o Bank of Providence. Mantidas todas as outras características, os clientes provavelmente preferirão o Jefferson Valley Bank, onde não correm o risco de entrar em uma fila muito mais lenta do que as outras.

Fazendo uma comparação subjetiva dos tempos de espera nos dois bancos, podemos ver a característica da variação. Passemos agora a algumas formas específicas de *medir* efetivamente a variação. Começaremos com a amplitude.

Amplitude

A **amplitude** de um conjunto de dados é a diferença entre o maior valor e o menor valor. Para calculá-lo, basta subtrairmos o menor valor do maior. Para o caso do Jefferson Valley Bank, a amplitude é de 7,7 − 6,5 = 1,2 min. Os tempos de espera no Bank of Providence têm uma amplitude de 5,8 min, o que sugere maior variação.

O cálculo da amplitude é bastante fácil, mas como ele depende apenas do menor e do maior valor, em geral não é tão bom quanto outras medidas de variação que levam em conta todos os valores. (Veja no Exercício 25 um exemplo em que a amplitude é enganosa.)

Desvio-Padrão e Variância

De modo geral, o desvio-padrão é a mais importante e mais útil medida de variação. Ao contrário da amplitude, o desvio-padrão leva em conta todos os valores, mas essa vantagem torna o cálculo mais difícil. Definimos a seguir o desvio-padrão, mas para entender perfeitamente esse conceito, é preciso lermos cuidadosamente o restante desta seção.

DEFINIÇÃO

O **desvio-padrão** de um conjunto de valores amostrais é uma medida da variação dos valores em relação à média. Calcula-se com o auxílio da Fórmula 2-4.

Fórmula 2-4 $s = \sqrt{\dfrac{\Sigma(x - \bar{x})^2}{n - 1}}$ desvio padrão amostral

Quase todas as calculadoras científicas e pacotes estatísticos são programados para calcular automaticamente o desvio-padrão. Na Seção 2-6 discutimos a utilização de calculadoras ecomputadores, mas é interessante o leitor consultar logo o manual de sua calculadora para ver o processo de cálculo que dá o desvio-padrão.

Por que definir uma medida de variação da maneira indicada na Fórmula 2-4? Ao medir a variação em um conjunto de dados amostrais, é razoável começarmos com os desvios dos valores em relação à média. Para determinado valor x, o valor do **desvio** é $x - \bar{x}$, que é a diferença entre o valor e a média. Mas a soma de todos esses desvios é sempre zero, o que na verdade nada significa para nós. Para termos uma estatística que realmente meça a variação (em lugar de ser sempre zero), poderíamos tomar a soma de valores absolutos, como em $\Sigma|x - \bar{x}|$. Determinando a média deste somatório, obtemos o **desvio médio** (ou **desvio absoluto**), dado pela seguinte expressão:

$$\text{Desvio médio} = \frac{\Sigma|x - \bar{x}|}{n}$$

Em vez de utilizar valores absolutos, podemos obter uma medida de variação ainda melhor, tomando os quadrados dos desvios $(x - \bar{x})$, que são não-negativos. Resulta que o desvio-padrão tem a mesma unidade de medida que os valores originais. Por exemplo, se os tempos de espera dos clientes são medidos em minutos, o desvio-padrão será expresso também em minutos. Com base na Fórmula 2-4, podemos estabelecer como se segue o processo de cálculo do desvio-padrão.

Processo para Determinar o Desvio-Padrão com a Fórmula 2-4

Passo 1: Achar a média $\bar{x}$ dos valores.

Passo 2: Subtrair a média de cada valor individual $(x - \bar{x})$.

Passo 3: Elevar ao quadrado cada uma das diferenças obtidas no Passo 2. [Este processo produz números da forma $(x - \bar{x})^2$].

Passo 4: Somar todos os quadrados obtidos no Passo 3, obtendo $\Sigma(x - \bar{x})^2$

Passo 5: Dividir o total do Passo 4 pelo número $(n - 1)$; isto é, 1 menos que o número total de observações.

Passo 6: Extrair a raiz quadrada do resultado do Passo 5.

Mais Ações, Menor Risco

Em seu livro *Investments*, os autores Zvi Bodie, Alex Kane e Alan Marcus afirmam que "o desvio-padrão médio dos ganhos proporcionados por uma carteira composta apenas de ações de **uma** única companhia é de 0,554. O risco médio de uma carteira diminui rapidamente na medida em que se diversificam as ações da carteira". Os autores observam que, com ações de 32 companhias, o desvio-padrão é de 0,325, indicando muito menor variação e risco. Salientam que com apenas uns poucos tipos de ações uma carteira tem um elevado grau de "risco específico", o que significa que o risco é atribuído ao pequeno número de ações em jogo. Com mais de 30 tipos de ação, há um risco específico muito pequeno; quase todo o risco é um "risco de mercado", atribuído ao mercado de ações como um todo. Os autores observam que esses princípios nada mais são do que a aplicação da bem conhecida lei das médias.

EXEMPLO Determine o desvio-padrão dos tempos de espera em guichês dos clientes do Jefferson Valley Bank. Esses tempos de espera (em minutos) são dados a seguir:

6,5 6,6 6,7 6,8 7,1 7,3 7,4 7,7 7,7 7,7

SOLUÇÃO Muitos estudantes acham fácil utilizar a função desvio-padrão embutida em suas calculadoras, mas recomendamos que o processo seja realmente entendido, seguindo os passos detalhados para o cálculo. (Ver Tabela 2-7, onde se executam os seguintes passos.)

Passo 1: Obtenha a média de 7,15, somando os valores e dividindo o total pelo número de valores:

$$\bar{x} = \frac{\Sigma x}{n} = \frac{71{,}5}{10} = 7{,}15 \text{ min}$$

Passo 2: Subtraia de cada valor a média 7,15, obtendo os seguintes valores de $(x - \bar{x})$: $-0{,}65$, $-0{,}55$, ..., $0{,}55$.

Passo 3: Eleve ao quadrado cada valor do Passo 2, obtendo os valores $(x - \bar{x})^2$: 0,4225; 0,3025; ...; 0,3025.

Passo 4: Some todos os valores precedentes, obtendo

$$\Sigma(x - \bar{x})^2 = 2{,}0450$$

Passo 5: Há $n = 10$ valores; divida, pois, por $9(= 10 - 1)$:

$$2{,}0450 \div 9 = 0{,}2272$$

Passo 6: Determine a raiz quadrada de 0,2272. O desvio-padrão é

$$\sqrt{0{,}2272} = 0{,}48 \text{ min}$$

Teoricamente, deveríamos dar aqui uma interpretação do desvio-padrão de 0,48 min, mas essa interpretação será dada mais adiante. Por ora, o leitor deve exercitar-se no cálculo de um desvio-padrão utilizando os tempos de espera no Bank of Providence. Com esses dados, verificará que o desvio-padrão é de 1,82 min. Embora a interpretação desses desvios-padrão seja dada mais adiante, podemos compará-los; verificaremos que o desvio-padrão dos tempos de espera no Jefferson Valley Bank (0,48 min) é muito menor do que o do caso do Bank of

TABELA 2-7 Cálculo do Desvio-Padrão para os Clientes do Banco Jefferson Valley

x	$x - \bar{x}$	$(x - \bar{x})^2$
6,5	−0,65	0,4225
6,6	−0,55	0,3025
6,7	−0,45	0,2025
6,8	−0,35	0,1225
7,1	−0,05	0,0025
7,3	0,15	0,0225
7,4	0,25	0,0625
7,7	0,55	0,3025
7,7	0,55	0,3025
7,7	0,55	0,3025
Totais: 71,5		2,0450

$$\bar{x} = \frac{71{,}5}{10} = 7{,}15 \text{ min} \qquad s = \sqrt{\frac{2{,}0450}{10-1}} = \sqrt{0{,}2272} = 0{,}48 \text{ min}$$

Providence (1,82 min). Isso reforça a nossa conclusão subjetiva, de que os tempos de espera no Jefferson Valley Bank têm variação muito menor do que os do Bank of Providence.

Em nossa definição, referimo-nos ao desvio-padrão de dados *amostrais*. Para o cálculo do desvio-padrão σ (minúscula grega sigma) de uma *população*, vale uma fórmula ligeiramente diferente: em lugar de dividirmos por $n - 1$, dividimos por N, tamanho da população, como se vê na expressão seguinte.

$$\sigma = \sqrt{\frac{\Sigma(x - \mu)^2}{N}} \quad \text{desvio-padrão populacional}$$

Por exemplo, se os 10 valores da Tabela 2-7 constituem uma população, o desvio-padrão é:

$$\sigma = \sqrt{\frac{\Sigma(x - \mu)^2}{N}} = \sqrt{\frac{2{,}0450}{10}} = 0{,}45 \text{ min}$$

Como em geral lidamos com dados amostrais, vamos utilizar a Fórmula 2-4, dividindo por $n - 1$. Muitas calculadoras dão o desvio-padrão, com a divisão por $n - 1$ correspondendo a uma tecla σ_{n-1} ou s, enquanto que a tecla σ_n ou σ corresponde a uma divisão por N. Por alguma razão, engenhosa mas estranha, as calculadoras utilizam diversas notações; as que seguem, entretanto, são as mais comuns em estatística. Essas notações compreendem referências à variância de um conjunto de valores; passamos agora a descrever essa medida de variação.

Notação

s denota o desvio-padrão de um conjunto de dados *amostrais*

σ denota o desvio-padrão de um conjunto de dados *populacionais*

s^2 é a variância de um conjunto de dados *amostrais*

σ^2 é a variância de um conjunto de dados *populacionais*

Nota: Em artigos de revistas e relatórios profissionais, costuma-se indicar o desvio-padrão por SD (*standard deviation*) e a variância por Var.

Omitindo a Etapa 6 (tomar a raiz quadrada) no processo de cálculo do desvio-padrão, obtemos a **variância**, definida na Fórmula 2-5.

Fórmula 2-5 $$s^2 = \frac{\Sigma(x - \bar{x})^2}{n - 1} \quad \text{variância amostral}$$

Analogamente, podemos expressar a variância populacional como

$$\sigma^2 = \frac{\Sigma(x - \mu)^2}{N} \quad \text{variância populacional}$$

Comparando as Fórmulas 2-4 e 2-5, vemos que a variância é o quadrado do desvio-padrão. Embora a variância venha a ser usada mais adiante, devemos inicialmente concentrar-nos no conceito de desvio-padrão, para bem apreender o significado dessa estatística. Uma dificuldade com a variância é que ela não é expressa nas mesmas unidades dos dados originais. Assim é que um conjunto de dados pode ter um desvio-padrão de \$3,00 e uma variância de 9,00 dólares quadrados. Como dólar quadrado é um conceito abstrato que não atingimos diretamente, a variância se nos afigura difícil de ser compreendida.

Regra do Arredondamento

Tal como na Seção 2-4, utilizamos a regra seguinte para arredondar resultados finais:

Tomar uma casa decimal a mais, em relação às que constam dos dados originais.

Devemos arredondar apenas o resultado final, e não resultados intermediários. Se, por alguma razão, tivermos de arredondar resultados intermediários, devemos trabalhar com pelo menos duas casas decimais além das que devem constar do resultado final.

Fórmula Abreviada e Dados Agrupados

Damos a seguir duas outras fórmulas para o desvio-padrão. Essas fórmulas não envolvem qualquer conceito diferente; são apenas versões distintas da Fórmula 2-4. Primeiro, a Fórmula 2-4 pode expressar-se na forma equivalente:

Fórmula 2-6 $$s = \sqrt{\frac{n(\Sigma x^2) - (\Sigma x)^2}{n(n - 1)}} \quad \text{Fórmula abreviada para o desvio padrão}$$

As Fórmulas 2-4 e 2-6 são equivalentes no sentido de que sempre dão os mesmos resultados. Poupamos ao leitor o trabalho algébrico para mostrar essa igualdade. A Fórmula 2-6 é chamada *fórmula abreviada*, porque tende a ser mais conveniente para uso com números extensos ou com grandes conjuntos de valores. A Fórmula 2-6 é usada em geral em calculadoras e programas de computador, porque exige apenas três registros de memória (para n, Σx e Σx^2), em lugar de um registro de memória separado para cada valor individual. A Fórmula 2-6 também elimina erros de arredondamentos intermediários, originados quando não se utiliza o valor exato da média. Não obstante, muitos professores preferem utilizar apenas a Fórmula 2-4 para o cálculo do desvio-padrão. Argumentam que a Fórmula 2-4 reforça o conceito de que *o desvio-padrão é um tipo de desvio médio*, enquanto a Fórmula 2-6 obscurece essa idéia. Outros professores não fazem qualquer objeção à Fórmula 2-6. Incluímos a fórmula abreviada para aqueles que desejem utilizá-la. Já apresentamos um exemplo de cálculo do desvio-padrão com a Fórmula 2-4; ilustraremos a seguir a aplicação da Fórmula 2-6.

EXEMPLO Calcule o desvio-padrão dos seguintes tempos de espera (em minutos) de clientes do Jefferson Valley Bank, aplicando a Fórmula 2-6:

6,5 6,6 6,7 6,8 7,1 7,3 7,4 7,7 7,7 7,7

SOLUÇÃO A Fórmula 2-6 exige a determinação dos valores de n, Σx e Σx^2. Como há 10 valores, temos $n = 10$. A soma dos 10 valores é 71,5 e, assim, $\Sigma x = 71{,}5$. Calcula-se como se segue a terceira componente necessária:

$$\begin{aligned}\Sigma x^2 &= 6{,}5^2 + 6{,}6^2 + 6{,}7^2 + \cdots + 7{,}7^2 \\ &= 42{,}25 + 43{,}56 + 44{,}89 + \cdots + 59{,}29 \\ &= 513{,}27\end{aligned}$$

Estamos em condições de aplicar a Fórmula 2-6 para calcular o valor do desvio-padrão.

$$s = \sqrt{\frac{n(\Sigma x^2) - (\Sigma x)^2}{n(n - 1)}} = \sqrt{\frac{10(513{,}27) - (71{,}5)^2}{10(10 - 1)}}$$

$$= \sqrt{\frac{20{,}45}{90}} = 0{,}4766783 = 0{,}48 \text{ min (arredondado)}$$

Pode-se estabelecer uma fórmula para o desvio-padrão quando os dados se apresentam resumidos em uma tabela de freqüências. O resultado é:

$$s = \sqrt{\frac{\Sigma f \cdot (x - \bar{x})^2}{n - 1}}$$

Daremos a esta fórmula uma expressão equivalente, que em geral simplifica os cálculos.

Fórmula 2-7

$$s = \sqrt{\frac{n[\Sigma (f \cdot x^2)] - [\Sigma (f \cdot x)]^2}{n(n-1)}} \quad \text{desvio-padrão para tabela de freqüência}$$

com: x = ponto médio da classe
f = freqüência da classe
n = tamanho da amostra (ou Σf = soma das freqüências)

EXEMPLO Aplique a Fórmula 2-7 para estimar o desvio-padrão das 175 cargas axiais das latas de alumínio da Tabela de Freqüências 2-2.

SOLUÇÃO A aplicação da Fórmula 2-7 exige a determinação dos valores de n, $\Sigma(f \cdot x)$ e $\Sigma(f \cdot x^2)$. Determinados esses valores, pela Tabela 2-8, podemos aplicar a Fórmula 2-7, como segue:

$$s = \sqrt{\frac{n[\Sigma (f \cdot x^2)] - [\Sigma (f \cdot x)]^2}{n(n-1)}} =$$

$$= \sqrt{\frac{175(12.579.173{,}75) - (46.757{,}5)^2}{175(175-1)}}$$

$$= \sqrt{\frac{15.091{,}600}{30.450}} = \sqrt{495{,}6190476} = 22{,}3 \text{ lb}$$

As 175 cargas axiais têm um desvio-padrão estimado em 22,3 lb. (O valor exato calculado com base no conjunto original de dados é 22,1 lb; a aproximação é, pois, bastante satisfatória.)

Podemos também utilizar uma calculadora TI-83 para calcular o desvio-padrão de dados condensados em uma tabela de freqüências. Introduzimos primeiro os pontos médios em L1, em seguida as freqüências em L2; utilizamos então `STAT`, `CALC` e `1-VarStats` e introduzimos L1 e L2 para obter os resultados que incluem a média e o desvio-padrão.

Para Entender o Desvio-padrão

Procuraremos aqui atribuir um sentido intuitivo ao desvio-padrão. De início, devemos ter em mente que o desvio-padrão mede a variação entre valores. Valores próximos uns dos outros originam desvios-padrão menores, enquanto valores muito afastados uns dos outros dão um desvio-padrão maior. Interrompamos a leitura e devotemos um momento ao estudo da Figura 2-9. Veremos que, quando os dados se dispersam, o valor do desvio-padrão aumenta.

Como a variação é um conceito relevante, e como o desvio-padrão tem grande importância na sua medida, abordaremos três maneiras diferentes de atribuir um sentido ao desvio-padrão. A primeira é uma regra prática que utiliza a amplitude para obter uma estimativa bastante rudimentar do desvio padrão. (Poderíamos melhorar a precisão dessa regra levando em conta fatores como o tamanho da amostra e a natureza da distribuição, mas, por ora, preferimos sacrificar a precisão em favor da simplicidade. Queremos uma regra simples que nos permita interpretar o valor do desvio-padrão; mais adiante estudaremos métodos que produzam resultados mais precisos.)

Regra Prática (desvio-padrão em termos da amplitude)

Para conjuntos de dados típicos, a amplitude mede aproximadamente 4 desvios-padrão ($4s$), de forma que podemos aproximar como segue o desvio-padrão:

$$\text{desvio-padrão} \approx \frac{\text{amplitude}}{4} \quad \text{regra prática}$$

Esta expressão dá uma estimativa razoável para o desvio-padrão, quando conhecemos os valores mínimo e máximo. Desde que conheçamos o desvio-padrão, podemos utilizá-lo para entender melhor os dados, fazendo estimativas dos valores mínimo e máximo como se segue:

mínimo $\approx$ **(média)** $-$ **2** $\times$ **(desvio-padrão)**
máximo $\approx$ **(média)** $+$ **2** $\times$ **(desvio-padrão)**

TABELA 2-8 Cálculo do Desvio-Padrão para uma Tabela de Freqüências

Carga Axial	Freqüência f	Ponto Médio da Classe x	$f \cdot x$	$f \cdot x^2$
200-209	9	204,5	1.840,5	376.382,25
210-219	3	214,5	643,5	138.030,75
220-229	5	224,5	1.122,5	252.001,25
230-239	4	234,5	938,0	219.961,00
240-249	4	244,5	978,0	239.121,00
250-259	14	254,5	3.563,0	906.783,50
260-269	32	264,5	8.464,0	2.238.728,00
270-279	52	274,5	14.274,0	3.918.213,00
280-289	38	284,5	10.811,0	3.075.729,50
290-299	14	294,5	4.123,0	1.214.223,50
Total	$\Sigma f = 175$		$\Sigma(f \cdot x) = 46.757{,}5$	$\Sigma(f \cdot x^2) = 12.579.173{,}75$

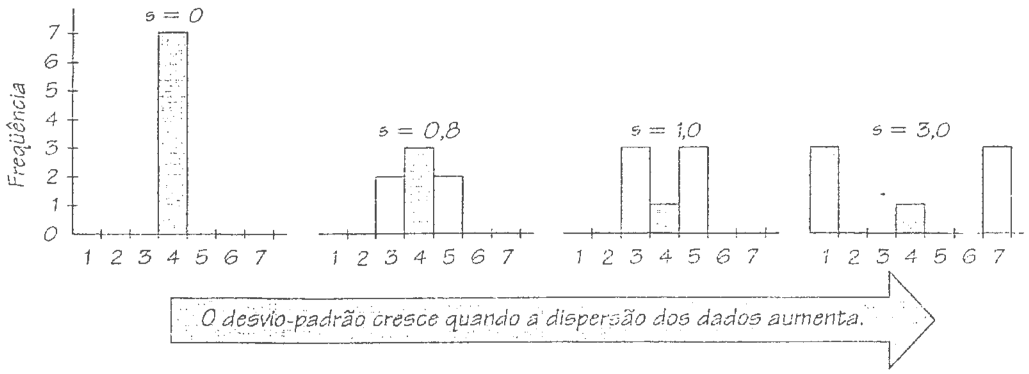


Fig. 2-9 Média idêntica, desvios-padrão diferentes.

Ao calcularmos um desvio-padrão com uma das Fórmulas 2-4 ou 2-6, podemos utilizar a regra prática como uma verificação do resultado obtido, mas não devemos esquecer que, embora a aproximação leve a uma vizinhança da resposta, ainda assim pode acusar grande diferença. Para os tempos de espera dos clientes do Jefferson Valley Bank (6,5; 6,6; 6,7; 6,8; 7,1; 7,3; 7,4; 7,7; 7,7; 7,7) calculamos o desvio-padrão pela Fórmula 2-6, obtendo $s = 0,48$. A amplitude desses valores é $7,7 - 6,5 = 1,2$, o que nos permite aplicar a regra prática para obter uma estimativa de s como segue:

$$s \approx \frac{\text{amplitude}}{4} = \frac{1,2}{4} = 0,3 \text{ min}$$

Ora, acabamos de ver que o desvio-padrão é realmente 0,48, de modo que a estimativa obtida pela regra prática (0,3) parece demasiadamente pequena. Todavia, nossa estimativa confirma que, de modo geral, estamos bem próximos do valor correto; sem dúvida, um valor como 7 para s se afiguraria incorreto.

Consistência no Correio

Pesquisa recente feita com 29.000 pessoas que utilizam o serviço postal dos EUA revelou que elas gostariam de maior *consistência* no tempo que uma carta leva para ser entregue. Ora, uma carta local pode levar um dia ou vários dias para ser entregue. O jornal *USA Today* registrou uma queixa comum: "Por favor, diga-me com quantos dias de antecedência eu devo postar um cartão de aniversário para minha mãe."

O nível de consistência pode ser medido pelo desvio-padrão dos tempos de entrega. Um desvio-padrão mais baixo revela maior consistência. O desvio-padrão é em geral um recurso criticamente importante para controlar a qualidade de bens e serviços.

EXEMPLO Com auxílio da regra prática, estime o desvio-padrão da amostra de 175 cargas axiais de latas de alumínio da Tabela 2-1.

SOLUÇÃO Utilizando a regra prática para estimar o desvio-padrão de dados amostrais, calculamos a amplitude e a dividimos por 4. Percorrendo a lista de valores, vemos que o menor é 200 e o maior é 297, de forma que a amplitude é $297 - 200 = 97$. O desvio-padrão s é estimado como segue:

$$s \approx \frac{\text{intervalo}}{4} = \frac{97}{4} = 24,3 \text{ lb}$$

Esse resultado está próximo do valor correto de 22,1, obtido com o cálculo do valor exato do desvio-padrão pela Fórmula 2-4 ou 2-6.

Como as cargas axiais das latas de alumínio da Tabela 2-1 têm uma média de 267,1, um desvio-padrão de 22,1 e uma distribuição como a da Figura 2-1, concluímos que essas latas podem facilmente suportar as pressões de 158 lb-165 lb aplicadas ao se fixarem as tampas no lugar. Recordemos, do enunciado do Problema do Capítulo, que essas latas têm uma espessura de 0,0109 in., que é inferior à espessura comumente adotada. Com base em nosso conhecimento das características importantes do conjunto de dados da Tabela 2-1, concluímos que é possível economizar utilizando essas latas menos espessas.

O exemplo precedente ilustra como utilizar dados sobre a amplitude, para estimar o desvio-padrão. O exemplo que se segue constitui uma ilustração particularmente importante de uma *interpretação* do desvio-padrão.

EXEMPLO A Gates Electronics Company fabrica barbeadores recarregáveis, sem fio, que têm vida média de 8,0 anos, com desvio-padrão de 3,0 anos. Utilizando a regra prática, estime a vida mais longa e a mais breve desses barbeadores.

SOLUÇÃO Estimamos a maior e a menor duração de vida pela regra prática, como se segue:

$$\begin{aligned} \text{mínimo} &\approx (\text{média}) - 2 \times (\text{desvio-padrão}) \\ &= 8,00 - 2(3,0) = 2,0 \text{ anos} \\ \text{máximo} &\approx (\text{média}) + 2 \times (\text{desvio-padrão}) \\ &= 8,0 + 2(3,0) = 14,0 \text{ anos} \end{aligned}$$

Podemos, pois, esperar que a maioria dos barbeadores em questão dure de 2,0 a 14,0 anos. Tenha em mente que esses resultados são estimativas grosseiras, mas, com o conhecimento da média e do desvio-padrão, estamos em condições de obter aproximações do menor e do maior valor, passando a entender melhor como os dados variam.

Regra Empírica (ou Regra 68-95-99) para os Dados

Outra regra que auxilia a interpretação do valor de um desvio-padrão é a **regra empírica**, aplicável *somente a conjuntos de dados com distribuição aproximadamente em forma de sino*, conforme a Figura 2-10. Essa figura mostra como a média e o desvio-padrão estão relacionados com a proporção dos dados que se enquadram em determinados limites. Assim é que, com uma distribuição em forma de sino, temos 95% dos seus valores a menos de dois desvios-padrão da média. A regra empírica costuma ser designada abreviadamente como a **regra 68-95-99.**

A Regra 68-95-99 para Dados com Distribuição em Forma de Sino

- Cerca de 68% dos valores estão a menos de 1 desvio-padrão a contar da média.
- Cerca de 95% dos valores estão a menos de 2 desvios-padrão a contar da média.
- Cerca de 99,7% dos valores estão a menos de 3 desvios-padrão a contar da média.

EXEMPLO Os QIs de um grupo de adultos apresentam distribuição em forma de sino com média 100 e desvio-padrão 15. Aplique a regra empírica para achar a porcentagem de adultos com QI entre 55 e 145.

SOLUÇÃO A chave para a resolução deste problema consiste em reconhecer que 55 e 145 estão, cada um, exatamente a três desvios-padrão da média. (Como o desvio-padrão é $s = 15$, decorre que $3s = 45$, de modo que 3 desvios-padrão abaixo da média são $100 - 45 = 55$, e 3 desvios-padrão acima da média são $100 + 45 = 145$.) A regra empírica afirma que 99,7% de todos os valores estão a menos de 3 desvios-padrão a contar da média, donde decorre que 99,7% dos adultos devem ter QI entre 55 e 145. Como os valores fora deste intervalo são bastante raros, uma pessoa com QI acima de 145 ou abaixo de 55 deve ser considerada excepcional.

Um terceiro conceito importante para compreendermos e interpretarmos o valor do desvio-padrão é o **teorema de Tchebichev**. A regra empírica precedente se aplica apenas a conjuntos de dados com distribuição em forma de sino. O teorema de Tchebichev se aplica a qualquer conjunto de dados, mas seus resultados são muito aproximados.

Teorema de Tchebichev

A proporção (ou fração) de *qualquer* conjunto de dados a menos de K desvios-padrão a contar da média é sempre *ao menos* $1 - 1/K^2$, onde K é um número positivo maior do que 1. Para $K = 2$ e $K = 3$, temos os seguintes resultados específicos:

- Ao menos 3/4 (ou 75%) de todos os valores estão no intervalo que vai de 2 desvios-padrão abaixo da média a 2 desvios-padrão acima da média ($\bar{x} - 2s$ a $\bar{x} + 2s$).
- Ao menos 8/9 (ou 89%) de todos os valores estão no intervalo que vai de 3 desvios-padrão abaixo da média até 3 desvios-padrão acima da média ($\bar{x} - 3s$ a $\bar{x} + 3s$).

Utilizando valores de QI com média 100 e desvio-padrão 15, o teorema de Tchebichev afirma que ao menos 75% dos valores estarão entre 70 e 130, e ao menos 89% dos valores estarão entre 55 e 145.

Após o estudo desta seção, deve estar claro para o leitor que o desvio-padrão é uma medida da variação entre os valores. O leitor deve ainda estar em condições de calcular o desvio-padrão para um conjunto de dados, interpretar os valores do desvio-padrão e reconhecer que, para um conjunto típico, é raro um valor do mesmo diferir da média por mais de 2 ou 3 desvios-padrão.

2-5 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-4, determine a amplitude, a variância e o desvio-padrão do conjunto de dados. (Os dados são os mesmos utilizados na Seção 2-4, onde determinamos medidas de tendência central. Aqui, trata-se de medidas de variação.)

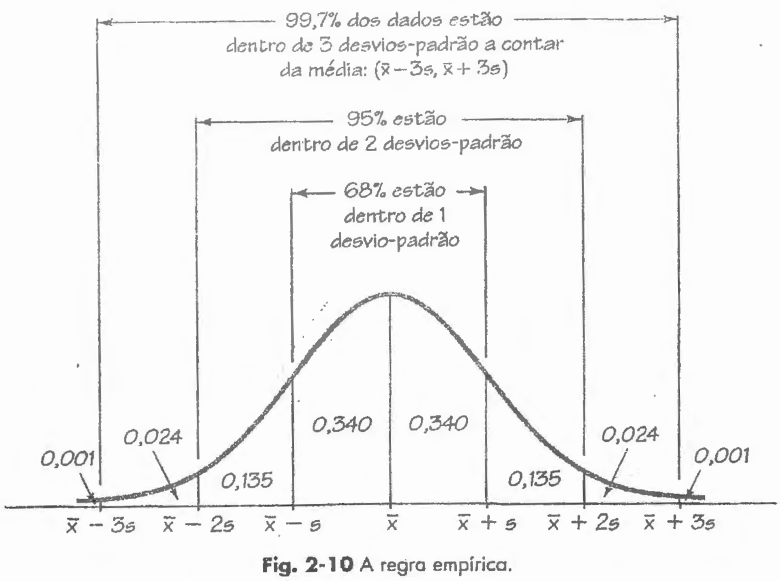


Fig. 2-10 A regra empírica.

1. Os valores a seguir são os pesos (em onças) de bifes constantes do cardápio de um restaurante como "bifes Porterhouse de 20 oz" (com base em dados coletados por um aluno do autor).

   17 20 21 18 20 20 20 18 19 19
   20 19 21 20 18 20 20 19 18 19

2. Algarismos escolhidos na loteria Pick Three de Maryland:

   0 7 3 6 2 7 6 6 6 3 8 1 7 8 7
   1 6 8 6 9 5 2 1 5 0 3 9 9 0 7

3. Resíduos de nitrato (em kg por hectare) como parte da chuva ácida em Massachusetts, de julho a setembro do últimos anos (com base em dados do Ministério da Agricultura dos EUA):

   6,40 5,21 4,66 5,24 6,96 5,53 8,23 6,80 5,78 6,00 5,41

4. Concentrações sangue-álcool de 15 motoristas envolvidos em acidentes fatais e condenados à prisão (Fonte: Dados do Ministério da Justiça dos EUA):

   0,27 0,17 0,17 0,16 0,13 0,24 0,29 0,24
   0,14 0,16 0,12 0,16 0,21 0,17 0,18

Nos Exercícios 5-8, determine a amplitude, a variância e o desvio-padrão para cada uma das duas amostras, e compare os dois conjuntos de resultados. (Na Seção 2-4 utilizamos esses mesmos dados.)

5. Tempos de espera de clientes no Jefferson Valley Bank (onde todos os clientes formam uma fila única) e no Bank of Providence (onde os clientes formam filas separadas para cada um dos três guichês). Esses conjuntos de dados já foram estudados nesta seção.

   Jefferson Valley: 6,5 6,6 6,7 6,8 7,1 7,3 7,4 7,7 7,7 7,7
   Providência: 4,2 5,4 5,8 6,2 6,7 7,7 7,7 8,5 9,3 10,0

6. Amostras das idades (em anos) de carros de alunos e de professores e funcionários de uma faculdade, obtidos na faculdade do autor.

   Estudantes: 10 4 5 2 9 7 8 8 16 4 13 12
   Profs. e funcs.: 7 10 4 13 23 2 7 6 6 3 9 4

7. Largura máxima de crânios de homens egípcios de 4000 aC a 150 aD (Fonte: Dados de *Ancient Races of the Thebaid*, por Thomson and Randall-Maciver):

   4000 a.C.: 131 119 138 125 129 126 131 132 126 128 128 131
   150 A.D.: 136 130 126 126 139 141 137 138 133 131 134 129

8. Pesos (em libras) de papel e plástico descartados em residências durante uma semana [Dados coletados no Projeto do Lixo da Universidade de Arizona]:

   Papel: 9,55 6,38 2,80 6,98 6,33 6,16 10,00 12,29
   Plástico: 2,19 2,10 1,41 0,63 0,92 1,40 1,74 2,87

Nos Exercícios 9-12, recorra aos dados do Apêndice B e calcule o desvio-padrão.

9. Conjunto 2, Apêndice B: temperaturas do corpo às 8 da manhã do dia 1
10. Conjunto 4, Apêndice B: conteúdo de nicotina em cigarros
11. Conjunto 3, Apêndice B: pesos de ursos
12. Conjunto 11, Apêndice B: pesos dos bombons M&M vermelhos

Nos Exercícios 13-16, determine o desvio-padrão dos dados resumidos na tabela de freqüências.

13. Os visitantes do Parque Nacional de Yellowstone (EUA) consideram uma erupção do gêiser Old Faithful uma atração imperdível. A tabela de freqüências resume os intervalos de tempo (em minutos) entre as erupções.

Tempo	Freqüência
40–49	8
50–59	44
60–69	23
70–79	6
80–89	107
90–99	11
100–109	1

14. Dá-se a seguir, numa tabela de freqüências, um resumo das idades de carros de alunos e de professores e funcionários da faculdade do autor. Determine o desvio-padrão de cada conjunto de dados. Com base nos resultados, há diferenças sensíveis entre as duas amostras? Em caso afirmativo, quais?

Idade	Estudantes	Professores/ Funcionários
0–2	23	30
3–5	33	47
6–8	63	36
9–11	68	30
12–14	19	8
15–17	10	0
18–20	1	0
21–23	0	1

15. A tabela de freqüências a seguir dá as velocidades desenvolvidas por motoristas multados na cidade de Poughkeepsie em um trecho onde a velocidade máxima é de 30 mi/h.

Velocidade	Freqüência
42–43	14
44–45	11
46–47	8
48–49	6
50–51	4
52–53	3
54–55	1
56–57	2
58–59	0
60–61	1

16. As companhias de seguro pesquisam continuamente as idades na morte e as causas de morte. Os dados se baseiam no estudo levado a efeito pela revista *Time* sobre as pessoas que morreram vitimadas por armas de fogo durante uma semana.

Idade na Morte	Freqüência
16–25	22
26–35	10
36–45	6
46–55	2
56–65	4
66–75	5
76–85	1

17. Se o leitor vai comprar uma bateria para substituir a do seu carro, preferirá uma que venha de uma população com $\sigma = 1$ mês ou uma que venha de uma população com $\sigma = 1$ ano? (Suponha que ambas as populações tenham mesma média e mesmo preço.) Justifique sua escolha.
18. Como administrador, o leitor deve comprar lâmpadas para um hospital. Escolheria as lâmpadas Ultralight, que têm vida média $\mu = 3000$ h e $\sigma = 200$ h, ou as lâmpadas Electrolyte, com $\mu = 3000$ h e $\sigma = 250$ h? Explique.
19. Aplique a regra prática para estimar o desvio-padrão das alturas de seus colegas da turma de estatística.
20. Aplique a regra prática para estimar o desvio-padrão das notas do último exame final de estatística.

2-5 Exercícios B: Além do Básico

21. Um teste de datilografia acusa notas com $\bar{x} = 80{,}0$ e $s = 10{,}0$, e um histograma mostra que a distribuição das notas tem a forma aproximada de um sino. Aplique a regra empírica para responder:
    a. Qual a porcentagem das notas entre 70 e 90?
    b. Qual a porcentagem das notas a menos de 20 pontos da média?
    c. Entre quais valores devem estar 99,7% das notas? (A média 80 deve estar a meio caminho entre esses dois valores.)
22. As alturas de mulheres adultas acusam média de 63,6 in. e desvio-padrão de 2,5 in. O que nos afirma o teorema de Tchebichev sobre as mulheres com altura entre 58,6 in. e 68,6 in.? Entre 56,1 in. e 77,1 in.?
23. a. Determine a amplitude e o desvio-padrão s da amostra seguinte de rendas (em dólares) de médicos autônomos (com base em dados da American Medical Association):

    108.000 236.000 179.000 206.000 236.000

    b. Como são afetados os resultados da parte (a) se se adiciona um valor constante k a cada renda?
    c. Se cada renda da parte (a) é multiplicada por uma constante k, como são afetados os resultados de (a)?
    d. Por vezes, os dados são transformados, substituindo-se cada valor x por $\log x$. Para os valores dados de x, determine se o desvio-padrão dos valores de $\log x$ é igual a $\log s$.
    e. Para os dados relativos a temperaturas do Conjunto 2 do Apêndice B (12 horas do dia 2), $\bar{x} = 98{,}20$°F e $s = 0{,}62$°F. Determine $\bar{x}$ e s para os dados, após transformar cada temperatura para a escala Celsius [*Sug.:* C = 5(F − 32)/9.]
24. Se considerarmos os valores 1, 2, 3, ..., n como uma população, o desvio-padrão pode ser calculado pela fórmula

$$\sigma = \sqrt{\frac{n^2 - 1}{12}}$$

Esta fórmula é equivalente à Fórmula 2-4, modificada pela divisão por n em lugar de $n - 1$, onde o conjunto de dados consiste nos valores 1, 2, 3, ..., n.

   a. Calcule o desvio-padrão da população 1, 2, 3, ..., 100.
   b. Ache uma expressão para o cálculo do desvio-padrão amostral s para os valores *amostrais* 1, 2, 3, ..., n.
   c. Os computadores e as calculadoras em geral utilizam um gerador de números aleatórios que produz valores entre 0,00000000 e 0,99999999. Com o decorrer do processo, todos os valores tendem a ocorrer com a mesma freqüência relativa. Determine a média e o desvio-padrão da *população* desses valores.
25. Dois grupos diferentes de uma turma de estatística fazem o mesmo teste-surpresa, com as notas relacionadas a seguir. Ache a amplitude e o desvio-padrão para cada grupo. Que conclusões sobre a variação nos dois grupos os valores da amplitude sugerem? Por que razão a amplitude é enganosa neste caso? Que conclusões sobre a variação nos dois grupos o desvio-padrão sugere?

Grupo 1:	1	20	20	20	20	20	20	20	20	20	20
Grupo 2:	2	3	4	5	6	14	15	16	17	18	19

26. a. Utiliza-se o **coeficiente de variação**, expresso como porcentagem, para descrever o desvio-padrão em relação à média. Esse coeficiente permite-nos comparar a variabilidade de conjuntos de dados com diferentes unidades de medida (como pés *versus* minutos), e se calcula como se segue:

$$\frac{s}{\bar{x}} \cdot 100 \quad \text{ou} \quad \frac{\sigma}{\mu} \cdot 100$$

Determine o coeficiente de variação para as seguintes idades de carros (em anos):

0 1 3 3 5 6 6 6 6 8 12

   b. Genichi Taguchi desenvolveu um processo de melhoria de qualidade e redução de custo de fabricação mediante uma combinação de engenharia e estatística. Um elemento fundamental no processo de Taguchi é a **razão sinal-para-ruído**. A maneira mais simples de calcular essa razão consiste em dividir a média pelo desvio-padrão. Determine a razão sinal-para-ruído para os dados amostrais da parte (a).
27. Na Seção 2-4, introduzimos o conceito geral de assimetria. A assimetria pode ser medida pelo **índice de assimetria de Pearson**:

$$I = \frac{3(\bar{x} - \text{mediana})}{s}$$

Se $I \geq 1{,}00$ ou $I \leq -1{,}00$, os dados podem ser considerados *significativamente assimétricos*. Ache o índice de assimetria de Pearson para as cargas axiais de latas de alumínio da Tabela 2-1, e determine então se existe assimetria significativa.
28. a. Uma amostra consiste em 6 valores que se situam entre 1 e 9 inclusive. Qual o maior valor possível do desvio-padrão?
    b. Para qualquer conjunto de n valores com desvio-padrão s, todo valor deve estar a menos de $s\sqrt{n-1}$ da média. Uma professora de estatística afirma que as notas de um teste em sua turma de 17 alunos tiveram média 75,0 e desvio-padrão 5,0. Kelly, que se julga a melhor aluna da turma, alega ter obtido nota 97. Pode ser verdadeira tal alegação?

2-6 Medidas de Posição

Vamos agora introduzir os escores z, que permitem comparar valores mais facilmente, através de sua padronização. Introduziremos também os quartis, percentis e decis, com os quais podemos entender melhor os dados, focalizando sua posição relativa em relação ao conjunto como um todo. Os quartis introduzidos aqui serão também utilizados nos diagramas em caixa (*boxplots*), a serem abordados na seção seguinte.

Escores z

Quase todos nós estamos familiarizados com os QIs, e reconhecemos que um QI de 102 é bastante comum, enquanto um QI de 170 é raro. Esse QI de 102 é bastante comum porque está próximo da média de 100, mas o QI de 170 é raro porque está bem acima de 100. Esta circunstância pode sugerir uma diferença entre os valores típicos e os valores raros, com base em sua diferença em relação à média ($x - \bar{x}$). Mas o vulto, ou tamanho, dessa diferença depende da escala que estamos utilizando. Com valores de QI, uma diferença de 2 pontos é insignificante, mas para médias de notas de uma faculdade uma diferença de 2 pontos entre 2,00 e 4,00 é altamente significativa, sobretudo para os pais dos alunos. Seria muito melhor se dispuséssemos de um padrão que não levasse em conta a escala utilizada. Com o valor, ou escore, padronizado, dividimos a diferença $x - \bar{x}$ (ou $x - \mu$) pelo desvio-padrão para chegarmos a esse resultado.

DEFINIÇÃO

O **escore padronizado**, ou escore z, é o número de desvios-padrão pelo qual um valor x dista da média (para mais ou para menos). Obtém-se como segue:

Amostra / População

$$z = \frac{x - \bar{x}}{s} \quad \text{ou} \quad z = \frac{x - \mu}{\sigma}$$

(Arredondar z para duas decimais.)

EXEMPLO As alturas da população de homens adultos têm média $\mu = 69{,}0$ in., desvio-padrão $\sigma = 2{,}8$ in. e distribuição em forma de sino. O jogador de basquete Michael Jordan ganhou reputação de gigante por suas proezas no jogo, mas com 78 in., ele pode ser considerado excepcionalmente alto, comparado com a população geral de homens adultos? Determine o escore z para a altura de 78 in.

SOLUÇÃO Como estamos lidando com parâmetros populacionais, o escore z se calcula como segue:

$$z = \frac{x - \mu}{\sigma} = \frac{78 - 69{,}0}{2{,}8} = 3{,}21$$

Podemos interpretar este resultado dizendo que a altura de Michael Jordan, de 78 in., está 3,21 desvios-padrão acima da média.

A importância dos escores z na estatística reside no fato de que eles permitem distinguir entre valores usuais e valores raros, ou incomuns. Consideramos usuais os valores cujos escores padronizados estão entre $-2{,}00$ e $2{,}00$, e incomuns os valores com escore z inferior a $-2{,}00$ ou superior a $2{,}00$. (Veja Figura 2-11.) A altura de Michael Jordan corresponde a um escore z de 3,21, que consideramos incomum, por ser superior a 2,00. Em comparação com a população geral, Jordan é excepcionalmente alto.

Nosso critério para classificar um escore z como incomum decorre da regra empírica e do teorema de Tchebichev. Recorde que, pela regra empírica, para dados com distribuição em forma de sino, cerca de 95% dos valores estão a menos de 2 desvios-padrão da média. (Veja Figura 2-10 da seção precedente.) Por outro lado, o teorema de Tchebichev afirma que, para qualquer conjunto de dados, ao menos 75% dos valores estão dentro de 2 desvios-padrão a contar da média.

Já vimos que os escores z são úteis para comparar escores de diferentes populações com médias distintas e desvios-padrão diferentes. O exemplo que segue ilustra essa aplicação dos escores z.

EXEMPLO Uma professora de estatística aplica dois testes diferentes a duas turmas do seu curso. Os resultados foram

Turma 1: $\bar{x} = 75$ e $s = 14$
Turma 2: $\bar{x} = 40$ e $s = 8$

Que nota é relativamente melhor: 82 no teste da Turma 1, ou 46 no da Turma 2?

SOLUÇÃO Não podemos comparar diretamente as notas 82 e 46 porque provêm de escalas diferentes. Transformamo-las, portanto, em escores z. Para o valor 82 da Turma 1, obtemos o escore z 0,50, porque

$$z = \frac{x - \bar{x}}{s} = \frac{82 - 75}{14} = 0{,}50$$

Para a nota 46 da Turma 2, o escore z correspondente é 0,75, porque

$$z = \frac{x - \bar{x}}{s} = \frac{46 - 40}{8} = 0{,}75$$

Isso significa que a nota 82 do teste da Turma 1 está 0,5 desvio-padrão acima da média, enquanto a nota 46 do teste da Turma 2 está 0,75 desvio-padrão acima da média. Isso implica que o resultado 46 do teste da Turma 2 é melhor, relativamente. Embora inferior a 82, a nota 46 tem melhor posição relativa no contexto dos outros resultados do teste. Mais adiante vamos utilizar amplamente os escores z.

Compra de Carro

Para a aquisição de um carro novo ou usado, uma boa referência é o grau de confiabilidade compilado e reportado pela revista *Consumer Reports*. Os dados relativos à freqüência de consertos se baseiam em 10 milhões de dados coletados de milhares de leitores. Os estatísticos analisam os dados em busca de padrões que conduzam a listas de carros confiáveis e carros que devem ser evitados. A presidente da Consumers Union, Rhoda Karpatkin, escreve: "Já que os números têm tanta importância em nosso trabalho, não é de surpreender que os estatísticos representem a chave desse processo."

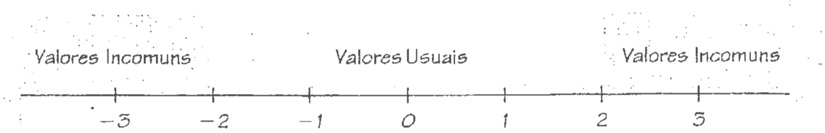


Fig. 2-11 Interpretação do escore z.
Valores com escores z inferiores a $z = 22{,}00$ ou superiores a $z = 2{,}00$ são considerados incomuns.

O exemplo precedente mostrou a eficácia dos escores z em medidas de comparação entre conjuntos diferentes de dados. Da mesma forma, os quartis, os decis e os percentis são medidas de posição convenientes para comparar valores dentro de um mesmo conjunto de dados, ou entre conjuntos diferentes.

Quartis, Decis e Percentis

Assim como a mediana divide os dados em duas partes iguais, os três **quartis**, denotados por Q_1, Q_2 e Q_3, dividem as observações *ordenadas* (dispostas em ordem crescente) em quatro partes iguais. Grosso modo, Q_1 separa os 25% inferiores dos 75% superiores dos valores ordenados; Q_2 é a mediana; e Q_3 separa os 75% inferiores dos 25% superiores dos dados. Mais precisamente, ao menos 25% dos dados serão no máximo iguais a Q_1, e ao menos 75% dos dados serão no mínimo iguais a Q_1. Ao menos 75% dos dados serão no máximo iguais a Q_3, enquanto ao menos 25% serão, no mínimo, iguais a Q_3.

Analogamente, há nove **decis**, denotados por $D_1, D_2, D_3, ..., D_9$, que dividem os dados em 10 grupos com cerca de 10% deles em cada grupo. Há, finalmente, 99 **percentis**, que dividem os dados em 100 grupos com cerca de 1% em cada grupo. (Os quartis, decis e percentis são exemplos de *fractis*, que dividem os dados em partes aproximadamente iguais.) Um estudante que se submeteu ao vestibular para ingresso em uma faculdade é informado de que está no 92.° percentil. Isso não significa, entretanto, que ele tenha obtido 92% no exame; indica, apenas, que qualquer que tenha sido a nota obtida, ela foi superior a 92% (e inferior a 8%) das notas de toda a turma. O 92.° percentil é, pois, uma excelente classificação em relação aos outros que fizeram o exame.

O processo de determinação do percentil correspondente a um determinado valor x é bastante simples, como se pode ver na expressão seguinte.

$$\text{percentil do valor } x = \frac{\text{número de valores inferiores a } x}{\text{número total de valores}} \cdot 100$$

EXEMPLO A Tabela 2-9 relaciona as 175 cargas axiais das latas de alumínio, ordenadas da mais baixa até a mais elevada. Determine o percentil correspondente a 241.

SOLUÇÃO Pela Tabela 2-9, vemos que há 21 valores inferiores a 241, de forma que

$$\text{percentil de } 241 = \frac{21}{175} \cdot 100 = 12$$

A carga axial de 241 é o 12.° percentil.

O exemplo precedente ilustra o processo de determinação do percentil correspondente a determinado valor. Para o processo inverso, há vários métodos diferentes para achar o valor correspondente a determinado percentil; o que vamos utilizar está esquematizado na Figura 2-12, em que é adotada a notação seguinte.

Notação

n	número de escores, ou valores, no conjunto de dados
k	percentil a ser utilizado
L	indicador que dá a *posição* de um escore
P_k	k^{mo} percentil

EXEMPLO Para as 175 cargas axiais de latas de alumínio da Tabela 2-9, determine o escore correspondente ao 25.° percentil; ou seja, determine o valor de P_{25}.

SOLUÇÃO Recorremos à Figura 2-12 e observamos que os dados já estão ordenados, do menor para o maior. Calculamos a seguir o indicador L como segue:

$$L = \left(\frac{k}{100}\right)n = \left(\frac{25}{100}\right) \cdot 175 = 43{,}75$$

Respondemos *não* à pergunta na Figura 2-12, se 43,75 é um número inteiro, e somos orientados a arredondar *L para cima,* ou seja, arredondar para 44. (Nesse processo em particular arredondamos L para o inteiro superior mais próximo, mas na maior parte das situações neste livro seguimos o processo geral de arredondamento.) O 25.° percentil, denotado por P_{25}, é o 44.° valor, ou escore, a contar do menor. Partindo, pois, do menor valor, 200, percorremos a lista até o 44.° valor, que é 262; assim, $P_{25} = 262$.

Suponha agora que queiramos achar o percentil correspondente a um escore de 262. Verificamos que há 41 valores abaixo de 262, não deixando de considerar cada valor individual, mesmo os que aparecem repetidos. Calculando o percentil correspondente a 262, obtemos $(41/175) \cdot 100 = 23$ (arredondado).

Custo do Risco

Há realmente um Índice de Custo do Risco (ICR) que levo em conta o custo de itens como óculos de Groucho Marx, entrada em clubes de comédia e 13 outros indicadores. Trata-se da mesma

TABELA 2-9 Valores Ordenados de Cargas Axiais de Latas de Alumínio

200	201	204	204	206	206	208	208	209	215	217	218	220	223	223
225	228	230	230.	234	236	241	242	242	248	250	251	251	252	252
254	256	256	256	257	257	258	259	259	260	261	262	262	262	262
262	263	263	263	263	263	264	265	265	265	266	267	267	268	268
268	268	268	268	268	268	268	269	269	269	269	270	270	270	270
270	270	270	270	271	271	272	272	272	272	272	273	273	273	273
273	273	274	274	274	274	275	275	275	275	276	276	276	276	276
277	277	277	277	277	277	277	277	278	278	278	278	278	278	278
279	279	279	280	280	280	281	281	281	281	282	282	282	282	282
282	283	283	283	283	283	283	284	284	284	284	285	285	285	286
286	286	286	287	287	288	289	289	289	289	289	290	290	290	291
291	292	292	292	293	293	294	295	295	297					

abordagem básica usada para estabelecer o Índice de Preços ao Consumidor (IPC), que se baseia em uma média ponderada de bens e serviços adquiridos por um consumidor típico. Enquanto valores padronizados e percentis permitem comparar diferentes valores, eles ignoram o elemento tempo. Índices como ICR e IPC permitem-nos comparar o valor de uma variável com seu valor em uma época de referência. O valor de um índice é o valor atual, dividido pelo valor de referência e multiplicado por 100.

Há aqui uma pequena discrepância: no exemplo precedente encontramos 262 para o 25.° percentil, mas no processo inverso, 262 corresponde ao 23.° percentil. À medida que aumenta o número de dados, tais discrepâncias diminuem. Poderíamos eliminá-las utilizando um processo mais complicado, que aplica a interpolação em lugar do arredondamento.

Em razão do tamanho da amostra no exemplo precedente, o indicador L calculado foi inicialmente 43,75, valor que foi arredondado para 44, porque o valor original de L não era inteiro. No próximo exemplo ilustramos um caso em que o valor original de L é um número inteiro. Essa condição nos levará para o ramo direito no fluxograma da Figura 2-12.

EXEMPLO Determine o 40.° percentil P_{40} das cargas axiais da Tabela 2-9.

SOLUÇÃO Seguindo o processo delineado na Figura 2-12 e notando que os dados já estão ordenados do menor para o maior, calculamos

$$L = \left(\frac{k}{100}\right)n = \left(\frac{40}{100}\right) \cdot 175 = 70 \quad \text{(exatamente)}$$

70 é um número inteiro, e a Figura 2-12 indica que P_{40} está a meio caminho entre os 70.° e 71.° valores. E como esses valores são ambos 269, concluímos que o 40.° percentil é 269.

Uma vez dominados os cálculos para os percentis, podemos seguir o mesmo processo para calcular os quartis e decis, levando em conta as relações indicadas na margem.

Utilizando essas relações, podemos ver que Q_1 é equivalente a P_{25}. Em um exemplo anterior, vimos que $P_{25} = 262$, e assim o primeiro quartil é $Q_1 = 262$. Se precisarmos achar o terceiro quartil, Q_3, basta reformular o problema para determinar P_{75} e proceder como indicado na Figura 2-12.

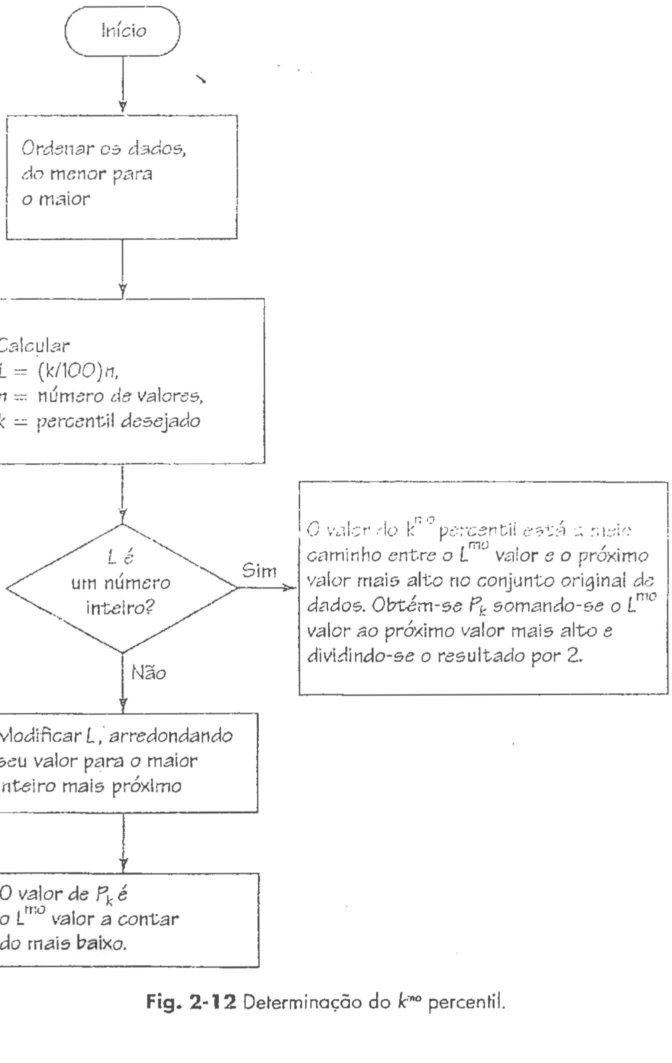


Fig. 2-12 Determinação do k^{mo} percentil.

Quartis	Decis
$Q_1 = P_{25}$	$D_1 = P_{10}$
$Q_2 = P_{50}$	$D_2 = P_{20}$
$Q_3 = P_{75}$	$\vdots$
	$D_9 = P_{90}$

Além das medidas de tendência central e de variação já introduzidas, costumamos definir outras estatísticas utilizando quartis, decis ou percentis, como segue:

intervalo interquartil $= Q_3 - Q_1$
intervalo semi-interquartil $= (Q_3 - Q_1)/2$
quartil médio $= (Q_1 + Q_3)/2$
amplitude de percentis 10-90 $= P_{90} - P_{10}$

Utilização de Calculadoras e Computadores na Estatística Descritiva

Ao lidarmos com grandes conjuntos de dados, é conveniente utilizarmos pacotes estatísticos a fim de obtermos resultados mais rápidos, fáceis e confiáveis. Os resultados que seguem, obtidos com STATDISK e Minitab, se baseiam nas 175 cargas axiais da Tabela 2-1; esses são exemplos de resultados que se obtêm quase com a mesma rapidez com que se introduzem os dados.

Podemos também utilizar as calculadoras para obter estatísticas descritivas. A maioria das calculadoras científicas dá pelo menos a média e o desvio-padrão. Com uma calculadora TI-83, devemos utilizar `STAT` e `Edit` para introduzir um conjunto de dados em uma coluna, como L1; em seguida, aplicar `STAT` e `CALC` para obter a opção `1-Var Stats`. Os resultados apresentados pela TI-83 incluem a média, a soma dos valores, a soma dos quadrados, o desvio-padrão, o número de valores (ou observações), o mínimo, o máximo, a mediana e os quartis. Como a TI-83 e o Minitab calculam os quartis de uma maneira ligeiramente diferente da adotada neste livro, pode haver algumas discrepâncias.

2-6 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-4, expresse todos os escores z com duas decimais.

1. Os homens adultos (nos EUA) têm altura média de 69,0 polegadas, com desvio-padrão de 2,8 polegadas. Determine os escores z correspondentes a:
   a. O jogador de basquete Mugsy Bogues que tem 5 pés e 3 in.
   b. O jogador de basquete Shaquille O'Neal, que tem 7 pés e 1 polegada.
   c. O autor, que é um jogador de golfe e tênis com 69,72 in.
2. Os carros dos estudantes na faculdade do autor têm idade média de 7,90 anos, com desvio-padrão de 3,67 anos. Determine os escores z para os carros com as seguintes idades:
   a. Um Corvette de 12 anos
   b. Uma Ferrari de 2 anos
   c. Um Porsche novo

STATDISK

File Edit Analysis Data Help

Cans109	
Sample Size, n	175
Mean, $\bar{x}$	267.11
Median	273.00
Midrange	248.50
RMS	268.02
Variance, s^2	488.95
St Dev, s	22.112
Mean Dev	16.019
Range	97.000
Minimum	200.00
1st Quartile	262.00
2nd Quartile	273.00
3rd Quartile	282.00
Maximum	297.00
$\sum x$	46745
$\sum x^2$	1257133

MINITAB

```
Variable      N     Mean   Median   Tr Mean    StDev  SE Mean
CANS109     175   267.11   273.00    269.15    22.11     1.67

Variable    Min      Max       Q1        Q3
CANS109  200.00   297.00   262.00    282.00
```

3. Os números de horas que os calouros passam estudando cada semana têm média de 7,06 h e desvio-padrão de 5,32 h (com base em dados de *The American Freshman*). Determine o escore z para um calouro que estuda 20 horas por semana.
4. Os tempos que os estudantes de curso secundário passam trabalhando em empregos cada semana têm média de 10,7 h e desvio-padrão de 11,2 h [com base em dados da National Federation of State High School Associations (Federação Nacional das Associações das Escolas Secundárias Estaduais)]. Determine o escore z correspondente a um estudante que trabalha 8 horas por semana.

Nos Exercícios 5-8, expresse todos os escores z com duas decimais. Considere fora do comum um escore z inferior a − 2,00 ou superior a 2,00.

5. A admissão ao Beanstalk Club é limitada a mulheres e homens muito altos. A exigência de altura mínima para as mulheres é 70 in. As alturas das mulheres têm média de 63,6 in. e desvio-padrão de 2,5 in. Ache o escore z correspondente a uma mulher com 70 in. de altura e determine se se trata de uma altura fora do comum.
6. Uma mulher escreveu a *Dear Abby*, alegando ter dado à luz uma criança 308 dias após uma visita do seu marido, que estava servindo na Marinha. Os tempos de duração da gravidez acusam uma média de 268 dias, com desvio-padrão de 15 dias. Determine o escore z correspondente a 308 dias. Esse prazo pode ser considerado fora do comum? Que se pode concluir?
7. Certa máquina automática aceita moedas de 25 cents (de dólar) que não fujam ao padrão comum. Ache o escore z para uma moeda de 25 cents que pesa 5,50 g. Essa moeda será aceita pela máquina? (Os pesos das moedas de 25 cents têm média de 5,67 gramas, com desvio-padrão de 0,070 gramas.)
8. Para os homens com idades entre 18 e 24 anos, os níveis de colesterol (em mg/100 ml) têm média de 178,1 e desvio-padrão de 40,7 [com base em dados do National Health Survey (Serviço Nacional de Saúde dos EUA)]. Determine o escore z para um homem, com idade entre 18 e 24 anos, que tem um nível de colesterol de 275,2 mg/100 ml. Esse nível pode ser considerado excepcionalmente elevado?
9. Qual dos dois escores abaixo acusa melhor posição relativa?

   a. Um escore de 60 em um teste com $\bar{x} = 50$ e $s = 5$
   b. Um escore de 250 em um teste com $\bar{x} = 200$ e $s = 20$.

10. Dois grupos semelhantes de estudantes fazem testes equivalentes de facilidade de linguagem. Qual dos resultados seguintes indica maior facilidade relativa de linguagem?

    a. Um escore de 65 em um teste com $\bar{x} = 70$ e $s = 10$
    b. Um escore de 455 em um teste com $\bar{x} = 500$ e $s = 80$.

11. Três candidatos a um emprego fazem testes equivalentes de pensamento crítico. Qual dos escores abaixo corresponde à posição relativa mais elevada?

    a. Um escore de 37 em um teste para o qual $\bar{x} = 28$ e $s = 6$
    b. Um escore de 398 em um teste para o qual $\bar{x} = 312$ e $s = 56$
    c. Um escore de 4,10 em um teste para o qual $\bar{x} = 2,75$ e $s = 0,92$

12. Três estudantes fazem testes equivalentes de senso de humor e, após terminada a risada, calculam-se seus escores. Qual é o escore relativo mais alto?

    a. Um escore de 2,7 em um teste com $\bar{x} = 3,2$ e $s = 1,1$
    b. Um escore de 27 em um teste em que $\bar{x} = 35$ e $s = 12$
    c. Um escore de 850 em um teste em que $\bar{x} = 921$ e $s = 87$

Nos Exercícios 13-16, utilize as 175 cargas axiais ordenadas da Tabela 2-9. Ache o percentil correspondente ao valor dado.

13. 254 **14.** 265 **15.** 277 **16.** 288

Nos Exercícios 17-24, utilize as 175 cargas axiais da Tabela 2-9 para achar o percentil, quartil ou decil indicado.

17. P_{70} **18.** P_{20} **19.** D_6 **20.** D_3
21. Q_3 **22.** Q_1 **23.** D_1 **24.** P_1

Nos Exercícios 25-28, com base nos pesos (em libras) de ursos do Conjunto de Dados 3, do Apêndice B, determine o percentil correspondente ao peso indicado.

25. 144 **26.** 212 **27.** 316 **28.** 90

Nos Exercícios 29-36, com base nos pesos (em libras) de ursos do Conjunto de Dados 3, do Apêndice B, determine o percentil, o quartil ou o decil indicado.

29. P_{85} **30.** P_{35} **31.** Q_1 **32.** Q_3
33. D_9 **34.** D_3 **35.** P_{50} **36.** P_{95}

2-6 Exercícios B: Além do Básico

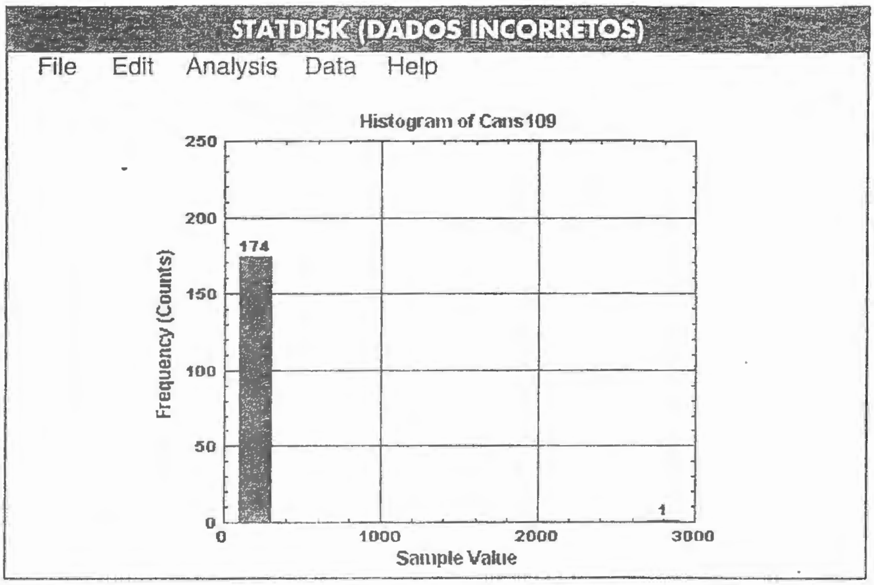
37. Tome por base as cargas axiais ordenadas da Tabela 2-9.

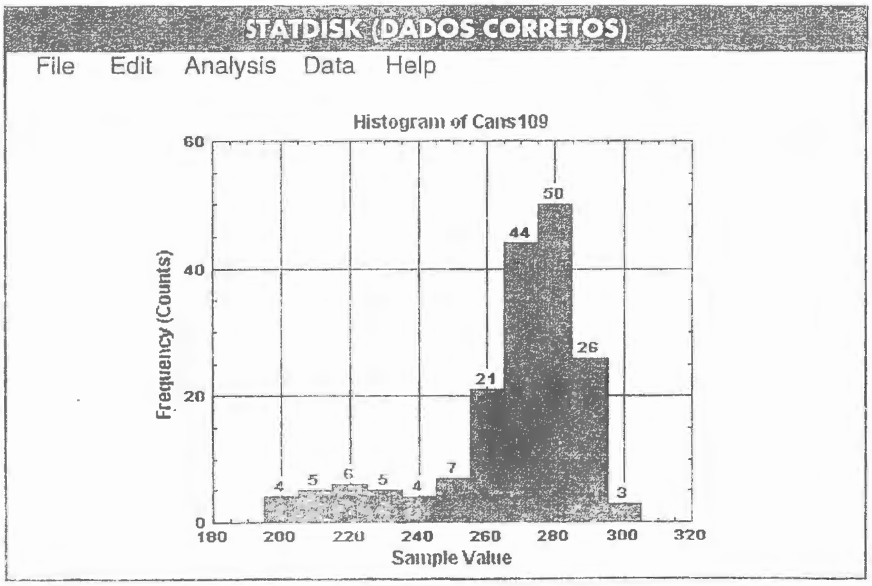
    a. Determine o intervalo interquartil.
    b. Determine o quartil médio.
    c. Determine a amplitude de percentis 10-90.
    d. $P_{50} = Q_2$? Em caso afirmativo, isso ocorre sempre?
    e. $Q_2 = (Q_1 + Q_3)/2$? Em caso afirmativo, isto ocorre sempre?

38. Ao determinar percentis utilizando a Figura 2-12, se o indicador L não é um número inteiro, arredondamo-lo para o maior inteiro mais próximo. Um processo alternativo consiste em interpolar, de modo que um indicador 23,75 conduza a um valor que está a 0,75 (ou 3/4) no caminho entre os 23.º e 24.º escores. Utilize esse método de interpolação para calcular P_{35}, Q_1 e D_3 para os pesos relacionados no Conjunto de Dados 3 do Apêndice B.
39. Para as 175 cargas axiais das latas da Tabela 2-1, a média é 267,1 e o desvio-padrão é 22,1. Ache os dois valores fronteira que separam os valores ordinários dos valores incomuns.
40. Com os escores 2, 5, 8, 9 e 16, primeiro calcule $\bar{x}$ e s; em seguida, substitua cada valor pelo escore z correspondente. (Não arredonde os escores z; tome tantas decimais quantas sua calculadora permitir.) Ache então a média e o desvio-padrão dos cinco escores z. Esses novos valores da média e do desvio-padrão serão obtidos para todo conjunto de escores z?

2-7 Análise Exploratória de Dados (EDA)

Às vezes observamos ou coletamos dados com um objetivo específico em vista — por exemplo, verificar a eficiência de um novo tratamento de insônia. Outras vezes, não há qualquer objetivo específico; apenas desejamos explorar os dados para ver o que eles nos revelam. Na exploração de dados, podemos aplicar muitas das técnicas já apresentadas neste capítulo. Recorde que, na Seção 2-1, relacionamos três importantes características dos dados: (1) natureza ou forma da distribuição; (2) um valor representativo; e (3) uma medida de variação. É imprescindível considerar a distribuição dos dados, porque ela pode afetar não só os métodos estatísticos a ser usados, como também as conclusões a que chegarmos. No espírito da **análise exploratória de dados**, não devemos apenas visualizar o histograma e achar que entendemos a natureza da distribuição — é preciso *explorar*. A título de exemplo, mostramos dois histogramas obtidos com o





STATDISK das 175 cargas axiais da Tabela 2-1. O primeiro histograma representa os 175 valores com uma alteração: o primeiro valor, 270, é registrado incorretamente como 2700. O segundo histograma está correto. Note o efeito acentuado que um simples erro em um dos 175 valores tem sobre a forma do histograma. Nesse caso, o valor extremo, incorreto, de 2700 causa séria distorção no histograma. Em outros casos, tais valores extremos (chamados *outliers*) podem ser corretos, mas podem dar uma idéia errônea da verdadeira natureza da distribuição quando ilustrada por um histograma. Sem uma exploração mais aprofundada dos dados, podemos tirar conclusões seriamente errôneas dos histogramas.

Com EDA, dá-se ênfase à exploração original, com os objetivos de simplificar a descrição dos dados e obter uma visão mais profunda da sua natureza. Adiante, nesta seção, fazemos uma comparação entre EDA e a estatística tradicional em três áreas principais da estatística.

Análise Exploratória de Dados	*Estatística Tradicional*
Explora os dados em um nível preliminar	*Confirma* conclusões finais sobre os dados
Poucas (ou talvez nenhuma) hipóteses são feitas sobre os dados	Tipicamente, exige hipóteses muito importantes sobre os dados
Costuma exigir cálculos e gráficos relativamente simples	Em geral, os cálculos são complexos e os gráficos desnecessários.

Na Seção 2-3, estudamos os gráficos do tipo ramo-e-folhas, um dos instrumentos comumente utilizados em EDA. Introduziremos agora os diagramas em caixa (*boxplots*) que não foram abordados antes porque exigem quartis só estudados na seção precedente.

Diagramas em Caixa (*Boxplots*)

Os diagramas em caixa são convenientes para revelar tendências centrais, dispersão, distribuição dos dados e a presença de *outliers* (valores extremos). A construção de um diagrama em caixa exige que tenhamos o valor mínimo, o primeiro quartil Q_1, a mediana (ou segundo quartil Q_2), o terceiro quartil Q_3 e o valor máximo. Como as medianas revelam uma tendência central, ao passo que os quartis indicam a dispersão dos dados, os diagramas em caixa têm a vantagem de não serem tão sensíveis a valores extremos como outras medidas baseadas na média e no desvio-padrão. Por outro lado, os diagramas em caixa (*boxplots*) não dão informação tão detalhada quanto os histogramas ou os gráficos ramo-e-folhas, podendo não ser, assim, a melhor escolha quando lidamos com um único conjunto de dados. Os diagramas em caixa são, entretanto, mais convenientes na comparação de dois ou mais conjuntos de dados. Ao utilizarmos dois ou mais diagramas em caixa para comparar diferentes conjuntos de dados, é importante utilizarmos a mesma escala, de forma a possibilitar a comparação.

DEFINIÇÕES

O valor mínimo, o primeiro quartil Q_1, a mediana, o terceiro quartil Q_3 e o valor máximo constituem um **resumo de cinco números** de um conjunto de dados.

Um **diagrama em caixas (boxplot)** é um gráfico de dados que consiste em uma reta que se prolonga do menor ao maior valor, e um retângulo com retas traçadas no primeiro quartil Q_1, na mediana e no terceiro quartil Q_3.

EXEMPLO Com base nos dados sobre pulsação de fumantes (Conjuto de Dados 8 do Apêndice B),

a. Determine os valores que constituem o resumo de 5 números.
b. Construa um diagrama em caixa para esses valores.

SOLUÇÃO

a. O resumo de cinco números consiste no mínimo, Q_1, mediana, Q_2 e no máximo. Para determinar esses valores, devemos primeiro ordenar os dados do menor para o maior. Segue a lista ordenada dos 22 valores de pulsação de fumantes (Conjunto de Dados 8):

52 52 60 60 60 60 63 63 66 67 68
69 71 72 73 75 78 80 82 83 88 90

Nesta lista ordenada, é fácil identificar o mínimo 52 e o máximo 90. Com auxílio do fluxograma da Figura 2-12, vemos que o primeiro quartil Q_1 (ou P_{25}) é 60, que localizamos calculando $L = (25/100)22 = 5{,}5$, arredondado para 6. Q_1 é o sexto valor na lista ordenada, a saber, 60. A mediana é 68,5, que é o valor a meio caminho entre os 11.º e 12.º valores. Vemos também que $Q_3 = 78$, procurando na Figura 2-12 o 75.º percentil. O resumo de 5 números é, pois, 52, 60, 68,5, 78 e 90.

b. Na Figura 2-13 temos o diagrama em caixas para os dados. Utilizemos o mínimo (52) e o máximo (90) para determinar uma escala de valores, e a seguir marcamos os valores com base no resumo de cinco números

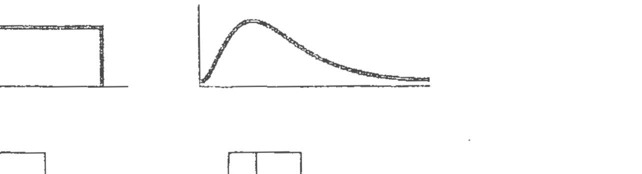


Fig. 2-13 Diagrama em caixas de pulsações (batidas por minuto) de fumantes.

Na Figura 2-14 exibimos alguns diagramas em caixas genéricos, juntamente com as formas usuais de distribuição.

Valores Extremos (*Outliers*)

No decorrer da determinação de um resumo de 5 números e da construção de um diagrama em caixas, torna-se fácil identificar *outliers* (ou valores extremos), que são valores extremamente raros, no sentido de que estão muito afastados da maioria dos dados. Ao explorarmos um conjunto de dados, não podemos deixar de considerar os *outliers*, porque eles podem revelar informações importantes. Consideremos, por exemplo, a lista completa de pulsações do Conjunto de Dados 8. Basta ordenarmos os valores para ver que os valores 8 e 15 são *outliers*. Tratam-se de valores realmente excepcionais ou são valores errados? Embora haja alguns estudantes cujas condições físicas podem ser descritas como letárgicas, é extremamente improvável que alguém com uma pulsação de 8 ou 15 seja capaz de entrar em uma

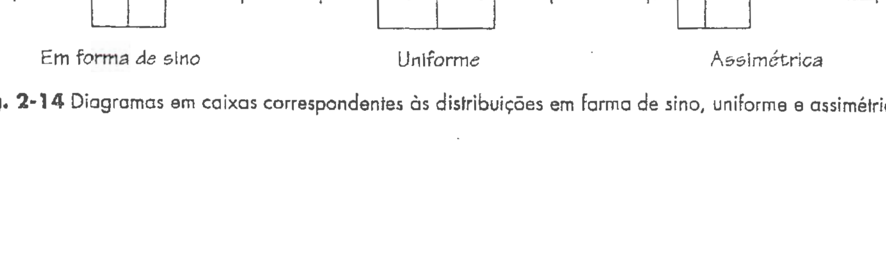


Fig. 2-14 Diagramas em caixas correspondentes às distribuições em forma de sino, uniforme e assimétrica.

sala de aula e sair dela por suas próprias forças. Concluímos, portanto, que 8 e 15 são erros, justificando-se a eliminação desses valores do conjunto. Devemos eliminar também a pulsação de 100? Não, porque esse valor não está demasiadamente distante dos outros, e provavelmente se refere a alguém excitado por estar em uma aula de estatística. De modo geral, devemos eliminar os *outliers* quando eles decorrem de erros óbvios; mas freqüentemente eles representam anomalias interessantes que merecem estudo mais detalhado. Na verdade, para alguns conjuntos de dados, os *outliers* são a característica mais importante. Um estudo sobre ovos e colesterol incluiu um homem que tinha consumido vários ovos por dia durante muitos anos. Sua taxa de consumo de ovos representava um *outlier*, mas o aspecto importante da questão é que o excesso de ovos não pareceu afetar seu nível de colesterol, que se manteve na média. Ao explorarmos dados, podemos estudar os efeitos dos *outliers* construindo gráficos e calculando medidas com eles e sem eles. (Veja Exercício 12, para uma forma de representar os *outliers* em diagramas em caixas.)

Utilização de Computadores e Calculadoras para Diagramas em Caixas

Podemos utilizar STATDISK, Minitab e a calculadora TI83 para criar diagramas em caixas. Com STATDISK, escolhemos o item `Data` do menu e utilizamos `Sample Editor` para introduzir os dados; *clicamos* `COPY`, escolhemos `Data/Boxplot` e *clicamos* `PASTE`; finalmente acionamos `Evaluate`. Com Minitab utilizamos as opções de `File/New Worksheet/Graph/Boxplot`. Os valores dos quartis calculados por Minitab e pela TI-83 podem diferir dos obtidos com a aplicação da Figura 2-12, de forma que os diagramas em caixas podem se apresentar ligeiramente diferentes.

Vimos que os *boxplots*, ou *diagramas em caixas*, são úteis para comparar conjuntos de dados; a figura a seguir apresenta os diagramas em caixa para as pulsações de fumantes e não-fumantes (Conjunto de Dados 8 do Apêndice B), feitos usando Minitab. Os *outliers* 8 e 15 foram excluídos.

Comparando os dois gráficos Minitab, vemos que não há diferenças substanciais. Os não-fumantes têm mais valores extremos, mas as medianas parecem coincidir, e a dispersão dos dados também é aproximadamente a mesma. Para o grupo de estudantes que faz estatística, parece que não há diferenças dignas de nota entre a pulsação dos fumantes e a dos não-fumantes.

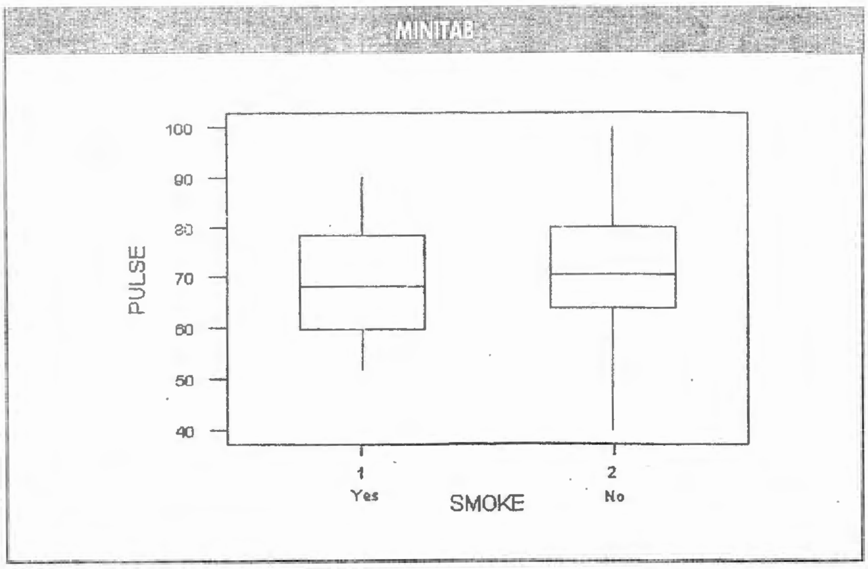
2-7 Exercícios A: Habilidades e Conceitos Básicos

Inclua valores do resumo de 5 números em todos os diagramas em caixas.

1. Considere os dados do Conjunto 4 do Apêndice B e construa um diagrama em caixas para o conteúdo de nicotina de cigarros.
2. Com base nos dados do Conjunto 4 do Apêndice B, construa um diagrama em caixas para o conteúdo de alcatrão dos cigarros.
3. Em "Ages of Oscar-Winning Best Actors and Actresses" na revista *Mathematics Teacher*, por Richard Brown e Gretchen Davis, utilizam-se diagramas em caixas, ou *boxplots*, para comparar as idades dos atores e das atrizes na ocasião em que receberam o Oscar. Relacionam-se adiante os 34 vencedores recentes de cada categoria. Compare os dois conjuntos de dados com auxílio de um diagrama em caixas.

Atores:	32	37	36	32	51	53	33	61	35	45	55	39
	76	37	42	40	32	60	38	56	48	48	40	
	43	62	43	42	44	41	56	39	46	31	47	
Atrizes:	50	44	35	80	26	28	41	21	61	38	49	33
	74	30	33	41	31	35	41	42	37	26	34	
	34	35	26	61	60	34	24	30	37	31	27	

4. Considere o Conjunto 8 do Apêndice B para estes dois conjuntos de dados: pulsações dos fumantes e pulsações dos não-fumantes. Construa um diagrama em caixas para cada conjunto. Com base nos resultados, parece haver diferença de pulsação entre os dois grupos? Em caso afirmativo, quanto? É este o resultado esperado? (Exclua os valores 8 e 15, que devem ser erros.)



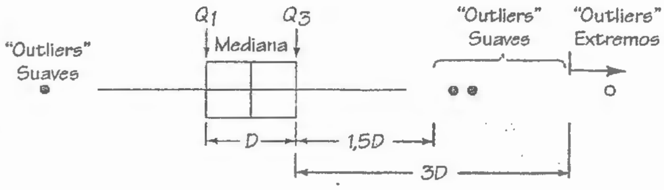
5. Considere o Conjunto 8 do Apêndice B para estes dois conjuntos de dados: taxas de pulsação para homens e para mulheres. Construa um diagrama em caixas para cada conjunto. Com base nos resultados, as taxas de pulsação dos dois conjuntos parecem ser diferentes? Em caso afirmativo, quanto? (Exclua os valores 8 e 15, que devem ser erros.)
6. Considere o Conjunto de Dados 10 do Apêndice B. Com auxílio de diagramas em caixas, compare os comprimentos dos filmes classificados R (restrito) com os dos filmes classificados não-R.
7. Considere o Conjunto de Dados 11 do Apêndice B. Com o auxílio de diagramas em caixas, compare os pesos dos bombons M&M vermelhos com os dos bombons M&M amarelos.
8. Considere o Conjunto de Dados 13 do Apêndice B. Construa um diagrama em caixas para os pesos das moedas de 25 cents. Compare a forma do gráfico resultante com as formas genéricas mostradas na Figura 2-14. Com base no diagrama em caixas, que podemos concluir sobre a natureza da distribuição?
9. Considere o Conjunto de Dados 12 do Apêndice B. Construa um diagrama em caixas para os 150 algarismos da Loteria "Pick Three" de Maryland. Compare a forma do gráfico resultante com as formas genéricas da Figura 2-14. Com base no gráfico, o resultado da loteria de Maryland parece estar de acordo com o resultado esperado?
10. Considere o Conjunto de Dados 1 do Apêndice B. Com auxílio de diagramas em caixas, compare os pesos do papel descartado com os pesos do plástico descartado.

2-7 Exercícios B: Além do Básico

11. O supervisor de manutenção de uma frota de carros deve comprar baterias de substituição de um de três fornecedores. Para isto, testa a durabilidade de amostras de baterias desses três fornecedores, registrando as vidas (em meses), conforme resumo nos diagramas em caixa a seguir, obtidos com Minitab. Qual desses gráficos corresponde à marca que vai adquirir? Por quê?
12. Os diagramas em caixas, ou *boxplots*, discutidos nesta seção costumam chamar-se diagramas *esqueletais*. No estudo dos *outliers*, convém introduzir uma modificação na construção dos diagramas em caixas, como segue:
    a. Calcular a diferença entre os quartis Q_3 e Q_1, denotando-a por D: $D = Q_3 - Q_1$.
    b. Desenhar normalmente a caixa com a mediana e os quartis, mas, ao prolongar as retas que se ramificam da caixa, caminhar apenas até os escores que estão a menos de $1{,}5D$ da mesma.
    c. Os ***outliers* suaves** são os valores que superam Q_3 em $1{,}5D$ a $3D$, ou estão $1{,}5D$ a $3D$ abaixo de Q_1. Marque os *outliers* suaves com pontos cheios.
    d. Os **"*outliers* extremos"** são escores que excedem Q_3 em mais de $3D$ ou estão a mais de $3D$ abaixo de Q_1. Marque os *outliers* extremos como pequenos círculos vazios.

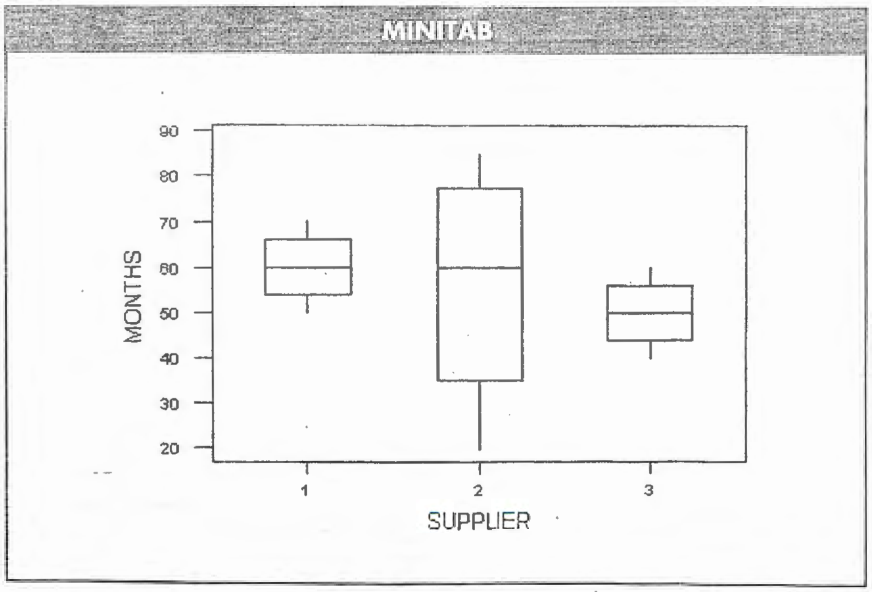
A figura que acompanha é um exemplo do diagrama em caixas descrito aqui. Utilize esse processo para construir o diagrama em caixas para os valores dados, identificando os *outliers* extremos e suaves:

3 15 17 18 21 21 22 25 27 30 38 49 68



Vocabulário

estatística descritiva
estatística inferencial
tabela de freqüência
freqüência
limite inferior de classe
limite superior de classe
fronteiras de classe
pontos médios de classe
amplitude de classe
freqüência relativa
tabela de freqüência relativa
freqüência acumulada
tabela de freqüência acumulada
histograma
histograma de freqüência relativa
gráfico por pontos
gráfico ramo-e-folhas
diagrama de Pareto
gráfico em setores
diagrama de dispersão
medida de tendência central
média aritmética
média
tamanho da amostra
mediana
moda
bimodal
multimodal



ponto médio
média ponderada
assimétrico
simétrico
negativamente assimétrico
positivamente assimétrico
amplitude
desvio-padrão
desvio
desvio médio (ou absoluto)
variância
regra prática (desvio-padrão em termos da amplitude)
regra empírica
regra 68-95-99
teorema de Tchebichev
escore padronizado
escore z
quartis
decis
percentis
análise exploratória de dados (EDA)
resumo dos 5 números
diagrama em caixas
boxplot
outlier

Revisão

O Capítulo 2 abordou principalmente métodos e técnicas para resumir, descrever, explorar e comparar dados. Vimos as três características mais importantes dos dados, a saber, (1) natureza ou forma da distribuição, (2) valor representativo, e (3) medida de variação. Essas características podem ser estudadas e descritas com os recursos do Capítulo 2. Especificamente, para determinado conjunto de dados, devemos saber

- Resumir os dados, construindo uma tabela de freqüências ou uma tabela de freqüências relativas (Seção 2-2)
- Apresentar visualmente a natureza da distribuição, construindo um histograma, um gráfico por pontos, um ramo-e-folhas, um gráfico em setores, ou um diagrama de Pareto (Seção 2-3)
- Calcular medidas de tendência central: média, mediana, moda e ponto médio (Seção 2-4)
- Calcular medidas de variação: desvio-padrão, variância e amplitude (Seção 2-5)
- Comparar valores individuais, utilizando escores z, quartis, decis ou percentis (Seção 2-6)
- Investigar e explorar a dispersão de dados, o centro de dados e a amplitude de valores, com a construção de um diagrama em caixas, ou *boxplot* (Seção 2-7)

É preciso não só calcular as tabelas, gráficos e medidas, mas também *compreender* e *interpretar* esses resultados. Assim é que devemos entender com clareza que o desvio-padrão é uma medida da variação dos dados, e saber utilizá-lo para distinguir entre valores usuais e valores incomuns.

Exercícios de Revisão

1. A NCAA estava estudando meios de acelerar o término dos jogos universitários de basquetebol. Dão-se abaixo os tempos (em segundos) decorridos para jogar os dois últimos minutos do tempo regulamentar em 60 jogos das quatro primeiras rodadas do campeonato NCAA de basquetebol (com base em dados publicados no *USA Today*). Tomando o tempo mínimo como limite inferior da primeira classe, construa uma tabela de freqüências com 9 classes.

756	587	929	871	378	503	564	1128	693	748
448	670	1023	335	540	853	852	495	666	474
443	325	514	404	820	915	793	778	627	483
861	337	292	1070	625	457	676	494	420	862
991	615	609	723	794	447	704	396	235	552
626	688	506	700	240	363	860	670	396	345

2. Construa uma tabela de freqüências relativas (com 9 classes) para os dados do Exercício 1.
3. Construa um histograma correspondente à tabela de freqüências do Exercício 1.
4. Para os dados do Exercício 1, determine (a) Q_1, (b) P_{45} e (c) o percentil correspondente ao tempo de 335 s.
5. Aplique a regra prática para estimar o desvio-padrão dos dados do Exercício 1.
6. Utilize a tabela de freqüências do Exercício 1 para achar a média e o desvio-padrão dos tempos.
7. Com os dados do Exercício 1, construa um gráfico ramo-e-folhas com 10 ramos.
8. Construa um diagrama em caixas (*boxplot*) para os dados do Exercício 1.
9. Dão-se a seguir os tempos (em segundos) decorridos entre a formulação do pedido e a entrega do prato em uma lanchonete McDonald's. Determine: (a) a média; (b) a mediana; (c) a moda; (d) o ponto médio; (e) a amplitude; (f) o desvio-padrão; (g) a variância:

   135 90 85 121 83 69 87 159 177 135 227

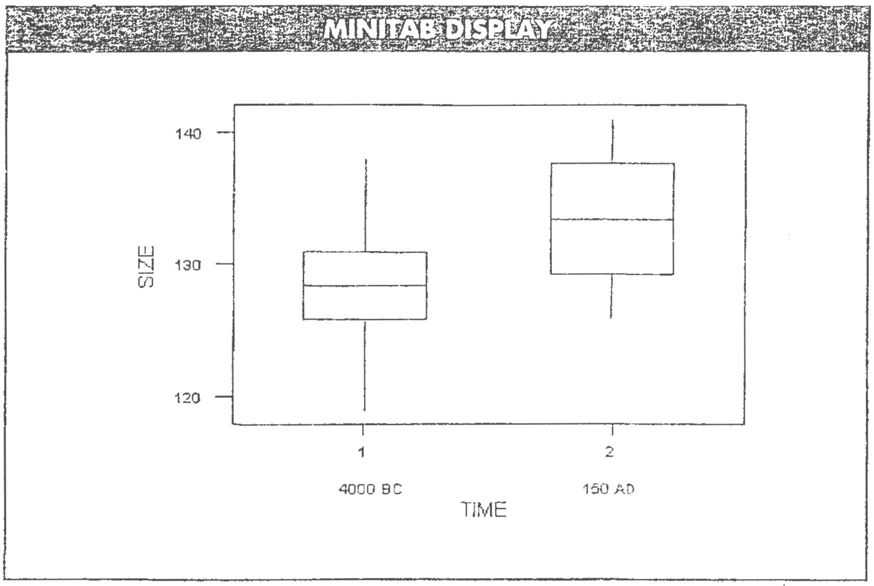
10. Dão-se abaixo as idades de presidentes dos EUA na ocasião da posse. Calcule: (a) a média; (b) a mediana; (c) a moda; (d) o ponto médio; (e) o intervalo; (f) o desvio-padrão; (g) a variância; (h) Q_1; (i) P_{30}; (j) D_7.

57	61	57	57	58	57	61	54	68	51	49	64	50	48
65	52	56	46	54	49	51	47	55	55	54	42	51	56
55	51	54	51	60	62	43	55	56	61	52	69	64	46

11. Os valores em um teste de percepção de profundidade acusam média 200 e desvio-padrão 40.
    a. Um valor de 260 pode ser considerado excepcionalmente alto? Explique.
    b. Qual o escore z correspondente a 185?
    c. Supondo que os escores tenham uma distribuição em forma de sino, que nos informa a regra empírica sobre a porcentagem de escores entre 120 e 280?
    d. Qual é a média, após adicionar 20 pontos a todos os escores?
    e. Qual é o desvio-padrão na hipótese d?
12. A tabela a seguir dá os tempos (em anos) que os estudantes de certa faculdade levaram para obter o grau de bacharel (a partir de dados do National Center for Education Statistics). Com base na tabela, calcule a média e o desvio-padrão. Podemos considerar como incomum o fato de um estudante levar 8 anos para concluir o bacharelado? Explique.

Tempo (anos)	Número
4	147
5	81
6	27
7	15
7,5 – 11,5	30

13. Construa o histograma de freqüências relativas para a tabela do Exercício 12.
14. Um psicólogo industrial deu a um empregado dois testes diferentes para medir o grau de satisfação no emprego. Qual resultado é melhor: um escore de 57 no primeiro teste, que teve média 72 e desvio-padrão 20, ou um escore de 450 no segundo teste, que acusou média 500 e desvio-padrão 80? Explique.
15. Considere os dois diagramas em caixas a seguir, obtidos com Minitab. O primeiro representa uma amostra de crânios de homens egípcios de cerca de 4000 a.C., enquanto o segundo representa uma amostra de crânios de homens egípcios de cerca de 150 a.D. (com base em dados de *Ancient Races of the Thebaid*, por Thomson and Randall-Maciver). Uma variação do tamanho das cabeças poderia sugerir mudanças sociais, como miscigenação com outras culturas. Comparando os dois gráficos, pode-se constatar variação na largura máxima dos crânios? Explique.



16. A Guarda Costeira dos EUA coletou dados sobre acidentes sérios com embarcações, categorizando-os conforme a seguir, com as respectivas freqüências dadas entre parênteses. Construa um diagrama de Pareto resumindo os dados.

    Colisão com outra embarcação (2203)
    Colisão com um objeto fixo (839)
    Encalhe (341)
    Queda de pessoa no mar (431)
    Soçobro (458)

Exercícios Cumulativos de Revisão

1. Dão-se a seguir os tempos (em horas) gastos em um dia com serviços de escritório por uma amostra de chefes de escritório (Fonte: Dados da Adia Personnel Services):

   3,7 2,9 3,4 0,0 1,5 1,8 2,3 2,4 1,0 2,0
   4,4 2,0 4,5 0,0 1,7 4,4 3,3 2,4 2,1 2,1

   a. Calcule a média, a mediana, a moda e o ponto médio.
   b. Calcule o desvio-padrão, a variância e a amplitude.
   c. Os dados provêm de uma população discreta ou contínua?
   d. Qual é o nível de mensuração desses valores? (Nominal, ordinal, intervalar, razão)
2. a. Um conjunto de dados está no nível nominal de mensuração, e desejamos obter um valor representativo dos dados. Qual das medidas seguintes é mais adequada: média, mediana, moda ou ponto médio? Por quê?
   b. Obtém-se uma amostra telefonando para os 250 primeiros assinantes da lista telefônica local. Que tipo de amostragem estamos usando? (aleatória, estratificada, sistemática, por conglomerado, de conveniência)
   c. Faz-se uma pesquisa abordando todas as pessoas que saem da cabine eleitoral em 50 zonas eleitorais selecionadas aleatoriamente. Que tipo de amostragem estamos utilizando? (aleatória, estratificada, sistemática, por conglomerado, de conveniência)
3. Anualmente o Ministério da Energia dos EUA publica um *Annual Energy Review* que inclui o consumo de energia *per capita* (em milhões de Btu) para cada um dos 50 estados. Calculando-se a média desses 50 valores, o resultado é o consumo médio de energia *per capita* para todos os 50 estados combinados? Em caso negativo, explique como calcularia o consumo médio *per capita* para os 50 estados em conjunto.

Projeto para Computador

Admite-se, de modo geral, que a temperatura média de um adulto sadio seja de 98,6°F. Com base no Conjunto de Dados 2 do Apêndice B, considere as temperaturas tomadas à meia-noite do segundo dia. Como o Conjunto de Dados 2 não está armazenado como um arquivo STATDISK ou Minitab, devemos utilizar STATDISK ou Minitab para introduzir as 106 temperaturas e salvá-las como um arquivo de nome BODYTEMP. Passamos então a obter um histograma, um diagrama em caixas, medidas de tendência central, medidas de variação, Q_1, Q_3, o mínimo e o máximo. Esses resultados permitem-nos descrever características importantes dos dados. Com base nessa amostra, que podemos concluir sobre a crença comum de que a temperatura média do corpo humano seja de 98,6°F? É este o resultado que esperávamos?

DOS DADOS PARA A DECISÃO

O Lixo e o Tamanho da População

Consideremos o Conjunto de Dados 1 do Apêndice B. Os dados se referem aos pesos de diferentes categorias de lixo de 62 residências, e foram coletados como parte do Garbage Project (Projeto do Lixo) na Universidade do Arizona. Há vários aspectos a considerar nesse conjunto de dados. No Capítulo 9 veremos se há alguma relação entre o tamanho da residência e a quantidade de lixo descartado, de forma que possamos predizer o tamanho da população de uma região analisando o lixo descartado. Por ora, vamos trabalhar com estatística descritiva baseada nos dados.

a. Construa um diagrama de Pareto e um gráfico em setores ilustrando os valores relativos dos pesos totais de resíduos de metal, papel, plástico, vidro, alimentos, jardinagem, tecidos e outros. (Em lugar de freqüências, utilizamos os pesos totais.) Com base nos resultados, que categorias parecem ser as maiores componentes da quantidade total de resíduos? Há alguma categoria isolada que se distinga como a maior componente?

b. Um gráfico em setores do *USA Today* mostra os resíduos de metal, papel, plástico, vidro, alimentos, jardinagem e outros com os percentuais de 14%, 38%, 18%, 2%, 4%, 11% e 13%, respectivamente. Esses percentuais se afiguram compatíveis com o Conjunto de Dados 1 do Apêndice B?

c. Determine, para cada categoria, a média e o desvio-padrão, e construa um histograma dos 62 pesos. Registre os resultados na tabela a seguir.

d. As quantidades de lixo descartado são dadas por peso. Muitas regiões têm serviço de coleta de lixo residencial feito por caminhões que comprimem o lixo, e as taxas do serviço se baseiam no peso. Sob essas condições, o volume do lixo tem importância para o problema de coleta na comunidade? Há outros fatores importantes? Quais?

e. Com base nos resultados precedentes, se fosse necessário desenvolver esforços de conservação ou reciclagem em virtude de a capacidade de coleta de resíduos em sua região estar quase esgotada, que providências tomaria?

	Metal	Papel	Plástico	Vidro	Alimentos	Jardinagem	Tecidos	Outros
Média								
Desvio-padrão								
Forma da distribuição								

ATIVIDADES EM GRUPO

1. *Atividade Extraclasse*: As Estimativas São Influenciadas por Números Âncora? No artigo "Weighing Anchors" na revista *Omni*, o autor John Rubin observou que, quando as pessoas estimam um valor, sua estimativa em geral é "ancorada" a (ou influenciada por) um número precedente, mesmo que esse número esteja totalmente desvinculado da grandeza que está sendo estimada. Para comprová-lo, pediu a diversas pessoas que estimassem rapidamente o valor de $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$. A resposta média foi 2250; mas, invertida a ordem dos números, a resposta média foi de 512. Rubin explica que, quando começamos nossos cálculos com números maiores (como $8 \times 7 \times 6$), nossas estimativas tendem a ser maiores. Observa que tanto 2250 como 512 estão muito abaixo do verdadeiro valor de 40.320. O artigo sugere que números irrelevantes podem influenciar avaliações de propriedades imóveis, de automóveis, ou estimativas da probabilidade de uma guerra nuclear.

Realize um experimento para testar essa teoria. Escolha algumas pessoas e peça-lhes que estimem rapidamente o valor de

$$8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$$

Selecione em seguida outro grupo de pessoas e peça-lhes que estimem rapidamente o valor de

$$1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8$$

Registre as estimativas, juntamente com a ordem utilizada. Planeje cuidadosamente o experimento de modo que as condições sejam uniformes e os dois grupos amostrais sejam selecionados com o mínimo possível de tendenciosidade. Não revele a teoria aos indivíduos até que eles tenham feito suas

ATIVIDADES EM GRUPO (Continuação)

estimativas. Compare os dois conjuntos de resultados amostrais com auxílio de métodos deste capítulo. Elabore um relatório datilografado que inclua os dados coletados, os métodos usados, o método de análise, quaisquer gráficos e/ ou estatísticas relevantes, e um resumo das conclusões. Inclua uma crítica das razões por que os resultados poderiam não ser corretos e indique maneiras de melhorar o experimento.

Uma variante do experimento precedente consiste em entrevistar pessoas sobre seus conhecimentos acerca da população do Quênia. Primeiro pergunte à metade das pessoas do grupo se elas acham que a população é superior ou inferior a 5 milhões; peça-lhes, em seguida, que estimem a população em um número efetivo. Pergunte à outra metade das pessoas se acham que a população é superior ou inferior a 80 milhões, e peça-lhes em seguida que estimem a população. (A população efetiva do Quênia é de 28 milhões.) Compare os dois conjuntos de resultados e identifique o efeito de "ancoragem" do número inicial mencionado aos indivíduos pesquisados.

2. *Atividade em Classe*: Em cada grupo de três ou quatro estudantes, ache o valor total das moedas em poder de cada um. Ache a média e o desvio-padrão do grupo, e permute essas estatísticas com os outros grupos. Utilizando as médias grupais como um conjunto individual de dados, calcule a média, o desvio-padrão e a forma da distribuição. Compare esses resultados com a média e o desvio-padrão achados originalmente no grupo.
3. *Atividades em Classe:* A seguir temos as idades de motociclistas mortos em acidentes de tráfego (com base em dados do Ministério dos Transportes dos EUA). Se seu objetivo é dramatizar os perigos das motos para os jovens, que recurso seria mais eficiente: histograma, diagrama de Pareto, gráfico em setores, gráfico por pontos, média, mediana, ...? Construa o gráfico e ache a estatística que melhor atende ao objetivo. É correto distorcer deliberadamente os dados se o objetivo é salvar vidas de motociclistas?

17	38	27	14	18	34	16	42	28	24	40	20	23	31
37	21	30	25	17	28	33	25	23	19	51	18	29	

entrevista

Anthony DiUglio

Analista Nuclear, Probabilistic Risk Assessment, Consolidated Edison Company of New York, Inc.

Anthony DiUglio trabalha no Probabilistic Risk Assessment (PRA) Group (Grupo de Avaliação de Risco Probabilistico) da Unidade Nº 2 (Indian Point) de geração nuclear da Consolidated Edison em Buchanan, Nova York. Em seu trabalho como Analista Nuclear, Tony estabelece probabilidades utilizadas para quantificar vários aspectos da avaliação do risco da usina. Tony é um ex-aluno do autor.

Quais são suas atribuições?

Na PRA preocupam-nos três questões básicas sobre o risco: o que pode acontecer, qual é a chance de acontecer, e quais são as conseqüências caso aconteça. Essas questões sobre o risco se aplicam ao funcionamento seguro, confiável e contínuo de nossa usina. Quando quantificamos o risco, obtemos números que são probabilidades. Se alguém sugere uma modificação no sistema de segurança da usina, analisamo-la do ponto de vista do risco. A modificação é melhor para o sistema? Afeta a operação da usina, ou coloca em risco a saúde pública e a segurança?

Como o Sr. utiliza a probabilidade e/ou a estatística?

Trata-se de recursos fundamentais. Nossa PRA exige que quantifiquemos as taxas de reparo específicas da usina para todos os componentes ligados à segurança. Ao estabelecer taxas de reparo de componentes para bombas e válvulas, recorremos a dados da indústria em geral (genéricos) e a dados específicos de nossa usina. Combinamos essas informações, sob incerteza, e chegamos a probabilidades de reparo específicas para as diversas componentes.

Como utiliza a probabilidade e/ou a estatística em outros departamentos na Indian Point?

Nosso Performance Department calcula diversos parâmetros da usina, como taxa de aquecimento, geração em megawatts, custo de geração por kilowatt etc. Esses parâmetros são todos obtidos com o auxílio da estatística. Os recursos estatísticos que utilizamos são tendência de dados, curvas normais, desvios padrão, histogramas etc. O Financial Planning utiliza amplamente a estatística ao projetar orçamentos e determinar suas restrições. Nossos previsores utilizam a teoria da probabilidade para predizer a demanda em diferentes épocas do ano (por exemplo, inverno e verão para um, três e cinco anos para a frente). Há tanta gente utilizando a estatística em seu trabalho cotidiano que a estatística é hoje um instrumento poderoso para engenheiros: planejadores, previsores, e para nós da Avaliação do Risco.

Em termos de estatística, quais seriam suas recomendações aos candidatos a emprego?

Eles devem ter um bom conhecimento de probabilidade, estatística e suas aplicações. Como PRA é ainda uma área relativamente nova, surgem freqüentemente problemas que até então não havíamos encontrado; assim, muitos deles exigem criatividade. Uma vez de posse dos intrumentos básicos, seu tempo é empregado eficientemente. Não podemos nos comunicar com eficiência a menos que utilizemos uma linguagem comum — e esta linguagem é a estatística.

Seu trabalho conseguiu convencer a opinião pública de que sua usina é segura?

A segurança é sempre nossa primeira preocupação. No início da década de 1980, houve uma série de reuniões públicas realizadas pela Nuclear Regulatory Commission (NRC) para discutir se nossa usina devia, ou não, continuar em operação. A Consolidated Edison afirmava que sua usina era segura, justificando-se a continuação das operações da mesma através do nosso PRA. Ao término daquelas reuniões, a NRC concordou com nossa posição, e continuamos a operar.

Quem foi seu melhor professor de matemática?

O professor Mario Triola.

Sua utilização da probabilidade e da estatística tem aumentado, diminuído ou permanecido constante?

Tem aumentado continuamente. Estamos muito envolvidos com os indicadores de desempenho da usina como parâmetros da eficiência operacional da usina. Com a PRA temos agora um instrumento que nos permite focalizar a atenção sobre as componentes e funções mais importantes da usina. No caso de três componentes necessitarem todas de manutenção, a PRA permite-nos identificar qual componente deve voltar ao serviço em primeiro lugar. Em engenharia, se temos diversas componentes que devem ser melhoradas, a PRA permite identificar qual delas deve ser melhorada primeiro. Podemos quantificar os efeitos e, assim, dirigir melhor nossos recursos, tornando a usina mais segura.

Rayane Andrade

Triola

3 Probabilidade

3-1 Aspectos Gerais

Identificam-se os objetivos do capítulo, abordando-se a importância da probabilidade, juntamente com seu papel nos métodos estatísticos básicos.

3-2 Fundamentos

Apresentam-se e ilustram-se as definições de probabilidade — a definição clássica e a definição como freqüência relativa. Dão-se métodos para achar probabilidades de eventos simples. Descreve-se a lei dos grandes números, definindo-se e ilustrando-se o complemento de um evento. Consideram-se as chances.

3-3 Regra da Adição

Dá-se a regra da adição como método para achar a probabilidade de ocorrência de um ou de outro evento (ou de ambos) na realização de um experimento. Definem-se os eventos mutuamente excludentes. Na aplicação da regra da adição, evita-se ou corrige-se a contagem dupla de eventos que não são mutuamente excludentes. Introduz-se a regra dos eventos complementares.

3-4 Regra da Multiplicação

Introduz-se a regra da multiplicação para achar a probabilidade de ocorrência de um evento em um experimento *e* outro evento em outro experimento. Definem-se os eventos independentes. Descreve-se a probabilidade de obter ao menos um resultado de um determinado evento. Define-se e ilustra-se a probabilidade condicional.

3-5 Probabilidades por Meio de Simulações

Freqüentemente, as probabilidades podem ser estimadas com base em simulações que retratam experimentos. Descrevem-se e ilustram-se métodos para criar simulações.

3-6 Contagem

Descrevem-se as seguintes técnicas importantes de contagem: o princípio fundamental da contagem, a regra do fatorial, a regra dos arranjos (quando todos os elementos são diferentes), a regra dos arranjos quando alguns elementos são idênticos e a regra das combinações. Utilizam-se esses processos fundamentais de contagem para determinar o número total de resultados.

Problema do Capítulo

Qual ocorrência parece ter maior chance para o leitor, ser atingido por um raio ou ganhar na loteria?

Muitos de nós agimos com base na chance da ocorrência de eventos. Alguns de nós viajamos em aviões, reconhecendo que, embora haja a chance de uma colisão com outro aparelho, essa chance é realmente mínima. Alguns aceleram seus automóveis durante uma tempestade, sabendo que podem ser atingidos por um raio, mas, novamente, a chance de tal evento é também mínima. Muitos compram bilhetes de loteria com a esperança de ganhar um prêmio, embora a chance de tal ocorrência seja muito pequena — mas sabemos realmente quão pequena? A chance de ganhar na loteria é realmente menor do que a chance de ser atingido por um raio? Neste capítulo, estudaremos a probabilidade ao determinar maneiras específicas de avaliar a chance de ocorrência de vários eventos — em particular, a chance de ser atingido por um raio e a de ganhar na loteria. Veremos então qual evento tem maior chance.

3-1 Aspectos Gerais

No Capítulo 2 apresentamos o método da *inferência estatística*, que se baseia na evidência amostral para formular inferências ou conclusões sobre toda uma população. As decisões inferenciais se baseiam em probabilidades — ou chances — de eventos. Suponha, por exemplo, que em uma verificação dos registros dos funcionários de sua faculdade fique constatado que os últimos 100 admitidos são do sexo masculino. Em uma política não-tendenciosa de admissão, a chance de os 100 primeiros admitidos serem todos homens é tão pequena, que somos levados a crer que se dá preferência aos homens. Este exemplo ilustra um importante princípio, que será a base de nosso raciocínio em vários capítulos futuros.

> **Se, sob determinada hipótese (tal como a contratação não-tendenciosa) a probabilidade de uma determinada amostra (como 100 homens contratados) é excepcionalmente pequena, concluímos que a hipótese provavelmente não é correta.**

Além de sua aplicação na metodologia estatística, a teoria da probabilidade vem adquirindo importância crescente como instrumento analítico em uma sociedade que é forçada a medir incertezas. Por exemplo, antes de ativar uma usina nuclear, devemos analisar a probabilidade de um acidente. Antes de armar um artefato nuclear, devemos analisar a probabilidade de uma detonação acidental. E antes de aumentar o limite de velocidade em nossas rodovias, devemos procurar estimar a probabilidade do aumento de acidentes fatais.

O objetivo principal deste capítulo é firmar um conhecimento sólido dos valores probabilísticos que serão utilizados em capítulos subseqüentes. Um objetivo secundário é desenvolver os conhecimentos básicos necessários para a resolução de problemas simples de probabilidade. Esses conhecimentos têm por si só grande valor como instrumentos de tomada de decisões que possibilitem melhor compreensão do nosso mundo.

3-2 Fundamentos

Ao lidarmos com problemas de probabilidade, vamos encontrar experimentos, eventos e a coleção de todos os resultados possíveis.

DEFINIÇÕES

Um **experimento** é qualquer processo que permite ao pesquisador fazer observações.

Um evento é uma coleção de resultados de um experimento.

Um evento **simples** é um resultado, ou um evento, que não comporta mais qualquer decomposição.

O **espaço amostral** de um experimento consiste em todos os eventos simples possíveis; ou seja, o espaço amostral consiste em todos os resultados que não comportam mais qualquer decomposição.

Exemplo: o arremesso de um dado é um *experimento*, e o resultado 3 é um *evento*. O resultado 3 é um *evento simples* porque não pode ser decomposto; e o *espaço amostral* consiste nesses eventos simples: 1, 2, 3, 4, 5, 6. Outro exemplo: o arremesso de um par de dados é um experimento, o resultado 7 é um evento, mas 7 não é um evento simples porque pode ser decomposto em eventos mais simples, como 3-4 e 6-1. Na jogada de um par de dados, o espaço amostral consiste em 36 eventos simples: 1-1, 1-2, ..., 6-6.

Embora não haja concordância universal sobre como definir a probabilidade de um evento, duas definições são correntes.

Apresentamos primeiro a notação básica, e passamos então às definições de probabilidade.

Notação para Probabilidades

P denota uma probabilidade. A, B, C denotam eventos específicos. $P(A)$ denota a probabilidade de ocorrência do evento A.

Regra 1: Aproximação da Probabilidade pela Freqüência Relativa

Realize (ou observe) um experimento um grande número de vezes e conte quantas vezes o evento A ocorre efetivamente. Então $P(A)$ é *estimada* como segue:

$$P(A) = \frac{\text{número de ocorrências de } A}{\text{número de repetições do experimento}}$$

Regra 2: Definição Clássica da Probabilidade

Suponha que um experimento tenha n eventos simples diferentes, cada um dos quais com *a mesma chance* de ocorrer. Se o evento A pode ocorrer em s dentre as n maneiras, então

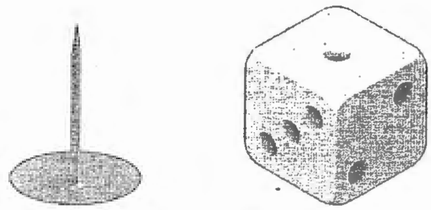
$$P(A) = \frac{\text{número de maneiras como } A \text{ pode ocorrer}}{\text{número de eventos simples diferentes}} = \frac{s}{n}$$

Qual é a Probabilidade?

Como interpretar termos como *provável*, *improvável* ou *extremamente improvável*? O Departamento de Aeronáutica dos EUA. (FAA) dá a seguinte interpretação. *Provável*: Uma probabilidade de 0,00001 ou mais para cada hora de vôo. Espera-se a ocorrência de tais eventos várias vezes durante a vida operacional de cada aeronave. *Improvável*: Uma probabilidade da ordem de 0,00001 ou menos. Tais eventos não são esperados no decorrer da vida operacional de uma aeronave de determinado tipo, mas podem ocorrer durante a vida operacional de todos os aviões daquele tipo. *Extremamente improvável*: Uma probabilidade da ordem de 0,000000001 ou menos. Tais eventos são tão improváveis que podem ser considerados como se jamais ocorressem.

É importante notar que *a definição clássica exige que os resultados tenham todos a mesma chance*. Se os resultados não têm todos a mesma chance, devemos apelar para a estimativa pela freqüência relativa. A Figura 3-1 ilustra essa importante distinção.

Ao calcular probabilidades pelo método da freqüência relativa (Regra 1), obtemos uma *aproximação* em lugar de um valor exato. À medida que o número de observações aumenta, as aproximações tendem a ficar cada vez mais próximas da probabilidade efetiva. Essa propriedade é enunciada como um teorema comumente conhecido como a *lei dos grandes números*.



(a) (b)

Fig. 3-1 Comparação entre freqüência relativa e abordagem clássica.

(a) Abordagem pela freqüência relativa (Regra 1): Ao procurar determinar P(tachinha cair com a ponta para cima) devemos repetir o experimento (jogar a tachinha) muitos vezes e determinar a razão do número de vezes que a ponta fica para cima para o número total de jogadas. Essa razão é a nossa estimativa da probabilidade.

(b) Abordagem clássica (Regra 2): Ao procurar determinar $P(2)$ com um dado equilibrado, cada uma das faces tem a mesma chance de aparecer,

$$P(2) = \frac{\text{número de possibilidades de ocorrência de 2}}{\text{número total de eventos simples}} = \frac{1}{6}$$

Lei dos Grandes Números

Se se repete um experimento um grande número de vezes, a probabilidade pela freqüência relativa (Regra 1) de um evento tende para a probabilidade teórica.

A lei dos grandes números afirma que a aproximação pela freqüência relativa (Regra 1) tende a melhorar quando o número de observações aumenta. Essa lei reflete uma noção bastante simples apoiada pelo senso comum: Uma estimativa probabilística baseada apenas em umas poucas observações pode apresentar grande divergência, mas com um número crescente de provas a estimativa tende a ser cada vez mais precisa. Por exemplo, se fazemos uma pesquisa entrevistando apenas algumas pessoas, os resultados podem acusar grande erro; mas se entrevistamos milhares de pessoas *selecionadas aleatoriamente*, os resultados amostrais estarão muito mais próximos dos verdadeiros valores populacionais.

A Figura 3-2 ilustra a lei dos grandes números através de resultados simulados em computador. Note que, à medida que aumenta o número de nascimentos, a proporção de meninas tende para 0,5.

Se atentarmos para as Regras 1 e 2, pode parecer que deveríamos utilizar sempre a Regra 2 quando um experimento tem resultados igualmente prováveis. Acontece freqüentemente, entretanto, que tais experimentos são tão complicados que a abordagem clássica (Regra 2) perde seu aspecto prático. Em lugar disso, podemos mais facilmente obter estimativas das probabilidades desejadas apelando para o processo da freqüência relativa (Regra 1). Em tais casos, as simulações costumam auxiliar. (Uma *simulação* de um experimento é um processo que se comporta da mesma maneira que o próprio experimento, produzindo assim resultados análogos.) Por exemplo, é muito mais fácil utilizar a Regra 1 para estimar a probabilidade de ganhar em um jogo de paciência — isto é, jogar um grande número de vezes (ou fazer uma simulação em um computador) do que efetuar os cálculos extremamente complexos exigidos pela Regra 2.

Os exemplos que seguem servem para ilustrar a aplicação das Regras 1 e 2. Em alguns deles, empregamos o termo *aleatório*.

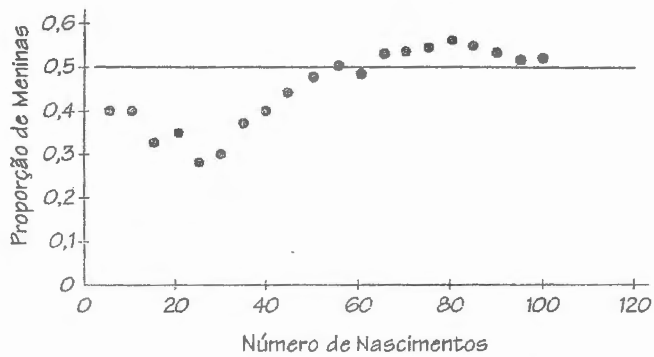


Fig. 3-2 Ilustração da Lei dos Grandes Números.

Recorde estas definições da Seção 1-4: Em uma **amostra aleatória de um elemento** de uma população, todos os elementos da mesma têm igual chance de ser escolhidos; uma amostra de n elementos é uma **amostra aleatória** (ou uma amostra aleatória simples) se é escolhida de tal maneira que toda amostra possível de n elementos da população tem a mesma chance de ser escolhida. O conceito geral de aleatoriedade é extremamente importante em estatística. Ao fazermos inferências baseadas em amostras, devemos ter um processo de amostragem que seja representativo, imparcial e não-tendencioso. Se uma amostra não é selecionada cuidadosamente, pode ser totalmente inútil.

EXEMPLO Determine a probabilidade de uma pessoa escolhida aleatoriamente ser atingida por um raio este ano.

SOLUÇÃO O espaço amostral consiste nestes dois eventos simples: A pessoa escolhida é atingida por um raio, ou não é. Como estes eventos simples não são igualmente prováveis, devemos apelar para uma aproximação por freqüência relativa, conforme a Regra 1. Não é prático realizar experimentos, mas podemos pesquisar eventos passados. Em um ano recente, 371 pessoas foram atingidas por um raio nos EUA. Em uma população de cerca de 260 milhões, a probabilidade de ser atingida por um raio em um ano é estimada em

$$\frac{371}{260.000.000} \approx \frac{1}{701.000}$$

O Vocabulário de Shakespeare

De acordo com Bradley Efron e Ronald Thisted, as obras de Shakespeare contêm 31.534 palavras diferentes. Com auxílio da

teoria das probabilidades, concluíram que Shakespeare conhecia ao menos outras 35.000 palavras que não empregou em suas obras. A estimativa do tamanho de uma população é um problema importante, encontrado freqüentemente em estudos de ecologia, mas o resultado apresentado aqui é outra aplicação interessante. [Veja "Estimating the Number of Unseen Species: How Many Words Did Shakespeare Know?" (Estimativa do Número de Espécies Não Vistas: Quantas palavras Shakespeare conhecia?"), *in Biometrika*, Vol. 63, Nº 3].

EXEMPLO Em um teste ACT (American College Test = Teste para Faculdade Americana) ou SAT (Scholastic Aptitude Test), uma questão típica de múltipla escolha tem 5 respostas possíveis. Respondendo à questão aleatoriamente, qual é a probabilidade de sua resposta estar errada?

SOLUÇÃO Há 5 resultados ou respostas possíveis e 4 maneiras de responder incorretamente a questão. A aleatoriedade implica que os resultados do espaço amostral são igualmente possíveis; pela abordagem clássica (Regra 2) obtemos

$$P\text{ (resposta errada)} = \frac{4}{5} = 0,8$$

Em problemas básicos de probabilidade do tipo que estamos considerando, é muito importante examinar cuidadosamente as informações de que dispomos e identificar corretamente o número total de resultados possíveis. Em alguns casos, temos esse número diretamente, mas em outros casos devemos manipular as informações para obtê-lo. No exemplo precedente, temos a informação de que o número total de resultados é 5, mas o exemplo que segue exige que calculemos o número total de resultados possíveis.

EXEMPLO A companhia de seguros American Casualty Company estudou as causas de morte por acidente doméstico e compilou um arquivo que consistia em 160 mortes causadas por quedas, 120 mortes causadas por envenenamento e 70 causadas por fogo e queimaduras. Selecionado aleatoriamente um desses casos, qual é a probabilidade de que a morte tenha sido causada por envenenamento?

SOLUÇÃO O número total de mortes por acidente é 160 + 120 + 70 = 350. Com a seleção aleatória, as 350 mortes são igualmente prováveis, e a Regra 2 se aplica como segue:

$$P\text{ (envenenamento)} = \frac{\text{número de mortes por envenenamento}}{\text{número total de mortes}} = \frac{120}{350} = 0,343$$

Há uma probabilidade de 0,343 de que, selecionada aleatoriamente uma das mortes, ela tenha sido causada por envenenamento.

1º 2º 3º

H-H-H

H-H-M

exatamente 2 Homens — H-M-H

H-M-M

M-H-H

M-H-M

M-M-H

M-M-M

EXEMPLO Determine a probabilidade de que um casal com três filhos tenha exatamente 2 meninos. Suponha que as probabilidades de menino e menina sejam as mesmas, e que o sexo de uma criança não seja influenciado pelo sexo de qualquer outra.

SOLUÇÃO Inicialmente relacionamos o espaço amostral que identifica os 8 resultados. Como esses resultados são igualmente prováveis, aplicamos a Regra 2. Dos 8 diferentes resultados possíveis, 3 correspondem a exatamente 2 meninos; assim,

$$P\text{ (2 meninos em 3 nascimentos)} = \frac{3}{8} = 0,375$$

Há uma probabilidade de 0,375 de que um casal com 3 filhos tenha exatamente 2 meninos.

EXEMPLO Ao escolher entre diversos fornecedores de computadores, um comprador deseja saber a probabilidade de um computador pessoal falhar durante os dois primeiros anos. Qual é essa probabilidade?

SOLUÇÃO Há apenas dois resultados: Um computador pessoal falha durante os dois primeiros anos, ou não falha. Como esses dois resultados não são igualmente prováveis, devemos recorrer à aproximação por freqüência relativa. Isso exige que, de algum modo, observemos um grande número de computadores pessoais. Uma pesquisa do *PC World* feita junto a 4000 possuidores de computadores pessoais revelou que 992 dos computadores falharam durante os dois primeiros anos. Com base em tais resultados, *estimamos* a probabilidade em 992/4000, ou 0,248.

EXEMPLO Selecionado um ano aleatoriamente, determine a probabilidade de o Dia de Ação de Graças cair (a) numa quarta-feira, (b) numa quinta-feira.

SOLUÇÃO

a. O Dia de Ação de Graças sempre cai na quarta quinta-feira do mês de novembro. É, pois, impossível aquele dia cair em uma quarta-feira. Quando um evento é impossível, dizemos que sua probabilidade é 0.

b. É certo que o Dia de Ação de Graças cairá em uma quinta-feira. Quando a ocorrência de um evento é certa, dizemos que sua probabilidade é 1.

Como qualquer evento imaginável é certo, impossível, ou se situa entre esses dois extremos, é razoável concluirmos que a probabilidade matemática de qualquer evento é 0,1 ou um número entre 0 e 1 (veja Figura 3-3).

- **A probabilidade de um evento impossível é 0.**
- **A probabilidade de um evento cuja ocorrência é certa é igual a 1.**
- **$0 \leq P(A) \leq 1$ para qualquer evento A.**

Na Figura 3-3, apresentamos à esquerda a escala de 0 a 1 e, à direita, as expressões mais familiares e comuns da verossimilhança.

Eventos Complementares

Eventualmente, devemos determinar a probabilidade de um evento A *não* ocorrer.

DEFINIÇÃO
O **complemento** de um evento A, denotado por $\overline{A}$, consiste em todos os resultados em que o evento A *não* ocorre.

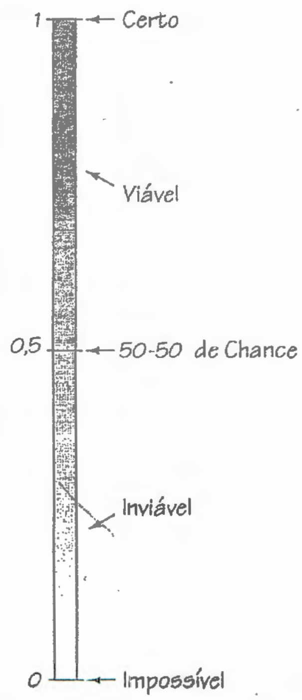


Fig. 3-3 Valores possíveis para probabilidades.

EXEMPLO A Nike Corporation deseja testar um novo material a ser usado na fabricação de tênis. Um grupo de teste consiste em 20 homens e 30 mulheres. Escolhida aleatoriamente uma pessoa desse grupo de teste, determine a probabilidade de *não* ser homem.

SOLUÇÃO Inicialmente, observemos que o espaço amostral total consiste em 50 pessoas. Em segundo lugar, como 20 dessas pessoas são homens, decorre que 30 são mulheres e, assim,

$$\begin{aligned} P(\text{não escolher um homem}) &= P(\overline{\text{homem}}) \\ &= P(\text{mulher}) \\ &= \frac{30}{50} = 0,6 \end{aligned}$$

Embora seja difícil estabelecer uma regra universal para o arredondamento de probabilidades, a orientação que segue se aplica à maioria dos problemas deste livro.

Arredondamento de Probabilidades

Ao expressarmos o valor de uma probabilidade, devemos dar a fração ordinária ou a expressão decimal *exata*, ou arredondar o resultado final para três algarismos significativos. (*Sug.*: Quando uma probabilidade não é uma fração simples como 2/3 ou 5/9, devemos expressá-la na forma decimal.)

Todos os algarismos de um número são significativos, menos os zeros incluídos para a localização correta da vírgula.

Coincidência?

John Adams e Thomas Jefferson (segundo e terceiro presidentes dos EUA) morreram ambos no dia 4 de julho de 1826. O Presidente Lincoln foi assassinado no Teatro Ford; o Presidente Kennedy foi assassinado em um carro Lincoln fabricado pela Ford Motor Company. Lincoln e Kennedy foram ambos sucedidos por vice-presidentes com o nome Johnson. Quatro anos antes do naufrágio do *Titanic*, uma novela descrevia o afundamento do *Titan*, um navio que se chocara com um iceberg; cf. *The Wreck of the Titanic Foretold?* (O Naufrágio do *Titanic* foi Previsto?), por Martin Gardner. Gardner afirma: "Na maior parte dos casos de coincidência surpreendente, é impossível fazer sequer uma estimativa grosseira de suas probabilidades."

EXEMPLOS

- A probabilidade de 0,0000128506 tem seis algarismos significativos (128506), e pode ser arredondada para três algarismos significativos como 0,0000129.
- A probabilidade 1/3 pode ser mantida em forma de fração ou arredondada em forma decimal para 0,333, mas não para 0,3.
- A probabilidade de "caras" na jogada de uma moeda pode expressar-se como 1/2 ou 0,5; como 0,5 é exata, não precisamos expressá-la como 0,500.
- A fração 7659/32785 é exata, mas seu valor não é óbvio; devemos, assim, expressá-la em forma decimal como 0,234.

Um conceito importante desta seção é a expressão matemática da probabilidade como um número entre 0 e 1. Esse tipo de expressão é fundamental em processos estatísticos; utilizá-lo-emos em todo o restante deste livro. Um resultado típico de computador, por exemplo, pode incluir uma expressão "valor-*P*" (*P*-value) como "significância inferior a 0,001." O significado dos valores-*P* será discutido mais adiante; mas trata-se essencialmente de probabilidades do tipo estudado nesta seção. Por ora, devemos entender que uma probabilidade de 0,001, equivalente a 1/1000, corresponde a um evento tão raro, que ocorre, em média, apenas uma vez em mil provas.

Probabilidades Subjetivas

Nesta seção apresentamos a abordagem pela freqüência relativa e a abordagem clássica como dois métodos formais para determinar probabilidades de eventos; entretanto, outra abordagem consiste simplesmente em conjecturar ("palpitar") ou estimar uma probabilidade. Essa técnica de "palpitar" já deve ser familiar aos que às vezes não estão tão preparados para um exame como deviam, mas é usada também por profissionais que estabelecem as chances para um cassino. As probabilidades subjetivas são usadas também pelas companhias de seguros, que estimam probabilidades para circunstâncias especiais, como a probabilidade de acidente com uma estrela do *rock*, que faz com que um *tour* seja cancelado. Uma probabilidade estimada com base no conhecimento de circunstâncias relevantes é chamada **probabilidade subjetiva**. Por exemplo, um apostador de Las Vegas estima em 0,05 a probabilidade de o New York Giants ganhar a Supercopa no próximo ano. Tal estimativa se baseia no conhecimento de fatores relevantes, como a capacidade do treinador e dos jogadores.

Chances

As expressões de verossimilhança são freqüentemente dadas em forma de *chances*, como 50:1 (ou "50 para 1"). Uma desvantagem séria das chances é que elas tornam muitos cálculos extremamente difíceis. Por isso, os estatísticos, matemáticos e cientistas preferem utilizar probabilidades. A vantagem das chances é que facilitam o manuseio com transferências de dinheiro associadas ao jogo, tendendo assim a ser usadas em cassinos, loterias e corridas de cavalos. Em primeiro lugar, devemos saber que a verossimilhança de um evento pode expressar-se em termos de chances contra ou a favor do evento.

Probabilidades Subjetivas no Turfe

Os pesquisadores estudaram a capacidade de os apostadores do turfe estabelecerem probabilidades subjetivas realistas. [Cf. "Racetrack Betting: Do Bettors Understand the Odds?" (Apostas no Turfe: Os Apostadores Compreendem as Chances?), por Brown, d'Amato e Gertner, *in* revista *Chance*, Vol. 7, N.º 3.] Após analisar os resultados de 4400 corridas, concluíram que embora os apostadores sobreestimem ligeiramente as probabilidades de ganho dos "azarões" e subestimem ligeiramente as probabilidades de ganho dos favoritos, seu desempenho geral é bastante bom. As probabilidades subjetivas foram calculadas com base nos prêmios, que, por sua vez, se baseiam nas quantias apostadas, e as probabilidades efetivas foram calculadas pelos resultados reais dos páreos.

7) n = 500 x = 430 $= \frac{430}{500} = 0{,}86 \to 86\%$, sim

8) n = 1066 x = 181 $\frac{181}{1066} = 0{,}169 \to 16\%$ / não

9) C → 1162 NC → 2468 3630 $P(C) = \frac{1162}{3630} = 0{,}320 \to 32\%$

DEFINIÇÃO

A **chance contra** a ocorrência do evento A é a razão $P(\bar{A})/P(A)$, comumente expressa na forma $a{:}b$ (ou "a para b"), com a e b inteiros, primos entre si.

A **chance a favor** do evento A é o inverso da chance contra aquele evento. Se a chance contra A é $a{:}b$, então a chance a favor do evento é $b{:}a$.

Como exemplo, se $P(A) = 2/5$, então

$$\text{chance contra } A = \frac{P(\bar{A})}{P(A)} = \frac{3/5}{2/5} = \frac{3}{2}$$

Escrevemos: 3:2, ou "3 para 2". Como a chance contra A é 3:2, a chance a favor de A é 2:3. Veja o Exercício 31 para a conversão de chances em probabilidades.

Nas apostas, a chance contra um evento representa a razão do ganho líquido para a quantia apostada.

Chance contra um evento A =
= (ganho líquido):(quantia apostada)

Suponhamos que uma aposta pague 50:1. Se a chance não é especificada como sendo a favor ou contra, trata-se provavelmente de chance contra a ocorrência do evento. Se, por um pequeno milagre, o leitor ganhasse esta aposta de 50:1, teria um lucro de \$50 para cada \$1 apostado. Assim é que, apostando \$2, o lucro líquido do apostador seria \$100; ele receberia um total de \$102, que inclui o ganho de \$100 e os \$2 apostados.

Apostas

Na loteria estadual típica (dos EUA), a "casa" leva uma vantagem de 65% a 70%, porque apenas 35% ou 40% do total apostado são devolvidos como prêmios. Nas corridas de cavalo, a vantagem da casa fica em torno de 15%. No cassino, a vantagem da casa é de 5,26% na roleta, 5,9% no vinte-e-um, 1,4% nos dados e 3% a 22% nas máquinas caça-níqueis. Alguns jogadores profissionais podem ganhar sistematicamente no vinte-e-um, utilizando técnicas complicadas de contagem de cartas. Eles sabem quando um baralho tem um número desproporcional de cartas de valor alto, e é aí então que fazem grandes apostas. Muitos cassinos reagem expulsando os contadores de cartas ou baralhando as cartas com maior freqüência.

3-2 Exercícios A: Habilidades e Conceitos Básicos

1. Quais dos valores abaixo *não podem* ser probabilidades?

$$0,\ 0{,}0001,\ -0{,}2,\ 3/2,\ 2/3,\ \sqrt{2},\ \sqrt{0{,}2}$$

2. a. Quanto é $P(A)$, se A é o evento "Fevereiro tem 30 dias este ano"? 0
   b. Quanto é $P(A)$, se A é o evento "Novembro tem 30 dias este ano"? 1
   c. Um espaço amostral consiste em 500 eventos separados, igualmente prováveis. Qual a probabilidade de cada um? 1/500
   d. Em um exame de admissão, cada questão tem 5 respostas possíveis. Respondendo aleatoriamente (por "palpite") a primeira questão, qual a probabilidade de acertar? 1/5
3. Determine a probabilidade do resultado "cara" ao jogar uma moeda. 1/2
4. Vimos nesta seção que o experimento que consiste em jogar um par de dados comporta 36 eventos simples que formam o espaço amostral: 1-1, 1-2, ..., 6-6. Determine a probabilidade de obter o total 4 no arremesso de um par de dados.

1-3
3-1
2-2
$\frac{3}{36} = 0{,}08$

5. Com base nos resultados amostrais do Conjunto de Dados 11 do Apêndice B, estime a probabilidade de um bombom M&M, escolhido aleatoriamente, ser vermelho.
6. Com base nos resultados amostrais do Conjunto de Dados 8 do Apêndice B, estime a probabilidade de um estudante de estatística selecionado aleatoriamente ter ao menos um cartão de crédito.
7. Um estudo de 500 vôos da American Airlines selecionados aleatoriamente mostrou que 430 chegaram no horário (com base em dados do Ministério dos Transportes). Qual é a probabilidade estimada de um vôo da American Airlines chegar no horário? Acha que é um resultado satisfatório?
8. A Kelly-Lynne Advertising Company está cogitando lançar uma campanha por computador junto aos jovens de 11 a 19 anos. Em uma pesquisa com 1066 desses jovens, 181 tinham um serviço de computador *on-line* em sua residência. Selecionado aleatoriamente um desses jovens, estime a probabilidade de ele ou ela ter acesso ao serviço *on-line* em sua residência. Aconselharia essa companhia a promover uma campanha publicitária por computador?
9. Em uma pesquisa entre estudantes de uma faculdade, 1162 afirmaram que "colavam" nos exames, enquanto 2468 afirmaram não "colar" [com base em dados do Josephson Institute of Ethics (Instituto Josephson de Ética)]. Selecionado aleatoriamente um desses estudantes, determine a probabilidade de ele ou ela ter "colado" em um exame.
10. Em um estudo efetuado em americanos de mais de 65 anos de idade, verificou-se que 255 tinham o mal de Alzheimer, enquanto 2302 não o tinham (com base em dados da Alzheimer Association). Escolhido aleatoriamente um americano de mais de 65 anos, qual a probabilidade estimada de ele (ou ela) ter o mal? Com base nessa probabilidade, acha que o mal de Alzheimer constitui uma preocupação para as pessoas com mais de 65 anos?
11. Em um estudo feito com doadores de sangue, 225 foram classificados como grupo O e 275 obtiveram classificação não-O [com base em dados do Greater New York Blood Program (Grande Programa de Sangue do Estado de Nova York)]. Qual a probabilidade estimada de uma pessoa ter sangue do grupo O?
12. Em uma pesquisa Nielsen em 3857 lares, constatou-se que 463 tinham sua televisão ligada no canal CBS na segunda-feira à noite entre 10:00 e 10:30h. Selecionada aleatoriamente uma casa, estime a probabilidade de ela estar ligada no CBS naquele instante.
13. a. Selecionada uma pessoa aleatoriamente, determine a probabilidade de ele ou ela fazer aniversário em 18 de outubro, que é Dia Nacional da Estatística no Japão. Ignore os anos bissextos.
    b. Determine a probabilidade de o aniversário de uma pessoa escolhida aleatoriamente cair em Novembro. Ignore os anos bissextos.
14. Em um estudo de reconhecimento de marca, 831 consumidores conheciam a Sopa Campbell, e 18 não a conheciam (com base em dados da Total Research Corporation). Com esses resultados, estime a probabilidade de um consumidor aleatório reconhecer a Sopa Campbell. Compare essa probabilidade com os valores típicos relativos a outras marcas.
15. Em uma pesquisa feita pela Bruskin-Goldring Research, perguntou-se aos entrevistados como deveria ser utilizado um bolo de frutas. Cento e trinta e dois responderam que deveria servir para calço de porta, e outros 880 indicaram outros usos, inclusive alimento de passarinho, aterro, e presente. Selecionado aleatoriamente um desses entrevistados, qual a probabilidade de obter alguém que utilize o bolo como calço de porta?*
16. O U.S.General Accounting Office (Departamento Geral de Contabilidade dos EUA) testou recentemente a IRS (Internal Revenue Service = Serviço de Receita Interna) quanto à correção das respostas a perguntas dos contribuintes. Em 1733 casos, a IRS se revelou correta 1107 vezes. Com esses resultados, estime a probabilidade de um

*N. do T.: O bolo de frutas é considerado muito pesado pelos americanos.

11) O → 225
no → 275
500
$P(O) = \frac{225}{500} = 0{,}45 \to 45\%$

consulta de um contribuinte aleatório ser respondida corretamente. Com base no resultado, diria que a IRS tem um bom desempenho respondendo corretamente consultas de contribuintes?

17. Dentre 400 motoristas selecionados aleatoriamente na faixa etária 20-24, 136 estiveram envolvidos em um acidente de carro no ano anterior (com base em dados do Conselho de Segurança Nacional — National Safety Council). Selecionado aleatoriamente um motorista naquela faixa etária, qual a probabilidade aproximada de ele (ou ela) se envolver em um acidente de carro no próximo ano? O valor resultante chega a constituir preocupação para os motoristas na faixa etária 20-24?

18. Os dados do Departamento de Estatística Judiciária (dos EUA) revelaram que, em uma amostra representativa de ladrões condenados, 76.000 foram para a cadeia, 25.000 cumpriram pena em liberdade e 2.000 receberam outras sentenças. Com esses resultados, estime a probabilidade de um ladrão condenado ir para a cadeia. Acha que o resultado é suficientemente elevado para reduzir os furtos?

19. Quando o antialérgico Seldane foi testado clinicamente, 70 pessoas experimentaram sonolência e 711 não (com base em dados da Merrell Dow Pharmaceuticals, Inc.). Com essa amostra, estime a probabilidade de um usuário de Seldane experimentar sonolência. Acha que a sonolência é um fator a ser levado em conta pelos usuários de Seldane?

20. De acordo com o Ministério dos Transportes dos EUA, a American Airlines transportou 59.377.306 passageiros no último ano. Nesse mesmo ano, 82.796 passageiros foram deliberadamente impedidos de embarcar, enquanto outros 1.664 passageiros foram involuntariamente impedidos. Determine a probabilidade de um passageiro selecionado aleatoriamente ser involuntariamente impedido de embarcar. Acha que uma pessoa deve se preocupar com a possibilidade de ser involuntariamente impedida de embarcar em sua viagem de férias?

21. A MasterCard International efetuou um estudo de fraudes em cartões de crédito; os resultados estão consubstanciados na tabela a seguir.

Tipo de Fraude	Número
Cartão roubado	243
Cartão falsificado	85
Pedido por correio/telefone	52
Outros	46

Selecionado aleatoriamente um caso de fraude nos casos resumidos na tabela, qual a probabilidade de a fraude resultar de um cartão falsificado?

22. Uma pesquisa Gallup originou os dados amostrais da tabela a seguir.

Escovadas por dia	Número
1	228
2	672
3	240

Selecionado aleatoriamente um dos entrevistados, qual a probabilidade de obter alguém que escove os dentes três vezes por dia, conforme recomendam os dentistas?

23. Um casal planeja ter 2 filhos.
 a. Relacione os diferentes resultados, de acordo com o sexo de cada criança. Suponha que esses resultados sejam igualmente prováveis.
 b. Determine a probabilidade de o casal ter 2 meninas.
 c. Determine a probabilidade de exatamente uma criança de cada sexo.

24. Um casal planeja ter 4 filhos.
 a. Relacione os 16 resultados distintos possíveis de acordo com o sexo de cada criança. Suponha que esses resultados sejam igualmente prováveis.
 b. Determine a probabilidade de serem todos meninas.
 c. Determine a probabilidade de haver ao menos uma criança de cada sexo.
 d. Determine a probabilidade de exatamente 2 crianças de cada sexo.

25. Em um teste com 3 questões do tipo verdadeiro/falso, um estudante que não está preparado deve responder cada uma aleatoriamente (por "palpite").
 a. Relacione os diferentes resultados possíveis.
 b. Qual é a probabilidade de responder corretamente todas as três questões?
 c. Qual é a probabilidade de "palpitar" incorretamente todas as três questões?
 d. Qual é a probabilidade de passar no teste "palpitando" corretamente ao menos 2 questões?

26. Ambos os pais têm o par de genes castanho/azul da cor dos olhos, e cada um deles contribui com um gene para um filho. Suponha que se o filho tem ao menos um gene castanho, essa cor dominará e os olhos serão castanhos. (Na realidade, a determinação da cor dos olhos é algo mais complexa.)
 a. Relacione os diferentes resultados possíveis, supondo-os igualmente prováveis.
 b. Qual a probabilidade de um filho desses pais ter o par de genes azul/azul?
 c. Qual a probabilidade de o filho ter olhos castanhos?

27. Determine a chance contra uma resposta correta em uma questão de múltipla escolha com cinco respostas possíveis.

28. Determine a chance contra a escolha aleatória de um canhoto, sabendo que 10% das pessoas são canhotas.

29. a. A probabilidade de um 7 em uma roleta é 1/38. Determine a chance contra 7.
 b. Apostando \$2 no número 7 na roleta e ganhando, o cassino lhe pagará \$72, que inclui os \$2 da aposta. Identifique primeiro o ganho líquido, e em seguida determine a chance usada para determinar o prêmio.
 c. Como explica a discrepância entre as chances na parte (a) e na parte (b)?

30. a. No jogo de dados em um cassino, pode-se apostar em que a próxima jogada de dois dados dê a soma 2. A probabilidade de obter a soma 2 é 1/36. Determine a chance contra o aparecimento da soma 2.
 b. Apostando \$5 no resultado 2 da próxima jogada dos dois dados, o apostador receberá \$155 (incluindo os \$5 da aposta) se ganhar. Identifique primeiro o ganho líquido e determine em seguida a chance usada para determinar o resultado.
 c. Como explica a discrepância entre as chances na parte (a) e na parte(b)?

3-2 Exercícios B: Além do Básico

$P(A) = \frac{3}{10+3} = \frac{3}{13}$

31. Se a chance contra o evento A é $a{:}b$, então $p(A) = b/(a + b)$. Determine a probabilidade de Horse Cents ganhar a próxima corrida, dado que a chance contra é de 10:3.

32. A chance contra Lazy Lady ganhar a próxima corrida é de 9:2. Determine a probabilidade de Lazy Lady ganhar o próximo páreo. (Veja Exercício 31.)

33. O gráfico ramo-e-folhas a seguir resume o tempo (em horas) que os gerentes gastam em serviços burocráticos em um dia (segundo dados da Adia Personnel Services). Com base nessa amostra, estime a probabilidade de um gerente selecionado aleatoriamente gastar mais de 2,0 horas por dia em serviços burocráticos.

0.	00
1.	0578
2.	00113449
3.	347
4.	445

34. Após coletar os escores de QI de centenas de indivíduos, constrói-se um diagrama em caixa (*boxplot*) com este resumo de 5 núme-

res: 82, 91, 100, 109, 118. Escolhido aleatoriamente um dos indivíduos, determine a probabilidade de seu QI ser superior a 109.

35. Na parte (a) do Exercício 13, foram ignorados os anos bissextos na determinação da probabilidade de um indivíduo selecionado aleatoriamente ter seu aniversário no dia 18 de outubro.

a. Recalcule essa probabilidade, sabendo que um ano bissexto ocorre a cada quatro anos. (Expresse sua resposta como uma fração exata.)

b. Os anos bissextos ocorrem em anos divisíveis por 4, com exceção de 3 de cada 4 anos centenários (anos terminados em 00). Os anos 1700, 1800 e 1900 não foram bissextos, mas 2000 será bissexto. Determine a probabilidade exata para este caso e expresse-a como uma fração.

36. a. Se duas moscas pousam em uma laranja, determine a probabilidade de elas pousarem no mesmo hemisfério.

b. Escolhem-se aleatoriamente dois pontos em uma vara retilínea. A vara é então quebrada nesses dois pontos. Determine a probabilidade de os 3 pedaços resultantes formarem um triângulo. (É um problema difícil.)

Probabilidades Que Desafiam a Intuição

Em certos casos, nossas estimativas subjetivas de probabilidades diferem acentuadamente das probabilidades efetivas. Eis um exemplo clássico: Se o leitor respira profundamente, há mais de 99% de chance de inalar uma molécula que tenha sido exalada no último suspiro de César. Nesse mesmo contexto mórbido e intuitivo, se a taça de cicuta fatal que matou Sócrates continha água em sua maior parte, então o próximo copo de água que o leitor beber provavelmente conterá uma daquelas mesmas moléculas. Eis outro exemplo menos mórbido, que pode ser verificado em aula: Em uma turma de 25 alunos, há uma chance superior a 50% de dois alunos terem a mesma data de aniversário.

3-3 Regra da Adição

O principal objetivo desta seção é introduzir a regra da adição para achar $P(A$ ou $B)$, a probabilidade de ocorrência do evento A, ou do evento B, ou de ambos, como resultado de um experimento. A palavra-chave aqui é a conjunção *ou*. (É o *ou* inclusivo, que significa um, ou outro, ou ambos, e que será utilizado em todo este texto. Afora o Exercício 27, não consideraremos o *ou* exclusivo, que significa um, ou outro, mas *não* ambos.) Na seção precedente consideramos eventos classificados como simples, porque envolviam apenas um resultado, geralmente denotado por A. Em muitas situações reais, entretanto, temos *eventos compostos*, tais como a escolha aleatória de um consumidor que é mulher ou que tem menos de 40 anos de idade.

DEFINIÇÃO

Um **evento composto** é qualquer evento que combina dois ou mais eventos simples.

Notação para a Regra da Adição

$P(A$ ou $B)$ = P(ocorrência de A, ou de B, ou de ambos)

Comecemos com um exemplo simples. Os indivíduos a serem pesquisados costumam ser escolhidos utilizando-se computadores para selecionar aleatoriamente os dois últimos algarismos de telefones. Se escolhemos aleatoriamente um dos dez algarismos 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 como último algarismo de um número de telefone, qual é a probabilidade de o número escolhido ser 0 ou 1? Aplicando a Regra 2 da Seção 3-2, vemos que o evento "obter 0 ou 1" pode ocorrer de duas maneiras diferentes, e o número total de resultados possíveis é 10; logo,

$$P(0 \text{ ou } 1) = \frac{2}{10} = 0{,}2$$

Esse exemplo parece sugerir uma regra geral, pela qual bastaria adicionar os números de resultados correspondentes a cada um dos eventos em questão. Antes de firmar essa regra, entretanto, consideremos outro exemplo — uma escolha aleatória de números do conjunto $\{0,1,2, ..., 9\}$, mas dessa vez vamos achar a probabilidade de obter um número ímpar ou um número superior a 6. Notemos que, dentre os 10 resultados possíveis, 5 são ímpares (1,3,5,7,9) e 3 são superiores a 6 (7,8,9). Ao contar o número de resultados ímpares ou superiores a 6, devemos ter cuidado para não contar um número mais de uma vez. Há, na realidade, seis resultados separados que são ímpares ou superiores a 6: 1,3,5,7,8,9. A probabilidade correta é, pois,

$$P(\text{ímpar ou superior a } 6) = \frac{6}{10} = 0{,}6$$

Eis a chave da questão: Ao determinar a probabilidade da ocorrência do evento A ou do evento B, devemos achar o total de maneiras como A pode ocorrer e o total de maneiras como B pode ocorrer, mas *de modo que nenhum resultado seja contado mais de uma vez*. Um método de cálculo consiste em combinar o número de ocorrências possíveis de A com o número de ocorrências possíveis de B e, se houver qualquer superposição, subtrair o número de resultados que são contados duas vezes, de acordo com a regra abaixo.

Regra Formal da Adição

$$P(A \text{ ou } B) = P(A) + P(B) - P(A \text{ e } B)$$

onde $P(A$ e $B)$ denota a probabilidade de ocorrência simultânea de A e B em um mesmo experimento.

Embora a regra formal da adição seja apresentada como uma fórmula, é preferível entender o espírito da regra e aplicá-la intuitivamente, como segue.

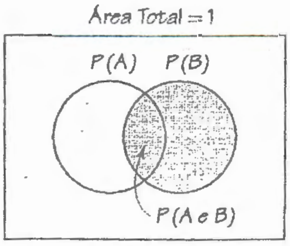


Fig. 3-4 Diagrama de Venn exibindo eventos que se superpõem.

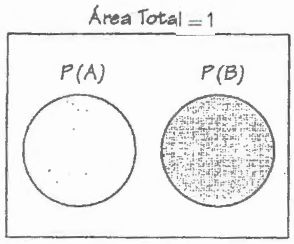


Fig. 3-5 Diagrama de Venn exibindo eventos que não se superpõem.

Regra Intuitiva da Adição

Para achar $P(A$ ou $B)$, somamos o número de ocorrências possíveis de A e o número de ocorrências possíveis de B, *adicionando esses números de tal modo que cada resultado seja contado apenas uma vez.* $P(A$ ou $B)$ é igual a essa soma, dividida pelo número total de resultados possíveis.

A Figura 3-4 mostra um diagrama de Venn que ilustra visualmente a regra formal da adição. Vemos ali que a probabilidade de A ou B é igual à probabilidade de A (círculo esquerdo) mais a probabilidade de B (círculo direito) menos a probabilidade de A e B (região comum aos dois círculos). Esta figura mostra que a adição pura e simples das áreas dos dois círculos acarreta uma dupla contagem da área do meio. Esse é o conceito básico subjacente da regra da adição. Em virtude da relação entre a regra da adição e o diagrama de Venn da Figura 3-4, é costume usarmos a notação $P(A \cup B)$ em lugar de $P(A$ ou $B)$. Analogamente, costuma-se utilizar a notação $P(A \cap B)$ em lugar de $P(A$ e $B)$. A regra formal da adição pode, pois, ser expressa como

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

A regra da adição se simplifica quando A e B não podem ocorrer simultaneamente, de forma que $P(A$ e $B)$ é zero. Pela Figura 3-5 vemos que, quando não há superposição de A e B, temos $P(A$ ou $B) = P(A) + P(B)$. A definição seguinte formaliza a ausência de superposição mostrada na Figura 3-5.

DEFINIÇÃO

Os eventos A e B dizem-se **mutuamente excludentes** se não podem ocorrer simultaneamente.

O fluxograma da Figura 3-6 mostra como os eventos mutuamente excludentes afetam a regra da adição.

Às vezes podemos ampliar não só nossa compreensão dos dados como nossa capacidade de resolver problemas, recolocando os dados em uma forma mais favorável. Consideremos esta afirmação. Em um teste com o antialérgico Seldane, 49 dos 781 usuários de Seldane experimentaram dores de cabeça, 49 dos 665 que usaram placebo experimentaram dores de cabeça, e 24 dos 626 indivíduos do grupo de controle experimentaram dores de cabeça. (Esses resultados se baseiam em dados de Merrell Dow Pharmaceutical, Inc.) Nessa afirmação, os dados são um tanto difíceis de ser compreendidos, mas se tornarão muito mais claros se forem reorganizados em forma tabular (veja Tabela 3-1). Os exemplos que seguem utilizam dados da Tabela 3-1.

EXEMPLO Se um dos 2072 indivíduos da Tabela 3-1 é escolhido aleatoriamente, determine a probabilidade de se obter alguém que fez uso de um placebo ou estava no grupo de controle.

SOLUÇÃO Podemos denotar a probabilidade procurada por P(placebo ou controle). A Tabela 3-1 mostra que os indivíduos que usaram o placebo e os do grupo de controle se excluem mutuamente. Ou seja, não há superposição desses dois grupos. Conseqüentemente, o número total de pessoas que utilizaram um placebo ou que estavam no grupo de controle é $665 + 626 = 1291$, e a probabilidade procurada é $1291/2072 = 0{,}623$.

$$P(\text{placebo ou controle}) = \frac{665}{2072} + \frac{626}{2072} = \frac{1291}{2072} = 0{,}623$$

EXEMPLO Escolhido aleatoriamente um dos 2072 indivíduos da Tabela 3-1, determine a probabilidade de obter alguém que tenha usado Seldane ou que não teve dor de cabeça.

SOLUÇÃO Podemos denotar a probabilidade procurada por P(Seldane ou não-dor de cabeça). A Tabela 3-1 mostra que há uma superposição entre o grupo dos usuários de Seldane e o grupo dos que não tiveram dor de cabeça. Ou seja, os dois eventos não são mutuamente excludentes; devemos, pois, ter cuidado de evitar dupla contagem ao calcularmos nossas so-

TABELA 3-1 Teste de Seldane

	Seldane	Placebo	Grupo de Controle	**Total**
Dor de cabeça	49	49	24	**122**
Não-dor de cabeça	732	616	602	**1950**
Total	**781**	**665**	**626**	**2072**

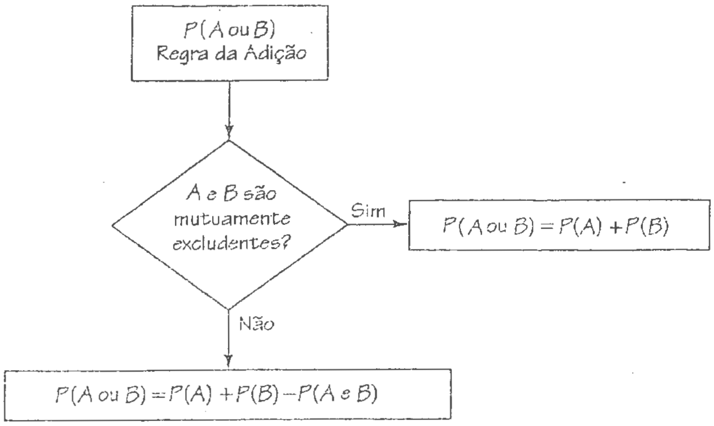


Fig. 3-6 Aplicação da regra da adição.

mas. A abordagem intuitiva consiste simplesmente em somar a coluna Seldane e a linha "sem dor de cabeça" de tal forma que o elemento 732 seja contado apenas uma vez.

P(Seldane ou sem dor de cabeça) =

$$= \frac{49 + 732 + 616 + 602}{2072} = \frac{1999}{2072} = 0{,}965$$

Chegaríamos ao mesmo resultado aplicando a regra formal da adição, como segue:

P(Seldane ou sem dor de cabeça) =
= P(Seldane) + P(sem dor de cabeça) − P(Seldane e sem dor de cabeça) =

$$= \frac{781}{2072} + \frac{1950}{2072} - \frac{732}{2072} = \frac{1999}{2072} = 0{,}965$$

Resumimos a seguir os pontos-chave desta seção.

1. Para calcular $P(A$ ou $B)$, começamos por associar a conjunção *ou* à adição.
2. Verificamos se os eventos A e B são mutuamente excludentes; em outras palavras, podem eles ocorrer simultaneamente? Se não são mutuamente excludentes (isto é, se podem ocorrer simultaneamente), devemos ter o cuidado de evitar (ou, ao menos, compensar) a contagem dupla ao somarmos as probabilidades. Se o leitor compreendeu a importância de evitar a contagem dupla ao calcular $P(A$ ou $B)$, não terá necessariamente de calcular $P(A) + P(B) - P(A$ e $B)$.

Os erros cometidos na aplicação da regra da adição freqüentemente envolvem contagem dupla; isto é, eventos que não são mutuamente excludentes são tratados como se o fossem. Uma indicação da presença de um tal erro é uma probabilidade total superior a 1. Mas nem sempre os erros que envolvem a regra da adição conduzem a uma probabilidade total superior a 1.

Palpite em Testes SAT?

Recomenda-se freqüentemente aos estudantes que se preparam para um teste de múltipla escolha, que não recorram ao "palpite"; mas isso não é necessariamente um bom conselho. Os testes padronizados de múltipla escolha tipicamente *compensam* o palpite, mas não o *penalizam*. Para questões com cinco escolhas de resposta, usualmente subtrai-se um quarto de ponto para cada resposta incorreta. Os princípios da probabilidade mostram que, a longo prazo, o simples "palpite" aleatório não aumenta nem diminui o resultado do exame. O estudante deve efetivamente "palpitar" se pode eliminar ao menos uma escolha ou se tem certo senso da resposta correta, mas deve evitar questões ardilosas com respostas atraentes. Também não deve perder muito tempo com essas questões.

Eventos Complementares

Na Seção 3-2 definimos o complemento do evento A e denotamo-lo por $\overline{A}$. A definição de eventos complementares implica que eles devem ser mutuamente excludentes, pois é impossível um evento ocorrer e *não* ocorrer simultaneamente. Outrossim, podemos ter certeza absoluta de que A ocorre ou não ocorre. Ou seja, um dos dois, A ou $\overline{A}$, deve ocorrer. Essas observações permitem-nos aplicar, como segue, a regra para eventos mutuamente excludentes:

$$P(A \text{ ou } \overline{A}) = P(A) + P(\overline{A}) = 1$$

Justificamos $P(A$ ou $\overline{A}) = P(A) + P(\overline{A})$ notando que A e $\overline{A}$ são mutuamente excludentes; e justificamos o total 1 pela certeza absoluta de que A deve ocorrer, ou não. Esse resultado da regra da adição conduz às três formas equivalentes que seguem:

Regra dos Eventos Complementares

$P(A) + P(\overline{A}) = 1$
$P(\overline{A}) = 1 - P(A)$
$P(A) = 1 - P(\overline{A})$

A primeira forma decorre diretamente de nosso resultado original. A segunda (veja Figura 3-7) e a terceira variantes envolvem manipulações muito simples da equação.

EXEMPLO Se P(chuva) = 0,4, determine P(não-chuva).

SOLUÇÃO Pela regra dos eventos complementares, temos

$$P(\text{não-chuva}) = 1 - P(\text{chuva}) = 1 - 0{,}4 = 0{,}6$$

Área Total = 1

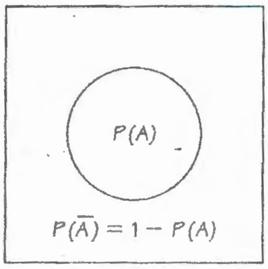


Fig. 3-7 Diagrama de Venn para o complemento do evento *A*.

Uma grande vantagem da *regra dos eventos complementares* é que sua aplicação pode simplificar consideravelmente certos problemas. Ilustraremos essa vantagem na seção seguinte.

3-3 Exercícios A: Habilidades e Conceitos Básicos

Para cada parte dos Exercícios 1 e 2, os dois eventos são mutuamente excludentes em um único experimento?

1. a. Escolha de um espectador de televisão do sexo masculino
   Escolha de alguém que raramente utiliza o controle remoto da TV N
   b. Escolha de um cidadão registrado no Partido Democrata
   Escolha de um cidadão contrário a todos os planos de bem-estar social N
   c. Girar uma roleta e obter o resultado 7 S
   Girar uma roleta e obter um número par
2. a. Adquirir um Corvette novo sem defeitos S
   Adquirir um carro com faróis inoperantes
   b. Escolher um curso de matemática N
   Escolher um curso que se afigure interessante
   c. Escolher uma pessoa com cabelo louro (natural ou artificial)
   Escolher uma pessoa careca S
3. a. Se $P(A) = 2/5$, determine $P(\bar{A})$. 3/5
   b. Com base em dados recentes do U.S. National Center for Health Statistics (Centro Nacional de Estatística da Saúde dos EUA), a probabilidade de uma criança ser menino é 0,513. Determine a probabilidade de uma criança ser menina. 0,487
4. a. Determine $P(\bar{A})$, dado que $P(A) = 0{,}228$.
   b. Com base em dados da National Conference of Bar Examiners (Conferência Nacional de Examinadores Forenses), se escolhermos aleatoriamente uma pessoa que se submete ao exame para exercício da advocacia, a probabilidade de obter alguém que seja aprovado é 0,57. Ache a probabilidade de escolher alguém que seja reprovado.
5. Ao jogar vinte-e-um com um único baralho de cartas no cassino Stardust, em Las Vegas, o apostador tira a primeira carta de um baralho bem baralhado. Qual é a probabilidade de obter (a) uma carta de paus ou um ás (b) um ás ou um 2?
6. Escolhida uma pessoa aleatoriamente, determine a probabilidade de seu aniversário não cair no dia 18 de outubro, que é o Dia Nacional da Estatística no Japão. Ignore os anos bissextos.
7. Com base na Tabela 3-1 desta seção, se escolhermos aleatoriamente um dos 2072 indivíduos, qual é a probabilidade de obter alguém que tomou Seldane ou usou um placebo?
8. Com base na Tabela 3-1 desta seção, se escolhermos aleatoriamente um dos 2072 indivíduos, qual é a probabilidade de obter alguém que utilizou um placebo ou teve dor de cabeça?
9. Os pesquisadores estão preocupados com o declínio do nível de cooperação por parte dos entrevistados em pesquisas. Um pesquisador aborda 84 pessoas na faixa etária 18-21 e constata que 73 respondem, enquanto 11 recusam responder. Quando são abordadas 275 pessoas na faixa etária 22-29, 255 respondem e 20 recusam responder [com base em dados de "Ouvi Você Bater Mas Você Não Pode Entrar" ("I Hear You Knocking but You Can't Come In"), por Fitzgerald and Fuller, *Sociological Methods and Research*, Vol. 11, N.º 1]. Suponha que 1 dos 539 indivíduos seja escolhido aleatoriamente. Determine a probabilidade de obter alguém na faixa etária 18-21 ou alguém que recuse responder.
10. Com base nos mesmos dados do Exercício 9, determine a probabilidade de obter alguém na faixa etária de 18-21 ou alguém que tenha respondido.
11. Os problemas de assédio sexual têm recebido muita atenção nos últimos anos. Em uma pesquisa, 420 trabalhadores (240 dos quais homens) consideram uma simples batida no ombro como uma forma de assédio sexual, enquanto que 580 trabalhadores (380 dos quais homens) não consideram isso como assédio (com base em dados de Bruskin/Goldring Research). Escolhido aleatoriamente um dos trabalhadores pesquisados, determine a probabilidade de obter alguém que não considere um simples tapa no ombro como uma forma de assédio sexual.
12. Com base nos mesmos dados do Exercício 11, determine a probabilidade de escolher aleatoriamente um homem ou alguém que não considere uma simples batida no ombro como uma forma de assédio sexual.
13. Um estudo de hábitos de fumantes compreende 200 casados (54 dos quais fumam), 100 divorciados (38 dos quais fumam) e 50 adultos que nunca se casaram (11 dos quais fumam)[com base em dados do Department of Health and Human Services (Departamento de Saúde e Serviços Humanos dos EUA)]. Escolhido aleatoriamente 1 indivíduo dessa amostra, determine a probabilidade de obter alguém divorciado ou fumante.
14. Com base nos dados do Exercício 13, determine a probabilidade de obter alguém que nunca se casou ou que não fume.

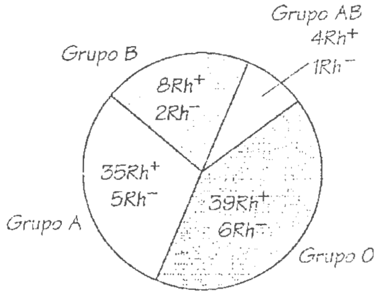
Nos Exercícios 15 e 16, use os dados da tabela a seguir, que resumem uma amostra de 200 tempos (em minutos) entre erupções do gêiser Old Faithful no Parque Nacional de Yellowstone.

Tempo	Freqüência
40−49	8
50−59	44
60−69	23
70−79	6
80−89	107
90−99	11
100−109	1

15. Os visitantes de Yellowstone naturalmente desejam assistir a uma erupção do Old Faithful, e assim o intervalo entre as erupções torna-se uma preocupação para os que não dispõem de muito tempo. Escolhido aleatoriamente um dos tempos da tabela, qual é a probabilidade de ser no mínimo de uma hora?
16. Escolhido aleatoriamente um dos tempos da tabela, qual é a probabilidade de ser no mínimo de 70 minutos, ou de estar entre 60 e 79 minutos?

Nos Exercícios 17-24, recorra à figura a seguir, que descreve os grupos sanguíneos e os tipos de Rh de 100 pessoas (com base em dados do Greater New York Blood Program). Suponha, em cada caso, que 1 dos 100 indivíduos seja selecionado aleatoriamente, e determine a probabilidade indicada.

17. P(não-grupo O)
18. P(não-tipo Rh^+)
19. P(grupo B ou tipo Rh^-)
20. P(grupo O ou grupo A)
21. P(tipo Rh^-)
22. P(grupo A ou tipo Rh^+)
23. P(grupo AB ou tipo Rh^-)
24. P(grupo A ou B ou tipo Rh^+)



3-3 Exercícios B: Além do Básico

25. a. Se $P(A$ ou $B) = 1/3$, $P(B) = 1/4$ e $P(A$ e $B) = 1/5$, determine $P(A)$.
    b. Se $P(A) = 0{,}4$ e $P(B) = 0{,}5$, que se pode dizer quanto a $P(A$ ou $B)$ se A e B são eventos mutuamente excludentes?
    c. Se $P(A) = 0{,}4$ e $P(B) = 0{,}5$, que se pode dizer quanto a $P(A$ ou $B)$, se A e B *não* são mutuamente excludentes?
26. Se A e B são mutuamente excludentes e B e C também o são, os eventos A e C devem ser mutuamente excludentes? Dê um exemplo que confirme sua resposta.
27. Como se modifica a regra da adição, se utilizamos *ou exclusivo* em lugar de *ou inclusivo*? Recorde que *ou exclusivo* significa um ou outro, mas não ambos.
28. Dado que $P(A$ ou $B) = P(A) + P(B) - P(A$ e $B)$, estabeleça uma regra formal para $P(A$ ou B ou $C)$. (*Sug.:* Trace um diagrama de Venn.)

3-4 Regra da Multiplicação

O objetivo da Seção 3-3 foi estabelecer uma regra para calcular $P(A$ ou $B)$, a probabilidade de uma prova em um experimento ter o resultado A, ou o resultado B, ou ambos. Nesta seção, temos em vista estabelecer uma regra para calcular $P(A$ e $B)$, a probabilidade de o evento A ocorrer em uma primeira prova, e o evento B ocorrer em uma segunda prova. Na Seção 3-3, associamos a conjunção *ou* à adição; nesta seção, vamos associar a conjunção *e* à multiplicação. Veremos que $P(A$ e $B)$ envolve multiplicação de probabilidades, e que às vezes é preciso ajustar a probabilidade do evento B de modo a refletir o resultado do evento A.

A teoria da probabilidade tem extensa aplicação em análise e planejamento de testes padronizados como SAT, ACT, LSAT (para direito) e MCAT (para medicina). Para facilidade de correção, tais testes utilizam tipicamente questões do tipo verdadeiro/falso ou questões de múltipla escolha. Suponhamos que a primeira questão em um teste seja do tipo verdadeiro/falso, enquanto a segunda questão é do tipo múltipla escolha, com 5 respostas possíveis (a,b,c,d,e). Utilizaremos as duas questões seguintes. Tente resolvê-las!

1. Verdadeiro ou falso: O fumo é uma das principais causas do câncer.
2. O coeficiente de correlação de Pearson é assim chamado em homenagem a
   a. Karl Marx
   b. Carl Friedrich Gauss
   c. Karl Pearson
   d. Carly Simon
   e. Mario Triola

Quando se julgam testes padronizados, faz-se em geral uma compensação pelos "palpites"; assim, vamos determinar a probabilidade de que, se alguém palpita as respostas a ambas as questões, tanto a primeira como a segunda respostas estejam corretas. Uma forma de determinar essa probabilidade consiste em listar o espaço amostral como segue:

V,a	V,b	V,c	V,d	V,e
F,a	F,b	F,c	F,d	F,e

Se as respostas são aleatórias, então os 10 resultados possíveis são igualmente prováveis. As respostas corretas são V e c, e assim

$$P(\text{ambas corretas}) = P(\text{V e c}) = \frac{1}{10} = 0{,}1$$

Considerando as respostas individuais de V e c, respectivamente, vemos que, com suposições aleatórias, temos $P(\text{V}) = 1/2$ e $P(\text{c}) = 1/5$. Como $1/10$ é o produto de $1/2$ e $1/5$, vemos que $P(\text{V e c}) = P(\text{V}) \cdot P(\text{c})$. Isso sugere que, de modo geral, $P(A \text{ e } B) = P(A) \cdot P(B)$, mas consideremos outro exemplo, antes de firmar esta generalização.

Notemos por ora que os diagramas em árvore podem auxiliar-nos na determinação do número de resultados possíveis em um espaço amostral. Um **diagrama em árvore** é uma representação pictórica dos resultados possíveis de um experimento, consistindo em segmentos retilíneos que emanam de um ponto de partida. Tais diagramas são úteis para a contagem do número de resultados possíveis, desde que o número de possibilidades não seja demasiadamente grande. O diagrama em árvore da Figura 3-8 resume os resultados das questões verdadeiro/falso e múltipla escolha. Pela Figura 3-8 vemos que, se ambas as respostas são aleatórias (palpi-

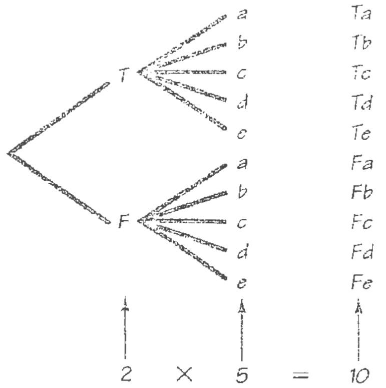


Fig. 3-8 Diagrama em árvore para respostas de um teste.

tes), todos os 10 ramos são igualmente prováveis e a probabilidade de obter a resposta correta (V,c) é 1/10. Para cada resposta da primeira questão, há cinco respostas da segunda. O número total de resultados é 5 tomado 2 vezes, ou seja, 10. O diagrama em árvore da Figura 3-8 justifica o uso da multiplicação.

Nosso primeiro exemplo de questões verdadeiro/falso e múltipla escolha sugeriu que $P(A \text{ e } B) = P(A) \cdot P(B)$, mas o próximo exemplo introduzirá outro elemento importante. Este exemplo envolve um baralho comum, de forma que o contexto deve ser familiar à maioria dos leitores, mas os princípios utilizados podem ser aplicados a circunstâncias mais significativas.

Amostragem Composta

O Exército Americano resolveu fazer um teste de sífilis, aplicando em cada recruta um teste individual de sangue, que foi analisado separadamente. Um pesquisador sugeriu que se misturassem pares de amostras de sangue. Após testados esses pares misturados, os recrutas sifilíticos podiam ser identificados, testando-se as poucas amostras de sangue que estavam nos pares que acusaram resultado positivo. Reduziu-se o número total de análises emparelhando-se espécimes de sangue. Por que então não colocá-los em grupos de três, quatro ou mais? Com auxílio da teoria das probabilidades, determinou-se o tamanho mais eficaz de grupo, elaborando-se uma teoria para detectar os defeitos em qualquer população. Essa técnica é conhecida como *amostragem composta*.

EXEMPLO Na extração de duas cartas de um baralho bem misturado, determine a probabilidade de que a primeira carta seja um ás e a segunda seja um rei. (Admita que a primeira carta extraída não seja reposta antes da extração da segunda carta.)

SOLUÇÃO Pela Regra 2 da Seção 3-2, achamos a probabilidade de obter um ás na primeira extração. Como há 4 ases nas 52 cartas distintas, temos $P(\text{ás}) = 4/52$. Para a segunda extração, suponha que tenhamos obtido um ás na primeira extração, de modo que temos agora 4 reis entre apenas 51 cartas, donde $P(\text{rei}) = 4/51$. A probabilidade de obter um ás na primeira extração *e* um rei na segunda é, pois,

$$P(\text{ás e rei}) = \frac{4}{52} \cdot \frac{4}{51} = 0{,}00603$$

Poderíamos justificar melhor esse resultado listando o espaço amostral ou fazendo um diagrama em árvore. Mas o espaço amostral tem 2652 possibilidades diferentes, e o diagrama em árvore tem 2652 ramos. Obviamente, este exemplo seria por demais laborioso, mas os resultados mostrariam que, dentre as 2652 possibilidades, há 16 casos que consistem em um ás seguido por um rei.

Esse exemplo ilustra o importante princípio de que a probabilidade do evento B deve levar em conta o fato de o evento A já ter ocorrido. Costuma-se expressar esse princípio com a seguinte notação.

Notação para a Regra da Multiplicação

$P(B|A)$ representa a probabilidade de ocorrência de B quando se sabe que o evento A já ocorreu. (Pode-se ler $B|A$ como "B dado A".)

No exemplo precedente, para achar a probabilidade de um ás na primeira extração e um rei na segunda, sem reposição da primeira carta extraída, temos

$$P(\text{ás na primeira extração}) = \frac{4}{52}$$

$$P(\text{rei}|\text{ás}) = \frac{4}{51}$$

onde $P(\text{rei}|\text{ás})$ denota a probabilidade de obter um rei na segunda extração, supondo que a primeira carta extraída tenha sido um ás.

Motores a Jato Independentes

Um jato de três motores decolou do Aeroporto Internacional de Miami com destino à América do Sul, mas um dos motores falhou logo após a decolagem. Enquanto o avião retornava à pista, os outros dois motores também falharam, mas o piloto conseguiu fazer uma aterrissagem segura. Com três motores *independentes*, a probabilidade de todos os três falharem simultaneamente é de apenas $0{,}0001^3$, ou seja, uma chance em um trilhão. As autoridades do Ministério da Aeronáutica americano constataram que um mesmo mecânico havia trocado o óleo nas três turbinas, colocando incorretamente os anéis de vedação da entrada de óleo. A utilização de três motores distintos *independentes* tem por objetivo aumentar a segurança, mas a interferência de um único mecânico tornou os motores dependentes. Os processos de manutenção exigem agora que os motores sejam vistoriados e ajustados por mecânicos diferentes.

DEFINIÇÕES

Dois eventos A e B são **independentes** se a ocorrência de um deles não afeta a probabilidade de ocorrência do outro. (Analogamente, vários eventos são independentes se a ocorrência de qualquer um deles não afeta as probabilidades de ocorrência dos outros.) Se A e B não são independentes, dizem-se **dependentes**.

Assim é que a jogada de uma moeda e a jogada de um dado são eventos *independentes*, porque o resultado da moeda não afeta a probabilidade do resultado do dado. Por outro lado, os eventos "conseguir dar partida no seu carro" e "chegar à aula no horário" são *dependentes*, porque o resultado da operação de dar partida no carro influi na probabilidade de chegar à aula no horário.

Com a notação e as definições precedentes, juntamente com os princípios ilustrados nos exemplos anteriores, resumimos o conceito-chave desta seção na Figura 3-9 e na *regra da multiplicação*.

Regra Formal da Multiplicação

$P(A \text{ e } B) = P(A) \cdot P(B)$ se A e B são *independentes*.
$P(A \text{ e } B) = P(A) \cdot P(B|A)$ se A e B são *dependentes*.

Regra Intuitiva da Multiplicação

Para determinar a probabilidade de ocorrência do evento A em uma prova e de ocorrência do evento B na próxima prova, devemos multiplicar a probabilidade de A pela probabilidade de B, não olvidando que a probabilidade do evento B deve levar em conta a ocorrência prévia do evento A.

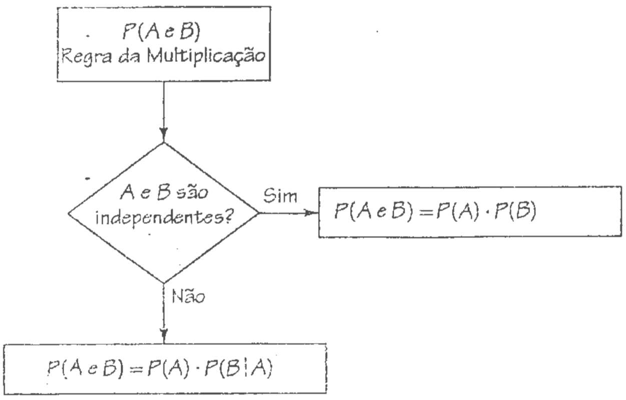


Fig. 3-9 Aplicação da Regra da Multiplicação.

A regra da multiplicação é extremamente importante em virtude de suas inúmeras aplicações. Uma área de aplicação envolve o teste de produtos, conforme exemplo a seguir.

Redundância ·

A confiabilidade de um sistema pode ser grandemente reforçado com a redundância, ou replicação de componentes críticos. Os aviões têm dois sistemas elétricos independentes, e as aeronaves usadas em vôo por instrumento têm tipicamente dois rádios. Damos a seguir o extrato de um artigo publicado na revista *Popular Science*: "Um avião construído em grande parte de fibras de carbono era o Lear Fan 2100, que devia levar dois transceptores de radar, porque se um único transceptor falhasse o avião se tornaria quase invisível na tela do radar." Essa redundância é uma aplicação da regra da multiplicação da teoria das probabilidades. Se um componente tem 0,001 de probabilidade de falha, a probabilidade de duas componentes idênticas falharem simultaneamente é de apenas 0,000001.

EXEMPLO A Detroit Auto Supply Company produz um lote de 50 filtros de combustível, dos quais 6 são defeituosos. (Os otimistas diriam que 44 são bons.) Escolhem-se aleatoriamente e testam-se dois filtros do lote. Determine a probabilidade de ambos serem bons, se os filtros são selecionados (a) com reposição, (b) sem reposição.

SOLUÇÃO

a. Se os filtros são escolhidos com reposição, as duas escolhas são independentes, porque o segundo evento não é afetado pelo primeiro resultado. Obtemos, portanto,

$$P(\text{primeiro e segundo bons}) = \frac{44}{50} \cdot \frac{44}{50} = 0{,}774$$

b. Se os filtros são escolhidos sem reposição, as duas escolhas são dependentes, porque o segundo evento é afetado pelo primeiro resultado. Temos:

$$P(\text{primeiro e segundo bons}) = \frac{44}{50} \cdot \frac{43}{49} = 0{,}772$$

Note que, nesse caso, ajustamos a segunda probabilidade de forma a levar em conta a escolha de um filtro bom na primeira seleção. Após selecionado um filtro bom na primeira vez, teremos 43 filtros bons entre os 49 restantes. Note também que, sem reposição, a chance de obter dois filtros bons é ligeiramente inferior. Ao estabelecermos um processo para testar lotes de produtos por amostragem, devemos utilizar amostras sem reposição, por duas razões: Primeiro, há menor chance de obter apenas artigos perfeitos quando há alguns defeituosos; segundo, não faz sentido utilizar amostragem com reposição, porque há a possibilidade de selecionarmos o mesmo artigo mais de uma vez — o que é uma perda de trabalho.

Macacos Datilógrafos

Uma afirmação clássica é que um macaco, batendo ao acaso nas teclas de uma máquina de escrever, acabaria compondo a obra completa de Shakespeare, admitindo-se que continuasse datilografando indefinidamente, século após século. Para tal estimativa, aplicou-se a regra da multiplicação da teoria das probabilidades. Um resultado de 1.000.000.000.000.000.000.000.000.000.000.000.000 anos é considerado muito pequeno por alguns. Nesse mesmo espírito, Sir Arthur Eddington escreveu este poema: "Havia uma vez um macaco inteligente, que sempre tocava um baixo, e que disse: 'Parece que, em bilhões de anos, acabarei compondo uma melodia.'

Até aqui abordamos o caso de dois eventos, mas a regra da multiplicação estende-se facilmente a três ou mais eventos. De modo geral, a probabilidade de qualquer seqüência de eventos independentes é simplesmente o produto das probabilidades correspondentes. Assim é que a probabilidade de obter três "caras" em três jogadas de uma moeda é $(0{,}5)(0{,}5)(0{,}5) = 0{,}125$. A regra da multiplicação também se aplica ao caso de vários eventos dependentes; basta ajustarmos convenientemente as probabilidades. Por exemplo, a probabilidade de obtermos três ases em três extrações de cartas de um baralho, sem reposição, é dada por

$$\frac{4}{52} \cdot \frac{3}{51} \cdot \frac{2}{50} = 0{,}000181$$

Neste último exemplo envolvendo três ases, consideramos os eventos como dependentes porque as escolhas foram feitas sem reposição. Todavia, é prática comum considerar os eventos como independentes quando se extraem pequenas amostras de grandes populações. (Em tais casos, é raro extrairmos o mesmo item duas vezes.) Uma orientação comum consiste em supor independência sempre que o tamanho da amostra não supere 5% do tamanho da população. Quando um entrevistador pesquisa 1200 adultos de uma população de milhões, supõe independência, mesmo que a amostragem se faça sem reposição.

A Probabilidade de "Ao Menos Um"

A regra da multiplicação e a regra dos complementos podem ser conjugadas para simplificar consideravelmente certos tipos de problemas, como a determinação da probabilidade de que, em várias tentativas, *ao menos 1* tenha um resultado especificado. Em tais casos, devem ficar bem claros os conceitos:

- "Ao menos 1" é equivalente a "1 ou mais".
- O complemento de "obter ao menos 1 item de determinado tipo" é "não obter nenhum item daquele tipo."

Suponhamos que um empregado na cidade de San Francisco precise falar com 1 de seus 5 colegas em sua casa. Admita que os 5 colegas sejam escolhas aleatórias de uma população, e que 39,5% dos números de telefone de San Francisco não estejam na lista (com base em dados da Survey Sampling, Inc.). Devemos determinar a probabilidade de que ao menos 1 dos 5 colegas de trabalho do nosso empregado tenha seu número de telefone na lista. Veja as interpretações que seguem.

Ao menos 1 número na lista = 1 ou mais números na lista
O complemento de "Ao menos 1 número na lista" = "nenhum número na lista" (ou **"todos os números são não-listados"**)

A resolução direta deste problema é complexa, mas a solução do exemplo seguinte dá uma abordagem simples, indireta.

EXEMPLO Determine a probabilidade de ao menos 1 dentre 5 empregados em San Francisco ter o número de telefone na lista (podendo, portanto, ser chamado). Suponha que os números de telefone sejam independentes e que, em San Francisco, 39,5% dos números não estejam na lista.

SOLUÇÃO

Etapa 1: Represente por um símbolo a probabilidade desejada.
Em nosso caso, seja L = ao menos 1 número na lista, dentre os números dos 5 empregados.

Etapa 2: Identificar o complemento do evento indicado em 1.

$\overline{L}$ = *nenhum* número na lista dentre os 5 empregados
= 5 números não-listados dentre os 5 empregados

Etapa 3: Determinar a probabilidade do complemento da Etapa 2.

$P(\overline{L}) = P$(5 números não listados entre 5 empregados)
$= (0{,}395)(0{,}395)(0{,}395)(0{,}395)(0{,}395) =$
$= (0{,}395)^5 = 0{,}00962$

Etapa 4: Determinar a probabilidade do evento considerado, subtraindo de 1 a probabilidade do complemento.

$$P(L) = 1 - P(\overline{L}) = 1 - 0{,}00962 = 0{,}990$$

Há, pois 0,990 de probabilidade de ao menos 1 dos empregados ter seu número na lista, podendo, portanto, ser contactado.

EXEMPLO Determinar a probabilidade de ao menos 1 menina se um casal planeja ter 3 filhos. Admita que as probabilidades de menino e menina sejam iguais, e que o sexo de qualquer filho não seja influenciado pelo sexo dos que o precedem.

SOLUÇÃO

Etapa 1: Represente por um símbolo a probabilidade desejada. Em nosso caso, A = ao menos 1 menina em 3 filhos.

Etapa 2: Identificar o complemento do evento da Etapa 1.

$\overline{A}$ = *não* obter ao menos 1 menina em 3 filhos
= obter 3 meninos em 3 filhos

Etapa 3: Determinar a probabilidade do complemento da Etapa 2:

$P(\overline{A}) = P$(3 meninos em 3 filhos) $= (0{,}5)(0{,}5)(0{,}5)$
$= (0{,}5)^3 = 0{,}125.$

Etapa 4: Determinar a probabilidade do evento desejado, subtraindo de 1 a probabilidade do complemento.

$$P(A) = 1 - P(\overline{A}) = 1 - 0{,}125 = 0{,}875$$

Há, assim, uma probabilidade de 0,875 de ao menos 1 menina em 3 filhos.

Probabilidade Condicional

A regra da multiplicação para eventos dependentes pode expressar-se formalmente como $P(A \text{ e } B) = P(A) \cdot P(B|A)$. É fácil resolver algebricamente essa equação em relação a $P(B|A)$; basta dividir ambos os membros da equação por $P(A)$. O resultado é chamado *probabilidade condicional* de ocorrência do evento B, dado que o evento A já ocorreu.

DEFINIÇÃO

A **probabilidade condicional** de B dado A é a probabilidade de ocorrência do evento B, sabido que o evento A já ocorreu. Pode ser determinada dividindo-se a probabilidade de ocorrência de ambos os eventos A e B pela probabilidade do evento A; como se mostra a seguir.

$$P(B|A) = \frac{P(A \text{ e } B)}{P(A)}$$

Essa fórmula é a expressão formal da probabilidade condicional, mas podemos também adotar a abordagem intuitiva:

Abordagem Intuitiva da Probabilidade Condicional

Podemos determinar a probabilidade condicional de B dado A supondo que A já tenha ocorrido e, sob essa hipótese, calculando a probabilidade de ocorrência do evento B.

Teste de independ.

Na regra da multiplicação para eventos dependentes, se $P(B|A) = P(B)$, então a ocorrência do evento A não influi na probabilidade do evento B. Esse fato costuma ser usado como teste de independência. Se $P(B|A) = P(B)$, então A e B são eventos independentes; mas se $P(B|A) \neq P(B)$, então A e B são eventos dependentes. Outro teste de independência envolve a igualdade de $P(A \text{ e } B)$ e $P(A) \cdot P(B)$. Se essas probabilidades são iguais, os eventos A e B são independentes. Se $P(A \text{ e } B) \neq P(A) \cdot P(B)$, então A e B são eventos dependentes. Esses resultados se resumem como segue:

Dois eventos A e B são *independentes* se
$P(B|A) = P(B)$
ou $P(A \text{ e } B) = P(A) \cdot P(B)$
Dois eventos A e B são *dependentes* se
$P(B|A) \neq P(B)$
ou $P(A \text{ e } B) \neq P(A).P(B)$

Por exemplo, se $P(B|A) = 0{,}2$ e $P(B) = 0{,}2$, então $P(B|A) = P(B)$ e concluímos que A e B são eventos independentes. Como $P(B|A)$

$= P(B)$, concluímos que a probabilidade do evento B não é afetada pela ocorrência do evento A; essa é a definição de independência. Entretanto, se $P(B|A) = 0{,}5$ e $P(B) = 0{,}6$, então $P(B|A) \neq P(B)$ e concluímos que A e B são dependentes. Aqui, os valores diferentes de $P(B|A)$ e de $P(B)$ mostram que a probabilidade do evento B é afetada pela ocorrência do evento A e, assim, A e B são eventos dependentes.

Condenados por Probabilidade

Uma testemunha descreveu uma ladra de Los Angeles como uma mulher caucasiana com cabelos louros penteados em forma de rabo de cavalo, que fugira em um carro amarelo dirigido por um homem afro-americano com bigode e barba. Janet e Malcolm Collins coincidiam com as características descritas, e foram condenados com base na hipótese de que há apenas 1 chance em 12 milhões de um casal ter essas mesmas características. A probabilidade de um carro amarelo era de 1/10, e as outras probabilidades foram estimadas em 1/4, 1/10, 1/3, 1/4 e 1/1000. As sentenças foram posteriormente revertidas, quando se observou que não havia evidência que apoiasse as probabilidades estimadas ou a independência dos eventos. Entretanto, como o casal não foi selecionado aleatoriamente, cometeu-se sério erro em não considerar a probabilidade de outros casais com as mesmas características estarem na mesma área.

EXEMPLO Com referência à Tabela 3-2, admita que todas as escolhas envolvam os 2000 indivíduos representados na tabela e determine:

a. Se uma pessoa é selecionada aleatoriamente, qual é a probabilidade de ela ter sido vítima de um estranho, dado que foi escolhida uma vítima de furto?
b. Escolhida uma vítima de assalto, qual a probabilidade de o criminoso ser um estranho?

O Teorema de Bayes

Thomas Bayes (1702-1761) afirmou que as probabilidades devem ser revistas quando conhecemos algo mais sobre os eventos. Eis uma forma do Teorema de Bayes:

$$P(A|B) = \frac{P(A) \cdot P(B|A)}{P(A) \cdot P(B|A) + P(\bar{A}) \cdot P(B|\bar{A})}$$

Suponha que 60% dos chips do computador de uma companhia sejam produzidos pela fábrica A e 40% por outra fábrica (denotada por $\bar{A}$). Para um chip escolhido aleatoriamente, a probabilidade de provir da fábrica A é 0,60. Suponha que um chip se revele defeituoso, e que as taxas de defeito nas duas fábricas sejam de 35% para A e 25% para $\bar{A}$. Com auxílio da fórmula acima, podemos determinar que a probabilidade de o chip defeituoso provir da fábrica A é 0,677.

TABELA 3-2 Relação entre Criminoso e Vítima

	Homicídio	Furto	Assalto	Totais
Estranho	12	379	727	**1118**
Conhecido ou parente	39	106	642	**787**
Ignorado	18	20	57	**95**
Totais	**69**	**505**	**1426**	**2000**

SOLUÇÃO

a. Queremos $P(\text{estranho}|\text{furto})$. Se a pessoa selecionada foi vítima de furto, estamos lidando com as 505 pessoas da segunda coluna de valores. Dessas 505, 379 foram vítimas de estranhos. Portanto,

$$P(\text{estranho}|\text{furto}) = \frac{379}{505} = 0{,}750$$

Podemos chegar ao mesmo resultado com a abordagem formal:

$$P(\text{estranho}|\text{furto}) = \frac{P(\text{furto e estranho})}{P(\text{furto})} = \frac{379/2000}{505/2000} = 0{,}750$$

b. Aqui, queremos $P(\text{estranho}|\text{assalto})$. Se a pessoa selecionada foi vítima de assalto, está entre as 1426 pessoas da terceira coluna. Dessas 1426 pessoas, 727 foram vitimadas por estranhos, de forma que

$$P(\text{estranho}|\text{assalto}) = \frac{727}{1426} = 0{,}510$$

Novamente aqui, chegamos ao mesmo resultado por uma abordagem formal:

$$P(\text{estranho}|\text{assalto}) = \frac{P(\text{assalto e estranho})}{P(\text{assalto})} = \frac{727/2000}{1426/2000} = 0{,}510$$

Comparando os resultados das partes (a) e (b), vemos que a probabilidade de a pessoa ser vítima de um estranho é muito diferente, conforme se trate de furto ou de assalto. Há, assim, uma dependência entre o tipo de crime e a relação entre criminoso e vítima.

3-4 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1 e 2, para cada par de eventos classifique-os como independentes ou dependentes. Alguns dos outros exercícios se baseiam em conceitos de seções anteriores deste capítulo.

1. a. Assistir a aulas de estatística D
   Passar em um curso de estatística
   b. Furar um pneu no trajeto para a aula I
   Acordar tarde demais para as aulas
   c. Eventos A e B, com $P(A) = 0{,}40$, $P(B) = 0{,}60$ e $P(A \text{ e } B) = 0{,}20$ D
2. a. Encontrar seu forno de microondas com defeito
   Encontrar seu detector de fumaça a bateria com defeito
   b. Encontrar a lâmpada de sua cozinha queimada
   Encontrar seu refrigerador com defeito
   c. Eventos A e B, tais que $P(A) = 0{,}90$, $P(B) = 0{,}80$ e $P(A \text{ e } B) = 0{,}72$
3. Dez por cento das pessoas são canhotas. Qual é a probabilidade de selecionar aleatoriamente 2 pessoas canhotas?
4. Determine a probabilidade de responder corretamente, por "palpite", as duas primeiras questões de um teste se
   a. As 2 primeiras questões são do tipo verdadeiro/falso
   b. As 2 primeiras questões são do tipo múltipla escolha com 5 possibilidades.
5. Determine a probabilidade de tirar 4 ases consecutivos na extração, sem reposição, de 4 cartas de um baralho.

6. Determine a probabilidade de ao menos 1 menina em um casal com 5 filhos. Admita que menino e menina sejam igualmente prováveis, e que o sexo de um dos filhos seja independente do sexo dos outros.
7. Um estudante tem dificuldade com o mau funcionamento de despertadores. Em lugar de utilizar 1 despertador, ele decide utilizar 3. Qual a probabilidade de ao menos 1 despertador funcionar, se cada despertador tem 98% de chance de funcionar?
8. Já vimos que jogar um par de dados comporta 36 resultados possíveis: 1-1, 1-2, ..., 6-6.
   a. Qual a probabilidade de um 7?
   b. Se você acaba de entrar em um jogo de dados em casa de amigos, e se a pessoa que trouxe os dados obtém oito 7 consecutivos, qual é a sua conclusão? Por quê?
9. Quatro estudantes que chegaram atrasados para o exame deram a clássica desculpa do pneu furado. No teste substitutivo, o instrutor pede que os estudantes identifiquem o pneu que furou. Se não houve realmente nenhum pneu furado e os estudantes responderam na base do "palpite", qual a probabilidade de todos eles escolherem o mesmo pneu?
10. Três firmas que trabalham com o mesmo auditor escolhem, independente e aleatoriamente, um mês para a realização da auditoria anual. Qual a probabilidade de os três meses escolhidos serem diferentes?
11. Um gerente de controle de qualidade utiliza equipamento de teste para detectar modems de computador defeituosos. Retiram-se aleatoriamente 3 modems diferentes de um grupo onde há 12 defeituosos e 18 sem defeito. Qual a probabilidade (a) de todos os 3 serem defeituosos; (b) de ao menos um dos modems escolhidos ser defeituoso?
12. Em seu trajeto para a aula, um estudante deve passar por dois sinais de tráfego que operam independentemente. Para cada sinal, a probabilidade de "verde" é 0,4. Se ele deve encontrar os dois sinais abertos para chegar a tempo na aula, qual a probabilidade de não se atrasar?
13. O IRS (Internal Revenue Service — Serviço de Receita Interna dos EUA) reporta que, de todos os contribuintes auditados, 70% acabam tendo que pagar diferença de imposto. Um auditor novo selecionou aleatoriamente 12 declarações de imposto, auditou-as e se gabou de ter cobrado diferença de imposto de todos os 12 contribuintes. Qual a probabilidade de isso realmente ocorrer? Com base no resultado, é possível que ele esteja dizendo a verdade?
14. Um casal atraiu a atenção da imprensa pelo fato de seus três filhos terem nascido no mesmo dia 4 de julho de anos diferentes. Ignorando os anos bissextos, determine a probabilidade de que três pessoas selecionadas aleatoriamente tenham nascido no dia 4 de julho.
15. Escolhida aleatoriamente uma pessoa dentre as que morreram há poucos anos, há uma probabilidade de 0,0478 de que a morte tenha sido causada por acidente (conforme dados do *Statistical Abstract of the United States* — Resumo Estatístico dos EUA). Um detetive de Baltimore tem suspeitas quanto às mortes de 5 pessoas, classificadas como acidentais. Determine a probabilidade de que, dentre cinco mortes selecionadas aleatoriamente, todas tenham sido causadas por acidente.
16. Em um método de amostragem, seleciona-se uma amostra aleatória, sem reposição, e todo o lote é rejeitado se há ao menos um defeito. A Niko Electronics Company acaba de fabricar 5000 CDs, 3% dos quais apresentam algum defeito. Escolhidos e testados 10 dos CDs, qual é a probabilidade de todo o lote ser rejeitado?
17. Um gerente pode identificar roubos cometidos por empregados verificando amostras de despachos efetuados pelos mesmos. De 36 empregados, 2 estão roubando. Se o gerente verifica 4 empregados selecionados aleatoriamente, qual é a probabilidade de nenhum dos ladrões ser identificado?
18. Em um caso ocorrido em Riverhead, New York, 9 vítimas diferentes de crimes ouviram a gravação das vozes de 5 homens diferentes. Todas as 9 vítimas identificaram a mesma voz como a voz do criminoso. Se as identificações de voz foram feitas aleatoriamente, determine a probabilidade de todas as nove vítimas indicarem a mesma pessoa. Isso constitui uma dúvida razoável?
19. Um avaliador de sinistros de uma companhia de seguros suspeita de 4 irmãos que reportaram um carro furtado cada um em diferentes regiões de Houston. Se a taxa anual de furto de carros em Houston é de 4,5%, determine a probabilidade de que, em 4 carros escolhidos aleatoriamente, todos tenham sido furtados em determinado ano. (Há 970.000 carros em Houston.) O que é que o resultado sugere?
20. Uma relação aprovada de jurados contém 20 mulheres e 20 homens. Determine a probabilidade de que, em uma escolha aleatória de 12 dessas pessoas, obtenhamos um júri composto só de homens. Nessas circunstâncias, se o acusado é condenado por um tal júri, há evidência suficiente para sugerir que a escolha dos jurados não foi aleatória?
21. Um processo de exame de sangue se torna mais eficaz combinando-se amostras de espécimes de sangue. Se se combinam amostras de sangue de 5 pessoas e o resultado do exame da mistura é negativo, podemos afirmar que todas as 5 amostras individuais são negativas. Determine a probabilidade de um resultado positivo para 5 amostras combinadas em uma única mistura, supondo que a probabilidade de o teste de uma amostra individual ser positivo é de 0,015.
22. Um empregado afirma que um novo processo de fabricação de videocassete é melhor, porque a taxa de defeitos é inferior a 5% (que era a taxa de defeitos no passado). Fabricados 20 videocassetes pelo novo processo, não se constata qualquer defeito. Supondo que o novo método tenha a mesma taxa de defeito de 5% verificada no passado, determine a probabilidade de não aparecer qualquer defeito entre os 20 videocassetes. Esse resultado constitui evidência suficiente de que o novo processo é melhor?

Nos Exercícios 23 e 24, utilize os dados da Tabela 3-2.

23. a. Determine a probabilidade de que, quando se escolhe 1 dos 2000 indivíduos, a pessoa escolhida tenha sido vitimada por um conhecido ou por um parente, sabendo-se que foi vítima de furto.
   b. Determine a probabilidade de que, quando se escolhe 1 dos 2000 indivíduos, a pessoa escolhida tenha sido furtada por um conhecido ou por um parente.
   c. Determine a probabilidade de que, quando se escolhe 1 dos 2000 indivíduos, a pessoa escolhida tenha sido roubada ou vitimada por um conhecido ou por um parente.
   d. Escolhidos aleatoriamente dois indivíduos diferentes, determine a probabilidade de ambos terem sido vítimas de furto.
24. a. Escolhida aleatoriamente uma das vítimas de crime constantes da tabela, determine a probabilidade de obter uma pessoa que tenha sido vitimada por alguém desconhecido ou que tenha sido vítima de homicídio.
   b. Escolhida aleatoriamente uma das vítimas de crime constantes da tabela, determine a probabilidade de obter alguém que tenha sido vítima de homicídio, dado que o criminoso é um estranho.
   c. Escolhida aleatoriamente uma das vítimas de crime constantes da tabela, determine a probabilidade de obter alguém que tenha sido vitimado por um estranho, dado que foi vítima de homicídio.

d. Escolhidos aleatoriamente dois indivíduos distintos, determine a probabilidade de ambos terem sido vítimas de criminosos desconhecidos.

Nos Exercícios 25-30, utilize a informação seguinte. O Departamento de Saúde do Estado de Nova York reporta uma taxa de 10% de incidência do vírus HIV na população considerada "de risco", e uma taxa de 0,3% de incidência de HIV para a população em geral. Os resultados dos testes de laboratório do vírus HIV são corretos 95% das vezes. Com base nesses resultados, se selecionamos aleatoriamente 5000 pessoas do grupo "de risco" e 20.000 pessoas da população geral, esperamos obter os resultados da tabela a seguir.

	Amostra Extraída da População de Risco		Amostra Extraída da População Geral	
Resultado do Teste de HIV	Positivo	Negativo	Positivo	Negativo
Infectado pelo vírus HIV	475	25	57	3
Não infectado pelo vírus HIV	225	4275	997	18.943

25. Escolhida aleatoriamente uma pessoa da população "de risco", qual é a probabilidade de estar infectada com o vírus HIV?
26. Escolhida aleatoriamente uma pessoa da população geral, qual é a probabilidade de seu teste de HIV dar resultado positivo?
27. Escolhida aleatoriamente uma pessoa da população de risco, qual é a probabilidade de seu teste ser positivo ou de ela estar infectada com HIV?
28. Escolhida aleatoriamente uma pessoa da população geral, qual é a probabilidade de seu teste ser positivo ou de ela estar infectada com HIV?
29. a. Considere apenas a amostra de risco e determine a probabilidade de uma pessoa ter o vírus HIV, dado que seu teste de HIV foi positivo.
    b. Considere apenas a população geral, e determine a probabilidade de uma pessoa ter o vírus HIV, dado que seu teste de HIV foi positivo.
    c. Compare os resultados das partes (a) e (b). Por que razão acha que o médico faz perguntas sobre o modo de vida do cliente durante uma consulta após um teste de HIV?
30. a. Considere apenas a amostra de risco e determine a probabilidade de uma pessoa não ter o vírus HIV, dado que o resultado do teste foi positivo.
    b. Considere apenas a amostra da população geral e determine a probabilidade de uma pessoa não ter o vírus HIV, dado que o resultado do teste foi positivo.
    c. Compare os resultados das partes (a) e (b). Se o leitor fosse o médico, como agiria no caso de um teste positivo de HIV para uma pessoa de cada grupo?

3-4 Exercícios B: Além do Básico

31. Determine a probabilidade de que, em 25 pessoas selecionadas aleatoriamente,
    a. Não haja duas com a mesma data de aniversário.
    b. Ao menos duas tenham a mesma data de aniversário.
32. a. Determine uma fórmula de não obter A ou B em um único experimento. Isto é, dê uma expressão para $P(\overline{A \text{ ou } B})$.
    b. Determine uma fórmula para a probabilidade de não obter A ou não obter B em uma única prova; isto é, dê uma expressão para $P(\overline{A} \text{ ou } \overline{B})$.
    c. Compare os resultados das partes (a) e (b). São os mesmos ou são diferentes?
33. Devemos extrair aleatoriamente duas cartas, sem reposição, de um baralho bem misturado. Determine a probabilidade de obter um 10 na primeira extração e uma carta de paus na segunda.

$\frac{4}{52} \times \frac{12}{51} = 0{,}076 \quad v\,0{,}235$

3-5 Probabilidades por Meio de Simulações

A determinação direta de probabilidades de eventos às vezes é muito difícil. Eventualmente os resultados, embora corretos, não são os que esperávamos. Em lugar de confiar exclusivamente nos princípios abstratos da teoria das probabilidades, a simulação pode vir em nosso auxílio.

DEFINIÇÃO

Uma **simulação** de um experimento é um processo que se comporta como o próprio experimento, produzindo resultados análogos.

EXEMPLO Em técnicas de teste sobre seleção de sexo, os pesquisadores médicos precisam conhecer probabilidades relacionadas com o sexo de nascituros. Admitindo que os sexos masculino e feminino tenham a mesma probabilidade, descreva um experimento que simule o sexo em nascimentos.

SOLUÇÃO Uma simulação consiste no simples arremesso de uma moeda, com "cara" representando *masculino* e "coroa" representando *feminino*. Outra abordagem consiste em utilizar um programa de computador, como STATDISK ou Minitab, para gerar 0s e 1s, com 0 representando *masculino* e 1 representando *feminino*. Tais números devem ser gerados de maneira que os 0s e os 1s sejam igualmente prováveis, como no caso de geradores de números aleatórios de uma distribuição uniforme. Mostra-se a seguir um resultado típico gerado por computador. Com base nesse resultado, podemos utilizar a aproximação da probabilidade por freqüência relativa para estimar $P(\text{masculino}) = 6/10$, porque há 6 homens entre os 10 nascimentos.

$$\begin{array}{cccccccccc} 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ \text{H} & \text{H} & \text{M} & \text{H} & \text{M} & \text{M} & \text{M} & \text{H} & \text{H} & \text{H} \end{array}$$

EXEMPLO O *problema do aniversário* é um exercício clássico em probabilidade. Trata-se de determinar a probabilidade de que, em uma turma de 25 estudantes, ao menos dois tenham a mesma data de aniversário. Ignorando os anos bissextos, descreva uma simulação do experimento que dê os aniversários de 25 estudantes em uma turma.

SOLUÇÃO Inicialmente, representemos as datas de aniversário como números inteiros de 1 a 365, como segue:

1 = 1 de Janeiro
2 = 2 de Janeiro
•
•
•
365 = 31 de Dezembro

Com esta representação, basta gerarmos inteiros entre 1 e 365, em lugar dos dias e meses separados. Utilizando qualquer fonte de inteiros igualmente prováveis (como os geradores aleatórios uniformes em STATDISK, Minitab ou a calculadora TI-83), podemos gerar uma relação de 25 inteiros aleatórios entre 1 e 365. Essa lista pode ser ordenada para facilitar a verificação da coincidência de "aniversários". Repetindo muitas vezes esse processo, podemos simular muitas turmas diferentes e estimar então a probabilidade de que, em uma turma de 25 estudantes, ao menos 2 tenham a mesma data de aniversário.

EXEMPLO A fabricante de telefones celulares Delmarva Communications Company vem experimentando uma taxa de 6% de defeitos. O controlador de qualidade sabe que os telefones são produzidos em lotes de 250 e que, em média, há 15 defeitos por lote. Ele deseja saber a variação típica do número de defeitos. Descreva uma simulação de 250 telefones celulares fabricados com uma taxa de 6% de incidência de defeitos.

SOLUÇÃO Com o STATDISK, Minitab, ou a calculadora TI-83, gere 250 inteiros, cada inteiro entre 1 e 100. Os inteiros 1,2,3,4,5,6 representarão os telefones defeituosos, enquanto 7,8,9, ..., 100 representarão os aparelhos perfeitos. Ordenados os 250 inteiros, torna-se fácil achar o número de "defeitos", que são os números entre 1 e 6, encontrados no começo da lista ordenada.

EXEMPLO Utilize Minitab para simular 500 jogadas de um par de dados e, com base nos resultados, estime $P(7)$.

SOLUÇÃO Podemos utilizar as opções Minitab `Calc/Random Data/Integer` para simular 500 jogadas de um único dado, com os resultados armazenados na coluna C1. As mesmas opções valem para um segundo dado, com os resultados armazenados na coluna C2. O comando `LET C3 = C1 + C2` cria uma coluna C3 que consiste nas 500 somas dos dois dados. Damos a seguir um resumo dos resultados de Minitab (obtidos com a aplicação de `Stat/Tables/Tally`). Por esses resultados, vemos que 7 ocorreram 84 vezes entre as 500 provas, e assim estimamos $P(7) = 84/500 = 0{,}168$. (A aplicação das regras da probabilidade dão $P(7) = 6/36 = 0{,}167$.)

C3	Contagem
2	13
3	29
4	34
5	56
6	73
7	84
8	61
9	55
10	45
11	32
12	18
N=	500

É de extrema importância construir a simulação com todo o cuidado, a fim, de que ela reproduza o mais fielmente possível as circunstâncias efetivas. Teríamos cometido sério erro no exemplo precedente se tivéssemos gerado os totais 2, 3, 4, ..., 12 correspondentes aos dois dados, como se esses totais fossem igualmente prováveis. Os resultados se assemelhariam a totais com os dados no sentido de que eles fossem inteiros entre 2 e 12, inclusive; deixando de simular cada dado individual e somar seus valores, não estaríamos imitando os verdadeiros dados. Um tal erro ocasionaria resultados bastante enganosos.

Além de permitir resolver problemas que, de outra forma, poderiam parecer insolúveis, as simulações podem ser utilizadas para verificar resultados de cálculos com probabilidades. Um problema que tem despertado muita atenção nos últimos anos é o *Problema de Monty Hall*, baseado no velho jogo de televisão "Let's Make a Deal" (Façamos um Negócio), patrocinado por Monty Hall. Suponha que um competidor tenha escolhido uma de três portas, após ser informado de que duas delas não escondem coisa alguma, mas que atrás da terceira está um Corvette vermelho novo. Em seguida, o patrocinador abre uma das portas que o competidor não escolheu, mostrando que não há coisa alguma atrás dela. Ele oferece então ao competidor a chance de ficar com a primeira escolha ou mudar para a porta que não foi aberta. Qual deveria ser a opção do competidor? A solução está longe de ser óbvia e não vamos calculá-la aqui, mas, com a teoria das probabilidades, é possível mostrar que o competidor deveria mudar sua escolha, porque a probabilidade de ganhar seria então de 2/3. Uma alternativa para este cálculo teórico consiste em simular o jogo com um amigo. A simulação deverá mostrar que mudar a escolha é melhor do que manter a primeira escolha, porque o competidor ganhará 2/3 das vezes. De acordo com a revista *Chance*, as escolas de administração em instituições como Harvard e Stanford utilizam este problema para melhorar a capacidade de decisão dos estudantes.

3-5 Exercícios A: Habilidades e Conceitos Básicos

1. Simule um experimento para registrar o número de meninas em famílias de 3 filhos, utilizando os resultados da Loteria Pick Three de Maryland, do Conjunto de Dados 12 do Apêndice B. Cada linha de 3 algarismos representará os 3 filhos em uma família, um número par representará um homem e um número ímpar representará uma mulher. Com base nas 50 famílias simuladas, qual é a probabilidade estimada de obter uma família com 3 meninas, quando se escolhe aleatoriamente uma família com 3 filhos? Compare o resultado simulado com a probabilidade efetiva, que é 0,125.
2. Considere os resultados da Loteria Pick Three de Maryland, do Conjunto de Dados 12 do Apêndice B, ignorando o terceiro algarismo em cada linha. Represente um homem por um algarismo par, e uma mulher por um algarismo ímpar. Com base nas 50 famílias simuladas, qual é a probabilidade estimada de obter uma família com 2 meninas, quando se escolhe aleatoriamente uma família com dois filhos? Compare o resultado simulado com a probabilidade real de 0,25.

3. A Telektronic Company fabrica telefones celulares em lotes de três e vem experimentando uma taxa global de defeitos de 10%. Recorrendo aos resultados da loteria Pick Three de Maryland, do Conjunto de Dados 12 do Apêndice B, represente por 0 um aparelho defeituoso, e por 1,2,3, ..., 9 os aparelhos sem defeito. Represente por linha de três algarismos um lote simulado de telefones celulares e utilize os 50 lotes simulados para estimar a probabilidade de ao menos um defeito em um lote. Compare a probabilidade estimada com o resultado teórico de 0,271.
4. Refaça o Exercício 3 se a taxa global de defeitos é de 20%. Represente por 0 e 1 os aparelhos defeituosos. Compare a probabilidade estimada com o resultado teórico de 0,488.
5. Sabemos que, quando se joga um dado equilibrado, a probabilidade de obter 1 é 1/6, ou 0,167. Qual é a probabilidade estimada quando se simula a jogada de um dado com os elementos da loteria Pick Three de Maryland (Conjunto de Dados 12 do Apêndice B)? (*Sug.:* Represente com os 150 algarismos as jogadas simuladas de um dado, ignorando quaisquer resultados que não sejam 1, 2, 3, 4, 5 ou 6.)
6. Um estudante responde ao acaso cada uma das três questões de um teste do tipo verdadeiro/falso. Utilize os dados da Loteria de Maryland Pick Three (Conjunto de Dados 12 do Apêndice B) para estimar a probabilidade de exatamente uma resposta certa entre as três.

3-5 Exercícios B: Além do Básico

7. Com os dados da loteria Pick Three de Maryland (Conjunto 12 do Apêndice B), simule 50 famílias com 3 filhos cada. Represente os meninos por números pares e as meninas por números ímpares.
   a. Ache o número médio de meninas em uma família.
   b. Ache o desvio-padrão do número de meninas.
8. O segundo exemplo desta seção descreve um método para simular uma turma com 25 aniversários. Use Minitab, ou STATDISK, ou uma calculadora TI-83, ou qualquer outra fonte de números entre 1 e 365 (como uma lista telefônica) para simular uma turma. Ordene os resultados e verifique se há ao menos dois aniversários coincidentes. Descreva detalhadamente o processo utilizado. No caso de ter utilizado Minitab ou STATDISK, obtenha uma cópia impressa dos resultados. (Veja também o Projeto para Computador no final deste capítulo.)

3-6 Contagem

Consideremos um problema de probabilidade cogitado seriamente por milhões de americanos esperançosos. Quais são as chances de ganhar na loteria? Na loteria do estado de Nova York, devemos escolher 6 números entre 1 e 54. Se sair a mesma combinação de 6 números escolhida, o apostador ganhará milhões de dólares. Há ainda alguns prêmios menores, mas são relativamente insignificantes. Poderíamos aplicar a Regra 2 da Seção 3-2 para achar a probabilidade de ganhar nessa loteria. Essa regra, que exige resultados igualmente prováveis, afirma que a probabilidade de um evento A se obtém aplicando $P(A) = s/n$, onde s é o número de maneiras como A pode ocorrer e n é o número total de resultados possíveis. Na loteria do estado de Nova York, há apenas uma maneira de ganhar o grande prêmio: Escolher a mesma combinação de 6 números que sai na loteria. Sabendo que há apenas uma maneira de ganhar, passemos a determinar o número total de resultados, isto é, quantas combinações de 6 números são possíveis. Relacionar todas as possibilidades uma a uma exigiria cerca de 4 anos de trabalho. Poderíamos construir um diagrama em árvore, mas teria 120 milhas de altura e violaria as regras do espaço aéreo. Devemos encontrar um meio mais prático de calcular o número total de possibilidades. Nesta seção introduziremos métodos eficientes para determinar tais números. Voltaremos ao problema da loteria após apresentarmos alguns princípios básicos. Começamos com a *regra fundamental da contagem*.

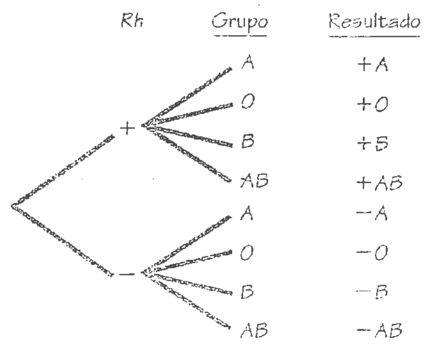


Fig. 3-10 Diagrama em árvore para tipos sanguíneos/fatores Rh.

Regra Fundamental da Contagem
Dados dois eventos, o primeiro dos quais pode ocorrer de m maneiras distintas e o segundo pode ocorrer de n maneiras distintas, então os dois eventos conjuntamente podem ocorrer de $m \cdot n$ maneiras distintas.

Por exemplo, se um médico laboratorista deve escolher aleatoriamente 1 dentre os 2 tipos de Rh (positivo, negativo) e 1 dos 4 grupos sanguíneos (A, O, B, AB), o número total de possibilidades é 2.4 = 8. Podemos ver a razão da multiplicação na Figura 3-10, onde ilustramos as diferentes possibilidades por meio de um diagrama em árvore. A regra fundamental da contagem

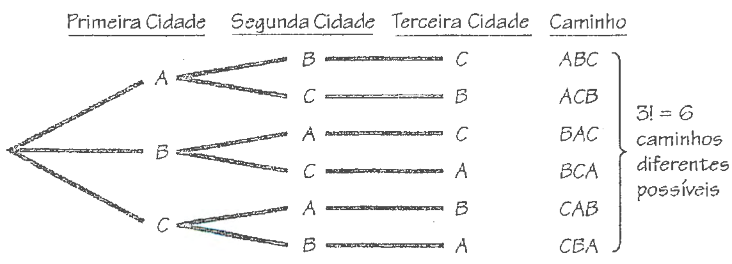


Fig. 3-11 Diagrama em árvore para roteamentos.

se estende facilmente a situações que envolvem mais de três eventos, conforme ilustrado no exemplo a seguir.

EXEMPLO Ao planejarmos um computador, se definimos um *byte* como uma seqüência de 8 bits, e cada bit deve ser 0 ou 1, quantos bytes diferentes são possíveis? (Costuma-se usar um byte para representar um caráter individual, como uma letra, um algarismo, ou um símbolo de pontuação. Por exemplo, em um sistema de codificação a letra A é representada por 01000001.)

SOLUÇÃO Como cada bit só pode ocorrer de duas maneiras (0 ou 1), e temos uma seqüência de 8 bits, o número total de possibilidades distintas é dado por

$$2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 = 256$$

Há, pois, 256 bytes distintos possíveis.

EXEMPLO Ao planejar pesquisas, os entrevistadores procuram minimizar o efeito causado pela ordem em que as questões são apresentadas. (Isto porque algumas questões influenciam as respostas das questões que seguem.) Se o Gallup planeja fazer uma pesquisa junto a consumidores formulando 5 questões aos entrevistados, quantas versões distintas da pesquisa são necessárias de modo a incluir todas as ordenações?

SOLUÇÃO Para qualquer pesquisa em particular, há 5 escolhas possíveis para a primeira questão, 4 escolhas restantes para a segunda questão, 3 escolhas para a terceira questão, 2 para a quarta e apenas 1 escolha para a quinta questão. O número total de arranjos possíveis é, pois,

$$5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$$

Isto é, o Gallup necessitaria de 120 versões da pesquisa para incluir todas as ordenações possíveis.

No exemplo precedente, vimos que é possível dispor 5 questões em 5.4.3.2.1 = 120 ordens possíveis. Essa solução particular pode ser generalizada utilizando a definição do símbolo ! e a seguinte *regra do fatorial.*

Notação

O símbolo **fatorial** ! denota o produto dos inteiros positivos em ordem decrescente. Por exemplo, $4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$. Por definição, $0! = 1$. (Muitas calculadoras têm a tecla !)

Regra do Fatorial

Uma coleção de n objetos diferentes pode ser ordenada de $n!$ maneiras distintas. (Esta **regra do fatorial** traduz o fato de que o primeiro objeto pode ser escolhido de n maneiras diferentes, o segundo objeto pode ser escolhido de $n - 1$ maneiras distintas, e assim por diante.)

A Secretária Aleatória

Eis um problema clássico de probabilidade: Uma pessoa expede 50 cartas e envelopes diferentes para 50 pessoas distintas, mas as cartas são misturadas aleatoriamente antes de serem postas nos envelopes. Qual é a probabilidade de que ao menos uma carta vá para o envelope correto? Embora à primeira vista a probabilidade possa parecer pequena, é efetivamente de 0,632. Mesmo com um milhão de cartas e um milhão de envelopes, ela continua sendo 0,632. A demonstração ultrapassa de muito o âmbito deste livro.

EXEMPLO Os problemas de roteamento costumam envolver aplicações da regra do fatorial. A AT&T (American Telephone and Telegraph) deseja estabelecer roteamentos para chamadas telefônicas através das redes mais curtas. A Federal Express deseja efetuar suas entregas através das rotas mais curtas. Suponha que um vendedor de computadores deva visitar 3 cidades distintas denotadas por A, B e C. Quantos caminhos são possíveis?

SOLUÇÃO Pela regra do fatorial, vemos que as 3 diferentes cidades (A, B, C) podem ser dispostas de $3! = 6$ maneiras distintas. Na Figura 3-11, vemos que há 3 escolhas para a primeira cidade e 2 escolhas para a segunda. Com isto, resta apenas 1 escolha para a terceira cidade. O número de arranjos possíveis para as 3 cidades é, pois, $3 \cdot 2 \cdot 1 = 6$.

EXEMPLO Em virtude de seus bons resultados em um curso de estatística, o leitor foi contratado pelo Gallup e sua primeira atribuição é fazer uma pesquisa na capital de cada um dos 50 estados americanos. Ao planejar a viagem, o leitor quer determinar o número de caminhos distintos possíveis. Quantos são esses caminhos?

SOLUÇÃO Pela regra do fatorial, sabemos que 50 elementos podem ser ordenados de 50! maneiras diferentes. Ou seja, as 50 capitais de estados podem ser dispostas de 50! maneiras, de forma que o número de caminhos diferentes é 50!, ou

30,414,093,201,713,378,043,612,608,166,064,768,844,377, 641,658,960,512,000,000,000,000

Trata-se de um número descomunalmente grande, que bem merece o símbolo ! utilizado para fatoriais.

O exemplo precedente é uma variação de um problema clássico denominado *problema do caixeiro viajante*. É especialmente interessante porque o grande número de possibilidades mostra que não podemos utilizar um computador para calcular a distância de cada caminho. O tempo necessário para o mais rápido computador calcular o caminho mais curto possível é

1,000,000,000,000,000,000,000,000,000,000,000,000,000,000 séculos

que é demasiadamente longo! Vêm sendo estudadas outras maneiras eficientes para resolver tais problemas.

Com a regra da contagem do fatorial, determinamos quantas diferentes maneiras de dispor um número de objetos são possíveis em algum tipo de seqüência ordenada. A regra do fatorial nos diz quantos arranjos são possíveis quando se tomam todos os n elementos distintos de um conjunto. Às vezes, entretanto, desejamos selecionar apenas alguns dentre os n elementos. Se, em uma pesquisa em capitais dos estados, temos tempo para visitar apenas quatro capitais, o número de caminhos diferentes possíveis é $50 \cdot 49 \cdot 48 \cdot 47 = 5.527.200$. Outra maneira de obter este mesmo resultado é calcular

$$\frac{50!}{46!} = 50 \cdot 49 \cdot 48 \cdot 47 = 5.527.200$$

Vamos observar, nesse cálculo, que os fatoriais 46! e 50! (parte) se cancelam, restando apenas os fatores 50, 49, 48 e 47 no numerador. De modo geral, se dispomos de n elementos diferentes e queremos escolher r dentre eles, o número de arranjos possíveis é $n!/(n - r)!$, como em 50!/46!. Essa generalização é conhecida como *regra dos arranjos*.

A Segurança nos Números

Alguns hotéis abandonaram a tradicional chave dos quartos, substituindo-a por uma chave eletrônica com um código numérico. Um computador central muda o código de acesso a um quarto assim que um hóspede deixa o hotel. Uma chave eletrônica típica tem 32 posições diferentes, que ou são perfuradas ou permanecem intocadas. Essa configuração comporta 2^{32}, ou 4.294.967.296 códigos diferentes possíveis, não sendo, assim, prático fabricar um conjunto completo de chaves ou tentar forçar uma entrada ilegal por tentativa-e-erro.

Regra dos Arranjos (Quando os elementos são todos distintos)

O número de **arranjos** (ou seqüências) de r elementos escolhidos dentre n elementos (sem repetição) é

$$_nP_r = \frac{n!}{(n - r)!}$$

Algumas calculadoras calculam automaticamente os valores de $_nP_r$. Na TI-83, por exemplo, introduzimos o valor de n e acionamos MATH, PRB, $_nP_r$ e o valor de r. Mesmo que sua calculadora não tenha tal dispositivo, ainda assim é fácil calcular $n!/(n - r)!$ utilizando a tecla do fatorial identificada com !.

É importante termos em mente que a regra dos arranjos exige as seguintes condições:

- Devemos ter um total de n elementos *diferentes*. (A regra não se aplica se alguns dos elementos são idênticos.)
- Devemos selecionar r dentre os n elementos (sem repetição).
- Ordenações distintas dos mesmos elementos devem ser consideradas arranjos diferentes.

Quando empregamos os termos arranjos ou seqüências, está implícito que *a ordem deve ser levada em conta*. As letras ABC comportam seis arranjos distintos: ABC, ACB, BAC, BCA, CAB, CBA. (Mais adiante estudaremos as *combinações*, em que a ordem dos elementos não é levada em conta.) No exemplo que segue, pede-se para achar o número total de seqüências diferentes possíveis. Isso sugere a aplicação da regra dos arranjos.

EXEMPLO No planejamento de um programa noturno da rede de televisão NBC, devem ser escolhidos 6 shows dentre 30 disponíveis. Quantas programações diferentes são possíveis?

SOLUÇÃO Devemos selecionar $r = 6$ dentre $n = 30$ programas disponíveis. Aqui a ordem tem importância, porque os espectadores são outros mais tarde. Como a ordem influi, devemos calcular o número de arranjos, como segue:

$$_nP_r = \frac{n!}{(n - r)!} = \frac{30!}{(30 - 6)!} = 427.518.000$$

Como há um total de 427.518.000 programações, torna-se impossível considerar cada uma individualmente.

A regra dos arranjos pode ser encarada como uma extensão da regra fundamental da contagem. Aplicando-a ao exemplo precedente, temos o seguinte: Com 30 shows disponíveis, e com a estipulação de que devemos escolher 6 deles, sabemos que há 30 escolhas para o primeiro tempo, 29 escolhas para o segundo tempo e assim por diante. O número total de arranjos possíveis é, pois,

$$30 \cdot 29 \cdot 28 \cdot 27 \cdot 26 \cdot 25 = 427.518.000;$$

mas $30 \cdot 29 \cdot 28 \cdot 27 \cdot 26 \cdot 25$ é, na verdade, $30! \div 24!$ [ou $30! \div (30 - 6)!$]. De modo geral, quando selecionamos r dentre n objetos, o número de arranjos possíveis é $n! \div (n - r)!$, o que é expresso pela regra dos arranjos.

Às vezes devemos determinar o número de arranjos, quando alguns dos elementos são idênticos entre si. Aplica-se então a seguinte variante da regra dos arranjos.

Pesquisas Sensíveis

Os entrevistados às vezes relutam em responder perguntas sobre tópicos considerados delicados, como sexo, furto de um empregado etc. Stanley Warner (Universidade de York, Ontário) elaborou um esquema que leva a resultados mais precisos em tais casos. Como exemplo, pergunte a um empregado se ele já roubou no passado, pedindo-lhe também que jogue uma moeda. Os empregados devem responder **não** se eles não furtaram e a moeda dá "cara". Em caso contrário, devem responder **sim**. Os empregados tendem a ser honestos em suas respostas porque a jogada da moeda ajuda a preservar sua privacidade. Pode-se então recorrer à teoria das probabilidades a fim de obter resultados mais precisos.

Regra dos Arranjos (Quando alguns elementos são idênticos)

Se há n elementos com n_1 iguais, n_2 iguais, ..., n_k iguais, o número de **permutações** (arranjos com a totalidade dos elementos) de todos os n elementos é

$$\frac{n!}{n_1!\, n_2! \cdots n_k!}$$

EXEMPLO Os exemplos clássicos da regra dos arranjos são os que mostram que as letras da palavra *Mississippi* comportam 34.650 permutações distintas, enquanto as letras da palavra *statistics* comportam 50.400 permutações. Vamos considerar as letras *DDDDRRRRR* incluídas em uma discussão do teste de repetições para aleatoriedade (Seção 13-7). As letras representam uma seqüência de coca-colas tipos dieta (D) e regular (R). De quantas maneiras podemos dispor as letras *DDDDRRRRR*?

SOLUÇÃO Na seqüência *DDDDRRRRR* temos $n = 9$ elementos, com $n_1 = 4$ iguais e $n_2 = 5$ iguais. O número de permutações (arranjos com a totalidade dos elementos) é:

$$\frac{n!}{n_1!\, n_2!} = \frac{9!}{4!\, 5!} = \frac{362.880}{2880} = 126$$

Na Seção 13-7 levamos em conta a existência de 126 seqüências diferentes possíveis de *DDDDRRRRR*; agora vemos como se obtém esse resultado.

O exemplo precedente envolveu n elementos, cada um deles pertencente a uma de duas categorias. Quando há apenas duas categorias, podemos estipular que x elementos são iguais e os outros $n - x$ são também iguais, de modo que a fórmula das permutações se simplifica para

$$\frac{n!}{(n-x)!\,x!}$$

Esse resultado particular será utilizado em experimentos binomiais, a serem introduzidos na Seção 4-3.

Quando queremos selecionar r elementos de um conjunto de n elementos distintos *sem levar em conta a ordem*, estamos considerando *combinações*, e não arranjos. Ou seja, **quando contamos separadamente ordenações diferentes dos mesmos elementos, temos um problema de arranjos; mas quando as diferentes ordenações não são contadas separadamente, temos um problema de combinações**, e podemos aplicar a regra seguinte:

Regra das Combinações

O número de combinações de r elementos extraídos de um conjunto de n elementos diferentes é

$$_nC_r = \frac{n!}{(n-r)!\,r!}$$

Algumas calculadoras estão preparadas para calcular automaticamente $_nC_r$. Com a TI-83, por exemplo, introduzimos o valor de n, acionamos MATH, PRB, $_nC_r$, e introduzimos r.

É importante termos em mente que, ao aplicar a regra das combinações, valem as seguintes condições:

- Devemos ter um total de n elementos distintos.
- Devemos selecionar r dentre os n elementos (sem repetição).
- Devemos considerar como uma mesma combinação ordenamentos diferentes dos mesmos elementos.

Como a distinção entre a regra dos arranjos e a regra das combinações nem sempre é clara, damos a seguir um exemplo que enfatiza essa diferença.

Quantas Embaralhadas?

Após extensas pesquisas, o matemático de Harvard, Persi Diaconis, constatou serem necessárias sete embaralhadas para que se tenha uma mistura completa das cartas de um baralho. A mistura é completa no sentido de que todos os arranjos possíveis são igualmente prováveis. Mais de sete embaralhadas não têm efeito significativo, e menos de sete não são suficientes. Os crupiês dos cassinos raramente embaralham sete vezes, de modo que os baralhos não ficam misturados adequadamente. Alguns jogadores profissionais conseguiram tirar vantagem desse fato.

EXEMPLO O Conselho Curador da faculdade do autor tem 9 membros. Cada ano é eleito um comitê de três pessoas para supervisionar os prédios e o campus. São eleitos também, anualmente, um presidente, um vice-presidente e um secretário.

a. Na eleição do comitê de edifícios e campus, quantos comitês diferentes, compostos de 3 pessoas, podem ser formados?

b. Quando o Conselho elege o presidente, o vice-presidente e o secretário, quantas chapas são possíveis?

SOLUÇÃO Observe que a ordem é indiferente na eleição do comitê dos edifícios e campus. Mas na eleição dos dirigentes, as diferentes ordens são consideradas distintas.

a. Trata-se aqui do número de combinações de $r = 3$ pessoas a serem selecionadas dentre 9. Temos:

$$_9C_3 = \frac{n!}{(n-r)!\,r!} = \frac{9!}{(9-3)!\,3!} = \frac{362.880}{4320} = 84$$

b. Aqui, desejamos o número de seqüências (ou permutações) de $r = 3$ pessoas a serem escolhidas dentre as $n = 9$. Temos:

$$_9P_3 = \frac{n!}{(n-r)!} = \frac{9!}{(9-3)!} = \frac{362.880}{720} = 504$$

Há 84 possibilidades de formação de comitês diferentes compostos de 3 membros, mas há 504 chapas diferentes.

Voltaire Vence a Loteria

Em 1729,o filósofo francês Voltaire ficou rico elaborando um esquema para vencer a loteria de Paris. O governo havia instituído uma loteria para compensar a desvalorização das apólices municipais. Como a cidade acrescentou grandes quantias, resultou que o valor dos prêmios ultrapassava o preço de todos os bilhetes. Voltaire formou um grupo que comprava todos os bilhetes da loteria de um mês e ganhou durante mais de um ano. Um apostador da loteria do estado de Nova York tentou ganhar uma parcela de um prêmio excepcionalmente grande, resultante da falta de ganhadores em sorteios prévios. Ele pretendia emitir um cheque de 6.135.756,00 dólares abrangendo todas as combinações, mas o estado não aceitou, sob a alegação de que a natureza da loteria teria sido alterada.

As técnicas de contagem apresentadas nesta seção costumam ser usadas em problemas de probabilidades. Os exemplos seguintes ilustram tais aplicações.

EXEMPLO Na loteria do estado de Nova York, um apostador ganha o primeiro prêmio se acertar a combinação de 6 números extraídos do conjunto de 1 a 54. Determine a probabilidade de um apostador ganhar. (O jogador não precisa escolher os 6 números na mesma ordem em que são extraídos; a ordem é irrelevante.)

SOLUÇÃO Como devem ser extraídos 6 números diferentes de um total de 54 possibilidades distintas, o número total de combinações é

$$_{54}C_6 = \frac{54!}{(54-6)!\,6!} = \frac{54!}{48!\;6!} = 25.827.165$$

Jogando apenas em uma combinação, a probabilidade de o apostador ganhar é de 1/25.827.165.

No exemplo precedente, vimos que a probabilidade de ganhar na loteria de Nova York é de apenas 1/25.827.165. Na Seção 3-2, vimos que a probabilidade de ser atingido por um raio em um ano é de 1/701.000. A comparação dessas duas probabilidades mostra que, em determinado ano, há uma chance muito maior de ser atingido por um raio do que de ganhar na loteria com uma única aposta. Naturalmente, é possível aumentar a chance de ganhar na loteria adquirindo muitos bilhetes — uma estratégia que não parece aconselhável.

EXEMPLO Um despachante da UPS envia um caminhão a 8 localidades diferentes. Se a ordem das entregas é aleatória, determine a probabilidade de o percurso resultante ser o menor possível.

SOLUÇÃO Com 8 localidades, há 8!, ou 40.320 percursos possíveis. Entre essas 40.320 possibilidades diferentes, apenas dois percursos são mínimos (na verdade, o mesmo percurso, em direções diferentes). Portanto, há uma probabilidade de apenas 2/40.320, ou 1/20.160, ou 0,0000496 de o percurso escolhido ser o menor possível.

Regiões com Diferentes Códigos Telefônicos

Periodicamente, as companhias telefônicas dividem regiões com um único código de área em sub-regiões com dois ou mais códigos diferentes, porque o número crescente de áreas de fax e de internet praticamente exauriu os números que podem ser enquadrados em um único código. Um número de telefone de sete algarismos não pode começar com 0 nem 1, mas levando em conta todas as outras possibilidades, obtemos 8 · 10 · 10 · 10 · 10 · 10 · 10 = 8.000.000 de números distintos possíveis! Mesmo assim, a cidade de Nova York, tendo conseguido sobreviver 80 anos com um único código de área, 212, recentemente foi subdividida em duas áreas com códigos 212 e 718. Muitas outras regiões também foram subdivididas dessa forma.

Nesta seção apresentamos cinco processos diferentes de contagem. Ao decidir qual iremos aplicar, devemos levar em conta vários aspectos relevantes. O resumo a seguir pode ajudar.

- Há uma seqüência de eventos em que o primeiro pode ocorrer de m maneiras, o segundo pode ocorrer de n maneiras e assim por diante? Em caso afirmativo, aplique o princípio fundamental da contagem e multiplique m, n, etc.
- Há n objetos *distintos* e *todos* eles entrarão nos diferentes arranjos? Em caso afirmativo, aplique a regra do fatorial e calcule $n!$.
- Há n elementos *diferentes* e apenas *alguns* deles entrarão em arranjos distintos? Em caso afirmativo, calcule $_nP_r$.
- Há n elementos com alguns deles ***idênticos*** uns aos outros, e devemos achar o número total de arranjos (permutações) com todos esses n elementos? Em caso afirmativo, utilize a expressão seguinte, onde n_1 elementos são iguais, n_2 são iguais etc.

$$\frac{n!}{n_1!\,n_2!\,\cdots\,n_k!}$$

- Há n elementos diferentes, dos quais *alguns* devem ser escolhidos, não importando a ordem de escolha? Temos então as *combinações* dos n elementos tomados r de cada vez, e devemos calcular $_nC_r$.

3-6 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-16, calcule a expressão dada.

1. $6!$	2. $11!$	3. $100!/97!$	4. $85!/82!$
5. $(10-4)!$	6. $(90-87)!$	7. $_6C_4$	8. $_6P_4$
9. $_{12}P_9$	10. $_{10}C_9$	11. $_{40}C_6$	12. $_{40}P_5$
13. $_nC_0$	14. $_nP_0$	15. $_nP_n$	16. $_nC_n$

17. O autor utiliza um sistema de segurança doméstica ADT que tem um código consistindo em 4 algarismos (0, 1, ..., 9) que deve ser introduzido na seqüência correta. Os algarismos podem ser repetidos no código.
    a. Quantas possibilidades distintas existem?
    b. Se um ladrão leva 5 segundos para tentar um código, quanto tempo levaria para tentar todas as possibilidades?
18. Há 12 membros na diretoria do Hospital Geral de Newport.
    a. Se eles devem eleger um presidente, um primeiro vice-presidente, um segundo vice-presidente e um secretário, quantas chapas de candidatos são possíveis?
    b. Se devem formar um subcomitê de ética com 4 membros, quantas são as possibilidades?
19. Cada número de inscrição no seguro social é uma seqüência de 9 algarismos. Qual é a probabilidade de serem gerados aleatoriamente 9 algarismos, obtendo-se o seu número de inscrição?
20. Uma fechadura típica do tipo de "combinação" abre com a seqüência correta de 3 números entre 0 e 49. Quantas seqüências são possíveis? (Um número pode ser usado mais de uma vez.) Tais seqüências são efetivamente combinações ou arranjos?
21. Ao fazer um teste de repetições para aleatoriedade (Seção 13-7), obtém-se os sexos dos pesquisados na ordem consecutiva seguinte: HHHHHHHHHHHMMMMMMMMM (H = homem, M = mulher). De quantas maneiras podemos ordenar estas letras?
22. Um percurso de entregas da Federal Express deve incluir paradas em 5 cidades.
    a. Quantos percursos diferentes são possíveis?
    b. Se o percurso é escolhido aleatoriamente, qual é a probabilidade de as cidades serem escolhidas em ordem alfabética?
23. Ao tentar decifrar uma mensagem interceptada da Líbia, um técnico decide relacionar todos os arranjos possíveis da palavra MGBTQRS. Quantos arranjos distintos há?
24. a. Se um casal planeja ter 8 filhos (é possível), quantas seqüências de sexos são possíveis?

b. Se um casal tem 4 filhos e 4 filhas, quantas seqüências distintas de sexos são possíveis?
c. Com base nos resultados das partes (a) e (b), qual é a probabilidade de um casal com 8 filhos ter 4 meninos e 4 meninas?

25. A loteria do estado de Nova York e as loterias de alguns outros estados costumavam prever a escolha de 6 números entre 1 e 40 inclusive.
a. Quantas escolhas diferentes são possíveis?
b. Escolhendo 6 números, qual a probabilidade de ganhar selecionando os mesmos números sorteados?
c. Quais são as chances contra ganhar em tal loteria?

26. Considere a mesma loteria do Exercício 25. Qual é a probabilidade de ganhar, se as novas regras agora exigem que o apostador escolha os 6 números na mesma ordem em que são sorteados?

27. No *Directory of Tunes and Musical Themes*, de Denys Parson, relaciona-se as melodias de mais de 14.000 canções, de acordo com o esquema seguinte: A primeira nota de cada canção é representada por um asterisco (*), e as notas sucessivas são representadas por *R* (*repetir* a nota anterior), *U* para uma nota acima, ou *D* para uma nota abaixo. A Quinta Sinfonia de Beethoven começa com *RRD. As melodias clássicas são representadas por suas primeiras 16 notas. Com esse esquema, quantas melodias clássicas é possível representar?

28. O Departamento de Pesca solicitou auxílio à Bell Laboratories para determinar a rota mais curta para obter amostras de locais no Golfo do México. Quantas rotas diferentes são possíveis, se as amostras devem ser obtidas de 11 locais?

29. Deve-se formar um comitê de 4 membros selecionados dentre 50 agentes do FBI que não estejam trabalhando em um projeto especial.
a. Quantos comitês diferentes são possíveis?
b. Se as escolhas são aleatórias, qual é a probabilidade de obter os 4 agentes que estejam há mais tempo em serviço?

30. Começamos a suspeitar quando um pesquisador de genética seleciona aleatoriamente um grupo de 20 recém-nascidos e obtém consistentemente 10 meninos e 10 meninas. O pesquisador explica que é comum obtermos 10 meninos e 10 meninas em tais casos.
a. Selecionados aleatoriamente 20 recém-nascidos, quantas seqüências diferentes de sexo são possíveis?
b. De quantas maneiras 10 meninos e 10 meninas podem ser dispostos em seqüência?
c. Qual é a probabilidade de 10 meninos e 10 meninas em 20 crianças selecionadas aleatoriamente?
d. Com base nos resultados precedentes, concorda com a alegação do pesquisador, que é comum obter 10 meninos e 10 meninas em uma escolha aleatória de 20 crianças?

31. A Detroit Music Company adquiriu os direitos de 15 canções diferentes e planeja lançar um novo CD com 8 delas. Admitindo que a ordem das canções seja importante, quantos CDs diferentes são possíveis?

32. Em um departamento de montagem da Ford Motor Company, 8 peças diferentes devem ser montadas em um carro, mas a ordem de montagem é indiferente. O gerente decide determinar a seqüência mais eficiente, tentando todas as possibilidades. Quantas seqüências distintas são possíveis?

33. a. Quantos códigos CEP são possíveis, se cada código é uma seqüência de 5 algarismos?
b. Se um computador gera aleatoriamente 5 algarismos, qual é a probabilidade de gerar seu código CEP?

34. O diretor de programação da TV ABC decidiu apresentar shows de 30 minutos entre 8 horas e 10 horas da noite às segundas-feiras. Se ele dispõe de 22 shows, quantas programações são possíveis para esses intervalos nas segundas-feiras?

35. Em um caso de discriminação de idade contra a Darmin, Inc., ficou provado que, dos últimos 40 candidatos a emprego, apenas os 8 mais jovens foram contratados. Determine a probabilidade de escolher as 8 pessoas mais jovens em um grupo de 40 pessoas. Com base no resultado, parece estar havendo discriminação de idade?

36. Dá-se a seguir um trecho de *The Man Who Cast Two Shadows*, de Carol O'Connell: "A menina tinha apenas os números escritos a tinta na palma de sua mão ..., todos exceto os quatro últimos desapareceram em uma mancha de sangue ... Ela devia colocar as moedas nos telefones públicos e discar três números não tentados e em seguida os quatro que ela sabia. Se uma mulher atendesse, ela diria, 'Sou Kathy. Estou perdida'." Se cada chamada custa 25 centavos para Kathy e ela tenta todas as possibilidades exceto as que começam com 0 ou 1, qual é seu desembolso total?

37. Após testar a presença de radônio em 12 residências, um pesquisador começou a desconfiar do seu aparelho de teste, porque o nível de radônio medido em cada residência era sempre superior ao precedente. Isto é, os 12 resultados se dispunham em ordem ascendente. Se as residências foram selecionadas aleatoriamente, qual é a probabilidade desse arranjo particular? Com base no resultado, justifica-se sua desconfiança quanto ao aparelho?

38. Um trem de carga deve ter 12 carros de carvão, 5 carros carregados de madeira e 4 carros-tanque para transporte de combustível. Quantos arranjos distintos são possíveis?

39. O repórter Paul Wiseman de *USA Today* descreveu as velhas normas para códigos de área para telefone escrevendo sobre "códigos de área possíveis com 1 ou 0 como segundo algarismo. (Excluídos: códigos terminando em 00 ou 11, para chamadas gratuitas, serviços de emergência e outros casos especiais.)" Deveriam ser excluidos também os códigos começando com 0 ou 1. Quantos códigos diferentes seriam possíveis sob aquelas velhas normas?

40. Há 22 membros no comitê de orçamento designado dentre 100 membros do senado americano. Quantos comitês diferentes de 22 membros podem ser formados com os 100 senadores?

3-6 Exercícios B: Além do Básico

41. Uma regra comum de programação de computador é que os nomes de variáveis devem ter de 1 a 8 caracteres. O primeiro deve ser exclusivamente literal (qualquer uma das 26 letras), e os demais podem ser qualquer uma das 26 letras ou qualquer um dos 10 algarismos. Exemplos: A, BBB, e M3477K. Quantos nomes diferentes é possível formar?

42. a. Em uma reunião de cinco gerentes, se cada um deles troca aperto de mão com cada um dos outros exatamente uma vez, qual o número total de cumprimentos?
b. Se n gerentes trocam apertos de mão com cada um dos demais exatamente uma vez, qual o número total de cumprimentos?
c. De quantas maneiras 5 pessoas podem sentar-se em torno de uma mesa redonda? (Admita que, se todos se movem para a direita, a disposição dos assentos continua a mesma.)
d. De quantas maneiras n pessoas podem sentar-se em torno de uma mesa redonda?

43. Muitas calculadoras ou computadores não podem calcular diretamente valores de 70! ou superiores. Para n muito grande, $n!$ pode ser aproximado por $n! = 10^K$, onde o valor de K é dado por $K = (n + 0{,}5)\log n + 0{,}39908993 - 0{,}43429448n$.

a. Calcule 50! utilizando a tecla fatorial de uma calculadora e também pela aproximação dada acima.
b. O Departamento de Pesca solicitou ajuda a Bell Laboratories para determinar o menor percurso para obter amostras de 300 localidades no Golfo do México. Há 300! rotas diferentes possíveis. Calculando-se 300!, quantos algarismos aparecem no resultado?

44. Os computadores podem "pensar"? De acordo com o *teste de Turing*, pode-se admitir que um computador pensa se, quando uma pessoa se comunica com ele, ela crê estar se comunicando com outra pessoa, e não com um computador. Em um experimento no Boston's Computer Museum (Museu dos Computadores de Boston), cada um dentre 10 juízes se comunicou com 4 computadores e com 4 outras pessoas, devendo distinguir entre eles.
a. Suponha que o primeiro juiz não possa distinguir entre os 4 computadores e as 4 pessoas. Se esse juiz faz suposições aleatórias, qual é a probabilidade de identificar corretamente os 4 computadores e as 4 pessoas?
b. Suponha que nenhum dos 10 juízes possa distinguir entre computadores e pessoas e que, assim, fazem suposições aleatórias. Com base no resultado da parte (a), qual é a probabilidade de que todos os 10 juízes façam suposições corretas? (Este evento nos levaria a concluir que os computadores não podem "pensar", mas pelo critério de Turing podem.)

Vocabulário

experimento	evento composto
evento	regra da adição
evento simples	mutuamente excludentes
espaço amostral	regra dos eventos complementares
aproximação de uma probabilidade por freqüência relativa	diagrama em árvore
abordagem clássica da probabilidade	eventos independentes
lei dos grandes números	eventos dependentes
amostra aleatória de um elemento	regra da multiplicação
amostra aleatória	probabilidade condicional
complemento	simulação
probabilidade subjetiva	regra fundamental da contagem
chances contra	símbolo fatorial
chances a favor	regra do fatorial
	regra dos arranjos
	regra das combinações

Revisão

Neste capítulo introduzimos os conceitos básicos da teoria da probabilidade. Na Seção 3-2 apresentamos as definições e a notação básicas, inclusive a representação de eventos por letras como A. Definimos as probabilidades de eventos simples como

$$P(A) = \frac{\text{número de ocorrências de } A}{\text{número de repetições do experimento}} \quad \text{(freqüência relativa)}$$

$$P(A) = \frac{\text{número de ocorrências de } A}{\text{número de eventos simples diferentes}} = \frac{s}{n} \quad \text{(para resultados igualmente prováveis)}$$

Notamos que a probabilidade de um evento impossível é 0, a probabilidade de um evento certo é 1, e que, para qualquer evento A, $0 \le P(A) \le 1$. Também, $\overline{A}$ denota o complemento do evento A, ou seja, $\overline{A}$ indica que o evento A não ocorre.

Após a Seção 3-2, passamos a considerar eventos compostos, que envolvem mais de um evento. De modo geral, associamos a conjunção *ou* à adição, e a conjunção *e* à multiplicação. Tenhamos sempre em mente as seguintes considerações básicas.

- Se, em uma prova, queremos a probabilidade do evento A *ou* B, devemos aplicar a regra da adição, tendo, porém, o cuidado de não contar qualquer resultado mais de *uma vez*.
- Ao calcular a probabilidade de o evento A ocorrer em uma prova *e* o evento B ocorrer em outra prova, devemos aplicar a regra da multiplicação: Multiplicar a probabilidade do evento A pela probabilidade do evento B. *Cuidado:* Ao calcular a probabilidade do evento B, leve em conta o fato de o evento A já ter ocorrido.

Em alguns problemas de probabilidade, o maior obstáculo é determinar o número de resultados possíveis. A última seção deste capítulo abordou as seguintes técnicas de contagem, resumidas no final da Seção 3-6:

- Regra fundamental da contagem
- Regra do fatorial
- Regra dos arranjos (quando todos os elementos são diferentes)
- Regra dos arranjos (quando alguns elementos são iguais)
- Regra das combinações

Grande parte do assunto dos capítulos que seguem se refere a inferência estatística com base em probabilidades. Como exemplo da abordagem básica utilizada, consideremos o teste da afirmação de que uma moeda usada em um jogo de cara-ou-coroa é equilibrada. Se jogamos a moeda 10 vezes e obtemos 10 caras, podemos inferir desse resultado:

1. A moeda é equilibrada, e a seqüência de 10 caras seguidas é um acaso.
2. A moeda não é equilibrada (é viciada).

A decisão do estatístico sobre qual inferência é correta se baseia na probabilidade de obter 10 caras seguidas, a qual, neste caso, é tão pequena (1/1024) que a inferência de moeda viciada é a melhor escolha. Podemos ver aqui o papel importante desempenhado pela probabilidade nos métodos-padrão de inferência estatística.

Exercícios de Revisão

Nos Exercícios 1-8, utilize os dados da Tabela 3-3, que resumem resultados de um estudo de 1000 mortes, selecionadas aleatoriamente, de homens com idade de 45 a 64 (com base em dados de "Chartbook on Smoking, Tobacco and Health," USDHEW).

1. Se, dos 1000 indivíduos, 1 é selecionado aleatoriamente, determine a probabilidade de se obter um fumante.
2. Se, dos 1000 indivíduos, 1 é selecionado aleatoriamente, determine a probabilidade de se obter um fumante ou alguém que tenha morrido em conseqüência de doença cardíaca.
3. Escolhidos aleatoriamente dois indivíduos, determine a probabilidade de ambos terem morrido de câncer.
4. Escolhido aleatoriamente um indivíduo, determine a probabilidade de obter um não-fumante que tenha morrido de câncer.
5. Escolhido aleatoriamente um indivíduo, determine a probabilidade de obter alguém que tenha morrido de câncer ou de doença cardíaca.

TABELA 3-3

	Causa da Morte		
	Câncer	Doença Cardíaca	Outros
Fumante	135	310	205
Não-fumante	55	155	140

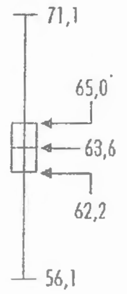
6. Escolhidos aleatoriamente três indivíduos diferentes, determine a probabilidade de serem todos fumantes.
7. Escolhido aleatoriamente um indivíduo, determine a probabilidade de se tratar de um fumante, dado que morreu de câncer.
8. Escolhido um indivíduo aleatoriamente, determine a probabilidade de obter alguém que tenha morrido de câncer, dado que se tratava de um fumante. O fumo e a incidência de câncer são eventos independentes? Por que sim ou por que não?
9. Ao delinear um processo de fabricação para um dispositivo de armazenamento de memória de computador, a configuração inicial tem um resultado positivo de 16%. Isto é, 16% dos dispositivos são aceitáveis, e 84% são defeituosos. Fabricados 12 desses dispositivos, qual a probabilidade de obter ao menos 1 que seja bom? Se é de grande importância obter ao menos 1 dispositivo bom para fins de teste, a probabilidade resultante é adequada?
10. Com base na experiência passada, um estudante que vai às aulas de carro sabe que, quando ele excede o limite de velocidade, tem uma chance de 2% de ser multado. Qual a probabilidade de não ser multado, se ele excede o limite em todos os 150 dias do ano letivo? Se esse estudante não pode arcar com o custo do aumento do seguro em função da multa, que decisão a probabilidade resultante sugere?
11. A diretoria do Jefferson Valley Bank tem 8 membros.
    a. Formado um comitê de 3 membros mediante seleção aleatória, qual a probabilidade de serem escolhidos os 3 mais ricos?
    b. Se a diretoria deve eleger um presidente, um vice-presidente e um secretário, quantas chapas são possíveis?
12. A New England Life Insurance Company emite apólices temporárias por 1 ano em nome de 12 homens, todos com 27 anos de idade. Com base em dados do Departamento de Saúde e Recursos Humanos, cada um deles tem 99,82% de chance de sobreviver por um ano. Qual é a probabilidade de todos os três sobreviverem um ano?
13. Ao apostar em *par* na roleta, há 38 resultados igualmente prováveis, mas somente 2, 4, 6, ..., 36 são ganhadores.
    a. Determine a probabilidade de ganhar ao apostar em *par*.
    b. Determine a chance contra o ganho ao apostar em *par*.
    c. Os cassinos pagam apostas vencedoras de acordo com chances descritas como 1:1. Qual é seu lucro líquido se aposta $5 em *par* e ganha?
14. Uma questão de um teste de história exige que 5 eventos sejam relacionados na ordem cronológica adequada. Escolhida uma ordem aleatória, qual a probabilidade de ser a ordem correta?
15. Um entrevistador afirma que 12 eleitores foram selecionados aleatoriamente de uma população de 200.000 eleitores (30% dos quais são republicanos), e todos os 12 eram republicanos. O entrevistador afirma que esse resultado pode ocorrer facilmente por pura chance. Determine a probabilidade de obter 12 republicanos quando se escolhem aleatoriamente 12 eleitores dessa população. Com base no resultado, a afirmação do entrevistador parece correta?
16. Em uma turma de estatística de 8 mulheres e 8 homens, formam-se aleatoriamente 2 grupos de 8 estudantes. Qual a probabilidade de todos serem mulheres no primeiro grupo e todos serem homens no segundo grupo? (*Sug.*: Determine o número de permutações de MMMMMMMMHHHHHHHH.)

Exercícios Cumulativos de Revisão

1. Utilize a tabela de freqüências a seguir, que dá a distribuição de idade dos americanos mortos por acidente (com base em dados do Conselho de Segurança Nacional).
    a. Supondo que a classe de "75 ou mais" tenha 80 como ponto médio, calcule a idade média dos americanos mortos por acidente.
    b. Com a mesma hipótese utilizada na parte (a), calcule o desvio-padrão das idades resumidas na tabela.
    c. Escolhida aleatoriamente uma das 95.277 idades, determine a probabilidade de ser inferior a 15 ou superior a 64.

Idade	Número
0–4	3.843
5–14	4.226
15–24	19.975
25–44	27.201
45–64	14.733
65–74	8.499
75 ou mais	16.800

    d. Escolhida aleatoriamente uma das 95.277 idades, determine a probabilidade de ser inferior a 15 ou estar entre 5 e 44.
    e. Escolhidas aleatoriamente duas idades da tabela, determine a probabilidade de ambas estarem entre 0 e 4 anos.
2. O diagrama em caixa a seguir ilustra as alturas (em polegadas) de um grande conjunto de mulheres selecionadas aleatoriamente.



    a. Escolhida aleatoriamente uma dessas mulheres, determine a probabilidade de sua altura estar entre 56,1 e 62,2 polegadas.
    b. Escolhida aleatoriamente uma dessas mulheres, determine a probabilidade de sua altura ser inferior a 62,2 ou superior a 63,6 polegadas.
    c. Escolhidas aleatoriamente duas mulheres, determine a probabilidade de ambas terem altura entre 62,2 e 63,6 polegadas.
3. Realizou-se uma pesquisa das relações entre os diâmetros, alturas e volumes de certo tipo de cerejeira na Allegheny National Forest, na Pennsylvania. O gráfico ramo-e-folhas abaixo representa os diâmetros (em centímetros) de uma amostra de 15 árvores utilizadas no estudo (com base em dados do *Minitab Student Handbook*).

10.	578
11.	00123447
12.	099
13.	3

    a. Determine o diâmetro médio.
    b. Determine o diâmetro mediano.
    c. Ache o desvio-padrão dos diâmetros.
    d. Ache a variância dos diâmetros.
    e. Que expressão identifica melhor o nível de mensuração dos diâmetros: nominal, ordinal, intervalar ou razão?
    f. Escolhido aleatoriamente um desses diâmetros, determine a probabilidade de ser inferior a 11,0 cm.
    g. Escolhidos aleatoriamente dois diâmetros diferentes, determine a probabilidade de serem ambos superiores a 11,9 cm.
    h. Selecionados aleatoriamente dois diâmetros, com reposição, determine a probabilidade de serem ambos superiores a 11,9 cm.
    i. Escolhido aleatoriamente um desses diâmetros, determine a probabilidade de ser inferior a 11,0 cm ou de estar entre 10,6 e 11,6 cm.

Projeto para Computador

É possível achar probabilidades por computador simulando-se um experimento. Recorde que uma simulação de um experimento é um processo que se comporta da mesma maneira que o próprio experimento, produzindo assim resultados semelhantes. No Exercício 31 da Seção 3-4 pedia-se para calcular a probabilidade de obter ao menos duas pessoas com mesma data de aniversário, em um grupo aleatório de 25 pessoas. Em lugar de fazer cálculos teóricos, utilizaremos STATDISK ou Minitab para simular o experimento. E em lugar de gerar datas de aniversário, geraremos números entre 1 e 365, que representam os diferentes aniversários possíveis. (Ignoraremos os anos bissextos.) Por exemplo, o número 5 gerado representa 5 de janeiro, e 364 representa 30 de dezembro. Podemos trabalhar com os próprios números gerados, não sendo necessário identificar o dia e o mês correspondentes. Gerados esses 25 números, podemos ordená-los, o que facilita ver se ao menos 2 deles são o mesmo "aniversário."

Utilize STATDISK ou Minitab para gerar 25 "aniversários", e ordene-os para ver se ao menos 2 coincidem. Registre o resultado. Repita o experimento até ter confiança em que sua probabilidade estimada está aproximadamente correta.

Para utilizar STATDISK, escolha primeiro `Data` na barra principal do menu; escolha então `Uniform Generator` e utilize `Format` para fixar em 0 o número de casas decimais (porque queremos gerar números inteiros). Passe a gerar um tamanho de amostra de 25, com um máximo de 365 e um mínimo de 1. Utilize a característica `Sample Editor's Format` para ordenar e apresentar os 25 aniversários simulados, para ver se ao menos 2 coincidem.

Para utilizar Minitab, selecione as opções `Calc/Random Data, Integer` e introduza 25 como número de linhas de dados, C1 como a coluna em que vai armazenar os resultados, 1 como valor mínimo, 365 como valor máximo, e clique `OK`. Selecione agora `Manip/Sort` e introduza C1 para a coluna, C1 para a coluna em que vai armazenar os resultados, C1 para a coluna pela qual os resultados serão selecionados, e clique `OK`. Os resultados ordenados devem aparecer na primeira coluna.

Pode-se usar também a calculadora TI-83 para este experimento. Acione `MATH`, selecione `PRB`, a seguir `randInt` e introduza `randInt (1,365,25)` para gerar 25 aniversários simulados. Os aniversários podem ser armazenados na lista L1 acionando `STO` e L1. Os dados podem ser ordenados (ou escolhidos) acionando-se `STAT`, e escolhendo-se `Sort A`, e introduzindo L1. Acione agora `STAT` e selecione `Edit` para visualizar os aniversários ordenados.

DOS DADOS À DECISÃO

Teste de Uso de Drogas em Candidatos a Emprego

De acordo com a American Management Association, a maioria das empresas dos EUA está fazendo teste do uso de drogas em alguns empregados e candidatos a emprego. O U.S. National Institute on Drug Abuse alega que cerca de 15% das pessoas na faixa etária de 18-25 consomem drogas ilegalmente. Allyn Clark, de 21 anos de idade, bacharel por uma universidade, candidatou-se a um emprego na Acton Paper Company, submeteu-se a um teste de drogas e não obteve o emprego. Clark suspeita que não tenha passado no teste de drogas, mesmo não sendo usuário. Ao investigar a situação no departamento de pessoal da companhia, constatou que o teste tem uma sensibilidade de 99%, e assim apenas 1% dos usuários dão resultado negativo. Outrossim, o teste tem 98% de especificidade, o que significa que apenas 2% dos não-usuários são incorretamente identificados como usuários. Allyn sentiu-se aliviado com essas cifras, porque pareciam traduzir um teste bastante confiável que costuma dar bons resultados — mas deveria mesmo estar aliviado? A companhia pode ter a certeza de que não está contratando usuários de drogas? A tabela a seguir apresenta os dados para Allyn e outros 1999 candidatos a emprego. Com base nesses resultados, determine a probabilidade de um "falso positivo", isto é, determine a probabilidade de selecionar aleatoriamente um dos indivíduos cujo teste foi positivo e obter alguém que não use drogas. Determine também a probabilidade de um "falso negativo", isto é, determine a probabilidade de selecionar aleatoriamente um dos indivíduos cujo teste foi negativo e obter alguém que seja usuário de drogas. As probabilidades desses resultados errados são suficientemente baixas para não preocuparem nem os candidatos a emprego nem a Acton Paper Company?

	Usuários	Não-usuários
Testes com resultado positivo	297	34
Testes com resultado negativo	3	1666

ATIVIDADES EM GRUPO

1. *Atividade na Classe:* Divida a turma em grupos de três ou quatro e estime P(2 meninas em 3 nascimentos) utilizando uma simulação com moedas. Descreva o processo exato utilizado e os resultados obtidos.
2. *Atividade na Classe:* Divida a turma em grupos de três ou quatro e utilize tachinhas para estimar a probabilidade de que, quando jogada, uma tachinha caia com a ponta para cima. Quantas tentativas são necessárias para dar um resultado que se afigure razoavelmente preciso?
3. *Atividade na Classe*: Divida a turma em grupos de três ou quatro. Em cada grupo, estabeleça uma probabilidade subjetiva do evento de uma mulher ser eleita presidente dos EUA em 2008. Os valores dos diversos grupos são aproximadamente os mesmos, ou são muito diferentes? A con-

cordância entre os grupos indicaria que os resultados são precisos?

4. *Atividade em Aula*: Cada estudante deve receber uma página diferente retirada de uma velha lista telefônica. Simule uma seleção de 25 aniversários utilizando os três últimos algarismos dos números dos telefones, escolhidos aleatoriamente (ignorando os que excederem 365). Após registrar os 25 aniversários, determine se há dois coincidentes. Os resultados das turmas podem ser combinados para formar uma estimativa da probabilidade de que, em 25 pessoas escolhidas aleatoriamente, ao menos duas tenham a mesma data de aniversário.
5. *Atividade Fora de Aula*: *O Método da Captura-Recaptura.* Os biólogos marinhos costumam utilizar o método da captura-recaptura para estimar o tamanho de uma população, como a de peixes em um lago. Esse método consiste em capturar uma amostra da população, etiquetar cada elemento da amostra e devolvê-lo à população. Coleta-se mais tarde uma segunda amostra e contam-se os elementos etiquetados entre a população total capturada. Como exemplo, suponha uma amostra de 50 peixes capturada e etiquetada. Suponha ainda que uma segunda amostra (capturada mais tarde) consista em 100 peixes com 20 deles etiquetados — o que sugere que, quando um peixe é capturado, a probabilidade de ser etiquetado é estimada em 0,20; isto é, 20% da população de peixes são etiquetados. Como a amostra original de 50 consistia em peixes todos etiquetados, podemos estimar em 250 peixes o tamanho da população ($50 \div 20/100 = 250$).

   Não é fácil capturar e recapturar peixes efetivamente, mas podemos simular um experimento utilizando uma coleção uniforme de elementos como bolinhas de mesma cor. Os elementos "capturados" na primeira amostra devem ser substituídos por elementos análogos de cor diferente. Ilustre o método da captura-recaptura planejando e realizando um tal experimento. Tomando um grande número de bolinhas coloridas, selecione uma amostra de 50. Substitua as bolinhas selecionadas por bolinhas de outra cor, restituindo-as à população original. Selecione uma segunda amostra e estime então o tamanho da população, comparando o resultado com o tamanho efetivo da população obtido pela contagem de todos os elementos.
6. *Atividade em Aula*: Divida a turma em grupos de dois. Na Seção 3-5 abordamos o problema "Monte Hall", que, de acordo com a revista *Chance*, foi utilizado no estudo da tomada de decisões nas faculdades de economia de Harvard e Stanford. O problema se baseia em um jogo de TV patrocinado por Monte Hall. Comece escolhendo um dos membros do time para servir como patrocinador. O outro membro do time é o concorrente, e há três portas numeradas 1, 2, 3. O patrocinador deve escolher *aleatoriamente* uma das portas e a escolha não deve ser revelada ao concorrente. Admita que o patrocinador tenha estipulado o prêmio de um Corvette novo, vermelho, atrás da porta que foi escolhida aleatoriamente; atrás das outras duas portas não há nada. O concorrente deve escolher uma das três portas. Após o concorrente revelar que porta escolheu, o patrocinador deve escolher uma porta "vazia" e informar o concorrente de que esta porta não esconde coisa alguma. O patrocinador deve agora oferecer ao concorrente a escolha de ficar com a porta original ou optar pela outra porta que não foi aberta. Após o concorrente ter comunicado sua decisão, o patrocinador deve anunciar que o concorrente ganhou (ou não ganhou) o Corvette. Registre o resultado juntamente com a decisão do concorrente (manter a porta, ou não). Repita o jogo 20 vezes, com o concorrente mantendo a porta 10 vezes e trocando a porta 10 vezes. Inverta então os papéis e jogue o jogo outras 20 vezes. Determine a proporção das vezes em que o jogo foi ganho mantendo-se a porta e a proporção das vezes em que foi ganho mudando-se a porta. Com base nos resultados, qual é a melhor estratégia: manter ou trocar?
7. *Atividade em Aula:* Divida a turma em grupos de dois, a fim de fazer um experimento destinado a mostrar uma abordagem de questões delicadas como uso de drogas, roubo ou atividade sexual. Para os fins desta atividade, utilizaremos a pergunta inócua: "Você nasceu entre 1 de janeiro e 31 de março?" Esperamos que 1/4 das respostas seja "sim", mas admitamos que a questão seja bastante delicada e as pessoas relutem em responder honestamente. Um membro do time (o "entrevistador") deve pedir ao outro (o "pesquisado") que jogue uma moeda e escreva "não" em um pedaço de papel se o pesquisado *não* nasceu entre 1 de janeiro e 31 de março *e* a moeda dá "cara"; se o pesquisado nasceu entre aquelas datas ou se a moeda dá "coroa", deve ser escrita a resposta "sim". Inverta os papéis de modo que as respostas sejam obtidas de cada time. Supõe-se que os pesquisados tendam a ser mais honestos porque a jogada da moeda protege sua privacidade. Combine todos os resultados e analise-os, a fim de determinar a proporção das pessoas nascidas entre 1 de janeiro e 31 de março. A precisão dos resultados pode ser verificada confrontando-os com as datas reais de nascimento. O experimento pode ser repetido com uma questão mais delicada, mas essa questão não é dada aqui, porque o autor já recebe correspondência suficiente.

entrevista

Barbara Carvalho

Diretora de Pesquisa do Marist College

Lee Miringoff

Diretor do Marist College Institute for Public Opinion

Barbara Carvalho e Lee Miringoff relatam os resultados de suas pesquisas em muitas entrevistas para a imprensa e a televisão, incluindo noticiári para NBC, CBS, ABC, FOX e a TV estatal. Lee Miringoff aparece regularmente no programa "Today" da NBC.

Que tipos de pesquisa fazem?

Fazemos pesquisa de interesse público. Pesquisamos assuntos de interesse público como índice de aprovação de autoridades da cidade de Nova York, do estado de Nova York e do país em geral. Não trabalhamos para partidos políticos, candidatos políticos ou grupos de *lobby*. Somos subvencionados independentemente pelo Marist College, e não recebemos qualquer outro subsídio que possa sugerir algum vínculo com qualquer grupo, sobre qualquer assunto em particular. Nosso programa é efetivamente um programa educacional, amplamente reconhecido porque os resultados são divulgados publicamente. Os repórteres passaram a depender de nossos resultados não só por sua precisão e profissionalismo, mas também porque sabem que nossas pesquisas são independentes, sem compromisso com qualquer fonte de notícias em particular, como acontece com muitas pesquisas.

Quem realiza suas entrevistas e quais são seus fundamentos?

Todas as nossas entrevistas são feitas por estudantes pagos, que são treinados na técnica de entrevistar e no assunto específico em que estão trabalhando. Os estudantes podem escolher o caminho da pesquisa de opinião pública e coleta de dados. Os estudantes de ciência política constituem um grupo natural, mas também contratamos muitos graduados em comunicação, bem como estudantes interessados em estatística, análise computacional, psicologia, economia, sociologia, administração e marketing.

Recomendaria a estatística a estudantes de áreas como história, governo ou ciências sociais?

Sem dúvida. Eles devem ter ao menos um curso de pesquisa relacionado com a análise estatística básica que lhes dê uma base para lidar com os números que vão encontrar — seja qual for o campo em que vão trabalhar. O estudo da estatística é importante para entender um aspecto do conhecimento e constitui uma chave para a abertura de outros caminhos a ser trilhados. A estatística permeia as disciplinas. Os estudantes fatalmente a encontrarão em algum ponto de suas carreiras. Pode ser na avaliação do seu trabalho ou do seu local de trabalho, como pode envolver aspectos de mercado ou promocionais. As pesquisas invadem hoje nossa cultura a tal ponto que o estudante irá encontrá-las em estágios posteriores de suas vidas, seja em suas carreiras ou simplesmente como cidadãos. O povo é hoje bombardeado com informações de pesquisas, sendo, pois, absolutamente vital que o cidadão esteja em condições de avaliar a precisão e o valor das mesmas.

Que conceitos da estatística utiliza?

A estatística entra em cena logo na amostragem, antes mesmo de chegarmos à análise dos dados. Com auxílio da estatística, determinamos o tamanho da amostra e formulamos uma estimativa do que seria estatisticamente significante. Na análise dos dados, aplicamos a estatística descritiva básica à maior parte de nossos estudos. Alguns estudos acadêmicos vão até a análise de regressão.

Como seleciona seus pesquisados?

Para uma pesquisa de âmbito estadual, escolhemos indivíduos na proporção dos registros de eleitores nos municípios. Diferentes municípios acusam diferentes taxas de recusa, e se fôssemos seleciono aleatoriamente as pessoas no estado como um todo, não teríamos um imagem uniforme de como o estado se apresenta. Fazemos a estratificação por municípios e utilizamos uma coleta aleatória de números, de modo a obter números listados e não-listados.

Qual seu tamanho típico de amostra?

Cerca de quatro indivíduos; mas eles são selecionadas com muito cuidado. Na realidade, poderiam ser de 400 a 1200 ou 1500. Se quiséssemos fazer uma análise de subgrupos dentro de nosso grupo populacional aumentaríamos o tamanho de nossa amostra de modo que tivéssemos subgrupos como homens *versus* mulheres, ou diferentes grupos regionais, ou diferentes grupos de renda.

O processo político sofre efetivamente influênci dos resultados de pesquisas?

Embora a maioria das pesquisas que o povo vê sejam pesquisas públicas, a realidade é que o processo político é influenciado por pesquisas privadas que o público nunca vê. Ninguém concorre a um posto elevado sem utilizar uma pesquisa privada.

Rayare Andrade

4 Triola

Distribuições de Probabilidade

4-1 Aspectos Gerais

Identificam-se os objetivos do capítulo. Descrevem-se as variáveis aleatórias e as distribuições de probabilidade em geral, examinando-se algumas distribuições especiais de probabilidade.

4-2 Variáveis Aleatórias

Apresentam-se nesta seção variáveis aleatórias e distribuições de probabilidade discretas e contínuas. Dão-se métodos para determinar a média, a variância e o desvio-padrão de uma distribuição de probabilidade. Define-se o valor esperado de uma distribuição.

4-3 Experimentos Binomiais

Definem-se os experimentos binomiais. Calculam-se probabilidades em experimentos binomiais com auxílio da fórmula binomial de probabilidade, de uma tabela de probabilidades, ou de um pacote estatístico.

4-4 Média, Variância e Desvio-padrão da Distribuição Binomial

Calculam-se a média, a variância e o desvio-padrão de uma distribuição binomial, discutindo-se também a interpretação desses valores.

4-5 A Distribuição de Poisson

Apresenta-se a distribuição de Poisson como outro exemplo especial e importante de uma distribuição discreta de probabilidade. Uma característica fundamental da distribuição de Poisson é que ela se aplica a ocorrências de um evento em um intervalo especificado de tempo, distância, área ou unidade semelhante.

Prof.: Jorge
Disc.: Estatística

Problema do Capítulo

Seriam os acidentes aéreos com jatos da USAir apenas uma coincidência?

Recentemente, os meios de comunicação deram grande cobertura ao fato de que jatos da USAir estavam envolvidos em quatro dentre sete acidentes aéreos graves consecutivos nos Estados Unidos. A USAir detém 20% das linhas domésticas. Se a USAir, detendo 20% das linhas, fosse tão segura quanto qualquer outra companhia de aviação, seria de esperar que a USAir tivesse 20% dos sete desastres ocorridos, ou seja, 1,4. Como a USAir teve quatro acidentes em lugar de apenas um ou dois, é lícito concluirmos que a USAir não é tão segura quanto as outras companhias, ou que o envolvimento da USAir é apenas uma coincidência? Essa conclusão depende da probabilidade de que os eventos ocorram por puro acaso. Vamos considerar as duas questões seguintes:

1. *Dado que a USAir detém 20% de todas as linhas domésticas e supondo que a USAir seja tão segura quanto qualquer outra companhia aérea e que os acidentes com avião sejam eventos independentes que ocorrem aleatoriamente, qual é a probabilidade de a USAir ter quatro dentre sete acidentes consecutivos?*
2. *Para decidir se a USAir não é segura ou se é vítima de coincidência, a probabilidade relevante é a descrita no item precedente? Em lugar de procurarmos a probabilidade de a USAir ter exatamente quatro dentre sete acidentes, há outra pergunta que melhor reflita o problema de saber se a USAir é tão segura quanto as outras?*

A primeira questão pode ser respondida facilmente com auxílio dos métodos apresentados neste capítulo. A segunda questão é mais difícil e exige estudo sério, mas é extremamente importante para identificar corretamente o evento que constitui a chave do problema. Abordaremos ambas as questões mais adiante neste capítulo.

4-1 Aspectos Gerais

No Capítulo 2 vimos que é possível explorar um conjunto de dados utilizando gráficos (como um histograma ou o *boxplot*), medidas de tendência central (como a média) e medidas de variação (como o desvio-padrão). No Capítulo 3 abordamos os princípios básicos da teoria das probabilidades. Neste capítulo, combinaremos esses conceitos ao estabelecermos distribuições de probabilidade que descrevem o que *provavelmente* acontecerá, em lugar do que efetivamente aconteceu. No Capítulo 2, construímos tabelas de freqüência e histogramas utilizando valores *observados*; neste capítulo, entretanto, vamos construir distribuições de probabilidade apresentando resultados possíveis juntamente com as freqüências relativas que *esperamos*, à vista do conhecimento de circunstâncias relevantes.

Suponha que o gerente de um cassino suspeite de fraude em uma mesa de dados. Ele pode comparar a distribuição de freqüências relativas dos resultados amostrais efetivos com um modelo teórico que descreva a distribuição de freqüências esperada com um dado equilibrado — que deve ter um histograma de freqüências relativas semelhante ao da Figura 4-1(a); já o histograma de freqüências relativas da Figura 4-1(b) espelha um dado viciado,

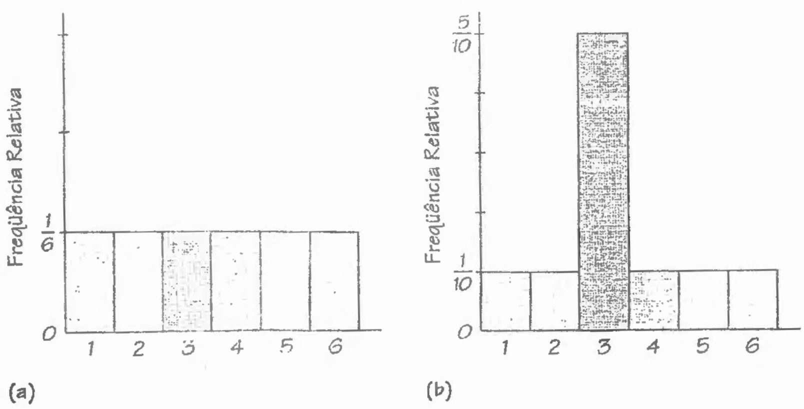


Fig. 4-1 Histograma de resultados de dados para (a) um dado equilibrado e (b) um dado viciado.

em que o aparecimento do 3 é favorecido. Se alguém é apanhado jogando com um dado como o da figura 4-1(b) certamente terá problema.

Na Figura 4-1(a), vemos freqüências relativas baseadas não em resultados efetivos, mas em nosso conhecimento das probabilidades dos resultados de um dado equilibrado. Essa figura representa uma distribuição de probabilidades que serve de modelo para a distribuição de freqüência de uma população teoricamente perfeita. Essencialmente, podemos descrever a tabela de freqüências e o histograma para um dado jogado um número infinito de vezes. Com tal conhecimento da população de resultados, estamos em condições de determinar características importantes, como a média e o desvio-padrão. O restante deste livro e o próprio âmago da inferência estatística se baseiam em algum conhecimento de distribuições de probabilidade. Examinaremos inicialmente o conceito de variável aleatória e a seguir consideraremos distribuições importantes que têm várias aplicações reais.

4-2 Variáveis Aleatórias

Nesta seção abordaremos os conceitos de variável aleatória, distribuição de probabilidade e processos para cálculo da média e do desvio-padrão de uma distribuição de probabilidade. Veremos que uma variável aleatória tem um número para cada resultado de um experimento e que uma distribuição de probabilidades associa uma probabilidade a cada resultado numérico de um experimento.

Muitas situações cotidianas podem ser usadas como experimentos que dão resultados correspondentes a algum valor, e tais situações podem ser descritas por uma variável aleatória.

DEFINIÇÃO

Uma **variável aleatória** é uma variável (geralmente representada por x) que tem um valor numérico único (determinado aleatoriamente) para cada resultado de um experimento.

Exemplos de variáveis aleatórias:

x = número de acidentes com aviões da USAir dentre sete acidentes aéreos selecionados aleatoriamente

x = número de mulheres entre 10 empregados recém-admitidos

x = número de alunos que não compareceram à aula de estatística hoje

x = altura de um adulto do sexo masculino selecionado aleatoriamente.

Empregamos o termo *variável aleatória* para descrever o valor que corresponde ao resultado de determinado experimento. A palavra *aleatória* indica que em geral só conhecemos aquele valor depois de o experimento ter sido realizado.

Profetas dos Lucros

Muitos livros e programas de computador pretendem ajudar a ganhar na loteria. Alguns se baseiam na teoria de que determinados números "devem" aparecer (devendo, por isso, ser escolhidos) porque não têm aparecido ultimamente; outros apelam para o fato de que alguns números são "frios" (devendo ser evitados), porque não têm aparecido com freqüência; outras ainda apelam para a astrologia, a numerologia ou sonhos. Como as combinações vencedoras na loteria são eventos independentes, tais teorias não têm valor algum. Uma abordagem válida consistiria em escolher números "raros", no sentido de que não são escolhidos por outras pessoas, de modo que, se ganhar, o apostador não tenha que repartir o "bolo" com muitos outros. Por essa razão, a combinação 1, 2, 3, 4, 5, 6 não é uma boa escolha, porque muitos a utilizam; enquanto 12, 17, 18, 33, 40, 46 é uma escolha muito melhor, pelo menos até ter sido divulgada neste livro...

EXEMPLO Um experimento consiste em selecionar aleatoriamente sete acidentes aéreos com vôos domésticos e contar os que envolvem aviões da USAir. Se a variável aleatória representa o número de acidentes com aviões da USAir, dentre sete ocorrências, esse experimento comporta os resultados 0, 1, 2, 3, 4, 5, 6, 7. (0 representa nenhum acidente com a USAir, 1 representa 1 acidente com a USAir e assim por diante.) A variável é aleatória no sentido de que só sabemos seu valor após havermos selecionado sete acidentes.

Na Seção 1-2 fizemos uma distinção entre dados discretos e dados contínuos. As variáveis aleatórias também podem ser discretas ou contínuas, e as duas definições que seguem são consistentes com as que foram dadas na Seção 1-2. Este capítulo aborda variáveis aleatórias discretas; as variáveis aleatórias contínuas serão objeto de capítulos seguintes.

DEFINIÇÕES

Uma **variável aleatória discreta** ou admite um número finito de valores ou tem uma quantidade enumerável de valores.

Uma **variável aleatória contínua** pode tomar um número infinito de valores, e esses valores podem ser associados a mensurações em uma escala contínua, de tal forma que não haja lacunas ou interrupções.

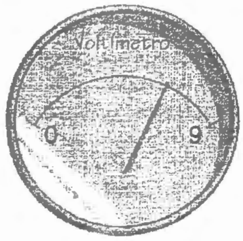
EXEMPLO

a. O número de espectadores que vêem um filme é um número inteiro, sendo, portanto, uma variável aleatória discreta. O dispositivo de contagem ilustrado na Figura 4-2(a) registra apenas números inteiros, podendo, assim, ser utilizado para obter valores de uma variável aleatória discreta.

b. A voltagem na pilha de um detector de fumaça pode ser qualquer valor entre 0 volts e 9 volts, sendo, por conseguinte, uma variável aleatória contínua. O voltímetro ilustrado na Figura 4-2(b) indica valores em uma escala contínua, gerando valores de uma variável aleatória contínua.

Além de identificar valores de uma variável aleatória, freqüentemente podemos atribuir uma probabilidade a cada um desses valores. Quando conhecemos todos os valores de uma variável

(a) Variável Aleatória Discreta: Contagem do número de espectadores.



(b) Variável Aleatória Contínua: Voltagem de uma pilha de detector de fumaça.

Fig. 4-2 Variáveis aleatórias discreta e contínua.

aleatória juntamente com suas respectivas probabilidades, temos uma distribuição de probabilidades, definida como segue.

DEFINIÇÃO

Uma **distribuição de probabilidades** dá a probabilidade de cada valor de uma variável aleatória.

EXEMPLO Suponha que a USAir detenha 20% de todas as linhas aéreas domésticas, e que todos os vôos tenham a mesma chance de um acidente. Se a variável aleatória x representa o número de acidentes com a USAir dentre sete acidentes escolhidos aleatoriamente, então a distribuição de probabilidades é dada pela Tabela 4-1.

TABELA 4-1 Distribuição de Probabilidade do Número de Acidentes com a USAir, dentre Sete Acidentes

x	$P(x)$
0	0,210
1	0,367
2	0,275
3	0,115
4	0,029
5	0,004
6	0+
7	0+

(Na Seção 4-3 veremos como obter as probabilidades da Tabela 4-1.) Por essa tabela, a probabilidade de 0 acidentes com a USAir (dentre sete acidentes) é 0,210; a probabilidade de um acidente é 0,367 etc. Os valores denotados por 0^+ representam probabilidades tão pequenas que equivalem a 0,000 quando arredondadas para três decimais. Preferimos não escrever 0,000 porque sugere (erroneamente) um evento impossível com probabilidade 0.

Há várias representações gráficas para uma distribuição de probabilidades; apresentaremos apenas o **histograma de probabilidade.** A Figura 4-3 é um histograma de probabilidades que se assemelha ao histograma de freqüências relativas do Capítulo 2, mas a escala vertical representa *probabilidades*, em lugar das correspondentes freqüências relativas.

Observe, na Figura 4-3, que, ao longo do eixo horizontal, os valores 0, 1, 2, ..., 7 estão localizados nos centros dos retângulos. Isso implica que os retângulos têm cada um a largura de 1 unidade, de modo que suas áreas são 0,210, 0,367 etc. Quando a área total de tal histograma de probabilidade é 1, as *probabilidades* são iguais às *áreas* dos retângulos correspondentes. No Capítulo 5 e nos capítulos seguintes veremos a importância e a utilidade dessa correspondência entre área e probabilidade.

Qualquer distribuição de probabilidades deve satisfazer as duas condições seguintes:

Condições para uma Distribuição de Probabilidades

1. $\Sigma P(x) = 1$, onde x toma todos os valores possíveis
2. $0 \leq P(x) \leq 1$ para todo x

A primeira condição afirma que a soma de todas as probabilidades individuais é 1 e se baseia na regra da adição para eventos mutuamente excludentes. Os valores da variável aleatória x representam todos os eventos possíveis em todo o espaço amostral, e assim temos a certeza (com probabilidade 1) de que um dos eventos ocorrerá. Aplicamos a regra simples da adição dos valores de $P(x)$ porque os diferentes valores de x correspondem a eventos que são mutuamente excludentes. Por exemplo, se escolhemos aleato-

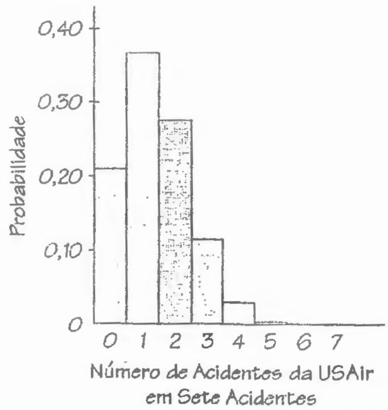


Fig. 4-3 Histograma de probabilidades para o número de desastres da USAir dentre sete desastres aéreos.

riamente sete acidentes aéreos e representamos por x o número de acidentes com a USAir, x não pode ser 4 e 5 ao mesmo tempo. Na Tabela 4-1 podemos ver que as probabilidades individuais têm efetivamente a soma 1. Outrossim, a regra de probabilidade (veja Seção 3-2) que afirma que $0 \leq P(A) \leq 1$ para qualquer evento A implica que $P(x)$ deve estar entre 0 e 1 para qualquer valor de x. Voltando à Tabela 4-1, vemos que cada valor de $P(x)$ está de fato entre 0 e 1. Como a Tabela 4-1 satisfaz duas dessas condições, é exemplo de uma distribuição de probabilidade. Uma distribuição de probabilidades pode ser dada por uma tabela, como a Tabela 4-1, ou por um gráfico, como na Figura 4-3, ou por uma fórmula, como nos dois exemplos seguintes.

EXEMPLO $P(x) = x/5$ (onde x toma os valores 0, 1, 2, 3) define uma distribuição de probabilidades?

SOLUÇÃO Para que fique definida uma distribuição de probabilidades, devem ser satisfeitas as duas condições anteriores. Ora,

$$\begin{aligned}\Sigma P(x) &= P(0) + P(1) + P(2) + P(3)\\ &= \frac{0}{5} + \frac{1}{5} + \frac{2}{5} + \frac{3}{5}\\ &= \frac{6}{5} \quad \text{(mostrando que } \Sigma P(x) \neq 1\text{)}\end{aligned}$$

Como a primeira condição não é satisfeita, concluímos que $P(x)$ dada neste exemplo não é uma distribuição de probabilidade.

EXEMPLO $P(x) = x/3$ (onde x pode ser 0, 1 ou 2) define uma distribuição de probabilidades?

SOLUÇÃO Para a função dada, temos que $P(0) = 0/3$, $P(1) = 1/3$ e $P(2) = 2/3$, e assim

1. $\Sigma P(x) = \frac{0}{3} + \frac{1}{3} + \frac{2}{3} = \frac{3}{3} = 1$
2. Cada um dos valores de $P(x)$ está entre 0 e 1 inclusive.

Como as duas condições são satisfeitas, a função $P(x)$ deste exemplo é uma distribuição de probabilidades.

Média, Variância e Desvio-Padrão

No Capítulo 2 vimos que há três características extremamente importantes de dados:

1. Valor representativo, como uma média
2. Medida de dispersão ou variação, como um desvio-padrão
3. Natureza ou forma da distribuição, como forma de sino.

O histograma de probabilidade permite-nos visualizar a natureza ou forma da distribuição. A média, a variância e o desvio-padrão traduzem outras características. Podemos achar a média, a variância e o desvio-padrão de uma distribuição de probabilidades aplicando as Fórmulas 4-1, 4-2, 4-3 e 4-4.

Fórmula 4-1 $\mu = \Sigma x \cdot P(x)$ **Média de uma distribuição de probabilidades**

Fórmula 4-2 $\sigma^2 = \Sigma[(x - \mu)^2 \cdot P(x)]$ **Variância para uma distribuição de probabilidades**

Fórmula 4-3 $\sigma^2 = [\Sigma x^2 \cdot P(x)] - \mu^2$ **Variância para uma distribuição de probabilidades**

Fórmula 4-4 $\sigma = \sqrt{[\Sigma x^2 \cdot P(x)] - \mu^2}$ **Desvio-padrão para uma distribuição de probabilidades**

Cuidado: A expressão $\Sigma x \cdot P(x)$ é a mesma que $\Sigma[x \cdot P(x)]$, ou seja, primeiro multiplique cada valor de x por sua probabilidade e em seguida adicione os resultados. Também, $\Sigma x^2 \cdot P(x)$ se calcula elevando-se cada valor de x ao quadrado, multiplicando-se cada quadrado pela probabilidade $P(x)$ correspondente e somando os resultados; isto é, $\Sigma x^2 \cdot P(x) = \Sigma[x^2 \cdot P(x)]$.

A TI-83 calcula a média e o desvio-padrão de uma distribuição de probabilidades. Introduza os valores de x na lista L1, introduza as probabilidades correspondentes na lista L2, selecione `STAT, CALC, 1-Var Stats` e introduza L1, L2 (com a vírgula). Após acionar a tecla ENTER, o valor exibido como $\bar{x}$ é efetivamente a média μ, e o valor mostrado com σx é o valor do desvio-padrão σ.

Regra de arredondamento para μ, σ^2 e σ

Ao utilizar as Fórmulas 4-1 a 4-4, aplique esta regra para arredondar os resultados:

Arredonde os resultados tomando uma decimal a mais além do número de casas decimais usadas na variável x. Se os valores de x são inteiros, arredonde μ, σ^2 e σ para uma decimal.

Às vezes é preciso aplicar uma regra de arredondamento diferente em virtude de circunstâncias especiais, como resultados que exigem mais casas decimais para terem sentido.

Ao calcularmos a média de uma distribuição de probabilidade, obtemos o valor médio que esperaríamos obter se pudéssemos repetir as provas indefinidamente. *Não obtemos* o valor que esperamos ocorrer com maior freqüência. Na realidade, obtemos em geral um valor médio que não pode ocorrer em nenhuma prova (como 1,5 meninas em 3 nascimentos). O desvio-padrão nos dá uma medida do quanto a distribuição de probabilidade se dispersa em torno da média. Um grande desvio-padrão reflete dispersão considerável, enquanto um desvio-padrão menor traduz menor variabilidade, com valores relativamente mais próximos da média. A regra prática da Seção 2.5 pode também auxiliar na interpretação do valor de um desvio-padrão. De acordo com essa regra, a maioria dos valores deve estar a menos de dois desvios-padrão da média; não é comum um valor diferir da média por mais de dois desvios-padrão.

EXEMPLO A Tabela 4-1 representa a distribuição de probabilidade do número de acidentes com a USAir, dentre sete acidentes selecionados aleatoriamente (supondo que a USAir detenha 20% dos vôos e que os acidentes sejam eventos independentes e aleatórios). Com a distribuição de probabilidade descrita na Tabela 4-1, suponha que repitamos o experimento que consiste em selecionar sete acidentes e que a cada vez achemos o número de acidentes com a USAir. Determine o número médio de acidentes com a USAir (entre sete), a variância e o desvio-padrão.

TABELA 4-2 Cálculo de m, σ^2 e σ para uma Distribuição de Probabilidade

x	$P(x)$	$x \cdot P(x)$	x^2	$x^2 \cdot P(x)$
0	0,210	0,000	0	0,000
1	0,367	0,367	1	0,367
2	0,275	0,550	4	1,100
3	0,115	0,345	9	1,035
4	0,029	0,116	16	0,464
5	0,004	0,020	25	0,100
6	0+	0,000	36	0,000
7	0+	0,000	49	0,000
Total	1,000 ↑ $\Sigma P(x)$	1,398 ↑ $\Sigma x \cdot P(x)$		3,066 ↑ $\Sigma x^2 \cdot P(x)$

SOLUÇÃO Na Tabela 4-2, as duas colunas à esquerda descrevem a distribuição de probabilidade dada anteriormente na Tabela 4-1. As três colunas à direita foram criadas para possibilitar os cálculos necessários.

Com as Fórmulas 4-1 e 4-3 e os resultados da tabela, obtemos

$$\mu = \Sigma x \cdot P(x) = 1{,}398 = 1{,}4 \text{ colisões (arredondado)}$$

$$\sigma^2 = [\Sigma x^2 \cdot P(x)] - \mu^2$$
$$= 3{,}066 - 1{,}398^2 = 1{,}111596$$
$$= 1{,}1 \text{ colisões (arredondado)}$$

O desvio-padrão é a raiz quadrada da variância:

$$\sigma = \sqrt{1{,}111596} = 1{,}054323 = 1{,}1 \text{ colisões (arredondado)}$$

Sabemos agora que, entre sete acidentes aéreos, o número médio de acidentes com a USAir é 1,4, a variância é 1,1 "acidentes ao quadrado" e o desvio-padrão é 1,1 acidentes. Aplicando a regra prática dada na Seção 2-5, podemos concluir que, na maior parte das vezes, a USAir deve ter de 0 a 3,6 acidentes dentre sete escolhidos aleatoriamente. (Recorde que, por essa regra prática, podemos obter estimativas de valores mínimo e máximo partindo da média de 1,4 e somando e subtraindo 2,2, que é o dobro do desvio-padrão.)

Por que trabalhar com as Fórmulas 4-1 a 4-4? Uma distribuição de probabilidades é na verdade um modelo de distribuição de freqüências de uma população teoricamente perfeita. A distribuição de probabilidades é como uma distribuição de freqüências relativas baseada em dados que se comportam de modo perfeito, sem as imperfeições da amostra. Como a distribuição de probabilidades permite-nos predizer os resultados populacionais, podemos determinar a média, a variância e o desvio-padrão. A Fórmula 4-1 desempenha a mesma função que a fórmula da média de uma tabela de freqüências. (Recorde que f representa a freqüência de uma classe e N representa o tamanho da população.) Reescrevendo a fórmula da média de uma tabela de freqüências de modo que ela se aplique a uma população, e mudando sua forma, temos

$$\mu = \frac{\Sigma(f \cdot x)}{N} = \sum \frac{f \cdot x}{N} = \sum x \cdot \frac{f}{N} = \Sigma x \cdot P(x)$$

Na fração f/N, o valor de f é a freqüência com que o valor x ocorre e N é o tamanho da população, de forma que f/N é a probabilidade do valor x.

Raciocínio análogo permite-nos tomar a fórmula da variância do Capítulo 2 e aplicá-la a uma variável aleatória para uma distribuição de probabilidades; o resultado é a Fórmula 4-2. A Fórmula 4-3 é uma versão abreviada que dá sempre o mesmo resultado que a Fórmula 4-2. Embora a Fórmula 4-3 tenha a vantagem do manuseio mais fácil, a Fórmula 4-2 é mais fácil de ser entendida diretamente. Com base na Fórmula 4-2, podemos expressar o desvio-padrão como

$$\sigma = \sqrt{\Sigma(x-\mu)^2 \cdot P(x)}$$

ou como a forma equivalente dada pela Fórmula 4-4.

Valor Esperado

A média de uma variável aleatória discreta é o resultado médio teórico de um número infinito de provas. Podemos encarar essa média como o *valor esperado* no sentido de que é o valor médio que esperaríamos obter se as provas se prolongassem indefinidamente. As aplicações do valor esperado (também chamado *esperança* ou *esperança matemática*) são extensas e variadas e desempenham papel de extrema importância em uma área de aplicação chamada *teoria da decisão*. (Para um estudo da teoria da decisão, veja *Business Statistics*, de Triola e Franklin.)

DEFINIÇÃO

O **valor esperado** de uma variável aleatória discreta é denotado por E e representa o valor médio dos resultados. É dado por $\Sigma x \cdot P(x)$:

$$E = \Sigma x \cdot P(x)$$

Pela Fórmula 4-1, vemos que $E = \mu$. Isto é, a média de uma variável aleatória discreta coincide com seu valor esperado. Repita 5 vezes o experimento da jogada de uma moeda; o número *médio* de caras é 2,5; ao jogarmos uma moeda cinco vezes, o *valor esperado* do número de caras é também 2,5.

EXEMPLO Considere o jogo de números praticado há muitos anos por organizações ligadas ao crime e agora legalizado por muitos governos organizados — assim como também por alguns governos não muito bem organizados. Em geral conhecido como "Escolha três" (*Pick three*), o apostador aposta em três números, que deverão coincidir com os números sorteados. O ganho típico é de 499 para 1, o que significa que para cada $1 apostado o jogador recebe $500; o retorno líquido é, pois, de $499. Suponha o leitor que apostou $1 no número 327. Qual é o valor esperado de seu ganho ou perda?

SOLUÇÃO Para essa aposta há dois resultados simples: ou o leitor ganha, ou perde. Como o número escolhido foi 327, e como há 1000 possibilidades (de 000 a 999), a probabilidade de ganhar é de 1/1000 (ou 0,001) e a probabilidade de perder é 999/1000 (ou 0,999). A Tabela 4-3 resume a situação.

TABELA 4-3 O jogo dos números

Evento	x	$P(x)$	$x \cdot P(x)$
Ganha	$499	0,001	$0,499
Perde	− $1	0,999	− $0,999
Total			− $0,50 (ou − 50¢)

Pela Tabela 4-3, vemos que, para uma aposta de $1 no jogo dos números, o valor esperado é

$$E = \Sigma x \cdot P(x) = -50 \text{ centavos.}$$

Isso significa que, a longo prazo, para cada aposta de $1 podemos esperar perder em média 50 centavos. Não se trata de um programa muito seguro de investimento.

No exemplo precedente, um jogador perde $1 ou ganha $499; nunca haverá uma perda de 50 centavos, como o valor esperado de −50 centavos poderia sugerir. Esse valor esperado é uma média para uma longa seqüência de apostas feitas. Mesmo que pretendamos fazer apenas uma aposta, o valor esperado de −50 centavos mostra que não se trata de um bom negócio. O ganho potencial é mais do que superado pela perda potencial.

Nesta seção, vimos que uma variável aleatória tem um valor numérico associado a cada resultado de um experimento aleatório, e que uma distribuição de probabilidade tem uma probabilidade associada a cada valor de uma variável aleatória. Estudamos métodos de determinação da média, da variância e do desvio-padrão de uma distribuição de probabilidades. Vimos que o valor esperado de uma variável aleatória coincide efetivamente com a média. Vimos ainda que as loterias não são investimentos aconselháveis.

Os Telefones Celulares Podem Causar Câncer no Cérebro?

Na revista *Discover*, o matemático John Allen Paulos cita o caso de um processo movido por um cidadão, sob a alegação de que o câncer no cérebro de sua esposa fora causado pelo uso de um telefone celular. Esse processo despertou considerável atenção na imprensa e no TV e causou uma baixa na cotação das ações das empresas fabricantes de telefones celulares. Paulos concluiu que, com uma taxa anual de 0,007% de incidência de câncer no cérebro e com 10 milhões de usuários de telefones celulares, podemos esperar, cada ano, cerca de 700 casos de câncer cerebral entre os usuários desses aparelhos. A conclusão de Paulos foi: "Como apenas uns poucos casos chamaram a atenção do público, deveríamos concluir que os telefones celulares poderiam até mesmo eliminar o câncer cerebral. Absurdo, sem dúvida, mas não mais absurdo do que o raciocínio que motivou a histeria original." Paulos cita esse caso como um exemplo "dos obstáculos psicológicos à compreensão racional da estatística".

4-2 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-4, identifique a variável aleatória como discreta ou contínua.

1. O peso de um livro escolhido aleatoriamente.
2. O custo de um livro escolhido aleatoriamente. > C
3. O número de ovos que uma galinha põe. > D
4. A quantidade de leite ordenhado de uma vaca.

Nos Exercícios 5-12, determine se é dada uma distribuição de probabilidade. Nos casos em que não é descrita uma distribuição de probabilidade, identifique a condição que não é satisfeita. E quando for descrita uma distribuição de probabilidade, determine sua média, variância e desvio-padrão.

5. Ao escolher aleatoriamente um colega de cela condenado por dirigir alcoolizado (DWI), a distribuição de probabilidade do número x de sentenças anteriores em casos de DWI é dada na tabela a seguir (baseada em dados do Ministério da Justiça dos EUA).

x	$P(x)$
0	0,512
1	0,301
2	0,132
3	0,055

6. Se sua faculdade contrata os 4 próximos funcionários sem distinção de sexo e o conjunto de candidatos é grande, com números iguais de homens e mulheres, a tabela a seguir dá a distribuição de probabilidade do número x de mulheres contratadas.

x	$P(x)$
0	0,0625
1	0,2500
2	0,3750
3	0,2500
4	0,0625

7. A Associação de Cardiologia de Newport planeja abrir um escritório de consulta telefônica com 8 empregados. Ao planejar a área de estacionamento para esse escritório, é preciso saber quantos funcionários dirigirão seus próprios carros. De acordo com a Hertz Corporation, 69% de todos os funcionários utilizam seus próprios carros; a tabela a seguir descreve, assim, a distribuição de probabilidades do número de funcionários (dentre oito selecionados aleatoriamente) que utilizam seus próprios carros.

x	$P(x)$
0	0,000
1	0,002
2	0,012
3	0,053
4	0,147
5	0,261

8. Ao avaliar riscos de crédito, o Jefferson Valley Bank investiga o número de cartões de crédito que a pessoa tem. Com x sendo o número de cartões de crédito que os adultos possuem, a tabela a seguir dá a distribuição de probabilidades para um conjunto de solicitantes (com base em dados da Maritz Marketing Research, Inc.).

x	$P(x)$
0	0,26
1	0,16
2	0,12
3	0,09
4	0,07
5	0,09
6	0,07
7	0,14

9. Para resolver uma questão de paternidade, fazem-se testes de sangue em duas pessoas diferentes. Se x é o número dos que têm sangue do grupo A, então x pode ser 0, 1 ou 2, e as probabilidades correspondentes são 0,36, 0,48 e 0,16, respectivamente (com base em dados do Programa de Sangue de Nova York).
10. A Baltimore Computer House afirma que as probabilidades de vender 0, 1, 2, 3 e 4 microcomputadores em um dia são 0,240, 0,370, 0,205, 0,075 e 0,080, respectivamente.
11. Relaciona-se a seguir o número de jantares que os americanos típicos preparam em uma semana, juntamente com as respectivas probabilidades (com base em dados de Millward Brown, citados em *USA Today*): 0 (0,08); 1 (0,05); 2 (0,10); 3 (0,13); 4 (0,15); 5 (0,21); 6 (0,09); 7 (0,19).
12. Um estudo da tendenciosidade quanto ao sexo nos meios de comunicação envolve a escolha de pessoas que figuram como personagens em *shows* da tarde da TV. As pessoas são selecionadas aleatoriamente em grupos de quatro, registrando-se o número de mulheres. As probabilidades de obter 0, 1, 2, 3 e 4 mulheres são 0,334, 0,421, 0,200, 0,042 e 0,003, respectivamente (com base em dados do *USA Today*).
13. Ao apostar em um cassino $5 no número 7 da roleta, tem-se uma probabilidade de 1/38 de ganhar $175 e uma probabilidade de 37/38 de perder $5. Qual é o valor esperado? Em um número muito grande de apostas, quanto se perde para cada dólar apostado?
14. Quando jogamos $5 em um cassino na *pass line* do jogo de dados, há uma probabilidade de 244/495 de ganhar $5 e uma probabilidade de 251/495 de perder $5. Qual é o valor esperado? Em um grande número de jogadas, quanto perdemos para cada dólar apostado?
15. Uma mulher de 27 anos decide contratar uma apólice de seguro de vida de $100.000,00 por 1 ano, pagando um prêmio de $156. A probabilidade de ela sobreviver 1 ano é de 0,9995 (com base em dados do Ministério da Saúde e Recursos Humanos dos EUA e da AFT Group Life Insurance). Qual é seu valor esperado para a apólice de seguro?
16. O *Reader's Digest* lançou um concurso (*sweepstake*), relacionando os prêmios com as respectivas chances de ganhar: $5.000.000 (1 chance em 201.000.000), $150.000 (1 chance em 201.000.000), $100,000 (1 chance em 201.000.000), $25.000 (1 chance em 100.500.000), $10.000 (1 chance em 50.250.000), $5.000 (1 chance em 25.125.000), $200 (1 chance em 8.040.000), $125 (1 chance em 1.005.000) e um relógio no valor de $89 (1 chance em 3774).
    a. Determine o valor esperado do ganho para uma aposta.
    b. Determine o valor esperado se o custo para participar desse *sweepstake* é o preço de um selo do correio.
17. A variável aleatória x representa o número de meninas em uma família de 3 filhos. [*Sugestão:* Admitindo que menino e menina sejam igualmente prováveis, obtemos $P(2) = 3/8$ atentando para este espaço amostral: HHH, HHM, HMH, HMM, MHH, MHM, MMH, MMM.] Determine a média, a variância e o desvio-padrão da variável aleatória x. Aplique a regra prática (da Seção 2-5) para obter uma aproximação dos valores mínimo e máximo de x.
18. A variável aleatória x representa o número de meninos em uma família com 4 filhos. (Veja Exercício 17.) Determine a média, a variância e o desvio-padrão da variável aleatória x. Aplique a regra prática (Seção 2-5) para obter uma aproximação dos valores mínimo e máximo de x.
19. A Menlo Park Electronics Company fabrica interruptores para sinais de tráfego. Um lote de 10 interruptores tem 2 defeituosos. Escolhidos aleatoriamente 2 interruptores desse lote (sem reposição), represente pela variável aleatória x o número de interruptores defeituosos. Determine a média, a variância e o desvio-padrão da variável aleatória x.
20. Uma turma de estatística compreende 3 canhotos e 24 destros. Selecionam-se aleatoriamente dois estudantes diferentes para um projeto de coleta de dados, representando-se por x o número de estudantes canhotos escolhidos. Calcule a média, a variância e o desvio-padrão da variável aleatória x. [*Sugestão:* Aplique a regra da multiplicação das probabilidades para achar primeiro $P(0)$ e $P(2)$].

4-2 Exercícios B: Além do Básico

21. Em cada caso, determine se a função dada é uma distribuição de probabilidade.
    a. $P(x) = 1/2^x$ onde $x = 1, 2, 3, \ldots$
    b. $P(x) = 1/2x$ onde $x = 1, 2, 3, \ldots$
    c. $P(x) = 3/[4(3-x)!\,x!]$ onde $x = 0, 1, 2, 3,$
    d. $P(x) = 0{,}4(0{,}6)^{x-1}$ onde $x = 1, 2, 3, \ldots$
22. A média e o desvio-padrão de uma variável aleatória x são 5,0 e 2,0, respectivamente. Determine a média e o desvio-padrão das seguintes variáveis aleatórias:
    a. $3 + x$
    b. $3x$
    c. $3x + 4$
23. Selecionam-se aleatoriamente os algarismos (0, 1, 2,..., 9) para números de telefone em pesquisas. A variável aleatória x é o algarismo escolhido.
    a. Ache a média e o desvio-padrão de x.
    b. Ache o escore z para cada um dos valores possíveis de x; determine então a média e o desvio-padrão da população de escores z
24. Suponha que a variável aleatória discreta x possa tomar os valores 1, 2, ..., n, e que esses valores sejam igualmente prováveis.
    a. Mostre que $\mu = (n + 1)/2$.
    b. Mostre que $\sigma^2 = (n^2 - 1)/12$.
    c. Um experimento consiste em escolher aleatoriamente um número inteiro entre 1 e 50; a variável aleatória x é o valor do número escolhido. Determine a média e o desvio-padrão de x.
    (*Sugestão:* $1 + 2 + 3 + \cdots + n = n(n + 1)/2$.
    $1^2 + 2^2 + 3^2 + \cdots + n^2 = n(n + 1)(2n + 1)/6$.)

4-3 Experimentos Binomiais

Na Seção 4-2, vimos que uma variável aleatória associa um valor numérico a cada resultado de um experimento aleatório e uma distribuição de probabilidade associa uma probabilidade a cada valor de uma variável aleatória. Na maioria dos exemplos e exercícios da Seção 4-2, davam-se probabilidades para os valores da variável aleatória; nesta seção, veremos como determinar as probabilidades para uma categoria importante de distribuição de probabilidades: os *experimentos binomiais*. Os experimentos binomiais têm a característica de apresentarem exatamente dois resultados complementares: em processos industriais, as peças falham ou não falham. Na medicina, um paciente sobrevive um ano, ou morre. Em propaganda, um consumidor reconhece um produto, ou não.

DEFINIÇÃO

Um **experimento binomial** é um experimento que satisfaz as seguintes condições:

1. O experimento deve comportar um *número fixo de provas*.
2. As provas devem ser independentes. (O resultado de qualquer prova não afeta as probabilidades das outras provas.)
3. Cada prova deve ter todos os resultados classificados em *duas categorias*.
4. As probabilidades devem permanecer *constantes* para cada prova.

$p = 0{,}10$
$n = 15$
$x = 3$
$q = 0{,}9$

Se fazemos um experimento binomial, a distribuição da variável aleatória x é chamada uma *distribuição de probabilidade binomial* (ou *distribuição binomial*). Usa-se comumente a seguinte notação.

Notação para a Distribuição Binomial

S e F (sucesso e falha) denotam as duas categorias possíveis de todos os resultados; p e q denotam as probabilidades de S e F, respectivamente; assim,

$$P(S) = p$$
$$P(F) = 1 - p = q$$

- n denota o número fixo de provas
- x denota um número específico de sucessos em n provas, podendo ser qualquer inteiro entre 0 e n, inclusive.
- p denota a probabilidade de sucesso em *uma* das n provas.
- q denota a probabilidade de falha em *uma* das n provas.
- $P(x)$ denota a probabilidade de obter exatamente x sucessos em n provas.

A palavra *sucesso*, tal como usada aqui, é arbitrária, e não descreve necessariamente um resultado desejado. Qualquer uma das duas categorias possíveis pode ser chamada um sucesso S, desde que a probabilidade correspondente seja identificada como p. (Pode-se sempre determinar o valor de q subtraindo-se p de 1; se $p = 0{,}95$, então $q = 1 - 0{,}95 = 0{,}05$.) Uma vez designada uma categoria como o sucesso S, devemos ter a certeza de que p é a probabilidade de sucesso e x é o número de sucessos. Isto é, os valores de p e x devem referir-se à mesma categoria designada como sucesso.

Uma aplicação muito comum da estatística envolve a amostragem sem reposição, como o teste de produtos manufaturados ou a realização de uma pesquisa. Estritamente falando, a amostragem sem reposição envolve eventos dependentes, o que viola a segunda condição da definição precedente. Entretanto, a seguinte regra empírica se baseia no fato de que, se a amostra é muito pequena em relação ao tamanho da população, a diferença entre os resultados será desprezível, no caso de considerarmos as provas independentes quando elas são, na verdade, dependentes.

Na amostragem sem reposição, os eventos podem ser considerados independentes se o tamanho da amostra não excede 5% do tamanho da população. (Isto é, $n \leq 0{,}05N$.)

Você Tem Medo de Voar?

Muitos tendem a superestimar a probabilidade de morte em conseqüência de eventos que dão manchete, como desastres de avião, crimes ou atentados terroristas. Pesquisas estatísticos mostram que, embora a década passada tenha sido a mais segura para a aviação nos EUA, o povo está se mostrando mais temeroso do que no passado em relação às viagens aéreas. Ao viajar de Nova York para a Califórnia, sua chance de morrer é de 1 em 11 milhões para um jato comercial, 1 em 900.000 para um trem, 1 em 14.000 para um automóvel.

EXEMPLO Dado que 10% das pessoas são canhotas, suponha que queiramos achar a probabilidade de obter exatamente 3 estudantes canhotos em uma turma de 15 estudantes. (Algumas carteiras são adaptadas para estudantes canhotos, e a probabilidade resultante poderia afetar o número de tais carteiras a serem encomendadas para as salas de aula.) $P(x) = 0{,}129$

a. Trata-se de um experimento binomial? s

b. Em caso afirmativo, identifique os valores de n, x, p e q.

SOLUÇÃO

a. O experimento verifica as condições de um experimento binomial, como mostramos a seguir.

1. O número de provas (15) é fixo.
2. As provas são independentes, porque o fato de um estudante ser canhoto ou destro não afeta a probabilidade de outro estudante ser canhoto.
3. Cada prova tem duas categorias de resultados: o estudante é canhoto ou não é.
4. A probabilidade (0,10) permanece constante para os diferentes estudantes.

b. Tendo concluído que o experimento é binomial, passemos a identificar os valores de n, x, p e q:

1. Com 15 estudantes em uma turma, temos $n = 15$.
2. Queremos 3 estudantes canhotos (sucessos), e assim $x = 3$.
3. A probabilidade de um estudante ser canhoto (sucesso) é 0,1 e, assim, $p = 0{,}1$.
4. A probabilidade de falha (não-canhoto) é 0,9, logo $q = 0{,}9$.

Enfatizamos novamente que x e p devem ambos referir-se ao mesmo conceito de "sucesso". Neste exemplo, denotamos por x o número desejado de estudantes canhotos, e p é a probabilidade de obter um estudante canhoto; x e p se referem, assim, ao mesmo conceito de sucesso.

Nesta seção vamos apresentar três métodos para determinar probabilidades em um experimento binomial. O primeiro método envolve cálculos que utilizam a *fórmula da probabilidade binomial*, e constitui a base para os outros dois métodos. O segundo método envolve a utilização da Tabela A-1, e o terceiro método apela para pacotes estatísticos. Vamos descrever os três métodos, ilustrá-los e então dar os fundamentos de cada um.

Método 1: Utilização da Fórmula da Probabilidade Binomial Em um experimento binomial, as probabilidades podem ser calculadas utilizando-se a fórmula da probabilidade binomial

Fórmula 4-5

$$P(x) = \frac{n!}{(n-x)!x!} \cdot p^x \cdot q^{n-x} \quad \text{para } x = 0, 1, 2, \ldots, n$$

com
n = número de provas
x = número de sucessos em n provas
p = probabilidade de sucesso em qualquer prova
q = probabilidade de falha em qualquer prova ($q = 1 - p$)

O símbolo fatorial !, introduzido na Seção 3-6, denota o produto de fatores decrescentes. Eis dois exemplos de fatorial: $3! = 1.2.3 = 6$ e $0! = 1$ (por definição). Muitas calculadoras têm a tecla de fatorial, assim como a tecla $_nC_r$ que podem simplificar os cálculos. Em calculadoras com a tecla ${}_nC_r$ utilize a versão seguinte da fórmula da probabilidade binomial (onde n, x, p e q são definidos como na Fórmula 4-5):

$$P(x) = {}_nC_x \cdot p^x \cdot q^{n-x}$$

A calculadora TI-83 tem um programa para o cálculo de probabilidades binomiais. Abordaremos o uso da TI-83 no Método 3 para cálculo de probabilidades binomiais.

Alguns autores utilizam $1 - p$ em lugar de q na fórmula da probabilidade binomial, conforme se segue:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

EXEMPLO Aplicando a fórmula da probabilidade binomial, determine a probabilidade de obter 3 estudantes canhotos em uma turma de 15 estudantes, dado que 10% da população são canhotos. Isto é, determine $P(3)$, se $n = 15$, $x = 3$, $p = 0,1$ e $q = 0,9$.

SOLUÇÃO Levando os valores dados de n, x, p e q na fórmula de probabilidade binomial (Fórmula 4-5), obtemos

$$\begin{aligned} P(3) &= \frac{15!}{(15-3)!3!} \cdot 0,1^3 \cdot 0,9^{15-3} \\ &= \frac{15!}{12!3!} \cdot 0,001 \cdot 0,282429536 \\ &= (455)(0,001)(0,282429536) = 0,129 \end{aligned}$$

A probabilidade de exatamente 3 dentre os 15 estudantes serem canhotos é de 0,129.

Sugestão para cálculo: Ao calcular uma probabilidade com a fórmula de probabilidade binomial, é conveniente calcular $n!/(n-x)!x!$, p^x e q^{n-x}, e multiplicar os três fatores. Não faça muitos arredondamentos ao calcular esses fatores; arredonde somente o resultado final.

Método 2: Utilize a Tabela A-1 do Apêndice A Em alguns casos, podemos achar facilmente probabilidades binomiais simplesmente recorrendo à Tabela A-1 no Apêndice A. Localize primeiro n e

Da Tabela A-1:

n	x	p 0,10
15	0	206
	1	343
	2	267
	3	129
	4	043
	5	010
	6	002
	7	0+
	8	0+
	9	0+
	10	0+
	11	0+
	12	0+
	13	0+
	14	0+
	15	0+

Probabilidade binominal $n = 15$ e $p = 0,10$

x	$P(x)$
0	0,206
1	0,343
2	0,267
3	0,129
4	0,043
5	0,010
6	0,002
7	0+
8	0+
9	0+
10	0+
11	0+
12	0+
13	0+
14	0+
15	0+

o valor correspondente de x desejado. A esta altura, deve-se isolar uma linha de números. Faça coincidir essa linha com a probabilidade adequada p utilizando a coluna através do topo. O número assim isolado representa a probabilidade desejada (sem a vírgula decimal no início). Uma probabilidade extremamente pequena como 0,000000345 é indicada como 0^+.

Mostramos a seguir parte da Tabela A-1. Quando $n = 15$ e $p = 0,10$, em um experimento binomial, as probabilidades de 0, 1, 2, ..., 15 sucessos são 0,206, 0,343, 0,267, ..., 0^+, respectivamente.

EXEMPLO No exemplo precedente utilizamos a fórmula da probabilidade binomial para achar a probabilidade de 3 sucessos, dado que $n = 15$, $x = 3$, $p = 0,1$ e $q = 0,9$. Recorrendo à parte da Tabela A-1 mostrada anteriormente, calcule:

a. A probabilidade de exatamente 3 sucessos
b. A probabilidade de *ao menos* 3 sucessos

SOLUÇÃO

a. A parte da Tabela A-1 apresentada mostra que, quando $n = 15$ e $p = 0,1$, $P(3) = 0,129$, que é o mesmo valor obtido com a fórmula binomial no exemplo precedente.
b. $$\begin{aligned} P(\text{ao menos } 3) &= P(3 \text{ ou } 4 \text{ ou } 5 \text{ ou} \ldots \text{ ou } 15) \\ &= P(3) + P(4) + P(5) + \cdots + P(15) \\ &= 0,129 + 0,043 + 0,010 + \cdots + 0 \\ &= 0,184 \end{aligned}$$

Na parte (b) da solução acima, se quiséssemos achar a probabilidade P (ao menos 3) aplicando a fórmula da probabilidade binomial deveríamos aplicar a fórmula 13 vezes para calcular 13 probabilidades diferentes, somando-as então. [Um processo abreviado consistiria em calcular $P(0)$, $P(1)$ e $P(2)$ e subtrair sua soma de 1, mas mesmo este processo abreviado exigiria muito mais tempo do que a consulta à Tabela A-1.] Dada esta possibilidade de escolha entre a fórmula e a tabela, tem sentido utilizar a tabela. Note, entretanto, que a Tabela A-1 dá apenas alguns valores de n, assim como um número limitado de valores de p, de forma que nem sempre podemos utilizá-la; devemos então recorrer à fórmula da probabilidade binomial ou a um pacote de software.

Método 3: Utilização de um Software de Computador ou de uma Calculadora TI-83 Muitos pacotes estatísticos de computador incluem uma opção para gerar probabilidades binomiais; essa característica está hoje incluída também nas calculadoras TI-83. Dão-se a seguir amostras de resultados obtidos em um experimento binomial com $n = 15$ e $p = 0,1$.

Com STATDISK, selecione `Analysis` do menu principal; selecione então a opção `Binomial Probabilities`. Introduza os valores desejados de n e p, e aparecerá toda a distribuição de probabilidades.

Com Minitab, introduza primeiro uma coluna C1 de valores de x cujas probabilidades deseja (como 0, 1, 2, 3, 4, 5, 6, 7 e 8); selecione então `Calc` do menu principal e passe a selecionar os itens do submenu de `Probability Distributions` e `Binomial`. Introduza o número de provas, a probabilidade de sucesso e C1 para a coluna de entradas; clique OK.

Se estiver trabalhando com uma calculadora TI-83 para achar probabilidades binomiais, acione `2nd VARS` (para obter `DISTR`, que denota "distribuições"); selecione então a opção identificada como `binompdf(`. Complete a entrada de `binompdf(n, p,`

x) com os valores específicos de n, p e x, e acione ENTER; o resultado será a probabilidade procurada $P(x)$.

Damos a seguir uma boa estratégia para escolher o melhor método de determinação de probabilidades binomiais.

1. Utilize um programa de software de computador ou uma calculadora TI-83.
2. Se não dispuser de um computador nem de uma calculadora TI-83, recorra à Tabela A-1, se possível.
3. Na ausência de um computador ou de uma TI-83, se as probabilidades não constam da Tabela A-1, aplique então a fórmula da probabilidade binomial.

A Estatística Tem Realmente Algum Valor?

Com uma venda anual superior a \$17 bilhões, a Motorola é uma das maiores produtoras de equipamento eletrônico, inclusive telefones celulares, telefones sem fio, *pagers*, rádios transceptores e módulos controladores de transmissão em carros. Nos últimos cinco anos, a Motorola economizou aproximadamente \$2,5 bilhões, implementando um plano de melhoria de qualidade que utiliza extensamente métodos estatísticos. Seus *pagers* e telefones celulares estão sendo fabricados com uma taxa projetada de defeitos de 0,00034%. A Motorola visa a um objetivo popularmente conhecido como o nível de qualidade dos "seis sigmas", que corresponde a menos de 3,4 defeitos por milhão de unidades produzidas. A Motorola constatou que a aplicação de métodos estatísticos é imprescindível para a sobrevivência em um mercado cada vez mais competitivo.

Na Seção 4-2 apresentamos a Tabela 4-1 como um exemplo de distribuição de probabilidades para acidentes com a USAir, abordado no início do capítulo. No próximo exemplo mostraremos como essas probabilidades foram obtidas.

EXEMPLO No início deste capítulo vimos que a USAir detinha 20% dos vôos domésticos e estava envolvida em quatro de cada sete acidentes aéreos consecutivos nos EUA. Supondo que os acidentes aéreos sejam eventos independentes e aleatórios, e admitindo ainda que a USAir seja tão segura quanto as outras companhias de aviação, determine a proba-

p = 0,2
n = 7
x = 4

= 0,029

STATDISK BINOMIAL PROBABILITIES

File Edit Analysis Data Help

Binomial Probabilities

Num Trials, n: 15
Success Prob, p: 0.1

Evaluate Help Plot

Mean	1.500000
St Dev	1.161895
Variance	1.350000

x	P(x)	P(x≤)	P(x≥)
0	0.20589	0.20589	1.00000
1	0.34315	0.54904	0.79411
2	0.26690	0.81594	0.45096
3	0.12851	0.94444	0.18406
4	0.04284	0.98728	0.05556
5	0.01047	0.99775	0.01272
6	0.00194	0.99969	0.00225
7	0.00028	0.99997	0.00031
8	0.00003	1.00000	0.00003
9	0.00000	1.00000	0.00000
10	0.00000	1.00000	0.00000
11	0.00000	1.00000	0.00000
12	0.00000	1.00000	0.00000
13	0.00000	1.00000	0.00000

MINITAB BINOMIAL PROBABILITIES

```
Binomial with n = 15 and p = 0.100000

        x        P( X = x)
     0.00          0.2059
     1.00          0.3432
     2.00          0.2669
     3.00          0.1285
     4.00          0.0428
     5.00          0.0105
     6.00          0.0019
     7.00          0.0003
     8.00          0.0000
     9.00          0.0000
    10.00          0.0000
    11.00          0.0000
    12.00          0.0000
    13.00          0.0000
    14.00          0.0000
    15.00          0.0000
```

bilidade de que, em sete acidentes aéreos, quatro ocorram com aviões da USAir.

SOLUÇÃO Trata-se de um experimento binomial, porque:

1. Temos um número fixo de provas (7).
2. Admite-se que as provas sejam independentes.
3. Há duas categorias: Cada acidente envolve, ou não envolve, um avião da USAir.
4. A probabilidade de um acidente envolver um avião da USAir (considerada "sucesso" neste experimento) é de 0,20 (porque a USAir detém 20% dos vôos domésticos), e permanece constante em cada prova. (Estamos admitindo que os acidentes sejam independentes e aleatórios.)

Identifiquemos inicialmente os valores de n, p, q e x:

$n = 7$	**Número de provas (acidentes)**
$p = 0{,}20$	**Probabilidade de sucesso (quando ocorre um acidente, envolve um avião da USAir)**
$q = 0{,}80$	**Probabilidade de falha (quando ocorre um acidente, não envolve um avião da USAir)**
$x = 4$	**Número de acidentes com a USAir, dentre sete acidentes**

Com esses valores particulares de n, p, q e x podemos obter a probabilidade desejada $P(4)$ na Tabela A-1, em lugar de aplicar a fórmula da probabilidade binomial. (Naturalmente, poderíamos aplicar a fórmula binomial, mas a consulta à tabela é mais rápida e menos sujeita a erro.) Recorrendo à Tabela A-1 com $n = 7$, $p = 0{,}20$ e $x = 4$, vemos que $P(4) = 0{,}029$; ou seja, há uma probabilidade de 0,029 de que, dentre sete acidentes aéreos, exatamente quatro ocorram com aviões da USAir.

Qual a Chance de um Asteróide Colidir com a Terra?

O astrônomo da NASA David Morrison afirmou que há cerca de 2000 asteróides cujas órbitas interceptam a órbita da Terra; não obstante, só encontramos 100 deles. É possível, pois, que um asteróide não-detectado venha a colidir com o nosso planeta, causando uma catástrofe global que destrua boa parte da vida. Morrison afirma que há uma probabilidade de 1/10.000 de que, no decorrer da vida de uma pessoa, haja um impacto de um asteróide, suficientemente violento para eliminar toda a safra de um ano, causando uma fome em massa. Alguns astrônomos recomendam um programa de 20 anos, com o objetivo de observar asteróides, visando a detectar em tempo hábil os que são perigosos. Essa detecção talvez nos permita alterar o curso de um asteróide, a fim de que ele não constitua uma ameaça fatal.

A pequena probabilidade (0,029) obtida no exemplo precedente poderia sugerir ser improvável a USAir envolver-se, por mero acaso, em quatro de cada sete acidentes; parece que a USAir não é tão segura quanto as outras companhias aéreas. Mas estaremos realmente formulando a pergunta correta? Há outra pergunta que melhor traduza o problema de saber se a USAir não é segura? Eis algumas outras possibilidades:

1. Qual é a probabilidade de a USAir ter *ao menos quatro* dentre sete acidentes?
2. Qual é a probabilidade de que qualquer companhia aérea (com 20% dos vôos domésticos) tenha quatro dentre sete acidentes?
3. Qual é a probabilidade de que qualquer companhia aérea (com 20% dos vôos domésticos) tenha ao menos quatro dentre sete acidentes?

Eliminamos a primeira questão, porque o problema real diz respeito à possibilidade de *qualquer* companhia aérea ter quatro dentre sete acidentes. Para esclarecer esse ponto, consideremos a loteria, em que a chance de determinada pessoa ganhar é extremamente pequena, mas a probabilidade de *alguém* ganhar é assaz alta. Quando alguém ganha, não concluímos que esse indivíduo tivesse melhor chance do que qualquer outro. Da mesma forma, um acidente com avião da USAir não significa necessariamente que essa companhia aérea não seja tão segura quanto as outras; tal como o ganho em uma loteria, o acidente poderia ser o resultado de eventos aleatórios que afetam todas as companhias aéreas da mesma forma. Devemos, portanto, achar a probabilidade de que *qualquer* companhia aérea tenha quatro dentre sete acidentes, supondo que cada companhia detenha 20% dos vôos.

Em seguida, devemos eliminar a segunda questão de nossa lista, porque se refere a *exatamente* quatro acidentes em sete. Nossa real preocupação não é a chance de obter um número específico de acidentes; repousa, ao contrário, na possibilidade de obter um resultado *ao menos tão extremo* como o observado. Trata-se de um conceito difícil, por isso vamos tentar esclarecê-lo com outro exemplo.

Suponha que estejamos jogando uma moeda, para determinar se ela é viciada em favor de "caras", e que obtenhamos 501 "caras" em 1000 jogadas. A intuição deve sugerir que a moeda não favorece "caras" (é equilibrada) porque é perfeitamente viável obtermos 501 caras em 1000 jogadas. Não obstante, a probabilidade de obtermos exatamente 501 caras em 1000 jogadas é na realidade bastante pequena: 0,0252. Essa pequena probabilidade reflete o fato de que, com 1000 jogadas, *qualquer* número *específico* de caras terá uma probabilidade muito baixa. Todavia, o resultado de 501 caras em 1000 jogadas não é *raro*, porque a probabilidade de obter *ao menos* 501 caras é alta: 0,488. Da mesma forma, estaríamos procurando a probabilidade de *ao menos quatro* acidentes, e não a probabilidade de exatamente quatro acidentes. Isto é, devemos procurar a probabilidade de que qualquer companhia aérea individualmente (com 20% dos vôos domésticos) tenha ao menos quatro dentre sete acidentes. Se tal probabilidade é muito pequena (como 0,05 ou menos), é razoável concluirmos que a USAir tem um problema; mas se é alta (superior a 0,05, por exemplo), é natural concluirmos que os acidentes com a USAir constituem uma coincidência. Primeiro determinamos a probabilidade de a USAir ter ao menos quatro acidentes dentre sete. Recorremos à Tabela A-1 (com $n = 7$, $p = 0{,}20$, $x = 4, 5, 6, 7$):

$$\begin{aligned} P(\text{USAir ter ao menos 4 em 7 acidentes}) &= P(4) + P(5) + P(6) + P(7) \\ &= 0{,}029 + 0{,}004 + 0^+ + 0^+ \\ &= 0{,}033 \end{aligned}$$

Suponhamos agora que haja cinco companhias aéreas A, B, C, D, E, cada uma com 20% dos vôos. Obtemos a probabilidade de que qualquer uma delas tenha ao menos quatro dentre sete acidentes, aplicando a regra da adição para eventos mutuamente excludentes. (Os eventos são mutuamente excludentes porque se uma companhia aérea tem ao menos quatro dentre sete acidentes, nenhuma outra companhia poderá também ter quatro dentre sete acidentes.)

$P(\text{A ou B ou C ou D ou E}) =$

$$= P(\text{A}) + P(\text{B}) + P(\text{C}) + P(\text{D}) + P(\text{E})$$

$$= 0{,}033 + 0{,}033 + 0{,}033 + 0{,}033 + 0{,}033$$
$$= 0{,}165$$

Para interpretar essa probabilidade, escolhamos arbitrariamente 0,05 como valor de separação, ou fronteira, que separa os resultados comuns dos resultados não considerados comuns. (Costuma-se usar o valor 0,05. No Capítulo 7 estudaremos outros valores de separação.) Como a probabilidade de 0,165 é maior do que o valor de separação de 0,05, concluímos que não é fora do comum uma companhia com 20% dos vôos domésticos ter ao menos quatro dentre sete acidentes. Concluímos assim que os acidentes com a USAir são uma coincidência, e não que a USAir tenha algum problema sério relacionado com a segurança. A análise não é simples, mas ilustra um problema importante. Devemos ter o máximo cuidado em formular questões que identifiquem corretamente o problema real.

Temos agora três métodos para achar probabilidades binomiais, mas tanto os elementos da Tabela A-1 como os programas estatísticos de software estabelecem suas probabilidades com base na fórmula de probabilidade binomial. Explicaremos agora como se estabelece essa fórmula.

No exemplo precedente, procuramos determinar a probabilidade de obter quatro sucessos em sete provas, com uma probabilidade de sucesso de 0,20 em cada prova. É correto raciocinar que a quatro sucessos em sete provas devem corresponder três falhas. Um erro comum consiste em determinar a probabilidade de quatro sucessos e três falhas como segue:

$$\underbrace{(0{,}20 \cdot 0{,}20 \cdot 0{,}20 \cdot 0{,}20)}_{\text{4 sucessos}} \cdot \underbrace{(0{,}80 \cdot 0{,}80 \cdot 0{,}80)}_{\text{3 falhas}} = 0{,}000819$$

Este cálculo é *errado* porque admite implicitamente que os *quatro primeiros* resultados devem ser sucessos, e os *três últimos* devem ser falhas. Ora, os quatro sucessos e três falhas podem ocorrer em *qualquer* ordem, e não na única ordem dada anteriormente. Na verdade, há 35 ordenações distintas de quatro sucessos e três falhas, cada uma com uma probabilidade de 0,000819, e assim a probabilidade correta é (35)(0,000819) = 0,029, conforme já havíamos encontrado. De modo geral, o número de maneiras em que podemos dispor x sucessos e $n - x$ falhas é dado pela Fórmula 4-6:

Fórmula 4-6 $\dfrac{n!}{(n-x)!\,x!}$ Número de resultados com exatamente x sucessos em n provas

A expressão dada na Fórmula 4-6 provém da Seção 3-6. (A Seção 3-6 não é necessária para a leitura deste capítulo.) Não deduziremos a Fórmula 4-6, mas o que ela representa deve ficar bem claro: ela dá o número de maneiras como podemos dispor x sucessos e $n - x$ falhas. Combinando a Fórmula 4-6 com a aplicação direta da regra de multiplicação para eventos independentes, temos como resultado a fórmula da probabilidade binomial

$$P(x) = \underbrace{\frac{n!}{(n-x)!\,x!}}_{\substack{\textbf{Número de resultados com}\\ \textbf{exatamente } x \textbf{ sucessos em}\\ n \textbf{ provas}}} \cdot \underbrace{p^x \cdot q^{n-x}}_{\substack{\textbf{Probabilidade de } x \textbf{ sucessos}\\ \textbf{em } n \textbf{ provas para determi-}\\ \textbf{nada ordem}}}$$

Para manter esta seção em perspectiva, tenha em mente que a fórmula da probabilidade binomial é apenas uma dentre muitas fórmulas de probabilidade que podem ser usadas em diferentes situações. É usada freqüentemente em aplicações como controle de qualidade, análise de eleitores, pesquisa médica, serviço de inteligência militar e propaganda. Embora o objetivo desta seção seja a distribuição binomial, podemos encontrar outras distribuições nos Exercícios 33-35 e na Seção 4-5.

4-3 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-8, determine se os experimentos são binomiais. Para os que não são, indique ao menos uma condição que não é satisfeita.

1. 50 jogadas de um dado
2. 200 jogadas de uma moeda equilibrada
3. 200 jogadas de uma moeda viciada
4. Pesquisa de 1000 consumidores americanos, perguntando se reconhecem a marca Nike
5. Fazer girar uma roleta 500 vezes
6. Pesquisar 1067 cidadãos, perguntando a cada um se votou na última eleição
7. Selecionar aleatoriamente (sem reposição) um grupo de 12 pneus diferentes de uma população de 30 pneus, dos quais 5 são defeituosos
8. Pesquisar 2000 espectadores de televisão para saber se recordam o nome de determinado produto após verem um comercial.

Nos Exercícios 9-12, suponha que, em um experimento binomial, uma prova se repita n vezes. Determine a probabilidade de x sucessos, dada a probabilidade p de sucesso em uma prova. (Utilize os valores dados de n, x e p e a Tabela A-1.)

0,243

9. $n = 3, x = 2, p = 0{,}9$
10. $n = 2, x = 0, p = 0{,}6$
11. $n = 8, x = 7, p = 0{,}99$
12. $n = 6, x = 1, p = 0{,}05$

(6) 0,075

Nos Exercícios 13-16, suponha que, em um experimento binomial, uma prova se repita n vezes. Determine a probabilidade de x sucessos, dada a probabilidade p de sucesso em uma prova. Utilize os valores dados de n, x e p e a fórmula de probabilidade binomial.

13. $n = 3, x = 2, p = 1/4$
14. $n = 6, x = 2, p = 1/3$
15. $n = 10, x = 4, p = 0{,}35$
16. $n = 8, x = 6, p = 0{,}85$

Nos Exercícios 17-20, utilize a seguir tabela obtida com Minitab. As probabilidades foram obtidas introduzindo-se os valores n = 8 e p = 0,77. Quando se escolhe aleatoriamente um jovem entre 11 e 19 anos, há uma probabilidade de 0,77 de ele utilizar videogame (com base em dados da Chilton Research Services). Em cada caso, suponha que são escolhidos 8 jovens aleatoriamente e determine a probabilidade pedida.

BINOMIAL WITH
N = 8 P = 0.77

K	P(X = K)
0	0.0000
1	0.0002
2	0.0025
3	0.0165
4	0.0689
5	0.1844
6	0.3087
7	0.2953
8	0.1236

17. A probabilidade de 6 jovens utilizarem o videogame.
18. A probabilidade de ao menos 5 jovens utilizarem o videogame.
19. A probabilidade de menos de 7 jovens utilizarem o videogame.
20. A probabilidade de mais de 4 jovens utilizarem o videogame.

Nos Exercícios 21-32, determine a probabilidade pedida.

21. Suponha que os nascimentos de menino e menina sejam igualmente prováveis e que o nascimento de qualquer criança não afete a probabilidade do sexo do próximo nascituro. Determine a probabilidade de:

a. Exatamente 4 meninas em 10 nascimentos.
b. Ao menos 4 meninas em 10 nascimentos.
c. Exatamente 8 meninas em 20 nascimentos.

22. De acordo com a Nielsen Media Research, 30% das televisões são sintonizadas no programa *NFL Monday Night Football* quando ele vai ao ar. Supondo que esse programa esteja sendo transmitido e que as televisões sejam escolhidas aleatoriamente, determine a probabilidade de:

a. 5 dentre 15 televisões estarem sintonizadas no *NFL Monday Night Footbal.*
b. Ao menos 5 dentre 15 televisões estarem sintonizadas no *NFL Monday Night Football.*
c. Exatamente 4 dentre 16 televisões estarem sintonizadas no *NFL Monday Night Football.*

23. A Mars, Inc. afirma que 20% de suas pastilhas de chocolate M&M são vermelhas. Determine a probabilidade de que, em 15 pastilhas M&M escolhidas aleatoriamente, exatamente 20% (ou seja, 3 pastilhas) sejam vermelhas.

24. Um artigo da revista *Time* relatou que, em Los Angeles, para cada 100 acidentes de carro em que há algum dano, em 99 casos ocorre algum ferimento. Se um estudo levado a efeito por uma seguradora começa com a seleção aleatória de 12 acidentes de carro em Los Angeles em que há algum dano, determine a probabilidade de ao menos em 11 deles haver ferimento.

25. Um teste de estatística consiste em 10 questões do tipo múltipla escolha, cada uma com 5 respostas possíveis. Para alguém que responda aleatoriamente (por palpite) todas as questões, determine a probabilidade de passar, se o percentual mínimo para aprovação é 60%. A probabilidade é suficientemente elevada para justificar o risco de tentar passar por palpite em lugar de estudar?

26. A Air America adota a política de vender 15 passagens para um avião que dispõe de apenas 14 assentos. (A experiência passada mostra que apenas 85% dos que reservam lugar comparecem efetivamente ao embarque.) Determine a probabilidade de não haver assentos suficientes no caso de a Air America vender 15 passagens.

27. De acordo com o Ministério da Justiça dos EUA, 5% de todos os lares americanos sofreram pelo menos um assalto no último ano, mas a polícia de Newport relata 4 casos de assalto em uma comunidade de 15 lares, no último ano. Com base na probabilidade de 4 ou mais assaltos em uma comunidade de 15 lares em um ano, pode-se dizer que aquela comunidade foi vítima apenas do acaso?

28. Uma pesquisa da *Computerworld* mostrou que 80% dos executivos de alto escalão utilizam microcomputadores em seu trabalho. A Telektronics Company planeja transferir 9 executivos para uma nova sede em Atlanta. Se há somente 7 microcomputadores disponíveis em Atlanta, determine a probabilidade de eles necessitarem de mais computadores. Essa probabilidade é suficientemente elevada para justificar a instalação de mais computadores?

29. Bill Connors, gerente de controle de qualidade da Menlo Park Electronics Company, sabe que sua companhia vem fabricando protetores contra oscilações de corrente elétrica com uma taxa de 10% de unidades defeituosas. Tomou, por isso, várias medidas para reduzir essa taxa de incidência de defeitos. Em um teste feito com 20 protetores selecionados aleatoriamente, encontrou-se apenas um defeituoso. Se a taxa de 10% de defeituosos se mantém, determine a probabilidade de que, em 20 unidades, no máximo uma apresente defeito. Com base no resultado, aparentemente as medidas adotadas estão surtindo efeito?

30. A Telektronic Company compra grandes lotes de lâmpadas fluorescentes e adota o seguinte método: selecionar aleatoriamente e testar 24 lâmpadas e aceitar todo o lote se no máximo uma não funcionar. Se determinado lote de lâmpadas tem efetivamente 4% de unidades defeituosas, qual é a probabilidade de todo o lote ser aceito?

31. Em um estudo de reconhecimento de marca, 95% dos consumidores reconheceram Coke (com base em dados da Total Research Corporation). Um pesquisador relata que em 15 consumidores selecionados aleatoriamente apenas 10 reconheceram o nome Coke. Determine a probabilidade de um número tão baixo; isto é, determine a probabilidade de obter no máximo 10 consumidores que reconhecem o nome Coke dentre 15 consumidores selecionados aleatoriamente. Com base no resultado, você acha que o resultado reportado pelo pesquisador possa ser conseqüência de mero acaso?

32. Após ser recusada para um emprego, Kim Kelly ficou sabendo que a Bellevue Advertising Company contratou apenas duas mulheres dentre os últimos 20 empregados novos. Ela sabe também que o número de candidatos é muito grande, com um número aproximadamente igual de homens e mulheres qualificados. Ajude-a a comprovar a acusação de discriminação sexual, determinando a probabilidade de obter no máximo duas mulheres quando se contratam 20 empregados, supondo que não haja discriminação quanto ao sexo. A probabilidade resultante confirma realmente a acusação?

4-3 Exercícios B: Além do Básico

33. Se um caso satisfaz todas as condições de um experimento binomial, exceto pelo fato de o número de provas não ser fixo, pode-se aplicar a **distribuição geométrica**. A probabilidade de obter o primeiro sucesso na x^{ma} prova é dada por $P(x) = p(1-p)^{x-1}$, onde p é a probabilidade de sucesso em uma prova. Suponha que a probabilidade de um componente de computador ser defeituoso é de 0,2. Determine a probabilidade de o primeiro defeito ocorrer no sétimo componente.

34. No caso de amostragem sem reposição de uma população finita, pequena, não devemos utilizar a distribuição binomial, porque os eventos não são independentes. Se a amostragem se faz sem reposição e os resultados comportam apenas dois tipos, utiliza-se a **distribuição hipergeométrica**. Se uma população tem A objetos de um tipo e os B objetos restantes são do outro tipo e se extraímos n objetos sem reposição, então a probabilidade de obter x objetos do tipo A e $n-x$ objetos do tipo B é

$$P(x) = \frac{A!}{(A-x)!x!} \cdot \frac{B!}{(B-n+x)!(n-x)!} \div \frac{(A+B)!}{(A+B-n)!n!}$$

Na Loto 54, um apostador escolhe 6 números de 1 a 54 (sem repetição), sorteando-se posteriormente uma combinação ganhadora. Determine a probabilidade de:

a. Acertar todos os 6 números ganhadores.
b. Acertar exatamente 5 dentre os 6 números ganhadores.
c. Acertar exatamente 3 dentre os 6 números ganhadores.
d. Não acertar qualquer número ganhador.

35. A distribuição binomial se aplica apenas a casos que envolvem 2 tipos de resultado, enquanto a **distribuição multinomial** envolve mais de 2 categorias. Suponha que tenhamos 3 tipos de resultados mutuamente excludentes denotados por A, B e C. Seja $P(A) = p_1$, $P(B) = p_2$ e $P(C) = p_3$. Em n provas independentes, a probabilidade de x_1 resultados do tipo A, x_2 resultados do tipo B e x_3 resultados do tipo C é dada por

$$\frac{n!}{(x_1!)(x_2!)(x_3!)} \cdot p_1^{x_1} \cdot p_2^{x_2} \cdot p_3^{x_3}$$

Um experimento de genética envolve 6 genótipos mutuamente excludentes identificados por A, B, C, D, E e F, todos igualmente prováveis. Testados 20 indivíduos, determine a probabilidade de obter exatamente 5 A's, 4 B's, 3 C's, 2 D's, 3 E's e 3 F's, desenvolvendo a expressão acima de forma que ela se aplique a 6 tipos de resultado, em lugar de apenas 3.

4-4 Média, Variância e Desvio-padrão da Distribuição Binomial

No Capítulo 2, exploramos coleções de dados reais e focalizamos três características importantes: (a) a medida de tendência central; (2) a medida de variação; (3) a natureza da distribuição. Um ponto fundamental deste capítulo é que as distribuições de probabilidade descrevem o que *provavelmente* acontecerá, e não o que efetivamente aconteceu. Na Seção 4-2, vimos como analisar distribuições de probabilidade através da média, do desvio-padrão e de um histograma de probabilidade. Em um experimento binomial, a distribuição da variável aleatória x é uma distribuição de probabilidade binomial que descreve o que provavelmente ocorrerá. Com as distribuições binomiais, é novamente importante investigar as três características importantes de tendência central, variação e natureza da distribuição. Na Seção 4-2, apresentamos as Fórmulas 4-1, 4-3 e 4-4 para determinar a média, a variância e o desvio-padrão de *qualquer* distribuição de probabilidade, e assim as fórmulas certamente se aplicam às distribuições binomiais. Mas essas fórmulas podem ser grandemente simplificadas no caso das distribuições binomiais, como vemos a seguir.

Para Qualquer Distribuição de Probabilidade

Fórmula 4-1 $\mu = \Sigma x \cdot P(x)$

Fórmula 4-3 $\sigma^2 = [\Sigma x^2 \cdot P(x)] - \mu^2$

Fórmula 4-4 $\sigma = \sqrt{[\Sigma x^2 \cdot P(x)] - \mu^2}$

Para a Distribuição Binomial

Fórmula 4-7 $\mu = n \cdot p$

Fórmula 4-8 $\sigma^2 = n \cdot p \cdot q$

Fórmula 4-9 $\sigma = \sqrt{n \cdot p \cdot q}$

A Fórmula 4-7 para a média tem um significado intuitivo. Se examinássemos 100 nascimentos, esperaríamos cerca de 50 meninas, e $n{\cdot}p$ nesse experimento é $(100)(1/2)$, ou 50. De modo geral, se consideramos p a proporção de sucessos, o produto $n{\cdot}p$ nos dará o número de sucessos esperado em n provas. A variância e o desvio-padrão já não se justificam com tanta facilidade, e preferimos omitir os complicados cálculos algébricos que levam às Fórmulas 4-8 e 4-9. Em lugar disso, o exemplo que se segue mostra que, para um experimento binomial, as Fórmulas 4-7, 4-8 e 4-9 produzirão os mesmos resultados que as Fórmulas 4-1, 4-3 e 4-4.

O Pára-quedismo É Seguro?

Cerca de 30 pessoas morrem a cada ano, dentre 100.000 que realizam 2,25 milhões de saltos de pára-quedas. Comparativamente, um ano típico registra cerca de 200 acidentes fatais com mergulho, 7000 afogamentos, 900 mortes em ciclismo, 800 mortes por queda de raio e 1150 mortes por picadas de abelha. Obviamente, essas cifras não significam que o pára-quedismo seja mais seguro do que o ciclismo ou a natação. Uma comparação correta deve utilizar *taxas* de ocorrência de casos fatais, e não apenas o número de mortes.

Com muita hesitação, o autor saltou de pára-quedas duas vezes, mas abandonou o esporte após sair da ampla zona de salto em ambas as vezes. Voou também em um planador, um balão de ar quente e um *blimp* Goodyear.

EXEMPLO No início deste capítulo, afirmamos que a USAir estava envolvida em quatro de sete acidentes. Na Seção 4-2 apresentamos a Tabela 4-1, com a distribuição de probabilidade do número de acidentes com a USAir, dentre sete acidentes escolhidos aleatoriamente, admitindo que a USAir detenha 20% dos vôos e que os acidentes sejam eventos independentes e aleatórios. Aplicamos então as Fórmulas 4-1, 4-3 e 4-4 (veja Tabela 4-2) para achar esses valores para a média, a variância e o desvio-padrão: $\mu = 1{,}4$ acidentes, $\sigma^2 = 1{,}1$ "acidentes ao quadrado", e $\sigma = 1{,}1$ acidentes. Na Seção 4-3 vimos que os acidentes com a USAir podem ser considerados um experimento binomial com $n = 7$, $p = 0{,}20$ e $q = 0{,}80$. Dados esses valores, aplique as Fórmulas 4-7, 4-8 e 4-9 para achar a média, a variância e o desvio-padrão. Verifique se os resultados são os mesmos que os obtidos com as Fórmulas 4-1, 4-3 e 4-4.

SOLUÇÃO Com os valores $n = 7$, $p = 0{,}20$ e $q = 0{,}80$, as Fórmulas 4-7, 4-8 e 4-9 dão os seguintes resultados:

$$\mu = n \cdot p = (7)(0{,}20) = 1{,}4$$
$$\sigma^2 = n \cdot p \cdot q = (7)(0{,}20)(0{,}80) = 1{,}1 \quad \text{(arredondado)}$$
$$\sigma = \sqrt{n \cdot p \cdot q}$$
$$= \sqrt{(7)(0{,}20)(0{,}80)} = \sqrt{1{,}12} = 1{,}1 \quad \text{(arredondado)}$$

Esses resultados mostram que, quando temos um experimento binomial, as Fórmulas 4-7, 4-8 e 4-9 dão os mesmos resultados que as Fórmulas 4-1, 4-3 e 4-4. (Comparando cuidadosamente os resultados, podemos notar pequenas discrepâncias nos valores não-arredondados, mas essas discrepâncias são devidas a erros de arredondamento na Tabela 4-2.)

Concentração de Doenças

Periodicamente, os meios de comunicação relatam casos de concentração de doenças em uma localidade. Uma comunidade de New Jersey acusou 13 casos de leucemia em 5 anos, quando a taxa normal seria de 1 caso em 10 anos. A pesquisa de tal concentração pode ser reveladora. Um estudo de concentração de casos de câncer nas proximidades de minas de asbesto na África levou à descoberta de que as fibras de asbesto podem ser cancerígenas.

Ao fazermos uma tal análise, devemos evitar o erro de criar artificialmente (ainda que de forma não-intencional) uma concentração, ou aglomerado, que se assemelharia a uma repartição arbitrária de zonas de influência eleitoral, visando a uma vitória.

EXEMPLO Alguns casais preferem ter filhos do sexo feminino, porque as mães são portadoras de um distúrbio recessivo que é herdado por 50% de seus filhos, mas por nenhuma de suas filhas. O método Ericsson de seleção de sexo tem uma taxa admitida de 75% de sucesso. Suponha que 100 casais utilizem o método Ericsson, com o resultado de que, dentre 100 recém-nascidos, há 75 meninas.

a. Supondo que o método de Ericsson não produza efeito, e admitindo que menino e menina sejam igualmente prováveis, determine a média e o desvio-padrão do número de meninas em um grupo de 100 crianças. (A hipótese de que menino e menina sejam igualmente prováveis não é precisamente correta, mas dá resultados bastante bons.)

b. Interprete os valores da parte (a) para determinar se o resultado de 75 meninas em 100 bebês confirma a alegação de eficiência do método de Ericsson.

SOLUÇÃO

a. Seja x a variável aleatória que representa o número de meninas em 100 nascimentos. Supondo que o método de Ericsson não produza efeito e que meninas e meninos sejam igualmente prováveis, temos $n = 100$, $p = 0{,}5$ e $q = 0{,}5$. Determinamos a média e o desvio-padrão com as Fórmulas 4-7 e 4-9:

$$\mu = n \cdot p = (100)(0.5) = 50$$

$$\sigma = \sqrt{n \cdot p \cdot q} = \sqrt{(100)(0{,}5)(0{,}5)} = 5$$

Para grupos de 100 casais com um filho cada um, o número médio de meninas é 50 e o desvio-padrão é 5.

b. Devemos agora interpretar os resultados para determinar se 75 meninas em 100 bebês é um resultado que pode ocorrer facilmente por chance, ou se é tão improvável que o método de Ericsson de seleção de sexo parece eficiente. Utilizaremos a regra prática da Seção 2-5.

De acordo com essa regra, as estimativas dos valores mínimo e máximo são as seguintes:

mínimo ≈ média − 2 × (desvio-padrão) = 50 − 2(5) = 40
máximo ≈ média + 2 × (desvio-padrão) = 50 + 2(5) = 60

A regra prática indica que os valores típicos estão provavelmente entre 40 e 60, e assim 75 meninas não parece um resultado devido unicamente ao acaso.

A regra prática se aplica apenas a distribuições em forma de sino, e a Figura 4-4 mostra que essa distribuição tem forma de sino. De acordo com a regra prática, 95% de todos os valores devem estar a menos de dois desvios-padrão da média, e 99,7% dos valores devem estar a menos de três desvios-padrão da média. O resultado de 75 meninas está a cinco desvios-padrão da média de 50. [Podemos calcular o escore z como segue: $z = (x - \mu)/\sigma = (75 - 50)/5 = 5$, ou podemos simplesmente observar que o valor 75 difere da média 50 por 25 unidades, que é equivalente a cinco desvios-padrão.)] Novamente, o resultado de 75 meninas não é usual e não parece ocorrer por puro acaso. Em lugar de concluir pela ocorrência de um evento extremamente raro, concluímos que o método de Ericsson parece eficiente para aumentar a probabilidade de um bebê ser menina.

É importante desenvolver habilidade técnica para calcular médias e desvios-padrão, mas é mais importante ainda desenvolver a capacidade de interpretar o significado de valores dessas medidas, conforme mostrado na parte (b) do exemplo precedente.

4-4 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-4, determine a média μ, a variância σ^2 e o desvio-padrão σ para os valores dados de n e p. Suponha satisfeitas as condições binomiais em cada caso.

1. $n = 64$, $p = 0{,}5$
2. $n = 150$, $p = 0{,}4$
3. $n = 1068$, $p = 1/4$
4. $n = 2001$, $p = 0{,}221$

Nos Exercícios 5-8, ache os valores indicados.

5. Vários estudantes não estão preparados para um teste do tipo verdadeiro/falso com 25 questões, e todos eles decidem responder "por palpite". Determine a média, a variância e o desvio-padrão do número de respostas corretas para cada estudante.
6. Em um teste de múltipla escolha com 50 questões, cada questão comporta quatro respostas possíveis a, b, c, d; apenas uma é correta. Calcule a média, a variância e o desvio-padrão do número de respostas corretas de um estudante que responde "por palpite" a todas as questões.

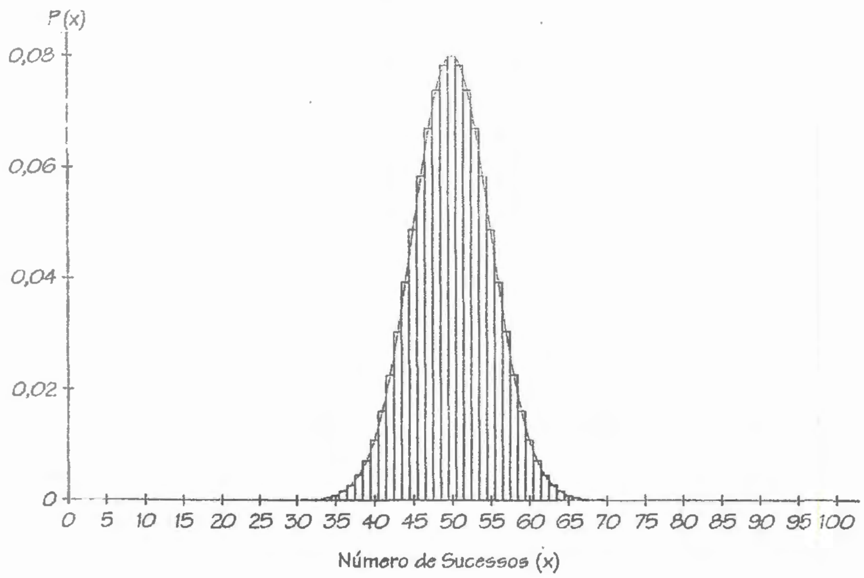


Fig. 4-4 Histograma de probabilidade para um experimento binomial com $n = 100$ e $p = 0{,}5$.

7. A probabilidade do 7 em uma roleta é 1/38. Em um experimento, a roleta é girada 500 vezes. Se esse experimento é repetido muitas vezes, determine a média e o desvio-padrão do número de 7s.
8. A probabilidade de ganhar na loteria do estado de Nova York é de 1/25.827.165. Determine a média e o desvio-padrão do número de ganhos para alguém que joga duas vezes por semana durante 50 anos (ou seja, 5200 vezes). (Expresse suas respostas com três algarismos significativos.)

Nos Exercícios 9-16, considere incomum qualquer resultado que difira da média por mais de dois desvios-padrão; isto é, os valores incomuns ou são inferiores a $\mu - 2\sigma$ ou são superiores a $\mu + 2\sigma$.

9. Em uma pesquisa sobre reconhecimento de marca, 95% dos consumidores reconheceram Coke (com base em dados da Total Research Corporation). Deve-se fazer uma nova pesquisa junto a 1200 consumidores selecionados aleatoriamente. Para tais grupos de 1200,
   a. Determine a média e o desvio-padrão do número dos que reconhecem a marca Coke.
   b. É incomum obter 1170 consumidores que reconhecem o nome Coke?
10. O Departamento de Saúde do Estado de Nova York relata uma taxa de 10% de incidência do vírus HIV para a população "de risco". Desenvolve-se em uma região uma intensa campanha educativa no sentido de reduzir essa taxa de 10%. Posto em prática o programa, faz-se um estudo subseqüente sobre 200 indivíduos do grupo de risco.
    a. Admitindo que o programa não tenha produzido efeito, determine a média e o desvio-padrão do número de casos de HIV em grupos de risco de 200 pessoas.
    b. Entre as 200 pessoas submetidas ao teste subseqüente, 7% (ou seja, 14 pessoas) tiveram resultado positivo no teste de HIV. Se o programa não produz efeito, essa taxa é excepcionalmente baixa? Este resultado sugere que o programa é eficaz?
11. A CIA, Central Intelligence Agency (Agência Central de Inteligência), analisa a freqüência das letras, tentando decifrar mensagens interceptadas. Em um texto padrão em língua inglesa, a letra *e* ocorre com uma freqüência relativa de 0,130.
    a. Determine a média e o desvio-padrão do número de vezes que a letra *e* deve aparecer em páginas-padrão de 2600 caracteres.
    b. Em uma mensagem interceptada enviada à Líbia, constatou-se a ocorrência da letra *e* 307 vezes em uma página com 2600 caracteres. Trata-se de um fato raro?
12. A Loja de Departamentos Newtower constatou uma taxa de 3,2% de queixas de clientes e decidiu reduzir essa taxa mediante um programa de treinamento de seus empregados. Ao fim do programa, observaram-se 850 clientes.
    a. Admitindo que o programa de treinamento não tenha produzido efeito, determine a média e o desvio-padrão do número de queixas nesses grupos de 850 clientes.
    b. No grupo de 850 clientes observados, 7 tiveram alguma queixa. Esse resultado é excepcional? O programa de treinamento parece ter sido eficaz?
13. De acordo com a Nielsen Media Research, Inc., 30% dos televisores são sintonizados na *NFL Monday Night Football* quando ele é transmitido. Suponha que esse programa esteja sendo transmitido e que sejam aleatoriamente escolhidos 4000 televisores.
    a. Para tais grupos de 4000, determine a média e o desvio-padrão do número de televisores sintonizados no *NFL Monday Night Football*.
    b. É fato incomum constatar que 1272 dentre os 4000 televisores estão sintonizados no *NFL Monday Night Football*? Qual é a causa provável de uma taxa tão superior a 30%?
14. A Mars, Inc., alega que 10% de suas pastilhas M&M são azuis; extrai-se uma amostra aleatória de 100 dessas pastilhas.
    a. Ache a média e o desvio-padrão do número de pastilhas azuis nesses grupos de 100.
    b. O Conjunto de Dados 11 do Apêndice B consiste em uma amostra aleatória de 100 M&Ms, em que 5 são azuis. Esse resultado pode ser considerado excepcional?
15. Um patologista sabe que 14,9% de todas as mortes podem ser atribuídas a infarto do miocárdio.
    a. Ache a média e o desvio-padrão do número dessas mortes que ocorrerão em uma região típica com 5000 mortes.
    b. Em certa região, examinam-se 5000 certidões de óbito, constatando-se 896 mortes por infarto do miocárdio. Há razões para preocupação? Por quê?
16. Um teste de percepção extra-sensorial envolve o reconhecimento de uma forma. Pede-se a 50 indivíduos de olhos vendados que identifiquem uma forma dentre as possibilidades de um quadrado, um círculo, um triângulo, uma estrela, um coração e o perfil do ex-presidente Millard Fillmore (1800-1874).
    a. Admitindo que todos os 50 indivíduos dêem respostas aleatórias, determine a média e o desvio-padrão do número de respostas corretas nesse grupo de 50.
    b. Se 12 das 50 respostas são corretas, esse resultado pode ter ocorrido por mera chance? O que podemos concluir?

4-4 Exercícios B: Além do Básico

17. A Providence Computer Supply Company sabe que 16% de seus computadores necessitarão de reparos sob garantia dentro de um mês da expedição. Em um mês típico, são expedidos 279 computadores.
    a. Se x é a variável aleatória que representa o número de computadores que exigem reparos sob garantia dentre os 279 computadores vendidos no mês, determine a média e o desvio-padrão de x.
    b. Para um mês típico em que são vendidos 279 computadores, qual seria um valor excepcionalmente baixo para o número de computadores que exigem reparo sob garantia dentro de um mês? Qual seria um valor excepcionalmente elevado? (Esses valores ajudam a determinar o número de técnicos necessários.)
18. a. Se uma empresa fabrica um produto com 80% de bons resultados (o que significa que 80% consistem em itens considerados bons), qual é o número mínimo de itens a serem produzidos para que haja no mínimo 99% de certeza de que a empresa produz pelo menos 5 itens bons?
    b. Se a empresa produz lotes de itens, cada um com o número mínimo determinado na parte (a), ache a média e o desvio-padrão do número de itens bons em tais lotes.

4-5 A Distribuição de Poisson

Se o leitor já esteve alguma vez em uma fila na Disney World, é provável que seu comportamento tenha sido analisado com uma distribuição de Poisson, que é uma distribuição de probabilidade freqüentemente utilizada como modelo matemático das chegadas de pessoas em uma fila. Outras aplicações incluem o estudo de acidentes com veículos, clientes chegando ao caixa de um supermercado, carros chegando a um posto de gasolina e usuários de computador ligados à Internet. A distribuição de Poisson se define como segue.

DEFINIÇÃO

A **distribuição de Poisson** é uma distribuição discreta de probabilidade, aplicável a ocorrências de um evento *em um intervalo especificado*. A variável aleatória x é o número de ocorrências do evento em um intervalo. O intervalo pode ser o tempo, a distância, a área, o volume ou outra unidade análoga. A probabilidade de o evento ocorrer x vezes em um intervalo é dada pela Fórmula 4-10.

Fórmula 4-10 $$P(x) = \frac{\mu^x \cdot e^{-\mu}}{x!} \quad \text{onde } e \approx 2{,}71828$$

A calculadora TI-83 tem um programa para achar probabilidades com o emprego da Fórmula 4-10. Acione `2nd VARS` (para obter `DISTR`, que denota "distribuições"); selecione então a opção identificada como `poissonpdf(`. Complete a introdução de `poissonpdf` (μ, x) com valores específicos de μ e x, acione em seguida ENTER. O STATDISK também calcula probabilidades para a distribuição de Poisson.

Eis alguns exemplos de uma variável aleatória de Poisson: número de carros que chegam a um posto de gasolina *durante um minuto*, número de aviões seqüestrados *em um dia*, e número de peças defeituosas substituídas em um Corvette *durante o primeiro ano* de garantia.

A Velocidade Pode Matar

Em 1987, a legislação federal americana permitiu que os estados elevassem de 55 mi/h para 65 mi/h o limite máximo de velocidade em certas rodovias interestaduais. Uma análise estatística mostrou que, no estado de Iowa, o aumento de 20% dos acidentes fatais foi conseqüência da elevação do limite da velocidade [veja "Evaluating the Impact of the 65 mph Maximum Speed Limit on Iowa Interstates" (Avaliando o Impacto do Limite Máximo de Velocidade de 65 mph nas estradas interestaduais de Iowa), de Ledolter e Chan, *The American Statistician*, Vol. 50, Nº 1)]. Esse estudo utilizou a distribuição de Poisson para determinar margens de erros de variações estimadas de casos fatais em rodovias interestaduais rurais, interestaduais urbanas, rurais principais e rurais secundárias. Com o auxílio da distribuição de Poisson, ficou constatada a significância estatística do aumento de casos fatais nas interestaduais rurais, mas as variações modestas em estradas que não alteraram seus limites de velocidade não foram significativas.

A distribuição de Poisson exige:

- Que a variável aleatória x seja o número de ocorrências de um evento *em um intervalo*.
- Que as ocorrências sejam *aleatórias*.
- Que as ocorrências sejam *independentes* umas das outras.
- Que as ocorrências sejam *distribuídas uniformemente* sobre o intervalo considerado.

A distribuição de Poisson tem os parâmetros:

- Média μ.
- Desvio-padrão $\sigma = \sqrt{\mu}$.

A distribuição de Poisson difere da binomial em dois aspectos importantes:

1. A distribuição binomial é afetada pelo tamanho amostral n e pela probabilidade p, enquanto a distribuição de Poisson é afetada apenas pela média μ.
2. Em uma distribuição binomial, os valores possíveis da variável aleatória x são 0, 1, 2, ..., n, enquanto em uma distribuição de Poisson os valores possíveis de x são 0, 1, ..., sem limite superior.

EXEMPLO Para fins de análise dos impactos de bombas V-1 na Segunda Guerra Mundial, o sul de Londres foi subdividido em 576 regiões com área de 0,25 km² cada. A área conjunta das 576 regiões foi atingida por 535 bombas. Escolhida aleatoriamente uma região, determine a probabilidade de ela ter sido atingida exatamente duas vezes.

SOLUÇÃO Aplica-se a distribuição de Poisson porque estamos em face de ocorrências de impactos de bomba no intervalo de uma região. O número médio de impactos por região é $535/576 = 0{,}929$. Como queremos a probabilidade de dois impactos em uma região, fazemos $x = 2$ e aplicamos a Fórmula 4-10:

$$P(x) = \frac{\mu^x \cdot e^{-\mu}}{x!} = \frac{0{,}929^2 \cdot 2{,}71828^{-0{,}929}}{2!}$$
$$= \frac{0{,}863 \cdot 0{,}395}{2} = 0{,}170$$

A probabilidade de uma região particular ser atingida exatamente duas vezes é, pois, $P(2) = 0{,}170$.

No exemplo precedente, podemos também calcular as probabilidades de 0, 1, 3, 4 e 5 impactos. (Paramos em $x = 5$ porque nenhuma região foi atingida mais do que cinco vezes, e as probabilidades de $x > 5$ são todas 0,000 quando arredondadas para três decimais.) A Tabela 4-4 dá essas probabilidades. Com as probabilidades assim calculadas, podemos achar o número esperado de regiões com 0 ataques, 1 ataque etc. Por exemplo, as 576 regiões têm, cada uma, uma probabilidade de 0,395 de não serem atingidas; o número esperado de regiões não-atingidas é, pois, $(576)(0{,}395) = 227{,}5$. (Observe que, ao calcular o valor esperado, utilizamos o número de *regiões*, e não de ataques.) Esse valor esperado está relacionado juntamente com os outros na terceira coluna da Tabela 4-4. A quarta coluna dá os resultados efetivos. Houve 229 regiões não-atingidas, 211 regiões atingidas uma vez etc. Comparando as freqüências *preditas* pela distribuição de Poisson (terceira coluna) com as freqüências *efetivas* (quarta coluna), concluímos que há uma concordância muito boa. Nesse caso, a distribuição de Poisson apresentou um bom resultado, predizendo resultados que ocorreram efetivamente. Na Seção 10-2 damos um processo estatístico para determinar se tais freqüências esperadas constituem um bom "ajuste" das freqüências efetivas.

Às vezes usa-se a distribuição de Poisson como aproximação da distribuição binomial, quando n é grande e p é pequeno. Uma regra empírica consiste em utilizar essa aproximação quando $n \geq 100$ e $np \leq 10$. Ao utilizarmos a distribuição de Poisson como aproximação da binomial, podemos achar o valor da média μ pela Fórmula 4-7: $\mu = n \cdot p$. (Veja Exercícios 13 e 14.)

TABELA 4-4 Ataques de Bomba V-1 em 576 Regiões do Sul de Londres

N.º de Ataques	Probabilidades	N.º Esperado de Regiões	N.º Efetivo de Regiões
0	0,395	227,5	229
1	0,367	211,4	211
2	0,170	97,9	93
3	0,053	30,5	35
4	0,012	6,9	7
5	0,002	1,2	1

Este capítulo apresentou uma diversidade de distribuições de probabilidade *discretas* diferentes, inclusive a binomial (Seções 4-3 e 4-4), a de Poisson (esta seção), a geométrica (Exercício 33 da Seção 4-3), a hipergeométrica (Exercício 34 na Seção 4-3) e a multinomial (Exercício 35, Seção 4-3). No capítulo seguinte, vamos focalizar uma distribuição extremamente importante — a distribuição *normal* de probabilidade, que é *contínua*, e não discreta.

4-5 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-4, suponha que a distribuição de Poisson tenha a média indicada e aplique a Fórmula 4-10 para achar a probabilidade do valor dado da variável aleatória x.

1. $\mu = 2, x = 3$ 0,180
2. $\mu = 4, x = 1$
3. $\mu = 0{,}845, x = 2$
4. $\mu = 0{,}250, x = 2$

Nos Exercícios 5-12, utilize a distribuição de Poisson para achar as probabilidades indicadas. $\mu = 2{,}25$

5. Está sendo planejado um novo hospital para Newtown, uma comunidade que ainda não tem hospital próprio. Se Newtown tem uma média de 2,25 nascimentos por dia, determine a probabilidade de que, em um dia, o número de nascimentos seja

   a. 0 b. 1 c. 4

6. A Scott Auto Park vende em média 0,5 carro por dia. Determine a probabilidade de que, em um dia qualquer, o número de carros vendidos seja

   a. 0 b. 1 c. 2

7. A Townsend Manufacturing Company acusa uma média semanal de 0,2 acidentes que exigem cuidados médicos. Determine a probabilidade de que, em uma semana qualquer, o número de acidentes que exigem cuidado médico seja a) 0,818

   a. 0 b. 1 c. 2 b) 0,163 c) 0,0164

8. Uma professora de estatística verifica que, quando ela marca uma hora do expediente para atendimento aos alunos, comparecem em média dois alunos. Determine a probabilidade de que, em uma dessas horas selecionada aleatoriamente, o número de alunos que comparecem seja

   a. 0 b. 2 c. 5

9. Um estudo cuidadoso de uma fita magnética de dados de computador mostra uma incidência de 2,0 defeitos para cada 500 pés de fita. Determine a probabilidade de mais de um defeito em 500 pés de fita selecionados aleatoriamente.

10. Em um ano recente, verificaram-se 46 casos de seqüestro de avião em todo o mundo (com base em dados do Departamento de Aviação dos EUA). Tomando um dia como intervalo exigido pela distribuição de Poisson, vemos que o número diário médio de seqüestros de avião pode ser estimado em $\mu = 46/365 = 0{,}126$. Se a ONU está cogitando organizar um grupo único de combate a seqüestro, é preciso conhecer as chances de seqüestros múltiplos em um dia. Tomando $\mu = 0{,}126$, calcule a probabilidade de o número de seqüestros (x) em um dia ser 0 ou 1. Um grupo único é suficiente?

11. Um exemplo clássico da distribuição de Poisson envolve o número de homens mortos por coice de cavalo no exército prussiano entre 1875 e 1894. Combinaram-se os dados de 14 corpos de exército para o período de 20 anos, e os 280 corpos-ano acusaram um total de 196 mortes. Após achar o número médio de mortes por corpo-ano, determine a probabilidade de que, em um corpo-ano aleatório, ocorram os seguintes números de mortes: b) 0,3476 e) 0,0050

    $\mu = \frac{196}{280} = 0{,}7$

    a. 0 b. 1 c. 2 d. 3 e. 4 c) 0,1217 d) 0,0284

    4966 Os resultados efetivos consistiram nas seguintes freqüências: 0 mortes (em 144 corpos-ano); 1 morte (em 91 corpos-ano); 2 mortes (em 32 corpos-ano); 3 mortes (em 11 corpos-ano); 4 mortes (em 2 corpos-ano). Compare os resultados efetivos com os resultados esperados pela distribuição de Poisson. A distribuição se revela um bom preditor dos resultados efetivos? sim

12. Em um ano recente, houve 116 mortes por homicídio em Richmond, Virginia [(com base em "A Classroom Note On the Poisson Distribution: A Model for Homicidal Deaths in Richmond, VA for 1991" (Uma Nota de Aula sobre a Distribuição de Poisson: Um Modelo para Mortes por Homicídio em Richmond, VA, em 1991), *Mathematics and Computer Education*, por Winston A. Richards).] Para um dia escolhido aleatoriamente, determine a probabilidade de o número de mortes por homicídio ser

    a. 0 b. 1 c. 2 d. 3 e. 4

    Compare as probabilidades calculadas com os resultados efetivos: 268 dias (0 homicídios); 79 dias (1 homicídio); 17 dias (2 homicídios); 1 dia (3 homicídios); não houve qualquer dia com mais de 3 homicídios.

4-5 Exercícios B: Além do Básico

13. Suponha que um experimento binomial comporte 15 provas, cada uma com 0,01 de probabilidade de sucesso. Determine a probabilidade de exatamente um sucesso nas 15 provas, utilizando (a) a Tabela A-1 e (b) a distribuição de Poisson como aproximação da distribuição binomial. Note que a regra empírica, que exige $n \geq 100$ e $np \leq 10$, sugere que a distribuição não é uma boa aproximação da distribuição binomial.

14. Dá-se a seguir um experimento binomial, onde o grande número de provas pode causar problemas sérios com muitas calculadoras. Supere esse obstáculo aproximando a distribuição binomial pela distribuição de Poisson.

    Apostando no 7 em uma rodada de roleta, temos uma probabilidade de 1/38 de ganhar. Suponha que apostemos no 7 em cada uma de 500 rodadas.

    a. Determine o número médio de ganhos em tais experimentos.
    b. Determine a probabilidade de o 7 ocorrer exatamente 13 vezes.
    c. Compare o resultado com a probabilidade de 0,111 obtida pela fórmula da probabilidade binomial.

Vocabulário

variável aleatória
variável aleatória discreta
variável aleatória contínua
distribuição de probabilidade
histograma de probabilidade
valor esperado
experimento binomial
fórmula da probabilidade binomial
distribuição de Poisson

Revisão

Os pontos centrais deste capítulo foram a variável aleatória e a distribuição de probabilidade. Nele abordamos exclusivamente distribuições discretas de probabilidade. (No Capítulo 5 trataremos das distribuições contínuas de probabilidade.) Discutimos os seguintes pontos-chave:

- Em um experimento que dá resultados numéricos, a *variável aleatória* tem valores numéricos que correspondem a diferentes resultados aleatórios de um experimento.
- Uma *distribuição de probabilidade* consiste em todos os valores de uma variável aleatória, juntamente com suas respectivas probabilidades. Qualquer distribuição de probabilidade deve satisfazer as duas condições: $\Sigma P(x) = 1$ e, para cada valor de x, $0 \leq P(x) \leq 1$.
- Podemos investigar as características importantes de uma *distribuição de probabilidade* calculando sua média (Fórmula 4-1) e seu desvio-padrão (Fórmula 4-4) e construindo seu histograma de probabilidade.

- Em um *experimento binomial*, podemos achar as probabilidades na Tabela A-1, podemos calculá-las com auxílio da Fórmula 4-5 ou de uma calculadora TI-83, ou podemos determiná-las por um programa de software, como STATDISK ou Minitab.
- Em um *experimento binomial*, acham-se facilmente a média e o desvio-padrão calculando os valores de $\mu = n \cdot p$ e $\sigma = \sqrt{n \cdot p \cdot q}$.
- Uma *distribuição de probabilidade de Poisson* se aplica a ocorrências de um evento em um intervalo especificado, e suas probabilidades podem ser calculadas pela Fórmula 4-10.

Exercícios de Revisão

1. a. O que é uma variável aleatória?
   b. O que é uma distribuição de probabilidade?
   c. Um estudo levado a efeito por uma associação de seguro sobre o uso de detetor de fumaça doméstico envolveu casas selecionadas aleatoriamente em grupos de quatro. A tabela a seguir relaciona valores e probabilidades de x, número de casas (em grupos de quatro) que têm detetores de fumaça instalados (com base em dados da National Fire Protection Association). Esta tabela descreve uma distribuição de probabilidade? Por quê?
   d. Supondo que a tabela a seguir descreva uma distribuição de probabilidade, calcule sua média.
   e. Supondo que a tabela a seguir descreva uma distribuição de probabilidade, calcule seu desvio-padrão.

x	$P(x)$
0	0,0004
1	0,0094
2	0,0870
3	0,3562
4	0,5470

2. Quinze por cento dos carros tipo esporte/compacto têm cor verde-escura (com base em dados da DuPont Automotive). Selecionados aleatoriamente 50 carros esporte/compacto,
   a. Qual é o número esperado de carros verde-escuros em um tal grupo de 50?
   b. Em tais grupos de 50, qual é o número médio de carros verde-escuros?
   c. Em tais grupos de 50, qual é o desvio-padrão do número de carros verde-escuros? c. 2,5
   d. É raro ter 15 carros verde-escuros em tal grupo? Por quê?
   e. Determine a probabilidade de exatamente 9 carros verde-escuros em um tal grupo de 50. e. 0,[illegible]
3. Trinta por cento dos estudantes de faculdades possuem videocassete (com base em dados da American Passage Media Corporation). A Telektronic Company produziu um videoteipe e enviou cópias a 10 estudantes de faculdade selecionados aleatoriamente, como parte de um programa-piloto de vendas.
   a. Determine a probabilidade de exatamente a metade dos 10 estudantes ter videocassete. 3. a. 0,103
   b. Determine a probabilidade de ao menos a metade dos 10 estudantes ter videocassete. b. 0,150
   c. Se os videoteipes são enviados a estudantes de faculdade selecionados aleatoriamente em muitos grupos diferentes de 10, determine a média e o desvio-padrão do número (entre 10) dos que têm videocassete.
4. A incapacidade de conviver com os outros é a razão citada em 17% dos casos de demissão de empregados (com base em dados da Robert Half International, Inc.). Preocupada com as condições de trabalho em sua empresa, a gerente de pessoal da Flint Fabric Company planeja investigar as 5 demissões de empregados ocorridas no ano passado. Supondo aplicável a taxa de 17%, determine a probabilidade de que, entre os 5 empregados, o número dos demitidos por incapacidade de conviver com os outros seja
   a. 0 **b.** 4 **c.** 5 **d.** ao menos 3
   (Uma vez identificadas as razões das demissões, essas probabilidades podem auxiliar a comparar a Flint Fabric Company com outras companhias.)
5. Considere os dados do Exercício 4. Represente pela variável aleatória x o número de empregados demitidos (entre os 5) por incapacidade de convivência com os outros.
   a. Determine o valor médio de x. 5. a. 0,85
   b. Ache o desvio-padrão da variável aleatória x. b. 0,84
   c. Considerando excepcionais quaisquer valores que distem da média por mais de dois desvios-padrão, é excepcional o fato de 4 empregados (entre 5) terem sido demitidos por incapacidade de convivência?
6. A Washington and Chang Trucking Company opera uma grande frota de caminhões. No ano passado, houve 84 casos de avaria.
   a. Determine o número diário médio de avarias. 6. a. μ = 0,230
   b. Determine a probabilidade de dois caminhões apresentarem avaria em um dia selecionado aleatoriamente. b. 0,0210

Exercícios Cumulativos de Revisão

1. A Sports Associates Vending Company fornece refrescos em um estádio de beisebol e deve se preparar para a possibilidade de uma Série Mundial de jogos. Na tabela de freqüência a seguir (baseada em resultados passados), x representa o número de jogos de beisebol necessários para completar uma Série Mundial.
   a. Construa a tabela de freqüências relativas correspondente.
   b. O resultado da parte (a) descreve uma distribuição de probabilidade? Por quê?
   c. Com base em resultados passados, qual é a probabilidade de a próxima Série Mundial ter ao menos 5 jogos? c. 0,854
   d. Selecionadas aleatoriamente duas séries diferentes incluídas na tabela, determine a probabilidade de ambas terem ao menos 7 jogos. d. 0,143
   e. Determine o número médio de jogos da Série Mundial incluídos na tabela. e. 58
   f. Determine o desvio-padrão do número de jogos da Série Mundial incluídos na tabela.
   g. Qual é o número esperado de jogos de uma Série Mundial? Se um vendedor fornece sanduíches para ambos os estádios envolvidos, e se em cada estádio são vendidos em média 30.000 sanduíches por jogo, qual o número esperado de sanduíches necessários?

x	f
4	13
5	22
6	20
7	34

2. Um cassino é flagrado tentando utilizar um par de dados viciados. No julgamento, ficou evidenciado que alguns pontos pretos eram escavados, enchidos com chumbo e repintados a fim de parecerem normais. Além da evidência física, os dados foram jogados no tribunal, com os seguintes resultados;

   12 8 9 12 12 9 8 7 12 10
   12 3 2 12 10 9 12 11 11 12

   Um perito em probabilidade afirma que, na jogada de dados equilibrados (honestos), a média deve ser 7,0, e o desvio-padrão deve ser 2,4.
   a. Determine a média e o desvio-padrão dos valores amostrais obtidos no julgamento.
   b. Com base nos resultados obtidos no julgamento, qual é a probabilidade de obter um 12? Compare esse resultado com a probabilidade de 1/36 (ou 0,0278) para dados equilibrados.

c. Se a probabilidade de obter 12 com dados equilibrados é 1/36, determine a probabilidade de obter ao menos um 12 em 20 jogadas de dados equilibrados.
d. Se o leitor fosse o advogado de defesa, como refutaria os resultados obtidos no tribunal?

Projeto para Computador

O vôo 2705 da American Air, de New York para San Francisco, tem assentos para 340 passageiros. Em média, 5% dos que reservam lugar não comparecem ao embarque, de forma que a American Air excede o limite de vendas, aceitando 350 reservas para os 340 assentos. Podemos analisar o sistema tratando-o como um experimento binomial com $n = 350$ e $p = 0,95$ (a probabilidade de que alguém com reserva realmente compareça).

Determine a probabilidade de que, para determinado vôo, o número de passageiros supere o de assentos. Isto é, determine a probabilidade de comparecimento de ao menos 341 passageiros com reserva. Em virtude do valor de n, não podemos utilizar a Tabela A-1; a fórmula binomial exigiria considerável tempo e seria extremamente trabalhosa. O melhor recurso é um *software* estatístico ou a calculadora TI-83.

Com STATDISK, selecione `Analysis` da barra principal do menu; selecione a opção `Binomial Probabilities`. Introduza então 350 como Número de Provas e 0,95 como probabilidade de sucesso em uma prova.

Com Minitab, crie primeiro uma coluna C1 com os inteiros de 0 a 350, utilizando as opções `Calc` e, em seguida, `Set Patterned Data`. Determine que o resultado seja armazenado na coluna C1; escolha então uma seqüência padronizada para começar com 0 e terminar com 350, com incrementos de 1. Clique `OK`. Selecione então `Calc, Probability Distributions` e `Binomial`. Escolha a opção `Cumulative Probability`, introduza 350 como número de provas e 0,95 como probabilidade de sucesso; introduza C1 para a coluna de entrada (*input*); clique `OK`. Interprete os resultados a fim de achar a probabilidade desejada.

Com uma calculadora TI-83, acione `2nd VARS` (para obter `DISTR` para distribuições); selecione então `binomcdf` e introduza `binomcdf (350,0,95,340)` para obter a probabilidade de 340 ou menos sucessos. A seleção de `binomcdf` chama a "função de densidade binomial acumulada", que dá a soma acumulada das probabilidades de $x = 0$ até o valor desejado inclusive.

DOS DADOS PARA A DECISÃO

Um Vôo Transatlântico com Dois Motores É Seguro?

A estatística desempenha seu melhor papel quando é usada de alguma forma em benefício da humanidade. As empresas utilizam a estatística para se tornar mais eficientes, aumentar o lucro dos acionistas e reduzir os preços. As agências reguladoras aplicam a estatística para garantir a segurança de empregados e clientes. Este exercício envolve uma situação em que tanto a eficiência do custo como a segurança dos passageiros são fatores críticos. Com novos projetos de aeronave e maior confiabilidade dos motores, as companhias de aviação pretendiam cobrir rotas transatlânticas com jatos de dois motores; mas a Federal Aviation Administration — FAA (Administração Federal da Aviação) exigia ao menos três motores para vôos transatlânticos. A diminuição dessa exigência era, naturalmente, de grande interesse para os fabricantes de jatos com dois motores (como o Boeing 767). Outrossim, os jatos com dois motores consomem a metade do combustível consumido pelos jatos com três ou quatro motores. Obviamente, o ponto crucial ao aprovar a pretensão é a probabilidade de um jato com dois motores proporcionar uma viagem transatlântica segura. Essa probabilidade devia ser comparada com a dos jatos de três e quatro motores. Um estudo dessa natureza deveria envolver a plena compreensão das probabilidades relacionadas.

A estimativa realística da probabilidade de um motor falhar em um vôo transatlântico é de 1/14.000. Utilize essa probabilidade e a fórmula de probabilidade binomial para achar as probabilidades de 0, 1, 2 e 3 falhas de motor para um jato de três motores e a probabilidade de falha de 0, 1 e 2 motores para um jato de dois motores. Em vista dos números em jogo, tome em todos os resultados tantas decimais quantas sua calculadora permitir. Resuma os resultados introduzindo as probabilidades constantes das duas tabelas.

Três Motores

x	$P(x)$
0	?
1	?
2	?
3	?

Dois Motores

x	$P(x)$
0	?
1	?
2	?

Utilize os resultados das tabelas e suponha que um vôo seja completado se ao menos um dos motores funcionar. Determine a probabilidade de um vôo seguro com um jato de três motores e a mesma probabilidade para um jato de dois motores. Faça um relatório para a Administração Federal de Aviação esboçando o problema-chave e inclua uma recomendação. Corrobore sua recomendação com resultados específicos.

ATIVIDADES EM GRUPO

1. *Atividade em Classe:* Divida a classe em grupos de três ou quatro. Seja a variável aleatória x o valor de uma moeda selecionada aleatoriamente entre as moedas no bolso de um estudante de estatística. Com base nas moedas pertencentes aos membros do grupo, construa uma tabela (análoga à Tabela 4-1), relacionando os valores possíveis de x juntamente com as respectivas probabilidades; ache então a média e o desvio-padrão. Qual a utilidade prática dos resultados?
2. *Atividade Extraclasse:* Divida a classe em grupos de três ou quatro. Faça uma pesquisa que inclua as três questões seguintes. (Como a primeira questão pode ferir suscetibilidades de algumas pessoas, utilize um processo que preserve a anonimidade.)

   - Quanto você pesa?
   - Introduza três algarismos aleatórios (0, 1, 2, 3, 4, 5, 6, 7, 8, 9). Um algarismo pode ser escolhido mais de uma vez.
   - Introduza os três últimos algarismos do seu número de inscrição no seguro social.

   Primeiro, produza uma relação consistindo no *último algarismo* de cada peso. Se os pesos são precisos, é de esperar que esses *últimos* algarismos tenham uma distribuição uniforme, de modo que cada algarismo tenha uma probabilidade de 1/10. Admitindo essa probabilidade, construa uma tabela que descreva a distribuição de probabilidade dos últimos algarismos. Em seguida, construa uma tabela de freqüências relativas com base na lista de *últimos algarismos* registrados. Compare a distribuição de probabilidade com a tabela de freqüências relativas. (Esse processo costuma ser usado como teste para determinar se as pessoas foram realmente pesadas ou se simplesmente relataram seus pesos.) Utilize processo análogo para analisar os algarismos aleatórios escolhidos. Esses algarismos parecem ser aleatórios? Finalmente, utilize o mesmo processo para analisar os algarismos nos números de inscrição no seguro social. Eles parecem ser aleatórios?

Rayane Andrade

5 Triola

A Distribuição Normal de Probabilidade

5-1 Aspectos Gerais

No Capítulo 4 abordamos distribuições discretas de probabilidade; neste capítulo vamos estudar a distribuição normal, que é contínua. Todo o capítulo é dedicado à distribuição normal, por se tratar da distribuição mais importante em estatística, e utilizada repetidamente no restante deste livro.

5-2 A Distribuição Normal Padronizada

Define-se a distribuição normal padronizada como uma distribuição normal de probabilidade com média $\mu = 0$ e desvio-padrão $\sigma = 1$. Damos nesta seção os métodos básicos para determinar probabilidades aplicando essa distribuição e para achar valores padronizados correspondentes a probabilidades dadas.

5-3 Distribuições Normais Não-Padronizadas: Determinação de Probabilidades

Utiliza-se o escore z (ou valor padronizado) para trabalhar com distribuições normais em que a média não é 0, ou o desvio-padrão não é 1, ou ambos.

5-4 Distribuições Normais Não-Padronizadas: Cálculo de Valores

Prossegue-se o trabalho com distribuições normais não-padronizadas através de métodos para determinação de valores que correspondam a probabilidades dadas.

5-5 O Teorema Central do Limite

Na medida em que o tamanho da amostra aumenta, mostra-se que a distribuição das médias amostrais tende para uma distribuição normal com média μ e desvio-padrão $\sigma/\sqrt{n}$, onde n é o tamanho da amostra, e μ e σ representam a média e o desvio-padrão da população, respectivamente. As idéias desta seção constituem o fundamento para os importantes conceitos introduzidos nos Capítulos 6 e 7.

5-6 A Distribuição Normal como Aproximação da Distribuição Binomial

Descrevemos e ilustramos a utilização da distribuição normal para estimar probabilidades em um experimento binomial.

Prof.: Jorge

Disc.: Estatística

Problema do Capítulo

Então, pretende ser um astronauta?

Se o leitor tem 6 pés e 8 polegadas de altura e pesa 350 lb, pode ser um excelente defensor de um time de rugby — mas não seria escolhido como astronauta. Recentemente, a cooperação entre os programas espaciais russo e americano focalizou o problema da altura dos astronautas. A espaçonave russa Soyuz leva astronautas ao espaço e serve como escape de emergência da estação espacial russa Mir, mas o projeto de construção da Soyuz exige que os astronautas tenham altura entre 64,5 e 72 polegadas, devendo também pesar menos de 188 lb. As espaçonaves americanas são muito mais espaçosas e exigem que os astronautas tenham altura entre 58,5 e 76 polegadas. Em razão das restrições mais severas impostas pela espaçonave Soyuz, a NASA se defrontou com uma falta de astronautas qualificados para participar dos programas conjuntos com a Rússia. Assim é que a NASA passou a cogitar de uma campanha de recrutamento com o objetivo de obter mais astronautas qualificados que satisfizessem as condições mais restritivas de altura e peso da Rússia.

As restrições quanto à altura são importantes na formação de equipes de astronautas, na seleção de dançarinos para um determinado show, na participação em certos brinquedos em parques de diversão e na admissão ao serviço militar. Em tais casos é de grande utilidade o conhecimento da distribuição não só das alturas como de outras medidas físicas importantes. Por exemplo, é possível calcularmos a percentagem dos homens americanos que têm entre 64,5 e 72,0 polegadas de altura? Ou das mulheres que se enquadram nessa faixa? Neste capítulo abordaremos métodos para calcular tais percentagens.

5-1 Aspectos Gerais

No Capítulo 4 introduzimos o conceito de *variável aleatória* como uma variável que toma um valor numérico único (determinado pela chance) para cada resultado de um experimento. Vimos que uma *distribuição de probabilidade* dá a probabilidade de cada valor da variável aleatória. O Capítulo 4 abrange apenas variáveis aleatórias *discretas*, como as de distribuições binomiais, que têm um número finito de valores possíveis. O número de moedas de 25 centavos produzidas em um dia pela Casa da Moeda dos EUA é um exemplo de uma distribuição discreta. Há também não poucas distribuições *contínuas* de probabilidade, como os pesos das moedas produzidas. As distribuições são discretas ou contínuas e podem ser descritas pela sua *forma* — como, por exemplo, forma de sino. Este capítulo aborda as distribuições normais — extremamente importantes porque ocorrem com grande freqüência nas aplicações. Alturas de mulheres adultas, pesos de homens adultos e notas de um teste de terceiro grau são apenas alguns exemplos de populações distribuídas normalmente.

DEFINIÇÃO

Uma variável aleatória contínua tem **distribuição normal** se essa distribuição é simétrica e apresenta a forma de um sino, conforme Figura 5-1; a distribuição se ajusta à equação dada na Fórmula 5-1.

Fórmula 5-1
$$y = \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}$$

Para muitos estudantes, a complexidade da Fórmula 5-1 é assustadora — mas felizmente não teremos necessidade de utilizá-la tal como se apresenta. O que a fórmula mostra é que qualquer distribuição normal é determinada por dois parâmetros: a média μ e o desvio-padrão σ. Dados valores específicos para μ e σ, podemos fazer o gráfico da Fórmula 5-1, como o faríamos para qualquer equação relacionando x e y; o resultado é uma distribuição de probabilidade com a forma de um sino. Veremos que esta distribuição normal tem inúmeras aplicações, e vamos utilizá-la com freqüência em todos os capítulos restantes.

5-2 A Distribuição Normal Padronizada

Iniciamos esta seção utilizando uma distribuição de probabilidade contínua para ilustrar uma importante correspondência entre área e probabilidade. É mais fácil visualizar essa correspondência com auxílio de uma distribuição como a ilustrada na Figura 5-2, que mostra a distribuição de probabilidade para tem-

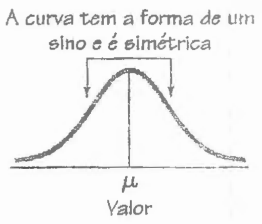


Fig. 5-1 A distribuição normal.

peraturas em um processo de fabricação. Essas temperaturas são controladas de modo a variarem entre 0°C e 5°C, com todos os valores igualmente prováveis. É tal a freqüência com que ocorre este tipo particular de distribuição que lhe foi dado o nome de *distribuição uniforme*.

DEFINIÇÃO

Uma **distribuição uniforme** é uma distribuição de probabilidade em que todos os valores da variável aleatória são igualmente prováveis.

Na Seção 4-2, identificamos duas condições para uma distribuição de probabilidade: (1) $\Sigma P(x) = 1$ e (2) $0 \leq P(x) \leq 1$ para todos os valores de x. Dissemos também, naquela seção que o gráfico de uma distribuição discreta de probabilidade é chamado de um histograma de probabilidade. O gráfico de uma distribuição contínua de probabilidade, tal como a Figura 5-2, é chamado *curva de densidade* e deve verificar duas propriedades análogas às condições para as distribuições discretas de probabilidade, dadas na Definição.

DEFINIÇÃO

Uma **curva de densidade** é o gráfico de uma distribuição contínua de probabilidade. Deve satisfazer as propriedades seguintes:

1. A área total sob a curva deve ser 1.
2. Todo ponto da curva deve ter uma altura vertical não inferior a 0.

As temperaturas ilustradas na Figura 5-2 são controladas de modo a permanecerem entre 0°C e 5°C, qualquer um desses valores sendo igualmente provável. Fixando em 0,2 a altura do retângulo da Figura 5-2, a área limitada será forçosamente igual a $5 \times 0{,}2 = 1$, como desejado. Esta propriedade (área = 1) facilita a resolução de problemas de probabilidade. Por exemplo, dada a distribuição uniforme da Figura 5-2, a probabilidade de escolher aleatoriamente uma temperatura entre 1°C e 4°C é 0,6, que é a área da região sombreada. Eis um ponto importante: **Uma curva de densidade é o gráfico de uma variável aleatória contínua, de forma que a área sob a curva é 1, estabelecendo-se uma correspondência entre área e probabilidade.** Na Figura 5-2, a probabilidade de um valor entre 1°C e 4°C é 0,6, que pode ser obtida determinando-se a área da região retangular correspondente com dimensões $0{,}2 \times 3$.

Nova Tecnologia para Coleta de Dados

Medir ou codificar produtos e registrar manualmente os resultados constitui uma forma de coleta de dados, mas a tecnologia nos dá hoje alternativas que não são tão suscetíveis de erro humano. Os supermercados utilizam leitoras de códigos de barras para fixar preços e analisar inventários e hábitos de compra. Os fabricantes estão utilizando em escala crescente dispositivos de "introdução direta de dados", como medidores eletrônicos ligados diretamente a um computador que registra os resultados. Uma terceira possibilidade é a entrada dos dados através da voz, por meio de um fone ligado a uma pequena caixa conectada à cintura da pessoa que fala. A caixa transmite os dados a um computador.

EXEMPLO Uma distribuição uniforme de temperaturas varia de 0°C a 5°C. Escolhida aleatoriamente uma temperatura, determine a probabilidade de ela ser superior a 1°C.

SOLUÇÃO Na Figura 5-3, as temperaturas superiores a 1°C são representadas pelo retângulo sombreado, cuja área é $(4)(0{,}2) = 0{,}8$. Como há aqui uma correspondência entre área e probabilidade, concluímos que há uma probabilidade de 0,8 de uma temperatura selecionada aleatoriamente ser superior a 1°C.

No caso da distribuição uniforme, a curva de densidade origina um retângulo, sendo, assim, fácil determinar qualquer área simplesmente multiplicando a largura pela altura. A curva de densidade de uma distribuição normal tem a forma mais complicada de um sino, conforme a Figura 5-1, de modo que é mais difícil achar áreas, mas o princípio básico é o mesmo: Há uma correspondência entre área e probabilidade.

Assim como há muitas distribuições uniformes diferentes, há também muitas distribuições normais diferentes, cada uma dependendo de dois parâmetros: a média populacional μ e o desvio-padrão populacional σ. A Figura 5-4 mostra as curvas de densidade para alturas de mulheres e homens adultos. Como os homens têm altura média maior, a curva de densidade dos homens está deslocada para a direita. E como os homens têm

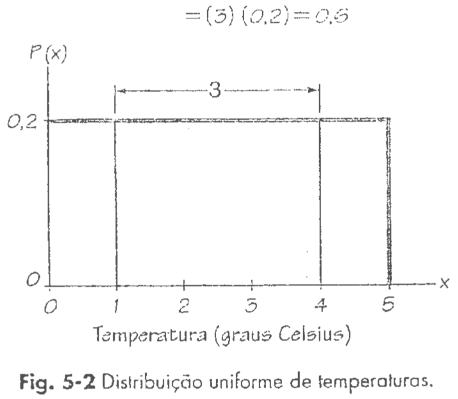


Fig. 5-2 Distribuição uniforme de temperaturas.

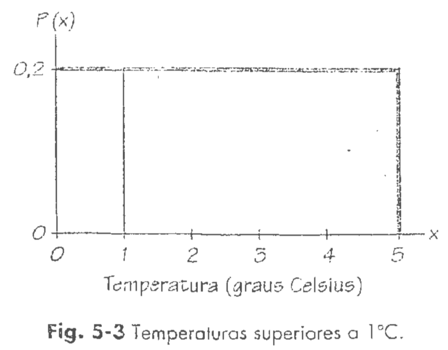


Fig. 5-3 Temperaturas superiores a 1°C.

um desvio-padrão ligeiramente maior, a curva de densidade dos homens é um pouco mais aberta. Dentre as infinitas possibilidades, há uma distribuição normal que tem especial interesse.

DEFINIÇÃO

> A **distribuição normal padronizada** é uma distribuição normal de probabilidade que tem média 0 e desvio-padrão 1. (Veja a Figura 5-5.)

Suponha que, por algum motivo, fôssemos forçados a fazer cálculos utilizando a Fórmula 5-1. Veríamos logo que os valores mais manejáveis de μ e σ são $\mu = 0$ e $\sigma = 1$. Fazendo $\mu = 0$ e $\sigma = 1$, os matemáticos calcularam áreas sob a curva. Como se pode ver na Figura 5-5, a área sob a curva, entre a média 0 e o valor 1, é 0,3413. Tenha em mente que a área total sob a curva é sempre 1; isto nos permite estabelecer a correspondência entre área e probabilidade, como fizemos no exemplo precedente com a distribuição uniforme.

Determinação de Probabilidades Quando São Dados os Escores z

A Figura 5-5 mostra que a área delimitada pela curva, o eixo horizontal e os valores 0 e 1 é igual a 0,3413. Embora a figura mostre apenas uma área, a Tabela A-2 no Apêndice A inclui áreas (ou probabilidades) para muitas regiões diferentes.

A Tabela A-2 dá a probabilidade correspondente à área sob a curva, delimitada à esquerda por uma reta vertical passando pela média 0 e à direita por uma reta vertical passando por um valor arbitrário denotado por z (Figura 5-6). (A calculadora TI-83 pode dar áreas do tipo achado na Tabela A-2; recorra à função identificada como `normalcdf`, que representa uma função de densidade normal acumulada.) Note que, ao utilizar a Tabela A-2, a parte em centésimos do escore z se encontra na linha superior. No exemplo seguinte, devemos achar a probabilidade associada a um valor entre 0 e 1,58. Comecemos com o valor de 1,58 para z, localizando 1,5 na coluna esquerda; a seguir tomamos o valor da linha de probabilidades, diretamente abaixo de 0,08, conforme mostra o excerto da Tabela A-2 a seguir:

z	...	0,08
.		.
.		.
.		.
1,5	...	0,4429

O valor 0,4429 da área (ou probabilidade) indica que há uma probabilidade de 0,4429 de escolher aleatoriamente um valor entre 0 e 1,58. É essencial termos em mente que a Tabela A-2 foi compilada apenas para a distribuição normal padronizada, que tem média 0 e desvio-padrão 1. (O caso não-padronizado será considerado nas seções seguintes.) É também essencial evitarmos confundir valores de z com áreas. (z dá o número de desvios-padrão que um valor dista da média.)

> **Escore z: *distância* ao longo da escala horizontal no gráfico; recorrer à *coluna mais à esquerda* (*e à linha superior*) da Tabela A-2.**
>
> **Área (ou probabilidade): *região* sob a curva; recorrer aos números no *corpo* da Tabela A-2.**

EXEMPLO A Precision Scientific Instrument Company fabrica termômetros que devem acusar a leitura de 0°C no ponto de congelamento da água. Testes feitos em uma grande amostra desses termômetros revelaram que, no ponto de congelamento da água, alguns acusavam valor inferior a 0° (denotado por um número negativo) e alguns acusavam valor superior a 0°C (denotado por um número positivo). Suponha que a leitura média seja 0°C e que o desvio-padrão das leituras seja 1,00°C. Admita também que a distribuição de freqüência dos erros se assemelhe a uma distribuição normal. Escolhido um termômetro aleatoriamente, determine a probabilidade de que, no ponto de congelamento da água, o termômetro marque entre 0° e +1,58°.

SOLUÇÃO A distribuição de probabilidade das leituras é uma distribuição normal padronizada porque as leituras são distribuídas normalmente com $\mu = 0$ e $\sigma = 1$. Devemos achar a área entre 0 e z (a área sombreada) na Figura 5-6 para $z = 1,58$. Pela Tabela A-2 vemos que essa área é 0,4429. A probabilidade de escolher aleatoriamente um termômetro com erro entre 0° e +1,58° é, portanto, 0,4429. Outra maneira de interpretar este resultado é concluir que 44,29% dos termômetros terão erros entre 0° e +1,58°.

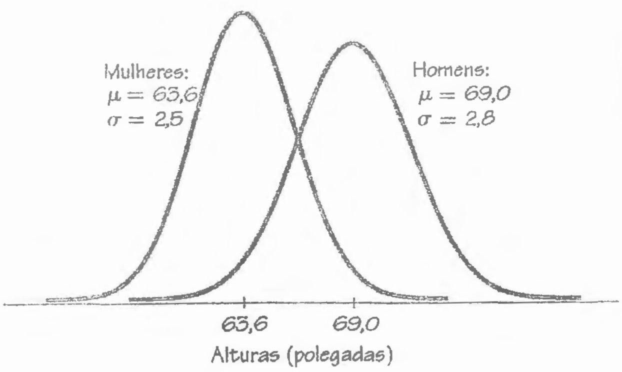


Fig. 5-4 Alturas de mulheres e homens adultos.

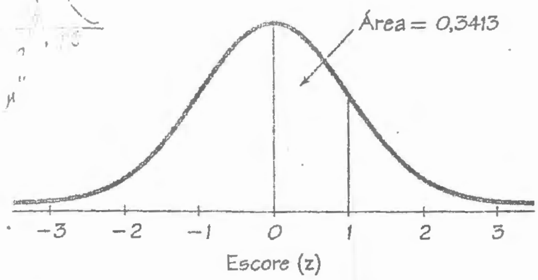
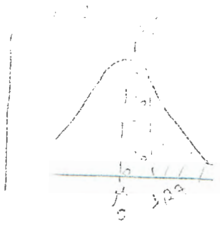


Fig. 5-5 Distribuição normal padronizada, com média $\mu = 0$ e desvio-padrão $\sigma = 1$.



EXEMPLO Com os termômetros do exemplo precedente, determine a probabilidade de selecionar aleatoriamente um termômetro que acuse (no ponto de congelamento da água) uma leitura entre −2,43° e 0°.

SOLUÇÃO Estamos interessados na região sombreada da Figura 5-7(a), mas a Tabela A-2 se aplica apenas a regiões à direita da média (0), como na Figura 5-7(b). Entretanto, comparando a área sombreada na Figura 5-7(a) com a área sombreada na Figura 5-7(b), vemos que essas duas áreas são idênticas, porque a curva de densidade é simétrica. Recorrendo à Tabela A-2, vemos facilmente que a área sombreada da Figura 5-7(b) é 0,4925, e assim a área sombreada da Figura 5-7(a) deve ser também 0,4925. Ou seja, a probabilidade de escolher um termômetro com erro entre −2,43° e 0° é 0,4925. Em outras palavras, 49,25% dos termômetros têm erro entre −2,43° e 0°.

A solução precedente ilustra um princípio importante:

Embora o escore z possa ser negativo, a área sob a curva (ou a probabilidade correspondente) nunca pode ser negativa.

Na Seção 2-5 apresentamos a regra prática, segundo a qual, para distribuições em forma de sino,

- cerca de 68% de todos os valores estão a menos de 1 desvio-padrão da média.
- cerca de 95% de todos os valores estão a menos de 2 desvios-padrão da média.
- cerca de 99,7% de todos os valores estão a menos de 3 desvios-padrão da média.

Se considerarmos a Figura 5-5 com $z = 1$, a Tabela A-2 mostra que a área sombreada é 0,3413. Decorre daí que a proporção de valores entre $z = -1$ e $z = 1$ será $0{,}3413 + 0{,}3413 = 0{,}6826$. Isto é, cerca de 68% de todos os valores estão a menos de 1 desvio-padrão da média. Cálculo análogo com $z = 2$ dá $0{,}4772 + 0{,}4772 = 0{,}9544$ (ou cerca de 95%) como a probabilidade de valores entre $z = -2$ e $z = 2$. Finalmente, a proporção de valores entre $z = -3$ e $z = 3$ é dada por $0{,}4987 + 0{,}4987 = 0{,}9974$ (ou cerca de 99,7%). Esses valores exatos correspondem, com boa aproximação, aos valores dados pela regra prática. Na realidade, os valores da regra prática foram achados diretamente a partir de probabilidades na Tabela A-2, tendo sido ligeiramente arredondados por questão de conveniência. A regra prática é chamada às vezes de *regra 68-95-99*; usando os valores exatos da Tabela A-2, seria chamada *68,26-95,44-99,74* — o que, entretanto, não soa tão bem.

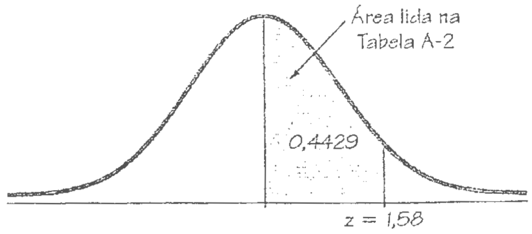


Fig. 5-6 Distribuição normal padronizada. A área da região sombreada delimitada pela média 0 e pelo número positivo z pode ser lida na Tabela A-2.

Como estamos lidando com a curva de densidade para uma distribuição de probabilidades, a área total sob a curva deve ser 1. Atentando para a Figura 5-8, vemos que uma reta vertical passando pela média 0 divide a área sob a curva em duas partes iguais, cada uma delas com área 0,5. No exemplo que segue utilizamos essa observação.

EXEMPLO Mais uma vez, faça uma escolha aleatória na mesma amostra de termômetros. Determine a probabilidade de que o termômetro escolhido acuse uma leitura superior a +1,27° (no ponto de congelamento da água).

SOLUÇÃO Novamente estamos em face de valores distribuídos normalmente com média 0° e desvio-padrão 1°. A probabilidade de escolher um termômetro que acuse leitura superior a +1,27° corresponde à área sombreada da Figura 5-8. Não podemos usar a Tabela A-2 para achar a área diretamente, mas podemos ver, ali, que $z = 1{,}27$ corresponde à área de 0,3980, conforme figura. Então, como a área à direita de zero é a metade da área total, ela é 0,5, e a área sombreada é $0{,}5 - 0{,}3980 = 0{,}1020$. Concluímos que há uma probabilidade de 0,1020 de escolher aleatoriamente um termômetro com leitura superior a +1,27°. Outra forma de interpretar este resultado é dizer que, num grande lote de termômetros escolhidos aleatoriamente e testados, 0,1020 (ou 10,20%) deles acusarão leitura superior a +1,27°.

Podemos determinar a área da região sombreada na Figura 5-8 por uma aplicação *indireta* da Tabela A-2. O exemplo que segue ilustra ainda outra aplicação indireta.

EXEMPLO Selecionado aleatoriamente um termômetro de nossa amostra, determine a probabilidade de acusar uma leitura entre 1,20° e 2,30° (no ponto de congelamento da água).

SOLUÇÃO A probabilidade de escolher um termômetro que acuse entre +1,20° e +2,30° corresponde à área sombreada da Figura 5-9. Mas a Tabela A-2 foi elaborada para regiões delimitadas à esquerda pela reta vertical em 0. Pela tabela, vemos que $z = 1{,}20$ corresponde a uma área de 0,3849, e que $z = 2{,}30$ corresponde a uma área de 0,4893, conforme a figura. Denotando por A a area da região sombreada, vemos, pela Figura 5-9, que

$$0{,}3849 + A = 0{,}4893$$

donde

$$A = 0{,}4893 - 0{,}3849 = 0{,}1044$$

Selecionado, pois, um termômetro aleatoriamente, a probabilidade de que ele acuse (no ponto de congelamento da água) leitura entre 1,20° e 2,30° é 0,1044.

O exemplo precedente conclui com a afirmação de que a probabilidade de uma leitura entre 1,20° e 2,30° é 0,1044. Tais probabilidades podem também ser expressas com a notação seguinte:

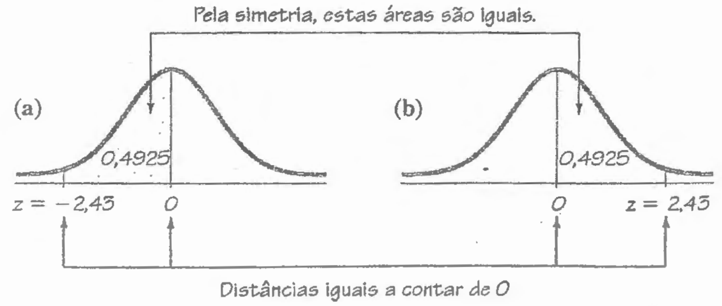


Fig. 5-7 Utilização da simetria para achar a área à esquerda da média.

Notação	
$P(a < z < b)$	denota a probabilidade de o valor de z estar entre a e b.
$P(z > a)$	denota a probabilidade de o valor de z ser maior do que a.
$P(z < a)$	denota a probabilidade de o valor de z ser menor do que a.

Com essa notação, podemos expressar o resultado do último exemplo como $P(1{,}20 < z < 2{,}30) = 0{,}1044$, o que afirma, em símbolos, que a probabilidade de o valor de z estar entre 1,20 e 2,30 é 0,1044. No caso de uma distribuição contínua de probabilidade, como a normal, a probabilidade de obter um valor *determinado* é 0; isto é, $P(z = a) = 0$. Assim, há uma probabilidade 0 de selecionar aleatoriamente alguém que tenha uma altura de exatamente 68,16243357 polegadas. Na distribuição normal, qualquer ponto individual na escala horizontal é representado não por uma região sob a curva, mas por uma reta vertical pelo ponto. Para $P(z = 1{,}33)$, temos uma reta vertical por $z = 1{,}33$; e essa reta vertical não tem área e, assim, $P(z = 1{,}33) = 0$. Para qualquer variável aleatória contínua, a probabilidade de um valor determinado é sempre zero, donde decorre que $P(a \leq z \leq b) = P(a < z < b)$. Decorre também que a probabilidade de obter um valor z, *no máximo* igual a b, é a mesma que a probabilidade de obter um valor z *menor do que b*. É importante interpretar corretamente expressões básicas como *no máximo*, *ao menos*, *mais de*, *não mais de* etc. As ilustrações da Figura 5-10 auxiliam na interpretação das expressões mais comuns.

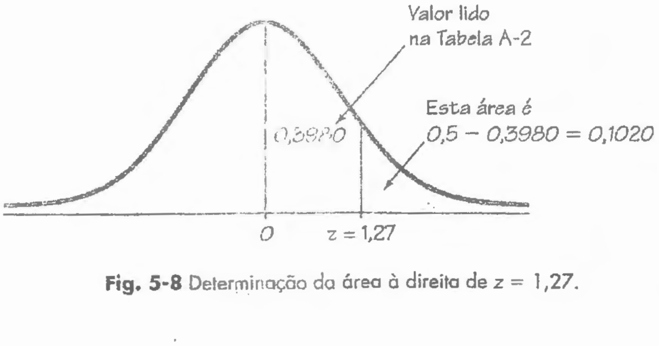


Fig. 5-8 Determinação da área à direita de z = 1,27.

Determinação dos Escores z Quando São Dadas Probabilidades

Até aqui, os exemplos desta seção que envolvem a distribuição normal padronizada têm seguido o mesmo esquema: Dado um valor, recorremos à Tabela A-2 para achar uma probabilidade. Mas há muitas outras circunstâncias em que conhecemos a probabilidade, e precisamos achar o valor de z correspondente. Em tais casos, é de suma importância evitar confusão entre valores de z e áreas. Recorde: os números da Tabela A-2 na coluna da extrema esquerda e através do topo são valores de z, que são *distâncias* ao longo da escala horizontal, enquanto os números no corpo da Tabela A-2 são *áreas* (ou probabilidades). Outrossim, os valores z à esquerda da linha do centro são sempre negativos (como na Figura 5-7a). Conhecida uma probabilidade, para determinar o valor z correspondente procedemos como segue:

1. Identificamos a probabilidade que representa uma área delimitada pela linha do centro; pode ser a probabilidade original, ou pode ser determinada por meio desta última. (Isto é, devemos ter a certeza de estar trabalhando com uma região delimitada pela linha do centro.)
2. Com a probabilidade que representa a área delimitada pela linha do centro, localizamos a probabilidade mais próxima no *corpo* da Tabela A-2 e identificamos o valor z correspondente.
3. Se o valor de z está à esquerda da linha do centro, consideramo-lo negativo.

EXEMPLO Utilize os mesmos termômetros com leituras de temperatura distribuídas normalmente com média 0°C e desvio-padrão 1°C. Determine a temperatura correspondente a P_{95}, o 95.º percentil. Isto é, determine a temperatura que separa os 95% da base dos 5% do topo. Veja a Figura 5-11.

SOLUÇÃO A Figura 5-11 mostra o valor de z correspondente ao 95.º percentil, que separa os 5% do topo dos 95% da base. Devemos recorrer à Tabela A-2 para achar esse valor de z utilizando uma região delimitada pela linha do centro (onde $\mu = 0$) em um lado, tal como a região sombreada de 0,45 na Figura 5-11. (Não esqueça que a Tabela A-2 dá diretamente apenas as áreas delimitadas à esquerda pela linha do centro e à direita pelo valor de z.) Procuramos primeiro a área 0,45 *no corpo da tabela* e determinamos então o valor z correspon-

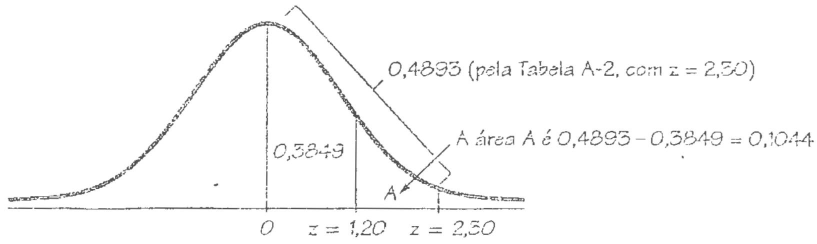


Fig. 5-9 Determinação da área entre $z = 1,20$ e $z = 2,30$.

dente. Na Tabela A-2, a área de 0,45 está entre os valores tabelares de 0,4495 e 0,4505, mas há um asterisco com uma nota especial indicando que 0,4500 corresponde ao valor 1,645 de z. Podemos assim concluir que o valor z na Figura 5-11 é 1,645, de modo que o 95.º percentil é a leitura de temperatura 1,645°C. Em um teste à temperatura de congelamento da água, 95% das leituras não superarão 1,645°C, e 5% delas não serão inferiores a 1,645°C.

Note que, na solução precedente, a Tabela A-2 levou a um valor z de 1,645, que está a meio caminho entre 1,64 e 1,65. Ao utilizar a Tabela A-2, podemos em geral evitar a interpolação simplesmente escolhendo o valor mais próximo. Todavia, há dois casos especiais que envolvem valores importantes pelo fato de serem utilizados com freqüência em uma grande diversidade de aplicações (veja a tabela que acompanha). Com exceção desses dois casos especiais, podemos escolher o valor mais próximo na tabela. (Se um valor procurado está a meio caminho entre dois valores da tabela, escolha o maior.) Outrossim, para valores de z superiores a 3,09, podemos tomar 0,4999 como aproximação da área correspondente.

Escore z	Área
1,645	0,4500
2,575	0,4950

EXEMPLO Com os mesmos termômetros, determine P_{10}, o 10.º percentil; isto é, determine a leitura de temperatura que separa os 10% inferiores dos 90% superiores de todas as temperaturas.

SOLUÇÃO Veja a Figura 5-12, onde o 10.º percentil é apresentado como o valor de z que separa os 10% inferiores dos 90% superiores. A Tabela A-2 é construída para áreas delimitadas pela linha do centro, de forma que nos referimos à área sombreada de 0,40 (correspondente a 50% − 10%). *No corpo da tabela*, escolhemos o valor mais próximo 0,3997 e constatamos que corresponde a $z = 1,28$. Todavia, como esse valor z está abaixo da média 0, deve ser negativo. O 10.º percentil é, pois, −1,28°C. Testados ao nível de congelamento, 10% das leituras do termômetro não superarão −1,28°C, e 90% das leituras não serão inferiores a −1,28°C.

Os exemplos desta seção foram elaborados de modo que a média 0 e o desvio-padrão 1 coincidissem exatamente com os

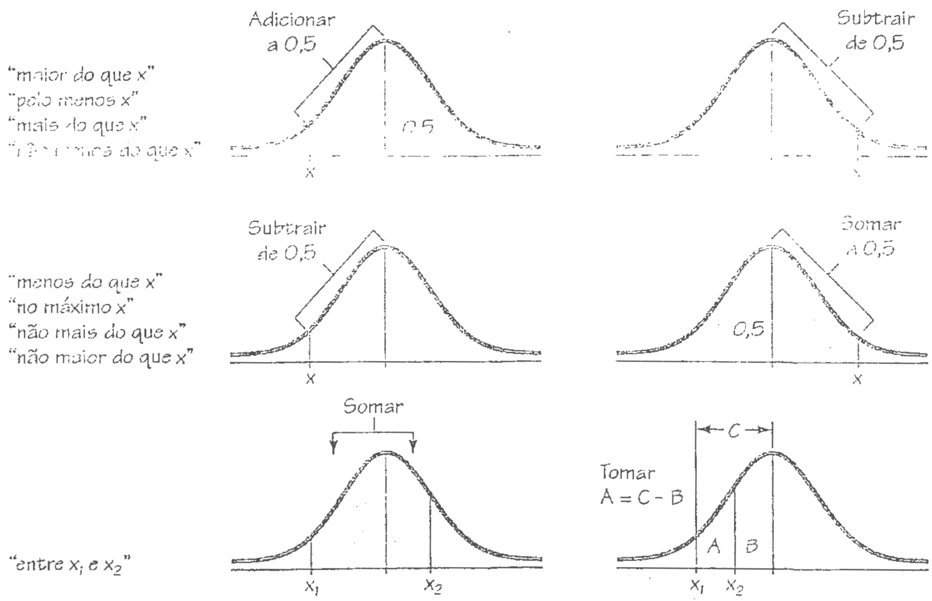


Fig. 5-10 Interpretação correta das áreas.

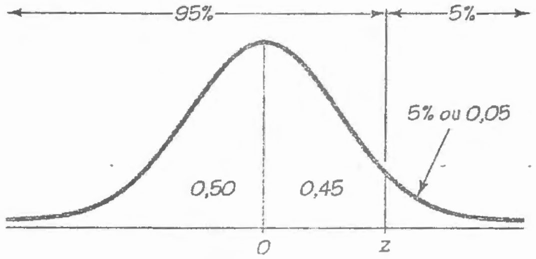


Fig. 5-11 Determinação do 95º percentil.

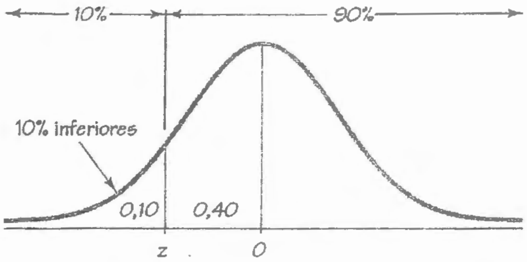


Fig. 5-12 Determinação do 10º percentil.

parâmetros da distribuição normal padronizada descrita na Tabela A-2. Na realidade, quase nunca ocorrem tais parâmetros convenientes, porque as distribuições normais típicas envolvem médias diferentes de 0 e desvios-padrão diferentes de 1. Na próxima seção introduziremos métodos para lidar com tais distribuições normais não-padronizadas.

5-2 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-4, com referência à distribuição contínua uniforme, ilustrada na Figura 5-2, suponha uma temperatura escolhida aleatoriamente e calcule a probabilidade de cada leitura em graus.

1. Maior do que 2
2. Menor do que 3
3. Entre 2 e 4
4. Entre 0,8 e 4,7

Nos Exercícios 5-24, suponha que as leituras dos termômetros tenham distribuição normal com média de 0° e desvio-padrão de 1,00°. Escolhe-se aleatoriamente e testa-se um termômetro. Em cada caso, faça um esboço e determine a probabilidade de cada leitura em graus.

5. Entre 0 e 3,00
6. Entre 0 e 1,96
7. Entre 0 e −2,33
8. Entre 0 e −1,28
9. Superior a 2,58
10. Inferior a −1,47
11. Inferior a −2,09
12. Superior a 0,25
13. Entre 1,34 e 2,67
14. Entre −1,72 e −0,31
15. Entre −2,22 e −1,11
16. Entre 0,89 e 1,78
17. Inferior a 0,08
18. Inferior a 3,01
19. Superior a −2,29
20. Superior a −1,05
21. Entre −1,99 e 2,01
22. Entre −0,07 e 2,19
23. Entre −1,00 e 4,00
24. Entre −5,00 e 2,00

Nos Exercícios 25-28, suponha que as leituras nos termômetros tenham distribuição normal com média de 0° e desvio-padrão de 1,00°. Determine a probabilidade indicada, sendo z a leitura em graus.

25. $P(z > 2{,}33)$
26. $P(2{,}00 < z < 2{,}50)$
27. $P(-3{,}00 < z < 2{,}00)$
28. $P(z < -1{,}44)$

P = 0,0099

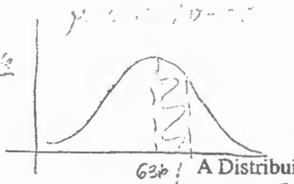
Nos Exercícios 29-36, suponha que as leituras nos termômetros tenham distribuição normal com média de 0° e desvio-padrão de 1,00°. Escolhe-se aleatoriamente e testa-se um termômetro. Em cada caso, faça um esboço e determine a leitura de temperatura correspondente à informação dada.

29. Determine P_{90}, o 90.º percentil. Esta é a temperatura que separa os 90% inferiores dos 10% superiores.
30. Determine P_{30}, o 30.º percentil.
31. Determine Q_1, a leitura correspondente ao primeiro quartil.
32. Determine D_1, a leitura correspondente ao primeiro decil.
33. Se 4% dos termômetros são rejeitados porque acusam leituras demasiadamente altas, enquanto todos os outros são aceitos, determine a leitura que separa os termômetros rejeitados dos termômetros restantes.
34. Se 8% dos termômetros são rejeitados porque acusam leituras demasiadamente baixas, enquanto todos os outros são aceitos, determine a leitura que separa os termômetros rejeitados dos termômetros restantes.
35. Um analista de controle de qualidade deseja examinar termômetros com leituras nos 2% inferiores. Que valor separa os 2% inferiores dos restantes?
36. Se 2,5% dos termômetros são rejeitados por acusarem leituras demasiadamente altas e outros 2,5% são rejeitados por acusarem leituras demasiadamente baixas, determine os dois valores que separam os termômetros rejeitados dos outros.

5-2 Exercícios B: Além do Básico

37. Suponha os escores z distribuídos normalmente com média 0 e desvio-padrão 1.
    a. Se $P(0 < z < a) = 0{,}3212$, determine *a*.
    b. Se $P(-b < z < b) = 0{,}3182$, determine *b*.
    c. Se $P(z > c) = 0{,}2358$, determine *c*.
    d. Se $P(z > d) = 0{,}7517$, determine *d*.
    e. Se $P(z < e) = 0{,}4090$, determine *e*.
38. Para uma distribuição normal padronizada, determine a percentagem dos dados que estão
    a. a menos de 1 desvio-padrão da média
    b. a menos de 1,96 desvios-padrão da média
    c. entre $\mu - 3\sigma$ e $\mu + 3\sigma$
    d. entre 1 desvio-padrão abaixo da média e 2 desvios-padrão acima da média.
    e. a mais de 2 desvios-padrão de distância da média
39. Na Fórmula 5-1, com $\mu = 0$ e $\sigma = 1$, e aproximando *e* por 2,7 e $\sqrt{2\pi}$ por 2,5, obtemos

$z = \frac{x-\mu}{\sigma} = \frac{68,6-63,6}{2,5} = 2,0$



63,6 | 68,6 = 2,0 = 0,4772
μ

$$y = \frac{2,7^{-x^2/2}}{2,5}$$

Com uma escala de 1 in. = 1 unidade no eixo-x e 1 in. = 1 unidade no eixo-y, faça o gráfico desta equação, após determinar as coordenadas y que correspondem às seguintes coordenadas x: $-4, -3, -2, -1, 0, 1, 2, 3$ e 4. Estime a área (em in.²) delimitada pela curva, pelo eixo-x e pelas retas verticais que passam por 0 e 1 no eixo-x. Compare este resultado com o valor na Tabela A-2.

40. Suponha que a variável aleatória x tenha distribuição de probabilidade contínua *uniforme* com $\mu = 0$ e $\sigma = 1$. Para essa variável aleatória, o mínimo é $-\sqrt{3}$ e o máximo é $\sqrt{3}$.

  a. Determine $P(x > 1)$.
  b. Determine $P(x > 1)$, admitindo incorretamente que a distribuição seja normal, e não uniforme.
  c. Compare o resultado correto da parte (a) com o resultado incorreto da parte (b). A utilização da distribuição errada afeta sensivelmente o resultado?

5-3 Distribuições Normais Não-Padronizadas: Determinação de Probabilidades

Embora a Seção 5-2 tenha introduzido métodos importantes para lidar com distribuições normais, os exemplos e exercícios ali incluídos em geral não são realistas, porque a maioria das populações distribuídas normalmente têm média diferente de zero, desvio-padrão diferente de 1, ou ambos. Nesta seção incluímos muitas distribuições normais não-padronizadas reais e importantes. Para padronizar casos não-padronizados, aplicamos a Fórmula 5-2:

Fórmula 5-2

$$z = \frac{x - \mu}{\sigma}$$

Valor dado — Média — DP

Veja a Figura 5-13, onde ilustramos o importante princípio que a área delimitada por um valor e pela média populacional é a mesma que a área delimitada pelo escore z correspondente e a média 0. Uma vez feita a conversão de um valor não-padronizado em um escore z, podemos utilizar a Tabela A-2 da mesma forma, como na Seção 5-2. Recomendamos, pois, o processo seguinte para achar probabilidades de valores de uma variável aleatória com distribuição de probabilidade normal:

1. Trace uma curva normal, assinale a média e outros valores de interesse, e *sombreie* a região que representa a probabilidade desejada.
2. Para cada valor x fronteira da região sombreada, aplique a Fórmula 5-2 para achar o escore z correspondente.
3. Recorra à Tabela A-2 para achar a área da região sombreada. Essa área é a probabilidade desejada.

No exemplo que segue utilizamos essas três etapas. Note que o exemplo se refere a uma altura de 68,6 in., que é exatamente 2 desvios-padrão acima da média, de forma que o escore z correspondente é 2,00. O escore z dá o número de desvios-padrão que um valor dista da média.

EXEMPLO As alturas das mulheres têm distribuição normal com média de 63,6 in. e desvio-padrão de 2,5 in. (com base em dados do National Health Survey — Serviço Nacional de Saúde dos EUA). Selecionada aleatoriamente uma mulher, determine a probabilidade de a sua altura estar entre 63,6 e 68,6 in.

SOLUÇÃO

Passo 1: Veja a Figura 5-14, onde introduzimos a média 63,6 e o valor 68,6, e sombreamos a área que representa a probabilidade desejada.

Passo 2: Para usar a Tabela A-2, devemos aplicar a Fórmula 5-2 para converter a distribuição não-padronizada de alturas em uma distribuição normal padronizada. A altura de 68,6 in. é convertida em um escore z como segue:

$$z = \frac{x - \mu}{\sigma} = \frac{68,6 - 63,6}{2,5} = \frac{5}{2,5} = 2,00$$

Este resultado mostra que a altura de 68,6 in. difere da média de 63,6 in. por 2,00 desvios-padrão.

Passo 3: Recorrendo à Tabela A-2, vemos que $z = 2,00$ corresponde a uma área de 0,4772.

Há, assim, uma probabilidade de 0,4772 de escolher uma mulher com altura entre 63,6 in. e 68,6 in. Em símbolos, temos:

$$P(63,6 < x < 68,6) = P(0 < z < 2,00) = 0,4772$$

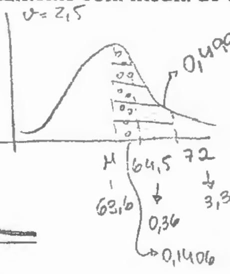
Outra forma de interpretar este resultado consiste em concluir que 47,72% das mulheres têm altura entre 63,6 in. e 68,6 in.

EXEMPLO No início deste capítulo, mencionamos que, para se adaptar a uma espaçonave russa Soyuz, um astronauta deve ter altura entre 64,5 in. e 72 in.

a. Determine a percentagem das mulheres americanas que satisfazem essa condição.
b. Entre 500 mulheres americanas selecionadas aleatoriamente, quantas satisfazem aquela condição?

SOLUÇÃO Tal como no exemplo precedente, estamos em face de alturas distribuídas normalmente com média de 63,6 in. e desvio-padrão de 2,5 in.

σ = 2,5



0,4999 a)

$P = 0,4999 - 0,1406 = 0,3593$

b) 179

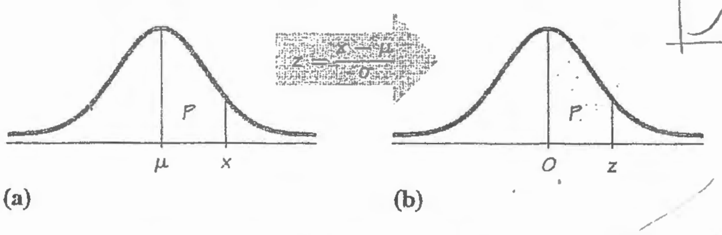


Fig. 5-13 Transformação de uma distribuição normal não-padronizada em uma distribuição normal padronizada.

$z_1 = \frac{x-\mu}{\sigma} = \frac{64,5-63,6}{2,5} = 0,36$

$z_2 = \frac{x-\mu}{\sigma} = \frac{72-63,6}{2,5} = 3,36$

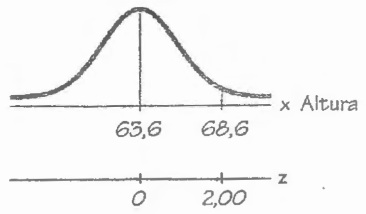


Fig. 5-14 Probabilidade de uma altura entre 63,6 in. e 68,6 in.

a. A região sombreada B na Figura 5-15 representa a proporção de mulheres que satisfazem a restrição de altura para a nave Soyuz, porque suas alturas estão entre 64,5 in. e 72 in. Não podemos achar diretamente aquela região sombreada, porque a Tabela A-2 não foi planejada para tais casos, mas podemos determiná-la indiretamente, aplicando os mesmos processos básicos apresentados na Seção 5-2. Eis como proceder: Determinamos a área sombreada B achando a diferença entre a região A e a área total das regiões A e B combinadas; ou seja,

$$B = (A \text{ e } B \text{ combinadas}) - A.$$

Para as áreas A e B combinadas:

$$z = \frac{x - \mu}{\sigma} = \frac{72 - 63{,}6}{2{,}5} = 3{,}36$$

Consultando a Tabela A-2, vemos que $z = 3{,}36$ está fora da tabela; tomamos, pois, 0,4999. (Veja a nota de rodapé que acompanha a Tabela A-2.)

Para a região A:

$$z = \frac{x - \mu}{\sigma} = \frac{64{,}5 - 63{,}6}{2{,}5} = 0{,}36$$

Consultando novamente a Tabela A-2, vemos que $z = 0{,}36$ corresponde a uma área de 0,1406. A região A tem, pois, área de 0,1406.

A área sombreada é a diferença entre 0,4999 e 0,1406:

$$\begin{aligned}\text{Área } B &= (\text{áreas de } A \text{ e } B \text{ combinadas}) - (\text{área } A)\\ &= 0{,}4999 - 0{,}1406 = 0{,}3593\end{aligned}$$

Se as mulheres americanas forem recrutadas como astronautas sem observância da altura, cerca de 35,93% delas satisfarão as exigências da Soyuz quanto à altura.

b. Entre 500 mulheres selecionadas aleatoriamente, esperamos que 35,93% delas satisfaçam a exigência de altura. O número efetivo é, pois,

$$(500)(0{,}3593) = 179{,}65 \text{ mulheres}$$

> **Filas**
>
> A teoria das filas é um ramo da matemática que se apóia na probabilidade e na estatística. O estudo das filas de espera é importante para negócios como supermercados, bancos, restaurantes *self-service*, companhias de aviação e parques de diversão. Os supermercados Grand Union procuraram manter no máximo três clientes em cada fila de caixa. O Wendy's introduziu o "Express Pak" para acelerar o atendimento de seus inúmeros clientes. A Disney faz longos estudos de filas em seus parques de diversão, a fim de manter os freqüentadores satisfeitos e planejar uma expansão ainda maior. O Bell Laboratories aplica a teoria das filas para otimizar o uso de redes telefônicas, e as fábricas utilizam-na no planejamento de linhas de produção mais eficientes.

EXEMPLO O exército americano exige que a altura das mulheres esteja entre 58 in. e 80 in. Determine a percentagem das mulheres que satisfazem esta exigência. Novamente, suponha que as mulheres tenham distribuição normal com média de 63,6 in. e desvio-padrão de 2,5 in.

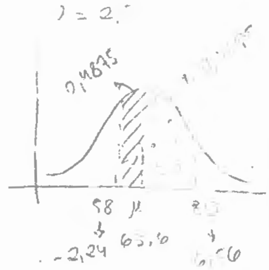
SOLUÇÃO A Figura 5-16 mostra a distribuição normal das alturas das mulheres, com a região sombreada representando alturas entre 58 in. e 80 in., conforme exigido pelo exército dos EUA. O método para achar a área da região sombreada envolve a sua divisão em duas partes A e B, conforme mostrado. Podemos utilizar a Fórmula 5-2 e a Tabela A-2 para achar as áreas dessas regiões separadamente, somando em seguida os resultados.

Para a área A somente:

$$z = \frac{x - \mu}{\sigma} = \frac{58{,}0 - 63{,}6}{2{,}5} = -2{,}24$$

Pela Tabela A-2, vemos que $z = -2{,}24$ corresponde a 0,4875, de forma que a área A é 0,4875.

Para a área B somente:



p = 0,9875

= 0,9874

$z_1 = \frac{x-\mu}{\sigma} = \frac{58-63,6}{2,5} = -2,24$

$z_2 = \frac{x-\mu}{\sigma} = \frac{80-63,6}{2,5} = 6,56$

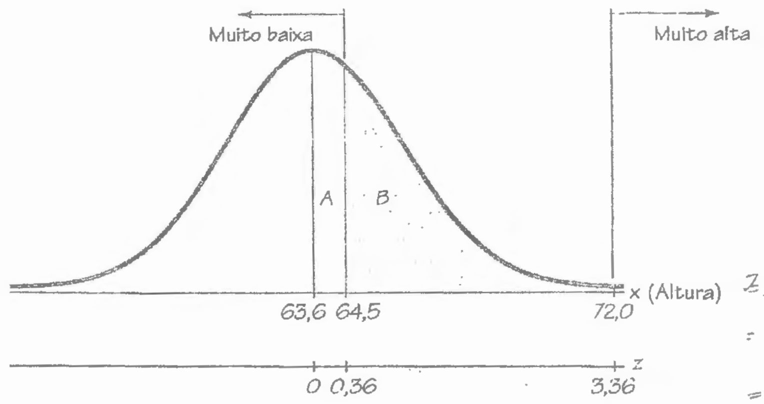


Fig. 5-15 Mulheres que se adaptam à espaçonave Soyuz.

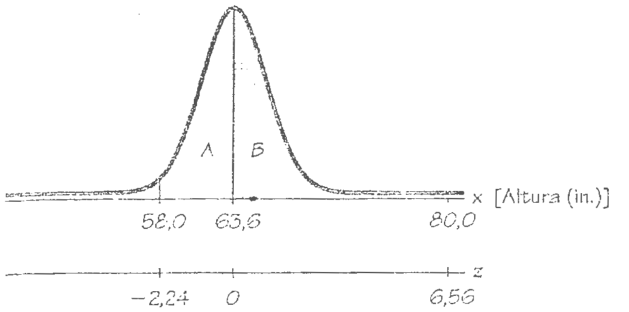


Fig. 5-16 Mulheres que satisfazem os requisitos de altura do exército americano.

$$z = \frac{80,0 - 63,6}{2,5} = 6,56$$

Embora a Tabela A-2 não contenha valores de z acima de 3,09, ela indica que, para tais valores, deveremos tomar 0,4999 como área. (Se necessário, podemos obter resultados mais precisos utilizando tabelas especiais ou programas de computador.) A área B é 0,4999.

A região sombreada consiste nas áreas A e B combinadas; assim,

Área das regiões A e B combinadas $= 0,4875 + 0,4999 = 0,9874$

A proporção de mulheres que satisfazem a exigência do exército americano é, pois, 0,9874. Como o problema pede uma percentagem, expressamos o resultado como 98,74%.

Repetindo com os homens o procedimento precedente, verificamos que 99,98% dos homens satisfazem a exigência quanto à altura. (Os homens têm alturas distribuídas normalmente com média de 69,0 in. e desvio-padrão de 2,8 in., e o exército americano exige alturas entre 60 in. e 80 in.) A percentagem de homens aceitáveis excede em muito pouco a percentagem de mulheres aceitáveis. Ambas as percentagens indicam uma rejeição muito pequena por motivo de altura.

Nesta seção ampliamos os conceitos da Seção 5-2 de forma a incluir distribuições de probabilidade normais não-padronizadas mais realistas. Mas todos os exemplos considerados até aqui são do mesmo tipo geral: acha-se uma probabilidade (ou percentagem) utilizando-se a distribuição normal (com valores relacionados na Tabela A-2), quando são dados os valores da média e do desvio-padrão populacionais, e outros valores de interesse. Em muitos casos práticos e reais, a probabilidade (ou percentagem) é conhecida: devemos achar os valores correspondentes. Na próxima seção abordaremos problemas deste tipo.

5-3 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-6, admita que as alturas das mulheres tenham distribuição normal com média $\mu = 63,6$ in. e desvio-padrão $\sigma = 2,5$ in.(com base em dados do National Health Survey). Admita também que uma mulher seja escolhida aleatoriamente. Trace um gráfico e ache a probabilidade pedida.

1. $P(63,6 \text{ in.} < x < 65,0 \text{ in.})$
2. $P(x < 70,0 \text{ in.})$
3. $P(x > 58,1 \text{ in.})$
4. $P(59,1 \text{ in.} < x < 66,6 \text{ in.})$
5. As alturas das dançarinas em um espetáculo no New York City's Radio City Music Hall devem estar entre 65,5 in. e 68,0 in. Escolhida aleatoriamente uma mulher, determine a probabilidade de ela poder ser uma dançarina nesse espetáculo.
6. O Beanstalk Club, uma organização social para pessoas de porte elevado, tem uma exigência de que as mulheres tenham ao menos 70 in. (ou 5 ft 10 in.) de altura. Cogita-se de abrir uma filial do Beanstalk Club em uma área metropolitana com 500.000 mulheres adultas.
   a. Determine a percentagem de mulheres adultas elegíveis para membro, por terem a altura mínima de 70 in.
   b. Entre as 500.000 mulheres adultas que vivem na área metropolitana, quantas podem ser candidatas ao Beanstalk Club?
   c. O leitor abriria uma filial do Beanstalk Club?
7. Os prazos de substituição de aparelhos de TV têm distribuição normal com média de 8,2 anos e desvio-padrão de 1,1 ano (com base em dados do "Getting Things Fixed", *Consumer Reports*). Determine a probabilidade de um aparelho de TV selecionado aleatoriamente acusar um tempo de substituição inferior a 7,0 anos.
8. Os prazos de substituição para CD players têm distribuição normal com média de 7,1 anos e desvio-padrão de 1,4 ano (com base em dados do "Getting Things Fixed", *Consumer Reports*). Determine a probabilidade de um CD player escolhido aleatoriamente ter um prazo de substituição inferior a 8,0 anos.
9. Supondo que os pesos do papel descartado semanalmente pelas residências tenham distribuição normal com média de 9,4 lb e desvio-padrão de 4,2 lb (com base em dados do Garbage Project da Universidade do Arizona), determine a probabilidade de escolher aleatoriamente uma residência que descarte entre 5,0 lb e 8,0 lb de papel em uma semana.
10. Com base nos resultados amostrais do Conjunto de Dados 2 do Apêndice B, suponha que as temperaturas do corpo humano tenham distribuição normal com média de 98,20°F e desvio-padrão de 0,62° F. Definindo como febre uma temperatura acima de 100° F, que percentagem de pessoas normais e sadias pode ser considerada como tendo febre? Essa percentagem sugere que o limite de 100° F é apropriado?
11. Uma aplicação clássica da distribuição normal é inspirada em uma carta a *Dear Abby*, em que uma esposa alegava ter dado à luz 308 dias após uma rápida visita de seu marido que estava servindo na Marinha. Os prazos da gravidez têm distribuição normal com média de 268 dias e desvio-padrão de 15 dias. Com base nessa informação, determine a probabilidade de uma gravidez durar 308 dias ou mais. Que é que o resultado sugere?
12. Os prazos de duração da gravidez têm distribuição normal com média de 268 dias e desvio-padrão de 15 dias. Definindo como *prematura* uma criança nascida com ao menos três semanas de antecipação, qual a percentagem das crianças nascidas prematuramente? (Essa informação é importante para os administradores de hospitais, que devem providenciar para ter à mão o equipamento necessário para atender às necessidades especiais dos prematuros.)
13. De acordo com a Opinion Research Corporation, os homens gastam em média 11,4 minutos no chuveiro. Suponha que esses tempos tenham distribuição normal com desvio-padrão de 1,8 min. Escolhido um homem aleatoriamente, determine a probabilidade de ele gastar ao menos 10,0 min no chuveiro.
14. De acordo com a International Mass Retail Association, as jovens com idade entre 13 e 17 anos gastam em média $31,20 em compras cada mês. Suponha que as importâncias desses gastos tenham distribuição normal com desvio-padrão de $8,27. Selecionada aleatoriamente uma jovem naquela faixa etária, qual é a probabilidade de ela gastar entre $35,00 e $40,00 em um mês?
15. Os escores de QI têm distribuição normal com média 100 e desvio-padrão 15. A Mensa é uma organização para pessoas com QI elevado, e a admissão exige um QI superior a 131,5.

a. Escolhida aleatoriamente uma pessoa, determine a probabilidade de ela satisfazer aquela exigência da Mensa.
b. Em uma região típica de 75.000 habitantes, quantos serão candidatos à Mensa?

16. Um subfornecedor da IBM foi contratado para fabricar substratos de cerâmica, utilizados para transmitir sinais entre chips de silício para computador. As especificações exigem uma resistência entre 1,500 ohm e 2,500 ohms, mas a população tem resistências distribuídas normalmente com média de 1,978 ohm e desvio-padrão de 0,172 ohm. Que percentagem dos substratos de cerâmica foge às especificações do fabricante? Esse processo de fabricação parece estar funcionando bem?

17. Os níveis de colesterol sérico em homens entre 18 e 24 anos de idade têm distribuição normal com média de 178,1 e desvio-padrão de 40,7. Todas as unidades são em mg/100 mL, e os dados se baseiam no National Health Survey. Escolhido aleatoriamente um homem entre 18 e 24 anos de idade, determine a probabilidade de seu nível de colesterol sérico estar entre 200 e 250.

18. Analisam-se medidas de crânios humanos de diferentes épocas, para determinar se variam com o tempo. Mede-se a largura máxima de crânios de homens egípcios que viveram por volta de 3300 a.C. Os resultados mostram que essas larguras têm distribuição normal com média de 132,6 mm e desvio-padrão de 5,4 mm (com base em dados do *Ancient Races of the Thebaid*, por Thomson e Randall-Maciver). Um arqueólogo descobre o crânio de um homem egípcio e a medida revela uma largura máxima de 119 mm. Determine a probabilidade de obter o valor de 119 mm ou menos para um crânio, selecionado aleatoriamente, do período de 3300 a.C. É provável que o crânio recentemente encontrado seja daquela época?

19. O Corpo de Fuzileiros Navais da Marinha dos EUA exige homens com altura entre 64 in. e 78 in. Determine a percentagem dos homens que satisfazem essa exigência. (O National Health Survey mostra que as alturas dos homens têm distribuição normal com média de 69,0 in. e desvio-padrão de 2,8 in.)

20. As máquinas "caça-níqueis" são fabricadas de modo que seus proprietários possam ajustar os pesos das moedas que são aceitas. Se são encontradas muitas moedas falsificadas, faz-se um ajuste para rejeitar mais moedas, com o efeito de que a maioria das moedas falsificadas é rejeitada juntamente com muitas moedas legítimas. Suponha que as moedas tenham pesos distribuídos normalmente com média de 5,67 g e desvio-padrão de 0,070 g. Se uma máquina "caça-níqueis" é ajustada para rejeitar moedas que pesem menos de 5,50 g ou mais de 5,80 g, qual é a percentagem de moedas legítimas rejeitadas?

5-3 Exercícios B: Além do Básico

Nos Exercícios 21-24, utilize o conjunto de dados do Apêndice B indicado.

a. Construa um histograma para determinar se o conjunto de dados tem distribuição normal.
b. Ache a média amostral $\bar{x}$ e o desvio-padrão amostral s.
c. Utilize a média amostral como estimativa da média populacional μ, o desvio amostral como estimativa do desvio-padrão populacional σ, e aplique os métodos desta seção para achar a probabilidade indicada.

21. Utilize a lista combinada de 100 pesos de bombons M&M do Conjunto de Dados 11 e estime a probabilidade de selecionar aleatoriamente um M&M que pese mais de 1,000 g.

22. Com os comprimentos de cabeças de ursos do Conjunto de Dados 3, estime a probabilidade de selecionar aleatoriamente um urso cuja cabeça tenha entre 12,0 in. e 13,0 in.

23. Com os dados de precipitação pluviométrica de Iowa (Conjunto de Dados 7), estime a probabilidade de selecionar aleatoriamente um ano com precipitação inferior a 40,0 in.

24. Com os pesos totais do lixo descartado (Conjunto de Dados 1), estime a probabilidade de selecionar aleatoriamente uma casa que descarte mais de 20,0 lb de lixo em uma semana.

5-4 Distribuições Normais Não-Padronizadas: Cálculo de Valores

Nesta seção abordamos problemas como este: Se as alturas das mulheres têm distribuição normal com $\mu = 63{,}6$ in. e $\sigma = 2{,}5$ in., determinar a altura que separa os 10% superiores. Neste problema é dada uma probabilidade (0,10) e devemos determinar a altura adequada x. Nos exemplos e exercícios da Seção 5-3, utilizamos um valor para achar uma probabilidade. Nesta seção seguiremos o caminho inverso — dada uma probabilidade, determinar o valor correspondente.

Ao considerarmos os problemas de determinação de valores quando são dadas as probabilidades, há três pontos importantes a respeito dos quais devemos acautelar-nos:

1. *Não confundir escores z com áreas.* Recorde que os escores z são *distâncias* ao longo da escala horizontal, mas as áreas representam *regiões* sob a curva normal. A Tabela A-2 relaciona os escores z na coluna à esquerda e através do topo, mas as áreas se encontram no corpo da tabela.
2. *Escolher o lado correto (direito/esquerdo) do gráfico.* Localiza-se à direita do gráfico um valor (escore) que separa os 10% superiores dos demais valores, mas um valor que separa os 10% inferiores está localizado à esquerda do gráfico.
3. *Um valor, ou escore, z deve ser negativo sempre que estiver à esquerda da linha central do 0.*

Tal como na Seção 5-3, os gráficos aqui são também de grande auxílio; recomendamo-los com empenho. Os gráficos estão incluídos como primeiro passo do processo seguinte:

1. Partindo de um esboço que tenha ao menos alguma semelhança com um sino, introduza a probabilidade (ou percentagem) dada na região apropriada do gráfico e identifique o(s) valor(es) x procurados.
2. Utilize a Tabela A-2 para achar o escore z correspondente à região delimitada por x e pela reta central do 0. Tenha em mente os seguintes detalhes:
   - Recorrer ao *corpo* da Tabela A-2 para achar a área mais próxima e, em seguida, identificar o valor z correspondente.
   - Considerar o valor z *negativo* se estiver à esquerda da reta central.
3. Utilizando a Fórmula 5-2, introduzir os valores de μ, σ e o valor de z encontrados no Passo 2; resolver então em relação a x. Com base na Fórmula 5-2, podemos resolver em relação a x como segue: $z = \frac{x-\mu}{\sigma}$

$$x = \mu + (z)\,(\sigma) \quad \text{(Outra forma da Fórmula 5-2)}$$

$\frac{x-\mu}{\sigma} = z$

$x - \mu = (z) \cdot \sigma$

4. Recorrer ao esboço da curva para verificar que a solução tem sentido no contexto do gráfico e no contexto do problema.

EXEMPLO Na Seção 5-3 vimos que as alturas das mulheres têm distribuição normal com média 63,6 in. e desvio-padrão 2,5 in. (com base em dados do Serviço Nacional de Saúde dos EUA). Determine o valor de P_{90}, isto é, determine a altura que separa os 90% inferiores dos 10% superiores.

SOLUÇÃO

Passo 1: Iniciamos com o gráfico da Figura 5-17. Introduzimos a média de 63,6, sombreamos a área que representa os 10% superiores e identificamos como x o valor desejado. A área entre 63,6 in. e x in. deve ser 40% da área total (porque a metade direita da área deve perfazer 50% do total). Como a área total é 1, a área que constitui 40% do total deve ser 0,4.

Passo 2: Recorremos à Tabela A-2, mas procuramos uma área de 0,4000 no corpo da tabela. (Lembre-se de que a Tabela A-2 relaciona áreas somente para as regiões delimitadas à esquerda pela média e à direita por algum valor, de modo que tomamos a área $0{,}90 - 0{,}50 = 0{,}40$, conforme mostra a Figura 5.17.) A área mais próxima de 0,4000 é 0,3997, que corresponde ao escore z de 1,28.

Passo 3: Com $z = 1{,}28$, $\mu = 63{,}6$ e $\sigma = 2{,}5$, resolvemos em relação a x, seja aplicando diretamente a Fórmula 5-2 ou a versão seguinte da mesma fórmula:

$$x = \mu + (z)(\sigma) = 63{,}6 + (1{,}28)(2{,}5) = 66{,}8$$

Passo 4: Fazendo $x = 66{,}8$ na Figura 5-17, vemos que esta solução é razoável, porque deve ser maior do que a média de 63,6.

A altura de 66,8 in. separa os 90% correspondentes às mulheres mais baixas dos 10% correspondentes às mais altas.

A solução precedente pode ser obtida com auxílio de uma calculadora TI-83. Acione 2nd/DISTR e utilize a função `invNorm` (normal inversa) com a probabilidade de 0,9, a média 63,6, e o desvio-padrão 2,5 introduzidos no seguinte formato: `invNorm (0.9, 63.6, 2.5)` para obter o resultado 66,8.

O Meio de Pesquisa Pode Afetar o Resultado

Em uma pesquisa entre católicos de Boston, EUA, as pessoas foram consultadas sobre se os anticoncepcionais deveriam ser vendidos a mulheres não-casadas. Em entrevistas pessoais, 44% dos entrevistados responderam *sim*. Mas em um grupo análogo contatado por correio ou telefone, 75% responderam *sim* à mesma pergunta.

O próximo exemplo inclui uma cilada que não está presente no exemplo precedente. Acompanhe atentamente o que acontece quando o escore desejado está *abaixo* da média, como neste exemplo.

EXEMPLO Em um estudo comparou-se o comportamento facial de esquizofrênicos não-paranóicos com o de um grupo de controle de pessoas normais. O grupo de controle foi cronometrado em relação ao piscar de olhos durante um período de 5 minutos, ou 300 segundos. Os tempos acusaram distribuição normal com média de 184 s e desvio-padrão de 55 s com base em dados de "Ethological Study of Facial Behavior in Nonparanoid and Paranoid Schizophrenic Patients" — Estudo Etológico de Comportamento Facial de Pacientes Esquizofrênicos Paranóicos e Não-paranóicos, por Pitman, Kolb, Orr e Singh, *Psychiatry*, 144:1. Como os resultados mostraram que os pacientes esquizofrênicos não-paranóicos acusavam tempos muito inferiores aos do grupo de controle, decidiu-se prosseguir a análise das pessoas do grupo de controle que estavam nos 5% inferiores. Para o grupo de controle, determine P_5, o 5.° percentil; ou seja, determine o tempo que separa os 5% inferiores dos restantes.

SOLUÇÃO

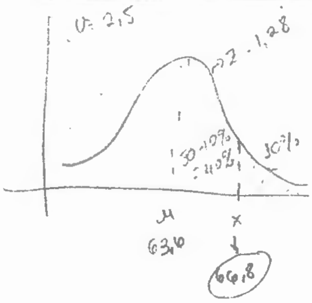
Passo 1: O gráfico aparece na Figura 5-18. Introduziu-se a média de 184, sombreou-se a região que representa os 5% inferiores e identificou-se o valor de x desejado.

Passo 2: A região delimitada por x e pela reta central é identificada como região A, e deve ter uma área de 0,45. (A área da cauda esquerda de 0,05 e a região A, devem, juntas, perfazer uma área de 0,50.) Recorrendo à Tabela A-2 procuramos no *corpo* da tabela uma área de 0,45, à qual corresponde o valor de 1,645. *Como o escore z é negativo sempre que estiver abaixo (à esquerda) da média, tomamos* $z = -1{,}645$.

Passo 3: Fazemos $z = -1{,}645$, $\mu = 184$ e $\sigma = 55$, e resolvemos em relação a x como segue:

$$x = \mu + (z)(\sigma) = 184 + (-1{,}645)(55) = 93{,}5.$$

Passo 4: A solução $x = 93{,}5$ s se afigura razoável na Figura 5-18, porque o valor de x deve estar abaixo da média de 184 s. (Se não tivéssemos tomado o valor de z negativo, teríamos obtido o resultado estranho



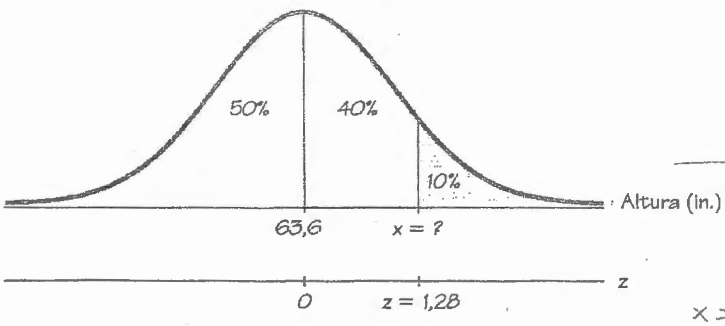
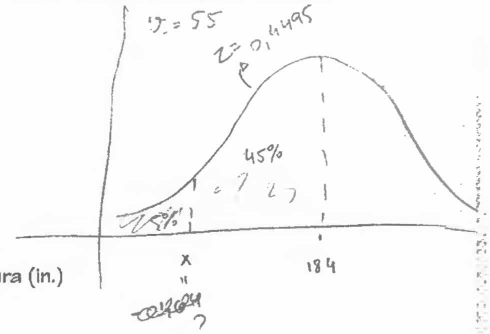


Fig. 5-17 Determinação de P_{90} para as alturas das mulheres.



$x = \mu + (z)(\sigma)$

$= 63{,}6 + (1{,}28)\cdot(2{,}5)$

$= 66{,}8$

$x = \mu + (z)\cdot(\sigma)$

$x = 184 + (-1{,}64)\cdot(55)$

$= 184 + (-90{,}2)$

$= 184 - 90{,}2$

$= 93{,}8$

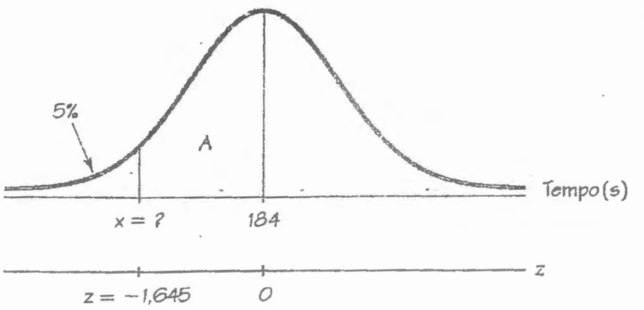


Fig. 5-18 Determinação do 5º percentil para os tempos de piscada de olhos.

de 274,5 s. Pela Figura 5-18, é óbvio que x não pode ser 274,5 s.)

Este resultado indica que $P_5 = 93{,}5$ s, isto é, 5% das vezes temos um valor inferior a 93,5 s. Querendo aprofundar a análise das pessoas nos 5% inferiores do grupo de controle, devemos escolher pessoas com tempos de piscadela inferiores a 93,5 s.

5-4 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-4, suponha que as mulheres tenham alturas distribuídas normalmente com média de 63,6 in. e desvio-padrão de 2,5 in. (com base em dados do National Health Survey dos EUA). Determine a altura correspondente ao percentil indicado.

1. P_{85} 2. P_{66} 3. P_{15} 4. P_{35}

5. Os tempos de substituição para aparelhos de TV têm distribuição normal com média de 8,2 anos e desvio-padrão de 1,1 ano (com base em dados do "Getting Things Fixed", *Consumer Reports*). Determine os tempos de substituição que separam os 20% superiores dos 80% inferiores. Este resultado tem utilidade para uma firma de assistência que deseje oferecer contratos de serviço de reparos para aparelhos de TV.
6. Os tempos de substituição para CD players têm distribuição normal com média de 7,1 anos e desvio-padrão de 1,4 ano (com base em dados do "Getting Thing Fixed", *Consumer Reports*). Determine o tempo de substituição que separa os 45% superiores dos 55% inferiores.
7. Os pesos do papel descartado semanalmente em residências têm distribuição normal com média de 9,4 lb e desvio-padrão de 4,2 lb (com base em dados do Garbage Project da Universidade do Arizona). Determine o peso que separa os 33% inferiores dos 67% superiores.
8. Com base nos resultados amostrais do Conjunto de Dados 2 do Apêndice B, suponha que as temperaturas do corpo humano tenham distribuição normal com média de 98,20°F e desvio-padrão de 0,62°F. Quais os dois níveis de temperatura que separam os 2% inferiores e os 2% superiores? Esses valores podem servir como limites razoáveis a serem usados para identificar pessoas que podem estar doentes.
9. Os prazos de duração da gravidez têm distribuição normal com média de 268 dias e desvio-padrão de 15 dias. Se definirmos como *prematura* uma criança cujo período de gestação esteja nos 4% inferiores, determine o prazo de gestação que separa as crianças prematuras das não-prematuras. Os bebês prematuros quase sempre exigem cuidados especiais, e este resultado pode ajudar os administradores de hospital a fazer o planejamento necessário.
10. De acordo com a Opinion Research Corporation (Companhia de Pesquisa de Opinião), os homens gastam em média 11,4 minutos no chuveiro. Suponha que os tempos se distribuam normalmente com desvio-padrão de 1,8 min. Determine os valores de Q_1 e Q_3.
11. Os QIs têm distribuição normal com média 100 e desvio-padrão 15. Definindo como gênio uma pessoa no 1% superior dos valores de QI, determine o valor que separa os gênios das pessoas comuns. Esse valor pode ser usado por uma companhia interessada em contratar "gênios".
12. Um subfornecedor fabricava substratos de cerâmica para a IBM. Esses dispositivos têm resistências distribuídas normalmente com média de 1,978 ohm e desvio-padrão de 0,172 ohm. Se as especificações exigidas devem ser modificadas de modo que 3% dos dispositivos sejam rejeitados por terem resistência muito baixa e 3% sejam rejeitados por terem resistência muito alta, determine os valores de separação para os dispositivos aceitáveis.
13. Os níveis de colesterol sérico nos homens com 18 a 24 anos de idade têm distribuição normal com média 178,1 e desvio-padrão 40,7. Todas as unidades são em mg/100 mL e os dados se baseiam no National Health Survey dos EUA.
    a. Escolhido aleatoriamente um homem entre 18 e 24 anos, determine a probabilidade de seu nível de colesterol sérico ser inferior a 200.
    b. Se um nível de colesterol sérico deve ser julgado muito alto se estiver nos 7% superiores, determine o nível de separação dos níveis demasiadamente altos.
14. Analisam-se as medidas de crânios humanos de diferentes épocas, para verificar se variam com o tempo. Mede-se a largura máxima de crânios de homens egípcios que viveram por volta de 3300 a.C. Os resultados mostram que essas larguras acusam distribuição normal com média de 132,6 mm e desvio-padrão de 5,4 mm (com base em dados do *Ancient Races of the Thebaid*, por Thomson e Randall-Maciver).
    a. Determine a probabilidade de obter um valor superior a 140 mm para um crânio daquela época selecionado aleatoriamente.
    b. Determine o valor de D_2, o segundo decil.
15. Para ingressar na Marinha dos EUA, uma mulher deve ter altura entre 58 in. e 73 in. Recorde que as alturas das mulheres têm distribuição normal com média de 63,6 in. e desvio-padrão de 2,5 in.
    a. Determine a percentagem das mulheres que satisfazem aquela exigência.
    b. Se mudarem as condições de admissão, de forma a excluir o 1% das mais baixas e o 1% das mais altas, determine as alturas aceitáveis.
16. As moedas de 25 centavos (de dólar) têm pesos distribuídos normalmente com média de 5,67 g e desvio-padrão de 0,070 g.
    a. Se uma máquina automática de refrescos é ajustada de modo a rejeitar moedas de 25 centavos com peso inferior a 5,53 g ou superior a 5,81 g, qual a percentagem de moedas legais rejeitadas?
    b. Determine os pesos de moedas legais aceitas se a máquina é reajustada de forma a rejeitar 1,5% das mais leves e 1,5% das mais pesadas.

5-4 Exercícios B: Além do Básico

17. No estudo de um conjunto de dados, a construção de um histograma revela que a distribuição é aproximadamente normal; constrói-se um *boxplot* com os seguintes valores de quartis: $Q_1 = 62$, $Q_2 = 70$, $Q_3 = 78$. Calcule o desvio-padrão.
18. Um professor dá um teste e obtém resultados distribuídos normalmente com média 50 e desvio-padrão 10. Se as notas são atribuídas segundo o esquema a seguir, determine os limites numéricos para cada conceito:

A: 10% superiores
B: Notas acima dos 70% inferiores e abaixo dos 10% superiores
C: Notas acima dos 30% inferiores e abaixo dos 30% superiores
D: Notas acima dos 10% inferiores e abaixo dos 70% superiores
F: 10% inferiores

19. De acordo com os dados da College Entrance Examination Board (Comissão de Exame Vestibular), a nota média do SAT de matemática é 475 e 17,0% das notas estão acima de 600. Determine o desvio-padrão e use o resultado para achar o 99.° percentil. (Admita que as notas sejam distribuídas normalmente.)
20. A Comissão de Exame Vestibular escreve que "para os Testes SAT, em dois terços das vezes, sua nota deve estar em um intervalo de 30 pontos acima ou abaixo de sua capacidade efetiva. Este intervalo é chamado erro-padrão da mensuração (SEM = standard error of measurement)." Use esta afirmação para estimar o desvio-padrão das notas de um indivíduo em um tal teste. (Admita que as notas tenham distribuição normal.)

5-5 O Teorema Central do Limite

Esta seção introduz o teorema central do limite, que é um dos conceitos mais importantes e mais úteis em estatística. Constitui o fundamento para a estimativa de parâmetros populacionais e para o teste de hipóteses — tópicos que serão amplamente estudados em capítulos subseqüentes. Antes de considerar esse teorema, procuraremos primeiro desenvolver uma compreensão intuitiva de uma de suas mais importantes conseqüências.

Na medida em que o tamanho da amostra aumenta, a distribuição amostral das médias amostrais tende para uma distribuição normal.

Ilustraremos o teorema central do limite utilizando os quatro últimos algarismos do Social Security Number — SSN (Número do Seguro Social) de cada um de 50 estudantes (veja a Tabela 5-1). Esses quatro números são aleatórios, ao contrário dos algarismos iniciais, que servem para identificar o estado e para dar outras informações. Combinando os quatro algarismos de cada estudante em uma grande coleção de 200 números, obtemos uma média $\bar{x}$ = 4,5 e um desvio-padrão s = 2,8, e uma distribuição aproximadamente uniforme, conforme gráfico da Figura 5-19. Vejamos agora o que acontece quando achamos as 50 médias amostrais, conforme Tabela 5-1. Embora a coleção original de dados tenha uma distribuição aproximadamente *uniforme*, as médias amostrais têm distribuição aproximadamente *normal*. Este conceito pode parecer um tanto confuso; aconselhamos o leitor a parar por um momento e estudar este parágrafo, até que fique claro seu ponto principal: O conjunto original de 200 números tem distribuição uniforme (porque os algarismos 0-9 ocorrem aproximadamente com a mesma freqüência), mas as 50 médias amostrais têm distribuição normal. É um fenômeno verdadeiramente fascinante e intrigante na estatística que, extraindo amostras de qualquer distribuição, possamos criar uma distribuição normal ou, ao menos, aproximadamente normal.

Em geral, a **distribuição amostral das médias amostrais** é a distribuição das médias amostrais quando extraímos repe-

TABELA 5-1

Algarismos do SSN				$\bar{x}$
1	8	6	4	4,75
5	3	3	6	4,25
9	8	8	8	8,25
5	1	2	5	3,25
9	3	3	5	5,00
4	2	6	2	3,50
7	7	1	6	5,25
9	1	5	4	4,75
5	3	3	9	5,00
7	8	4	1	5,00
0	5	6	1	3,00
9	8	2	2	5,25
6	1	5	7	4,75
8	1	3	0	3,00
5	9	6	9	7,25
6	2	3	4	3,75
7	4	0	7	4,50
5	7	5	6	5,75
4	1	5	7	4,25
1	2	0	6	2,25
4	0	2	8	3,50
3	1	2	5	2,75
0	3	4	0	1,75
1	5	1	0	1,75
9	7	4	0	5,00
7	3	1	1	3,00
9	1	1	3	3,50
8	6	5	9	7,00
5	6	4	1	4,00
9	3	9	5	6,50
6	0	7	3	4,00
8	2	9	6	6,25
0	2	8	6	4,00
2	0	9	7	4,50
5	8	9	0	5,50
6	5	4	9	6,00
4	8	7	6	6,25
7	1	2	0	2,50
2	9	5	0	4,00
8	3	2	2	3,75
2	7	1	6	4,00
6	7	7	1	5,25
2	3	3	9	4,25
2	4	7	5	4,50
5	4	3	7	4,75
0	4	3	8	3,75
2	5	8	6	5,25
7	1	3	4	3,75
8	3	7	0	4,50
5	6	6	7	6,00

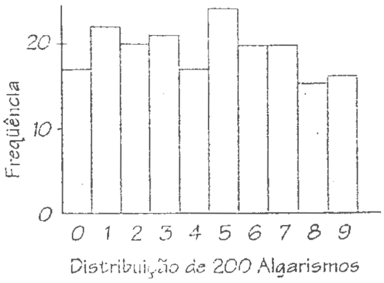


Fig. 5-19 Distribuição de 200 algarismos dos números de inscrição no Seguro Social (4 últimos algarismos) de 50 estudantes.

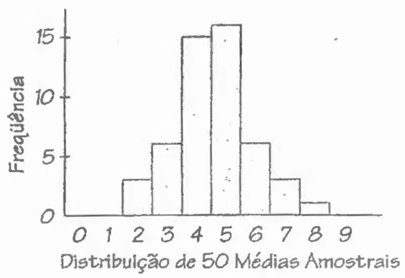


Fig. 5-20 Distribuição de 50 médias amostrais para 50 estudantes.

tidas amostras de mesmo tamanho, da mesma população. Em outras palavras, se extraímos amostras de mesmo tamanho da mesma população, calculamos suas médias e construímos um histograma dessas médias, esse histograma tende para a forma de um sino de uma distribuição normal. Isto é verdade independentemente da forma da distribuição da população original. As Figuras 5-19 e 5-20 se referem a algarismos específicos do registro do seguro social de 50 estudantes, mas a Figura 5-21 é uma ilustração geral que inclui também distribuições normais e assimétricas. Observações exatamente como estas é que levaram à formulação do teorema central do limite, que passamos a estudar.

Suponhamos que a variável x represente notas que podem ter, ou não, distribuição normal, e que a média dos valores x seja μ e o desvio-padrão seja σ. Suponha que coletemos amostras de tamanho n e calculemos as médias amostrais. O que sabemos sobre a coleção de todas as médias amostrais que obtemos repetindo esse experimento? O teorema central do limite nos diz que, na medida em que o tamanho n da amostra aumenta, a distribuição amostral das médias amostrais *tende* para uma distribuição normal com média μ e desvio-padrão $\sigma/\sqrt{n}$. A distribuição das médias amostrais tende para uma distribuição normal no sentido de que, quando n aumenta, a distribuição das médias amostrais se aproxima de uma distribuição normal. Esta conclusão não é óbvia intuitivamente; foi obtida após extensa pesquisa e análise.

Embora a demonstração formal rigorosa exija recursos da matemática avançada, ultrapassando o âmbito deste livro, podemos encontrar uma certa justificativa com base nos dados da Tabela 5-1. Se selecionamos aleatoriamente amostras de algarismos de uma população distribuída uniformemente com média $\mu = 4{,}5$, as médias amostrais resultantes também tenderão a centrar-se em torno de 4,5, de modo que as médias amostrais também têm média 4,5; a média das 50 médias amostrais da Tabela 5-1 é, de fato, 4,5. Inspecionando visualmente os 200 algarismos originais na Tabela 5-1, vemos que eles variam de 0 a 9, mas as 50 médias amostrais acusam *menor variação*, indo de 1,75 a 8,25. O conjunto original de 200 algarismos tem um desvio-padrão de 2,8, mas as 50 médias amostrais têm um desvio-padrão de 1,4, que é menor, conforme esperado. Enunciaremos a seguir o teorema e daremos exemplos de sua aplicação.

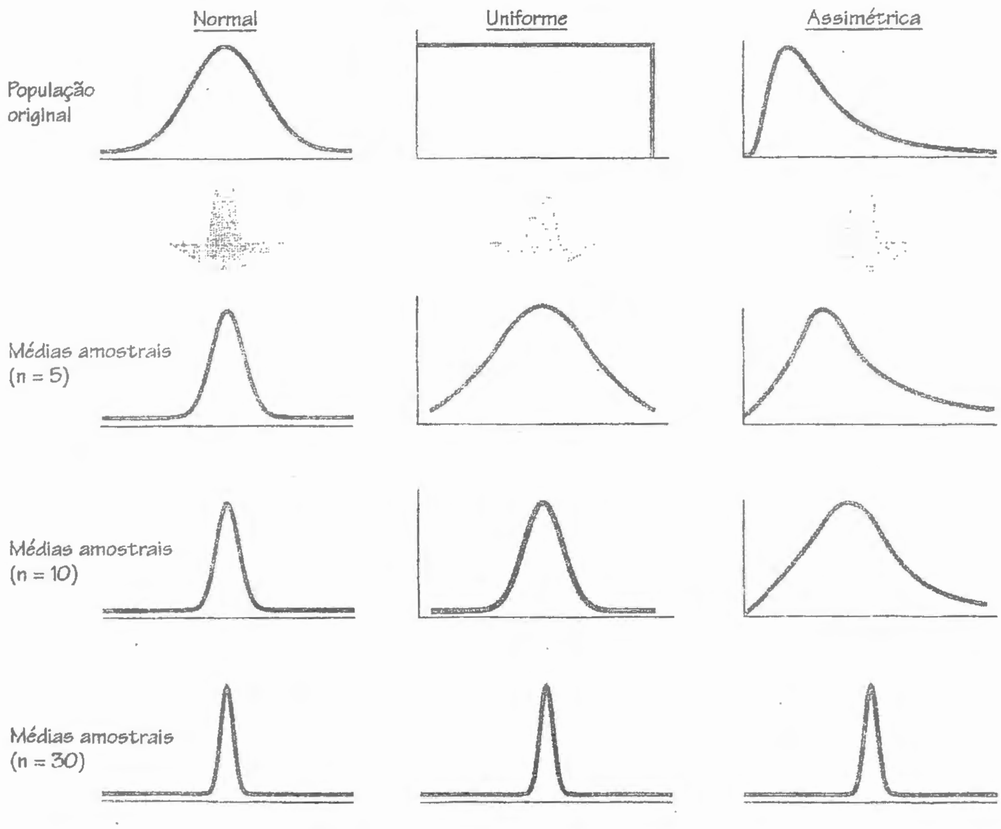


Fig. 5-21 Distribuição normal, uniforme e assimétrica.

Teorema Central do Limite

Dado:

1. A variável aleatória x tem distribuição (que pode ser normal, ou não), com média μ e desvio-padrão σ.
2. Amostras de tamanho n são extraídas aleatoriamente dessa população.

Conclusões:

1. Na medida em que o tamanho da amostra aumenta, a distribuição das médias amostrais $\bar{x}$ tende para uma distribuição normal.
2. A média das médias amostrais será a média populacional μ.
3. O desvio-padrão das médias amostrais será $\sigma/\sqrt{n}$.

Regras Práticas de Uso Comum:

1. Para amostras de tamanho $n > 30$, a distribuição das médias amostrais pode ser aproximada satisfatoriamente por uma distribuição normal. A aproximação melhora na medida em que aumenta o tamanho da amostra n.
2. Se a própria distribuição original tem distribuição normal, então as médias amostrais terão distribuição normal para *qualquer* tamanho amostral n.

O teorema central do limite envolve duas distribuições diferentes: a distribuição da população original e a distribuição das médias amostrais. Tal como em capítulos anteriores, utilizamos os símbolos μ e σ para denotar a média e o desvio-padrão da população original. Vamos agora introduzir uma nova notação para a média e o desvio-padrão da distribuição de médias amostrais.

Muitos problemas importantes e de ordem prática podem ser resolvidos com o teorema central do limite. No exemplo que segue, a parte (a) envolve um valor individual e, assim, utilizaremos os métodos apresentados na Seção 5-3; tais métodos aplicam-se à distribuição normal da variável aleatória x. A parte (b), entretanto, envolve a média de um grupo de 36 homens, e assim

Notação para o Teorema Central do Limite

Se extrairmos todas as amostras aleatórias possíveis, de tamanho n, de uma população com média μ e desvio-padrão σ, a média das médias amostrais se denota por $\mu_{\bar{x}}$; assim,

$$\mu_{\bar{x}} = \mu$$

Por sua vez, o desvio-padrão das médias amostrais se denota por $\sigma_{\bar{x}}$; então,

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

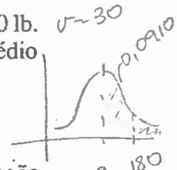
$\sigma_{\bar{x}}$ é comumente chamado **de erro-padrão da média.**

devemos aplicar o teorema central do limite ao lidarmos com a variável aleatória $\bar{x}$.

EXEMPLO Na engenharia humana e no projeto de produtos, freqüentemente é importante considerar os pesos das pessoas, de modo que não haja sobrecarga em aviões ou elevadores, as cadeiras não se quebrem, e não ocorram outros acontecimentos perigosos ou embaraçosos. Dado que a população de homens tem pesos distribuídos normalmente com média de 173 lb e desvio-padrão de 30 lb (com base em dados do National Health Survey dos EUA), determine a probabilidade de que

a. um homem escolhido aleatoriamente pese mais de 180 lb.
b. em 36 homens escolhidos aleatoriamente, o peso médio seja superior a 180 lb.

$\sigma \sim 30$



$0,23$

$P = 0,50 - 0,0910$
$= 0,409$

SOLUÇÃO

a. *Abordagem: Aplique os métodos apresentados na Seção 5-3* (porque estamos em face de um valor *individual* proveniente de uma população com distribuição normal). Procuramos a área da região sombreada na Figura 5-22(a).

$$z = \frac{x - \mu}{\sigma} = \frac{180 - 173}{30} = 0,23$$

Recorremos então à Tabela A-2, por onde vemos que a área daquela região A é 0,0910. A área da região sombreada é, pois, $0,5 - 0,0910 = 0,4090$. A probabilidade de o homem pesar mais de 180 lb é, portanto, 0,4090.

b. *Abordagem: Utilize o teorema central do limite* (porque estamos lidando agora com a *média para um grupo* de 36 valores, e não um valor individual). Como temos agora uma distribuição de médias amostrais, devemos usar os parâmetros $\mu_{\bar{x}}$ e $\sigma_{\bar{x}}$, que são calculados como segue:

$$\mu_{\bar{x}} = \mu = 173$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{30}{\sqrt{36}} = 5$$

Devemos determinar a área sombreada na Figura 5-22(b); o escore z de interesse se calcula como segue:

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{180 - 173}{\frac{30}{\sqrt{36}}} = \frac{7}{5} = 1,40$$

Na Tabela A-2, vemos que $z = 1,40$ corresponde a uma área de 0,4192, de forma que a região sombreada é $0,5 - 0,4192 = 0,0808$, que é a probabilidade de os 36 homens terem peso médio superior a 180 lbs.

Há uma probabilidade de 0,4090 de um homem pesar mais de 180 lb, mas a probabilidade de 36 homens terem peso médio superior a 180 lb é de apenas 0,0808. É muito mais fácil um único indivíduo se afastar da média, do que um grupo de 36 indivíduos. Um peso extremo entre os 36 perderá seu impacto quando considerado em conjunto com os outros 35 pesos.

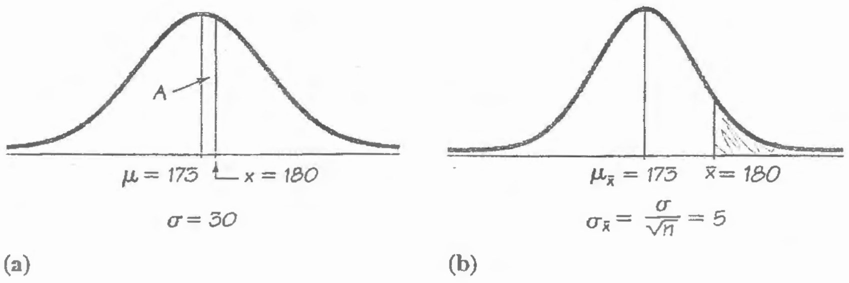


Fig. 5-22 Distribuição de (a) pesos de homens, (b) médias amostrais de pesos de 36 homens.

Um Profissional Fala sobre Erro de Amostragem

Em um artigo para a revista *Time*, Daniel Yankelovich comentou o erro de amostragem quase sempre reportado conjuntamente com resultados de pesquisas. Afirmou que o erro de amostragem se refere somente à imprecisão criada com a utilização de dados amostrais aleatórios para fazer inferências sobre uma população; o erro de amostragem não se refere a questões mal formuladas, tendenciosas ou emocionais. Ele disse: "O mais importante de tudo é que rótulos de advertência sobre erros de amostragem nada nos dizem sobre se o público está envolvido emocionalmente com a questão ou se deu suficiente atenção ao assunto. Esta é a mais séria fonte de má interpretação das pesquisas de opinião."

EXEMPLO Suponha que a população das temperaturas do corpo humano tenha média de 98,6°F, como se aceita em geral. Suponha também que o desvio-padrão da população seja de 0,62°F (com base em dados dos pesquisadores da Universidade de Maryland). Selecionada aleatoriamente uma amostra de tamanho $n = 106$, determine a probabilidade de obter uma média de 98,2°F ou menos. (Obteve-se realmente o valor 98,2°F; veja as temperaturas à meia-noite para o Dia 2 do Conjunto de Dados 2 do Apêndice B.)

SOLUÇÃO Não conhecemos a distribuição da população, mas como o tamanho da amostra $n = 106$ excede 30, aplicamos o teorema central do limite e concluímos que a distribuição das médias amostrais é normal com os parâmetros:

$$\mu_{\bar{x}} = \mu = 98{,}6$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0{,}62}{\sqrt{106}} = 0{,}06021972$$

A Figura 5-23 mostra a área sombreada (veja a cauda esquerda do gráfico) correspondente à probabilidade que desejamos. Tendo determinado os parâmetros que se aplicam à distribuição mostrada na Figura 5-23, podemos agora achar a área sombreada pelos mesmos processos utilizados na seção precedente. Começamos determinando o escore z:

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{98{,}2 - 98{,}6}{0{,}06021972} = -6{,}64$$

Pela Tabela A-2, vemos que $z = -6{,}64$ está fora da tabela, mas, para valores superiores a 3,09, tomamos a área de 0,4999. Concluímos, pois, que a região A na Figura 5-23 é 0,4999 e que a região sombreada é $0{,}5 - 0{,}4999 = 0{,}0001$. (Em tabelas mais precisas, essa área aparece como 0,000000001.)

Este resultado mostra que, se a média de nossas temperaturas é realmente 98,6°F, há uma probabilidade extremamente pequena de se obter uma média amostral de 98,2°F ou menos para 106 indivíduos selecionados aleatoriamente. Os pesquisadores da Universidade de Maryland obtiveram, de fato, essa média, para o que há duas explicações possíveis: ou a média populacional é realmente de 98,6°F e sua amostra representa um evento aleatório extremamente raro, ou a média populacional é efetivamente inferior a 98,6°F e sua amostra é típica. Como a probabilidade é tão pequena, parece mais razoável concluir que a média populacional é inferior a 98,6°F. Este é o tipo de raciocínio utilizado no *teste de hipóteses*, a ser introduzido no Capítulo 7. Por ora, devemos enfocar a aplicação do teorema central do limite para achar a probabilidade de 0,0001; mas devemos ter em mente que esse teorema será utilizado mais adiante para estabelecer importantes conceitos da estatística.

O Teorema Central do Limite

Em *The Cartoon Guide to Statistics*, de Gonick e Smith, os autores descrevem como segue o *Teorema Central do Limite*: "Os dados que são influenciados por muitos efeitos aleatórios pequenos e não relacionados têm distribuição aproximadamente normal. Isto explica por que a normal está em toda parte: flutuações do mercado de ações, pesos de estudantes, temperaturas médias anuais, notas SAT. Todos são o resultado de muitos efeitos diferentes." As alturas das pessoas, por exemplo, são resultado de fatores hereditários, fatores ambientais, nutrição, cuidados médicos, região geográfica e outras influências que, combinadas, produzem valores distribuídos normalmente.

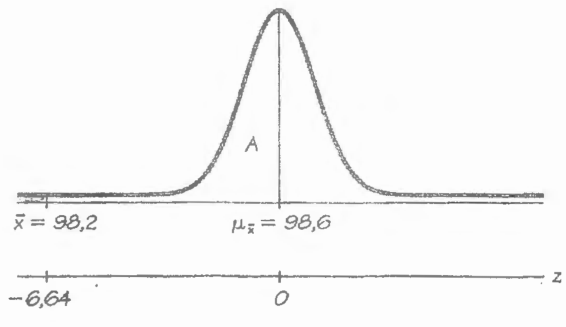
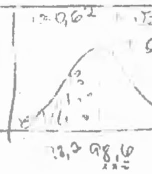


Fig. 5-23 Distribuição da temperatura média amostral do corpo ($n = 106$).

$\mu_{\bar{x}} = \mu$

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0{,}62}{\sqrt{106}} = 0{,}0602$

$n = 106$



$0{,}4999$

$P = 0{,}5 - 0{,}4999 = 0{,}0001$

$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{98{,}2 - 98{,}6}{0{,}0602} = -6{,}64$

Ao aplicar o Teorema Central do Limite, a utilização de $\sigma_{\bar{x}} = \sigma\sqrt{n}$ supõe que a população tenha um número infinito de elementos. No caso da amostragem com reposição (isto é, cada elemento extraído é devolvido à população antes da próxima extração), a população é, de fato, infinita. Entretanto, muitas aplicações realísticas envolvem amostragem sem reposição, de modo que as amostras sucessivas dependem dos resultados anteriores. Em processos industriais, os controladores de qualidade extraem itens de uma linha de produção finita sem os repor. Para tais populações finitas, devemos ajustar $\sigma_{\bar{x}}$. Eis uma regra empírica comum:

No caso de amostragem sem reposição, quando o tamanho n da amostra é superior a 5% do tamanho N da população finita (isto é, $n > 0{,}05N$), ajustamos o desvio-padrão da média amostral $\sigma_{\bar{x}}$ multiplicando-o pelo *fator de correção para população finita*:

$$\sqrt{\frac{N-n}{N-1}}$$

Com exceção dos Exercícios 21, 22 e 24, nos exemplos e exercícios desta seção supõe-se que o fator de correção para população finita não se aplica, porque a população é infinita ou porque o tamanho da amostra não excede 5% do tamanho da população.

É de suma importância notar que o teorema central do limite se aplica quando estamos em face de uma distribuição de médias amostrais e/ou o tamanho da amostra é superior a 30, ou a população original tem distribuição normal. Estabelecida a aplicabilidade do teorema central do limite, podemos determinar valores de $\mu_{\bar{x}} = \mu$ e $\sigma_{\bar{x}} = \sigma/\sqrt{n}$ e passar então a aplicar os métodos apresentados na seção precedente.

A importância do teorema central do limite decorre do fato de ele permitir que utilizemos os métodos básicos da distribuição normal em uma ampla diversidade de circunstâncias diferentes. No Capítulo 6, por exemplo, vamos aplicar o teorema ao utilizar dados amostrais para estimar médias populacionais. No Capítulo 7, aplicaremos o teorema na utilização de dados para testar afirmações sobre médias populacionais. Essas aplicações à estimação de parâmetros populacionais e ao teste de afirmações constituem uma parte extremamente importante da estatística — possibilitada pelo teorema central do limite.

5-5 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-4, suponha que as alturas das mulheres tenham média $\mu = 63{,}6$ in. e desvio-padrão $\sigma = 2{,}5$ in. (com base em dados do National Health Service dos EUA).

1. a. Escolhida aleatoriamente uma mulher, determine a probabilidade de sua altura estar entre 63,6 in. e 64,6 in.
   b. Selecionadas 36 mulheres, determine a probabilidade de sua altura média estar entre 63,6 in. e 64,6 in.
2. a. Escolhida aleatoriamente uma mulher, determine a probabilidade de sua altura ser superior a 63,0 in.
   b. Selecionadas aleatoriamente 100 mulheres, determine a probabilidade de sua altura média ser superior a 63,0 in.
3. a. Escolhida aleatoriamente 1 mulher, determine a probabilidade de sua altura ser superior a 64,0 in.
   b. Escolhidas aleatoriamente 50 mulheres, determine a probabilidade de suas alturas terem média superior a 64,0 in.
4. a. Selecionada aleatoriamente 1 mulher, determine a probabilidade de sua altura estar entre 63,0 in. e 65,0 in.
   b. Selecionadas aleatoriamente 75 mulheres, determine a probabilidade de suas alturas terem média entre 63,0 in. e 65,0 in.
5. Os prazos de substituição de aparelhos de TV têm distribuição normal com média de 8,2 anos e desvio-padrão de 1,1 ano (com base em dados do "Getting Things Fixed", *Consumer Reports*). Determine a probabilidade de 40 aparelhos de TV selecionados aleatoriamente terem prazo médio de substituição inferior a 8,0 anos.
6. Os prazos de substituição para CD players têm distribuição normal com média de 7,1 anos e desvio-padrão de 1,4 ano (com base em dados do "Getting Things Fixed", *Consumer Reports*). Determine a probabilidade de 45 CD players selecionados aleatorimente terem prazo de substituição superior a 7,0 anos.
7. De acordo com a Opinion Research Corporation, os homens gastam em média 11,4 minutos no chuveiro. Admita que os tempos tenham distribuição normal com desvio-padrão de 1,8 minuto. Selecionados aleatoriamente 33 homens, determine a probabilidade de que seus tempos no chuveiro tenham média entre 11,0 min e 12,0 min.
8. De acordo com a International Mass Retail Association, as jovens de 13 a 17 anos de idade gastam em compras uma média mensal de \$31,20. Suponha que essas importâncias tenham um desvio-padrão de \$8,27. Selecionadas aleatoriamente 85 jovens, qual é a probabilidade de que a média de suas compras mensais fique entre \$30,00 e \$33,00?
9. Para as mulheres na faixa etária de 18 a 24 anos, a pressão sistólica do sangue (em mm Hg) tem distribuição normal com média de 114,8 e desvio-padrão de 13,1 (com base em dados do National Health Survey dos EUA).
   a. Selecionada aleatoriamente uma mulher nessa faixa etária, determine a probabilidade de a sua pressão sistólica ser superior a 120.
   b. Selecionadas aleatoriamente 12 mulheres nessa faixa etária, determine a probabilidade de sua pressão sistólica média ser superior a 120.
   c. Dado que a parte (b) envolve uma amostra de tamanho não superior a 30, por que podemos usar o teorema central do limite?
10. As quantidades de precipitação anual no estado de Iowa aparentam ter distribuição normal com média de 32,473 in. e desvio-padrão de 5,601 in. (com base em dados do Ministério da Agricultura dos EUA).
    a. Escolhido um ano aleatoriamente, determine a probabilidade de a precipitação anual correspondente ser inferior a 29,000 in.
    b. Para uma década selecionada aleatoriamente, determine a probabilidade de a média das precipitações anuais ser inferior a 29,000 in.
    c. Como a parte (b) envolve uma amostra de tamanho não superior a 30, por que podemos aplicar o teorema central do limite?
11. As idades dos aviões comerciais dos EUA têm uma média de 13,0 anos e um desvio-padrão de 7,9 anos (com base em dados do Departamento de Aviação Civil dos EUA). Se a Administração Federal da Aviação seleciona aleatoriamente 35 aviões comerciais para um teste especial de resistência, determine a probabilidade de a idade média desse grupo de aviões ser superior a 15,0 anos.
12. Uma análise dos números de horas por semana que os calouros universitários (nos EUA) dedicam ao estudo acusa média de 7,06 horas e desvio-padrão de 5,32 horas (com base em dados do *The American Freshman*). Selecionados aleatoriamente 55 calouros, determine a probabilidade de seu tempo semanal médio de estudo exceder 7,00 horas.

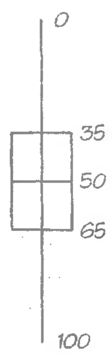
13. O gerador de números aleatórios de um computador típico produz números com uma distribuição uniforme entre 0 e 1, com média de 0,500 e desvio-padrão de 0,289. Gerados 45 números aleatórios, determine a probabilidade de sua média ser inferior a 0,565.
14. Realizou-se um estudo da utilização de cintos de segurança entre crianças envolvidas em acidentes de automóvel que exigiram hospitalização. Verificou-se que as crianças que não usavam nenhum dispositivo de segurança acusaram uma estada média de 7,37 dias em hospitais, com desvio-padrão de 0,79 dias [com base em dados de "Morbidity Among Pediatric Motor Vehicle Crash Victims: The Effectiveness of Seat Belts" (Morbidade entre Acidentes de Automóvel com Vítimas Infantis: A Eficácia dos Cintos de Segurança), por Osberg e Di Scala, *American Journal of Public Health*, Vol. 82, No. 3]. Selecionadas aleatoriamente 40 dessas crianças, determine a probabilidade de sua permanência média em hospital ser superior a 7,00 dias.
15. A cidade de Newport tem um serviço de coleta de lixo que acusa sobrecarga se a média do lixo das suas 4872 casas exceder 27,88 lb em uma semana. Os pesos totais têm distribuição normal com média de 27,44 lb e desvio-padrão de 12,46 lb (com base em dados do Projeto do Lixo da Universidade do Arizona). Qual é a proporção de semanas em que o serviço de coleta de lixo acusa sobrecarga? Trata-se de uma situação aceitável, ou devem-se tomar providências para corrigir um problema de sobrecarga no sistema?
16. Os testes verbais SAT têm distribuição normal com média de 430 e desvio-padrão de 120 (com base em dados do College Board ATP). Escolhem-se aleatoriamente testes verbais SAT dentre a população de estudantes que fizerem o curso preparatório na Tillman Training School. Admita que esse curso de treinamento não influa nas notas do teste.
    a. Escolhido aleatoriamente 1 estudante, determine a probabilidade de ele ter obtido uma nota superior a 440.
    b. Selecionados aleatoriamente 100 estudantes, determine a probabilidade de a nota média ser superior a 440.
    c. Se 100 estudantes da Tillman conseguem uma média amostral de 440, parece razoável concluir que o curso é eficiente porque os estudantes se saem melhor no SAT?
17. As durações da gravidez têm distribuição normal com média de 268 dias e desvio-padrão de 15 dias.
    a. Selecionada aleatoriamente uma mulher grávida, determine a probabilidade de a duração de sua gravidez ser inferior a 260 dias.
    b. Se 25 mulheres escolhidas aleatoriamente são submetidas a uma dieta especial a partir do dia em que engravidam, determine a probabilidade de os prazos de duração de sua gravidez terem média inferior a 260 dias (admitindo que a dieta não produza efeito).
    c. Se as 25 mulheres têm realmente média inferior a 260 dias, há razão de preocupação para os supervisores médicos?
18. Utilizando uma medida-padrão de satisfação com os salários, um estudo constata que os administradores de universidade têm uma média de 38,9 e um desvio-padrão de 12,4 [com base em dados de "Job Satisfaction Among Academic Administrators" (Satisfação com o Emprego entre Administradores Acadêmicos), por Glick, *Research in Higher Education*, Vol. 33, No. 5]. Um pesquisador seleciona aleatoriamente 150 administradores de faculdade e mede seus níveis de satisfação com o salário.
    a. Determine a probabilidade de a média ser superior a 42,0.
    b. Se uma amostra de 150 administradores acusa média de 42,0 ou mais, há razão para crer que essa amostra provenha de uma população com média superior a 38,9?
19. Os bombons M&M têm peso médio de 0,9147 g e desvio-padrão de 0,0369 g (com base em dados do Conjunto de Dados 11 do Apêndice B). Os bombons M&M usados naquele Conjunto de Dados provêm de um pacote contendo 1498 bombons, e o rótulo do pacote informa que o peso líquido é de 48,0 oz (3 lb), ou 1361 g. (Se cada pacote tem 1498 bombons, o peso médio deve exceder 1361/1498 = 0,9085 g para que o conteúdo líquido pese no mínimo 1361 g.)
    a. Selecionado aleatoriamente 1 bombom M&M, determine a probabilidade de pesar mais de 0,9085 g.
    b. Selecionados aleatoriamente 1498 bombons M&M, determine a probabilidade de seu peso médio ser no mínimo de 0,9085 g.
    c. À vista desses resultados, parece que a Mars Company esteja dando aos consumidores de M&M as quantidades indicadas no rótulo?
20. A população de pesos de homens tem distribuição normal com média de 173 lb e desvio-padrão de 30 lb (com base em dados do National Health Survey dos EUA). Um elevador do Clube Masculino de Dallas impõe o limite de 32 ocupantes, mas haverá uma sobrecarga se esses 32 ocupantes tiverem peso médio superior a 186 lb (dando um peso total de (32)(186) = 5952 lb). Se os 32 ocupantes homens resultam de uma seleção aleatória, determine a probabilidade de seu peso médio exceder 186 lb, ocasionando uma sobrecarga no elevador. Com base no resultado obtido, há razão para preocupação?

5-5 Exercícios B: Além do Básico

21. Refaça o Exercício 20, supondo que o tamanho da população seja $N = 500$ e que toda a amostragem seja feita sem reposição. (*Sugestão*: Veja o estudo do fator de correção para população finita.)
22. Uma *população* consiste nos valores 2, 3, 6, 8, 11, 18.
    a. Determine μ e σ.
    b. Relacione todas as amostras de tamanho $n = 2$ que podem ser obtidas sem reposição.
    c. Determine a população de todos os valores de $\bar{x}$ achando a média de cada amostra da parte (b).
    d. Ache a média $\mu_{\bar{x}}$ e o desvio-padrão $\sigma_{\bar{x}}$ para a população de médias amostrais da parte (c).
    e. Verifique que

$$\mu_{\bar{x}} = \mu \quad \text{e} \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

23. Um pesquisador de educação cria um índice de interesse acadêmico e obtém notas para uma amostra aleatória de 350 alunos de faculdade. O diagrama mostrado resume os resultados. Selecionados ao acaso 15 estudantes, determine a probabilidade de sua nota média ser superior a 55.



24. O fator de correção para população finita pode ser desprezado quando a amostragem se faz com reposição ou quando $n \leq 0{,}05N$. No caso de uma amostra (sem reposição) que representa 5% da população N, que é que os valores do fator de correção para populações finitas têm em comum para valores de $N \geq 600$?

5-6 A Distribuição Normal como Aproximação da Distribuição Binomial

Na Seção 4-3, introduzimos a *distribuição de probabilidade binomial*, que se aplica a uma variável aleatória discreta (e não a uma variável aleatória contínua, como a distribuição normal). Vimos que uma distribuição binomial deve satisfazer as quatro condições seguintes:

1. O experimento deve comportar um *número fixo de provas.*
2. As provas devem ser *independentes.*
3. Os resultados de cada prova devem ser classificados *em duas categorias.*
4. As probabilidades devem permanecer *constantes* em todas as provas.

A determinação de $P(x)$, a probabilidade de x sucessos em n provas, é um problema típico de probabilidade binomial. Na Seção 4-3 vimos como resolver problemas binomiais com um programa de computador, ou com o auxílio da Tabela A-1, ou com a fórmula da probabilidade binomial. Em muitos casos, entretanto, esses métodos não são práticos. Consideremos o problema binomial que consiste em determinar a probabilidade de obter ao menos 64 homens em 100 pessoas se-lecionadas aleatoriamente. Não podemos utilizar a Tabela A-1, porque não vai além de $n = 15$, e a fórmula não é prática, porque teríamos de aplicá-la 37 vezes (uma vez para cada inteiro de 64 a 100). Tais cálculos são do tipo que os matemáticos costumam chamar de "tediosos". Assim, em lugar deles, vamos introduzir um novo método pelo qual aproximamos a distribuição binomial por uma distribuição normal, simplificando grandemente os cálculos. Resumimos a seguir o ponto-chave desta seção.

Distribuição Normal como Aproximação da Binomial

Se $np \geq 5$ e $nq \geq 5$, então a variável aleatória binomial tem distribuição aproximadamente normal com média e desvio-padrão dados por

$$\mu = np$$

$$\sigma = \sqrt{npq}$$

Confiabilidade e Validade

A confiabilidade de dados se refere à consistência com que os resultados ocorrem, enquanto a validade dos dados diz respeito ao grau com que os dados medem o que deve ser medido. A confiabilidade de um teste de QI pode ser julgada comparando-se os valores de um teste feito em uma data, com os valores do mesmo teste realizado em outra ocasião. Para julgar a validade de um teste de QI, podemos comparar os valores do teste com outro indicador da inteligência, como, por exemplo, o desempenho acadêmico. Muitos críticos alegam que os testes de QI são confiáveis mas não são válidos; dão resultados consistentes, mas não medem a inteligência.

Reveja a Figura 4-4 (que se aplica a uma distribuição binomial com $n = 100$, $p = 0{,}5$ e $q = 0{,}5$) e note que o histograma dessa distribuição binomial particular tem aproximadamente a mesma forma que uma distribuição normal. A justificativa formal do emprego da distribuição normal como aproximação da binomial requer recursos matemáticos mais avançados, mas a Figura 4-4 constitui um argumento visual em favor dessa aproximação.

A aproximação pela normal envolve o processo a seguir, detalhado no fluxograma da Figura 5-24:

Processo de aproximação pela normal

1. Verificar se a distribuição de probabilidade binomial é aplicável.
2. Utilize (se possível) programas (como STATDISK ou Minitab) ou uma calculadora TI-83. (Na Seção 4-3, encontram-se os resultados obtidos com STATDISK e Minitab.)
3. Se não dispuser de programas, procure resolver o problema utilizando a Tabela A-1.
4. Se não puder utilizar a Tabela A-1, considere a fórmula da probabilidade binomial. Se puder resolver o problema *facilmente* com a fórmula binomial, aplique-a. Em caso contrário, passe à Etapa 5.
5. Certifique-se de que a distribuição normal é uma aproximação conveniente da binomial, verificando se $np \geq 5$ e $nq \geq 5$.
6. Determine os valores dos parâmetros μ e σ, calculando $\mu = np$ e $\sigma = \sqrt{npq}$.
7. Identifique o valor discreto x que representa o número de sucessos no experimento binomial. Modifique o valor *discreto* x substituindo-o pelo *intervalo* de $x - 0{,}5$ a $x + 0{,}5$. (Veja o tópico *Correções de Continuidade* mais adiante nesta seção.) Trace uma curva normal e introduza os valores de μ, σ, e/ou $x - 0{,}5$ ou $x + 0{,}5$, conforme apropriado.
8. Modifique x, substituindo-o por $x - 0{,}5$ ou $x + 0{,}5$, conforme adequado, e determine o escore z: $z = (x - \mu)/\sigma$.
9. Com o escore z achado em 8, consulte a Tabela A-1, para achar a área entre μ e $x - 0{,}5$ ou $x + 0{,}5$, conforme adequado. Identifique a área correspondente à probabilidade desejada.

Ela Ganhou Duas Vezes na Loteria!

Evelyn Marie Adams ganhou na Loteria de New Jersey duas vezes em quatro meses. Esse feliz acontecimento foi considerado uma incrível coincidência com uma chance de apenas 1 em 17 trilhões. Mas os matemáticos de Harvard, Persi Diaconis e Frederick Mosteller, observam que há 1 chance em 17 trilhões de uma determinada pessoa, com um bilhete de cada uma das duas loterias de New Jersey, ganhar ambas as vezes. Todavia, há cerca de 1 chance em 30 de que alguém nos Estados Unidos ganhe em uma loteria duas vezes em um período de quatro meses. Diaconis e Mosteller analisaram coincidências e concluíram que, "com uma amostra suficientemente grande, qualquer coisa pode acontecer."

O exemplo seguinte ilustra esse processo de aproximação pela normal.

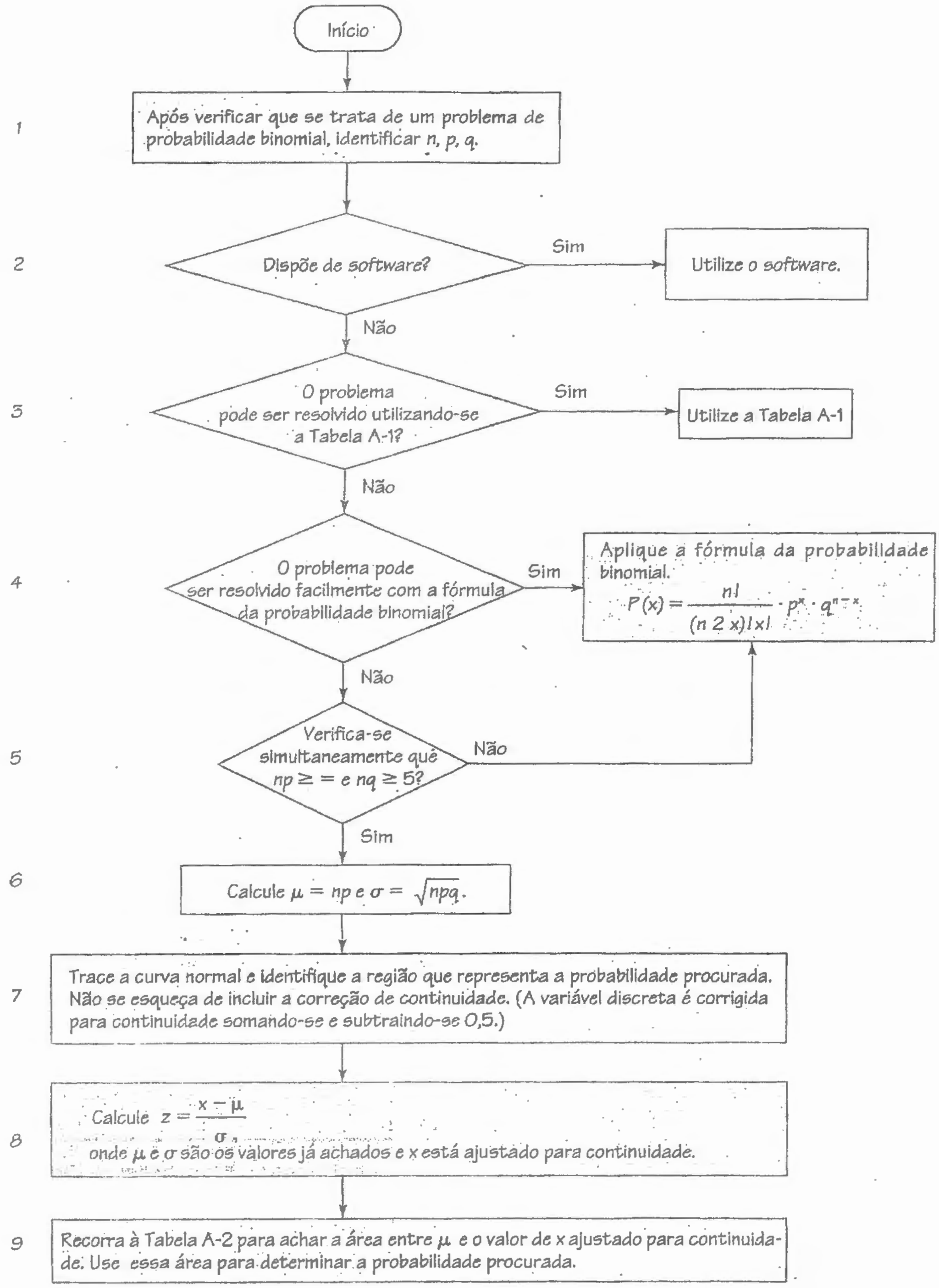


Fig. 5-24 Resolução de um problema de probabilidade binomial utilizando a aproximação normal.

EXEMPLO Suponha que o quadro administrativo de sua faculdade tenha igual número de candidatos e candidatas a emprego, e que 64 dos 100 funcionários recém-admitidos são homens. Estime a probabilidade de obter *pelo menos* 64 homens, se cada contratação é feita independentemente e sem qualquer discriminação quanto a sexo. (A probabilidade de obter *exatamente* 64 homens na verdade não nos diz nada, porque, em 100 provas, a probabilidade de qualquer número determinado de homens é por demais pequena. Em lugar disso, necessitamos da probabilidade de obter um resultado *pelo menos* tão extremo como o verificado.) Com base no resultado, parece que a faculdade está fazendo discriminação quanto ao sexo?

SOLUÇÃO

Passo 1: O problema envolve uma distribuição binomial com um número fixo de provas ($n = 100$), presumivelmente independentes, duas categorias (homem, mulher) de resultado para cada prova, e probabilidades que supomos permanecerem constantes.

Passo 2: Para os fins deste exemplo, admitiremos não dispor dos recursos de um computador.

Passo 3: Não podemos utilizar a Tabela A-1, porque o valor $n = 100$ excede o máximo $n = 15$ da tabela.

Passo 4: Poderíamos aplicar a fórmula, mas teríamos de fazê-lo para cada inteiro de 64 a 100. Preferimos evitar a aplicação da fórmula da probabilidade binomial 37 vezes.

Passo 5: Convimos que é razoável aproximar a distribuição binomial pela distribuição normal, porque $np \geq 5$ e $nq \geq 5$, conforme se vê a seguir:

$$np = (100)(0,5) = 50 \text{ (Portanto, } np \geq 5.)$$
$$nq = (100)(0,5) = 50 \text{ (Portanto, } nq \geq 5.)$$

Passo 6: Passamos a determinar os valores de μ e σ. Obtemos:

$$\mu = np = (100)(0,5) = 50$$

$$\sigma = \sqrt{npq} = \sqrt{(100)(0,5)(0,5} = (5)$$

Passo 7: Na Figura 5-25, mostramos o valor discreto 64 representado pela faixa vertical delimitada por 63,5 e 64,5. A Figura 5-26 mostra a área desejada: é a área sombreada no extremo da cauda direita do gráfico.

Passo 8: Não podemos achar diretamente a área sombreada na Figura 5-26; procuramos, portanto, a área da

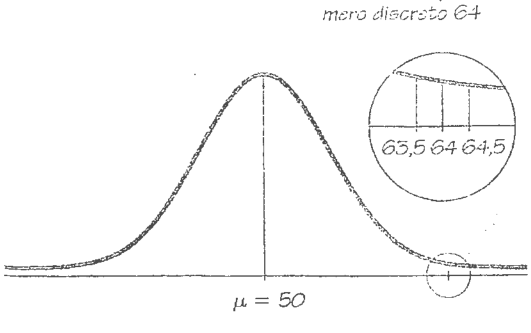


Fig. 5-25 Ilustração da correção de continuidade.

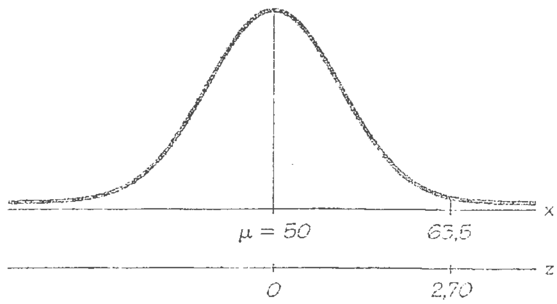


Fig. 5-26 Determinação da probabilidade de "pelo menos" 64 homens entre 100 novos empregados.

região delimitada por $\mu = 50$ e o valor 63,5. Essa área corresponde ao seguinte escore z:

$$z = \frac{x - \mu}{\sigma} = \frac{63,5 - 50}{5} = 2,70$$

Passo 9: Recorrendo à Tabela A-1, vemos que $z = 2,70$ corresponde a uma área de 0,4965 e, assim, a área procurada é $0,5 - 0,4965 = 0,0035$.

Supondo que homens e mulheres tenham igual chance de ser contratados, a probabilidade de pelo menos 64 homens em 100 novos empregados é de aproximadamente 0,0035. (Com os programas STATDISK e Minitab e a calculadora TI-83, obtemos a resposta 0,0033; a aproximação aqui é, pois, suficientemente boa.)

Como a probabilidade de obter ao menos 64 homens é tão pequena (0,0035), concluímos que ou ocorreu um evento bastante raro, ou a hipótese de homens e mulheres terem a mesma chance é incorreta. Parece que a faculdade faz discriminação quanto ao sexo. O raciocínio aqui desenvolvido será retomado com maior detalhe no Capítulo 7, onde abordamos o método formal de teste de hipótese. Por ora nos concentraremos no método de determinação da probabilidade com a técnica da aproximação normal.

Correções de Continuidade

Para muitos, não é fácil entender o Passo 7 do processo delineado. Baseia-se no conceito de correção de continuidade.

DEFINIÇÃO

Quando utilizamos a distribuição normal (que é contínua) como aproximação da distribuição binomial (discreta), aplica-se uma **correção de continuidade** a um número discreto x na distribuição binomial, representando o valor único x pelo *intervalo* de $x - 0,5$ a $x + 0,5$ (isto é, somamos e subtraímos 0,5).

As sugestões práticas seguintes auxiliarão o leitor no uso da correção de continuidade.

Processo para correções de continuidade

1. Ao utilizar a distribuição normal como aproximação da binomial, faça *sempre* a correção de continuidade.
2. Ao aplicar a correção de continuidade, identifique primeiro o número discreto x de importância para o problema de probabilidade binomial. Por exemplo, se desejamos a probabilidade de ao menos 64 homens em 100 pessoas selecionadas aleatoriamente, o número inteiro, discreto, de interesse é $x = 64$. Focalize sua atenção inicialmente no próprio valor x, ignorando temporariamente se deseja pelo menos x, mais de x, menos de x, ou qualquer um.
3. Faça o esboço de uma distribuição normal centrada em μ, e trace uma *faixa vertical* com base centrada em x. Marque o extremo esquerdo da faixa com o número $x - 0,5$ e o extremo direito com o número $x + 0,5$. Para $x = 64$, a faixa deve ir de 63,5 a 64,5. *Considere toda a área da faixa como representando a probabilidade do número discreto x.*
4. Determine agora se o próprio valor de x deve ser incluído na probabilidade desejada. (Por exemplo, "pelo menos x" inclui x, porém "mais de x" não inclui x.) Em seguida, determine se a probabilidade desejada se refere a pelo menos x, no máximo x, mais de x, menos de x, ou exatamente x. Sombreie a área à direita ou à esquerda da faixa, conforme apropriado; sombreie também o interior da própria faixa *se e somente se* o valor de x deve ser incluído. A região total sombreada corresponde à probabilidade que procuramos.

Para ver como este processo resulta em correções de continuidade, atente para os casos comuns ilustrados na Figura 5-27. Esses casos correspondem às seguintes afirmações.

Afirmação	Área
Pelo menos 64 (inclui 64 ou mais)	À *direita* de 63,5
Mais de 64 (não inclui 64)	À *direita* de 64,5
No máximo 64 (inclui 64 ou menos)	À *esquerda* de 64,5
Menos de 64 (não inclui 64)	À *esquerda* de 63,5
Exatamente 64	Entre 63,5 e 64,5

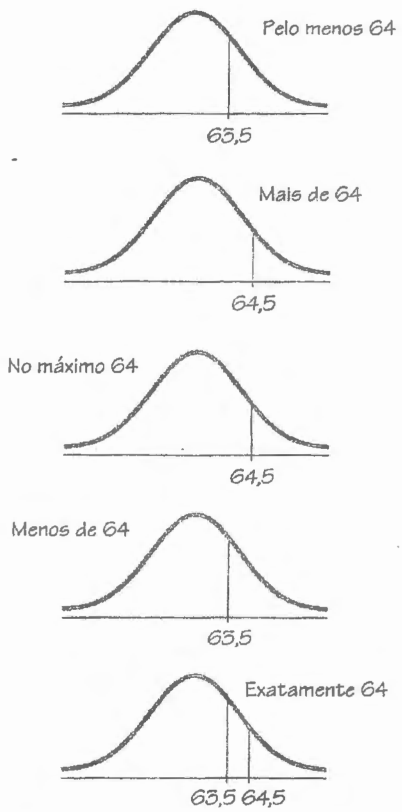


Fig. 5-27 Identificação da área correta.

EXEMPLO Cerca de 4,4% dos acidentes fatais com automóveis são causados por pneus defeituosos (com base em dados do National Safety Council — Conselho Nacional de Segurança dos EUA). Se um estudo de segurança em uma rodovia começa com a escolha aleatória de 750 casos de acidentes fatais com veículos motorizados, estime a probabilidade de exatamente 35 deles terem sido causados por pneus defeituosos

SOLUÇÃO A solução segue os passos indicados na Figura 5-24.

Passo 1: As condições descritas verificam os critérios para a distribuição binomial com $n = 750$, $p = 0,044$, $q = 0,956$ e $x = 35$. (O valor de q se obtém de $q = 1 - p = 1 - 0,044 = 0,956$.)

Passo 2: Para fins deste exemplo, admitimos não dispor de um computador.

Passo 3: Não podemos utilizar a Tabela A-1, porque $n = 750$ excede o valor máximo $n = 15$.

Passo 4: Pode-se aplicar a fórmula da probabilidade binomial, mas

$$P(35) = \frac{750!}{(750 - 35)!\,35!}\left(0,044^{35}\right)\left(0,956^{750-35}\right)$$

é difícil de calcular, porque muitas calculadoras não calculam 70! ou mais.

Passo 5: Neste problema

$$np = (750)(0,044) = 33,0 \text{ (Portanto } np \geq 5.)$$
$$nq = (750)(0,956) = 717,0 \text{ (Portanto } nq \geq 5.)$$

Como tanto np como nq são ambos no mínimo iguais a 5, concluímos que a aproximação normal para a binomial é satisfatória.

Passo 6: Obtemos os valores de μ e σ como segue:

$$\mu = np = (750)\,(0,044) = 33,0$$
$$\sigma = \sqrt{npq} = \sqrt{(750)(0,044)(0,956)} = 5,6167606$$

Passo 7: Esboçamos a curva normal exibida na Figura 5-28. A região sombreada da figura representa a probabilidade procurada. A aplicação da correção de continuidade resulta na representação de 35 pela região que vai de 34,5 a 35,5.

Passo 8: O formato da Tabela A-2 exige que primeiro achemos a probabilidade correspondente à região delimi-

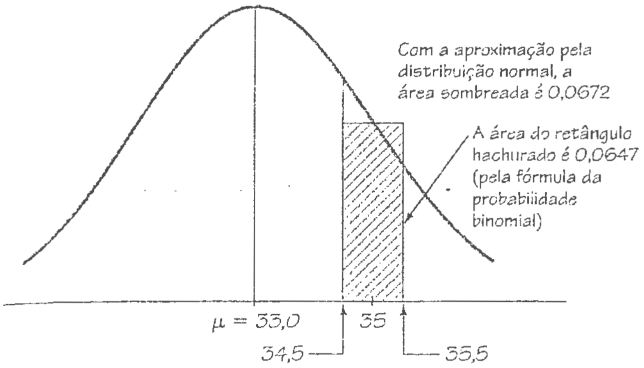


Fig. 5-28 Distribuição do número de acidentes fatais causados por pneus defeituosos.

tada à esquerda pela vertical que passa pela média 33,0 e à direita pela vertical que passa por 35,5. Portanto, um dos cálculos necessários neste passo é:

$$z = \frac{35,5 - 33,0}{5,6167606} = 0,45$$

Necessitamos também da probabilidade correspondente à região delimitada por 33,0 e 34,5; calculamos, pois

$$z = \frac{34,5 - 33,0}{5,6167606} = 0,27$$

Passo 9: Com auxílio da Tabela A-2, vemos que $z = 0,45$ corresponde a uma probabilidade de 0,1736 e $z = 0,27$ corresponde a uma probabilidade de 0,1064. Conseqüentemente, a região sombreada da Figura 5-28 representa uma probabilidade de $0,1736 - 0,1064 = 0,0672$. A probabilidade de obter exatamente 35 casos (em 750) de acidentes fatais causados por pneus defeituosos é de aproximadamente 0,0672.

Escolhendo Números em uma Loteria

Em uma loteria típico, o apostador escolhe seis números diferentes. Após uma extração aleatória, qualquer aposta com a combinação correta participa do prêmio. Como os números vencedores são escolhidos aleatoriamente, qualquer escolha de seis números tem a mesma chance que qualquer outra, mas algumas combinações são melhores do que outras. A combinação 1, 2, 3, 4, 5, 6 é uma escolha fraca, porque muitos tendem a escolhê-la. Em uma loteria da Flórida, com um prêmio de \$105 milhões, 52.000 bilhetes tinham a combinação 1, 2, 3, 4, 5, 6; se essa combinação tivesse sido sorteada, o prêmio teria sido apenas de \$1000. É aconselhável escolher combinações não procuradas por muitas pessoas. Evite combinações que formem qualquer padrão.

Se resolvermos o exemplo precedente utilizando STATDISK, Minitab ou uma calculadora TI-83, obteremos o resultado 0,0647, mas a aproximação normal deu o resultado 0,0672. A discrepância de 0,0025 é devida ao fato de que a aplicação da distribuição normal resulta em um valor *aproximado* que é a área da região sombreada na Figura 5-28, enquanto a área correta é um retângulo centrado acima de 35. (A Figura 5-28 ilustra essa discrepância.) A área do retângulo é 0,0647, mas a área da região sombreada aproximadora é 0,0672.

5-6 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-8, aplique a correção de continuidade e descreva a região da curva normal que corresponde à probabilidade indicada. Por exemplo, a probabilidade de "mais de 47 sucessos" corresponde à seguinte área da curva normal: a área à direita de 47,5.

1. Probabilidade de mais de 35 peças defeituosas.
2. Probabilidade de ao menos 175 meninas.
3. Probabilidade de menos de 42 respostas corretas em questões de múltipla escolha.
4. Probabilidade de exatamente 65 respostas corretas em questões do tipo verdadeiro/falso.
5. Probabilidade de, no máximo, 72 carros com freios defeituosos.
6. Probabilidade de o número de meninas estar entre 35 e 45 inclusive.
7. Probabilidade de o número de eleitores republicanos estar entre 125 e 150 inclusive.
8. Probabilidade de o número de pacientes com sangue do grupo A ser exatamente 34.

Nos Exercícios 9-12, faça o seguinte: (a) Ache a probabilidade binomial indicada utilizando a Tabela A-1 do Apêndice A. (b) Se $np \geq 5$ e $nq \geq 5$, estime também a probabilidade indicada utilizando a distribuição normal como aproximação da binomial; se $np < 5$ ou $nq < 5$, indique então que a aproximação normal não é adequada.

9. Com $n = 14$ e $p = 0,50$, determine $P(8)$.
10. Com $n = 10$ e $p = 0,40$, determine $P(7)$.
11. Com $n = 15$ e $p = 0,80$, determine P (pelo menos 8).
12. Com $n = 14$ e $p = 0,60$, determine P (menos de 9).
13. Estime a probabilidade de pelo menos 55 meninas em 100 nascimentos.
14. Estime a probabilidade de exatamente 32 meninos em 64 nascimentos.
15. Estime a probabilidade de aprovação em um teste de 50 questões do tipo verdadeiro/falso, se 60% (ou 30 respostas corretas) garantem a aprovação e se todas as respostas são dadas aleatoriamente (por "palpite").
16. Um teste de múltipla escolha consiste em 50 questões com respostas possíveis a, b, c, d, e. Estime a probabilidade de alcançar no máximo 30% corretas, se todas as respostas são dadas por "palpite".
17. Há 80% de chance de que um empregador verifique o nível de instrução de um candidato a emprego (com base em dados do Bureau of National Affairs, Inc.). Para 100 candidatos selecionados aleatoriamente, estime a probabilidade de *exatamente* 85 terem seus currículos examinados.
18. Susan Stein é a diretora de publicidade para um seriado policial transmitido no horário nobre da TV NBC; ela pretende convencer os anunciantes de que a audiência é muito maior do que o número dos que vêem o *show* na hora do programa. Uma pesquisa da Nielsen mostrou que 66% das pessoas que usam videocassete gravam os programas das redes mais importantes. Estime a probabilidade de que, entre 1000 *shows* gravados, selecionados aleatoriamente, pelo menos 700 sejam das redes mais importantes, como a diretora de vendas alega. Com base nos resultados, sua alegação parece plausível?
19. De acordo com um representante dos consumidores da Mars (a companhia de bombons), 10% de todos os bombons M&M são azuis. O Conjunto de Dados 11 do Apêndice B mostra que, entre 100 bombons M&M escolhidos, 5 são azuis. Estime a probabilidade de escolher aleatoriamente 100 bombons M&M e obter 5 ou menos azuis. Admita que a taxa de 10% de azuis seja correta. Com

base no resultado, pode ser considerado um evento raro o aparecimento de 5 ou menos bombons azuis, quando se escolhem aleatoriamente 100 bombons M&M?

20. Marc Taylor planeja fazer 200 apostas de $1 cada no número 7 da roleta. Em qualquer rodada, há uma probabilidade de 1/38 de dar o 7. Para que Marc saia com lucro, o número 7 deve aparecer pelo menos 6 vezes nas 200 rodadas. Estime a probabilidade de Marc terminar ganhando.
21. Com base em dados do Ministério da Justiça dos EUA, 16% das pessoas presas são mulheres. Selecionam-se aleatoriamente em um estado 400 casos de prisão. Estime a probabilidade de o número de mulheres ser 38 ou menos. Se esses casos de prisão incluem 38 ou menos mulheres, parece plausível que a taxa de prisão de mulheres nesse estado seja de 16%?
22. A Air America vem experimentando uma taxa de 7% de não-comparecimento a embarque em reservas feitas com antecipação. Em um projeto-piloto que exige que os passageiros confirmem as reservas, verificou-se que em 250 reservas antecipadas selecionadas aleatoriamente, há 4 não-comparecimentos. Admitindo que a exigência de confirmação não faça efeito, de forma que a taxa de 7% é mantida, estime a probabilidade de 4 ou menos não-comparecimentos entre 250 reservas selecionadas aleatoriamente. Com base nesse resultado, a exigência parece produzir efeito?
23. Atualmente, cerca de dois terços das companhias americanas fazem teste de uso de drogas em empregados recém-admitidos, e os resultados do teste em 3,8% dos empregados dão positivo (com base em dados da American Management Association). A Sigma Electronics Company testa 150 candidatos a emprego e constata que, em 10 deles, o teste foi positivo. Estime a probabilidade de 10 ou mais resultados positivos em 150 candidatos. Com base nesse valor, os 10 resultados positivos parecem uma cifra excepcionalmente elevada?
24. A American Airlines alega que 77,5% de seus vôos chegam no horário (com base em dados do Departamento de Transportes dos EUA). Em 50 vôos da American Airlines selecionados aleatoriamente, constatou-se que 34 (ou 68%) deles chegaram no horário. Estime a probabilidade de que, em 50 vôos da American Airlines, no máximo 34 cheguem no horário. Com base no resultado, parece plausível que a taxa de pontualidade de 77,5% alegada seja correta?
25. De acordo com a American Medical Association (Associação Médica Americana), 18,4% dos bacharéis fumam. Um estudo sobre condições de saúde começa com a escolha de 300 bacharéis, mas o número de fumantes é 72, superior ao esperado. O estudo está sendo questionado, porque o número de fumantes não parece corresponder à taxa global de 18,4% para toda a população de bacharéis. Estime a probabilidade de obter pelo menos 72 fumantes em uma amostra aleatória de 300. Com base no resultado, parece viável obter 72 fumantes aleatoriamente, ou haverá algo errado com a amostra?
26. Alguns casais têm características genéticas configuradas de modo que um quarto de todos os rebentos têm olhos azuis. Faz-se um estudo em 40 casais que se supõe terem aquelas características, com o resultado de que 8 em 40 rebentos têm olhos azuis. Estime a probabilidade de que, em 40 rebentos, 8 ou menos tenham olhos azuis. Com base nessa probabilidade, a taxa de um quarto parece correta?
27. O Providence Memorial Hospital está fazendo uma campanha de doações, porque seu estoque de sangue do grupo O está baixo; o hospital necessita de 177 doadores de sangue do grupo O. Se 400 voluntários doam sangue, estime a probabilidade de que haja pelo menos 177 com sangue do grupo O. Quarenta e cinco por cento dos americanos têm sangue do grupo O, de acordo com os dados fornecidos pelo Greater New York Blood Program (Grande Programa de Sangue de Nova York).
28. Vimos, na Seção 3-4, que algumas companhias controlam a qualidade utilizando um método de amostragem segundo o qual um lote inteiro de artigos é rejeitado se, em uma amostra aleatória, o número de defeitos é pelo menos igual a um número prefixado. A Dayton Machine Company compra ferrolhos em lotes de 5000 e rejeita um lote se, em uma amostra de 50, são encontrados pelo menos 2 defeitos. Estime a probabilidade de rejeitar um lote se o fabricante de ferrolhos trabalha com uma taxa de 10% de defeituosos.

5-6 Exercícios B: Além do Básico

29. Os prazos de substituição para aparelhos de TV têm distribuição normal com média de 8,2 anos e desvio-padrão de 1,1 ano (com base em dados do "Getting Things Fixed", *Consumer Reports*). Estime a probabilidade de que, para 250 aparelhos de TV selecionados aleatoriamente, pelo menos 15 deles tenham prazos de substituição superior a 10,0 anos.
30. Suponha que um jogador de beisebol tenha uma marca de 0,350, de modo que sua probabilidade de marca é 0,350. Admita também que suas tentativas de marcar sejam independentes entre elas.
    a. Determine a probabilidade de pelo menos uma marca em 4 tentativas em 1 jogo.
    b. Supondo que esse batedor bata 4 vezes em cada jogo, estime a probabilidade de obter um total de pelo menos 56 marcas em 56 jogos.
    c. Supondo que esse batedor bata 4 vezes em cada jogo, determine a probabilidade de pelo menos 1 marca em cada um dos 56 jogos consecutivos (recorde de Joe DiMaggio em 1941).
    d. Qual a média mínima de batidas necessária para que a probabilidade na parte (c) seja superior a 0,1?
31. Ache a diferença entre as respostas obtidas com a correção de continuidade e sem ela, em cada um dos seguintes casos. Que se pode concluir dos resultados?
    a. Estime a probabilidade de pelo menos 11 meninas em 20 nascimentos.
    b. Estime a probabilidade de pelo menos 22 meninas em 40 nascimentos.
    c. Estime a probabilidade de pelo menos 220 meninas em 400 nascimentos.
32. A Air America trabalha somente com reservas antecipadas e experimenta uma taxa de 7% de não-comparecimentos. Quantas reservas podem ser aceitas para um avião com 250 assentos, a fim de que haja pelo menos 0,95 de probabilidade de acomodar todos os que fizeram reserva?

Vocabulário

distribuição normal
distribuição uniforme
curva de densidade
distribuição normal padronizada
distribuição amostral de médias amostrais
teorema central do limite
erro-padrão da média
fator de correção para população finita
correção de continuidade

Revisão

O Capítulo 4 introduziu o conceito de distribuição de probabilidade, mas incluiu apenas tipos discretos. Neste capítulo foram introduzidas as distribuições contínuas de probabilidade, enfocando-se o tipo mais importante: a distribuição normal. As distribuições normais serão utilizadas extensivamente nos capítulos seguintes.

A distribuição normal, cujo gráfico se apresenta em forma de sino, pode ser descrita algebricamente por uma equação, cuja complexidade, entretanto, força-nos a lançar mão de uma tabela de valores (Tabela A-2). Essa tabela dá áreas correspondentes a regiões específicas sob a curva de distribuição normal padronizada, que tem média 0 e desvio-padrão 1. Essas áreas correspondem a valores de probabilidades.

Neste capítulo trabalhamos com processos padronizados para aplicar a Tabela A-2 a uma grande diversidade de situações, inclusive as que envolvem distribuições normais não-padronizadas (com média diferente de 0 ou desvio-padrão diferente de 1). Esses processos envolvem a utilização do valor, ou escore, padronizado: $z = (x - \mu)/\sigma$.

Na Seção 5-5 apresentamos os seguintes pontos importantes relacionados com o teorema central do limite:

1. A distribuição de médias amostrais tende para uma distribuição normal na medida em que o tamanho n da amostra aumenta.
2. A média das médias amostrais é a média populacional μ.
3. O desvio-padrão das médias amostrais é $\sigma/\sqrt{n}$.

Na Seção 5-6 vimos que às vezes é possível aproximar uma distribuição binomial por uma distribuição normal. Se $np \geq 5$ e $nq \geq 5$, simultaneamente, a variável aleatória binomial x tem distribuição aproximadamente normal, com média e desvio-padrão dados por $\mu = np$ e $\sigma = \sqrt{npq}$. Como a distribuição binomial de probabilidades se refere a dados discretos, enquanto a distribuição normal se refere a dados contínuos, aplicamos a correção de continuidade, que deve ser usada na aproximação da binomial pela normal.

Exercícios de Revisão

Nos Exercícios 1-4, suponha que os homens tenham alturas distribuídas normalmente com média de 69,0 in. e desvio-padrão de 2,8 in. (com base em dados do National Health Survey dos EUA).

1. Uma pessoa planeja abrir uma loja de roupas para homens. Para minimizar os custos iniciais, não manterá estoques de ternos nem para os 5% mais altos, nem para os 5% mais baixos. Determine as alturas mínima e máxima dos homens para os quais será mantido estoque de ternos.
2. O Beanstalk Club é uma organização social para pessoas de porte elevado. Os homens devem ter pelo menos 74 in. de altura para serem admitidos. Qual a percentagem dos homens que podem candidatar-se ao Beanstalk Club?
3. Se um homem é escolhido aleatoriamente, qual é a probabilidade de ele satisfazer as exigências de altura do Departamento de Polícia de Newport, que impõe altura entre 62,0 in. e 76,0 in. para os homens?
4. Selecionados aleatoriamente 45 homens, qual é a probabilidade de sua altura média estar entre 70,0 in. e 71,0 in.?
5. A revista *Entertainment Report* lança um *sweepstake* como parte de uma campanha para angariar novos assinantes. No passado, 26% que receberam material para adesão ao *sweepstake* acabaram entrando no concurso e assinando a revista (com base em dados reportados no *USA Today*). Estime a probabilidade de que, quando se envia o material de adesão a 500 residências selecionadas aleatoriamente, o número resultante de novos assinantes esteja entre 125 e 150, inclusive.
6. O Supermercado Gleason utiliza uma balança para pesar os produtos, e os erros são distribuídos normalmente com média de 0 oz e desvio-padrão de 1 oz. (Os erros podem ser positivos ou negativos.) Escolhe-se aleatoriamente e pesa-se um produto. Determine a probabilidade de que o erro
   a. esteja entre 0 e 1,25 oz
   b. seja superior a 0,50 oz
   c. seja superior a −1,08 oz
   d. esteja entre −0,50 oz e 1,50 oz
   e. esteja entre −1,00 oz e −0,25 oz.
7. As notas da parte de biologia do exame de admissão ao Medical College (EUA) têm distribuição normal com média de 8,0 e desvio-padrão de 2,6. Dentre os 600 candidatos que fizeram o exame, quantos podemos esperar que tenham nota entre 6,0 e 7,0?
8. No GRE (Graduate Record Examination — EUA) em economia, as notas têm distribuição normal com média 615 e desvio-padrão 107. Se o departamento de admissão de uma faculdade exige notas acima do 70.° percentil, determine a fronteira de separação.
9. Das residências americanas, 24% têm secretárias eletrônicas (com base em dados da U.S. Consumer Electronics Industry). Se uma campanha de propaganda por telefone envolve 2500 residências, estime a probabilidade de mais de 650 terem aquelas máquinas.
10. A Chemco Company fabrica pneus de automóveis cuja vida útil (em distância percorrida) tem distribuição normal com média de 35.600 milhas e desvio-padrão de 4275 milhas.
    a. Escolhido aleatoriamente um pneu, qual a probabilidade de durar 30.000 milhas?
    b. Escolhidos aleatoriamente 40 pneus, qual a probabilidade de suas vidas úteis terem média superior a 35.000 milhas?
    c. Se o fabricante deseja garantir os pneus de modo que apenas 3% deles precisem ser substituídos antes do número garantido de milhas, por quantas milhas os pneus devem ser garantidos?

Exercícios Cumulativos de Revisão

1. De acordo com dados da American Medical Association (Associação Médica Americana), 10% das pessoas são canhotas.
   a. Escolhidas três pessoas aleatoriamente, determine a probabilidade de todas elas serem canhotas.
   b. Escolhidas três pessoas aleatoriamente, determine a probabilidade de pelo menos uma delas ser canhota.
   c. Por que não podemos resolver o problema da parte (b) utilizando a aproximação normal da binomial?
   d. Selecionados aleatoriamente grupos de 50 pessoas, qual é o número médio de canhotos nesses grupos?
   e. Selecionados aleatoriamente grupos de 50 pessoas, qual é o desvio-padrão dos números de canhotos nesses grupos?
   f. Seria incomum obter 8 canhotos em 50 pessoas selecionadas aleatoriamente? Por quê?
2. Os valores amostrais dados a seguir são os tempos (em milissegundos) que o toca-discos do autor levou para completar uma revolução. Os tempos foram cronometrados por um programa de diagnóstico.

199,7	200,0	200,1	200,1	200,1	200,3	200,3	200,3	200,3	200,3
200,3	200,3	200,4	200,4	200,4	200,4	200,4	200,4	200,4	200,4
200,5	200,5	200,5	200,5	200,5	200,5	200,5	200,5	200,5	200,6
200,6	200,6	200,6	200,6	200,6	200,7	200,8	201,1	201,2	201,2

   a. Calcule a média $\bar{x}$ dos tempos desta amostra.
   b. Calcule a mediana desses tempos.
   c. Calcule a moda desses tempos.
   d. Calcule o desvio-padrão s dessa amostra.
   e. Converta o tempo de 200,5 ms em um escore z.
   f. Ache a percentagem real dos valores amostrais superiores a 201,0 ms.
   g. Admitindo uma distribuição normal, determine a percentagem *populacional* dos valores superiores a 201,0 ms. Use os valores amostrais de $\bar{x}$ e de s como estimativas de μ e σ.
   h. As especificações exigem tempos entre 197,0 ms e 202,0 ms. Com base nos resultados amostrais obtidos, o toca-discos parece estar girando a uma velocidade aceitável?

Projeto para Computador

Neste capítulo abordamos o problema da determinação da percentagem de mulheres com alturas entre 64,5 in. e 72 in., conforme exigência da espaçonave russa Soyuz. Vimos que 35,93% das mulheres têm alturas entre aqueles valores. A solução, dada na Seção 5-3, envolve cálculos teóricos baseados nas hipóteses de que as alturas das mulheres têm distribuição normal com média de 63,6 in. e desvio-padrão de 2,5 in. (com base em dados do National Health Service). Este projeto descreve um método de solução diferente, baseado na simulação seguinte: com auxílio de um computador ou de uma calculadora TI-83 geraremos 100 alturas de mulheres (de uma população distribuída normalmente com μ = 63,6 e σ = 2,5), e acharemos a percentagem das alturas simuladas que se situam entre 64,5 in. e 72 in. Descrevem-se a seguir os procedimentos para fazer essa simulação com STATDISK, Minitab e TI-83.

Para utilizar STATDISK: Selecionar `Data` da barra principal de menu, e escolher a opção de `Normal Generator`. Passar à geração de 100 valores com média 63,6 e desvio-padrão 2,5. (Utilizar a opção

Format para especificar 1 casa decimal.) Em seguida, ordenar os dados utilizando as opções Sample Editor e Format. Com esta lista ordenada, torna-se muito fácil contar o número de alturas entre 64,5 e 72. Dividir esse número por 100 para achar a percentagem das alturas simuladas adequadas à espaçonave russa Soyuz. Comparar o resultado com o valor teórico 35,93% achado na Seção 5-3.

Para utilizar Minitab: Selecionar as opções Calc, em seguida Random Data, em seguida Normal. Em seguida introduzir 100 como número de linhas, C1 para a coluna em que vão ser armazenados os dados, 63,6 para o valor da média e 2,5 para o valor do desvio-padrão. Selecionar então a opção Manip, em seguida Sort. A ordenação deve ser feita pela coluna C1 e a coluna ordenada armazenada em C1. Examinar os valores na coluna C1 e contar o número de alturas entre 64,5 e 72; dividir esse número por 100 para achar a percentagem de alturas simuladas adequadas à espaçonave russa Soyuz. Comparar o resultado com o valor teórico 35,93% encontrado na Seção 5-3.

Para utilizar uma calculadora TI-83: Acionar MATH, selecionar PRB, introduzir randNorm(63.6, 2.5, 100) para gerar 100 valores de uma população distribuída normalmente com $\mu = 63{,}6$ e $\sigma = 2{,}5$. Introduzir STO→L1 para armazenar os dados na lista L1. Acionar então STAT, introduzir SortA(L1) para ordenar os dados. Examinar os valores na lista L1 para achar o número de valores entre 64,5 e 72, dividir esse número por 100 para achar a percentagem de alturas simuladas adequadas à espaçonave Soyuz. Comparar o resultado com o valor teórico 35.93% encontrado na Seção 5-3.

DOS DADOS À DECISÃO

Como Desmascarar Quem Usa Moedas Falsas em Máquinas Automáticas

Os operadores de máquinas automáticas estão continuamente imaginando estratégias para combater o uso de moedas e notas falsas. Uma máquina coleta notas de dólar em uma bandeja e fornece troco para uso em outras máquinas. Uma vareta passa pelo centro da bandeja receptora e puxa a nota de dólar para uma caixa. Uma pessoa desonesta, porém criativa, descobriu que, usando uma nota de dólar com uma ranhura no centro, a máquina forneceria o troco mas a nota não seria retirada. Essa pessoa poderia então retirar todo o troco de uma máquina, passando em seguida a outra máquina.

O uso de fichas falsas também inferniza a vida dos operadores de máquinas automáticas. Vamos procurar desenvolver uma estratégia para minimizar as perdas decorrentes do uso de moedas falsas de 25 centavos. No Conjunto de Dados 13 do Apêndice B, temos uma amostra aleatória de pesos (em gramas) de moedas legais e legítimas de 25 centavos. Com base nessa amostra, suponhamos que, para a população de todas as moedas de 25 centavos, a distribuição seja normal com peso médio de 5,622 g e desvio-padrão de 0,068 g. Suponhamos também que tenha sido recolhido um lote de moedas falsas em máquinas automáticas no centro de Dallas. Uma análise dessas moedas falsificadas mostra que elas têm peso médio de 5,450 g, com desvio-padrão de 0,112 g. As máquinas automáticas são construídas de maneira a aceitar ou rejeitar as moedas de acordo com o peso. Os limites para aceitação das moedas podem ser modificados, mas a máquina normalmente aceita qualquer moeda com peso entre 5,500 g e 5,744 g. Determine a percentagem de moedas legítimas e de moedas falsas aceitas pela máquina.

Eis um dilema: Se restringimos demasiadamente os limites de peso das moedas, reduzimos a proporção de moedas legítimas aceitas e os clientes serão prejudicados porque muitas de suas moedas legítimas não serão aceitas. Se não restringimos suficientemente os limites de peso das moedas, muitas moedas falsas serão aceitas, ocasionando prejuízo. Quais as percentagens de moedas legítimas e de moedas falsas aceitas, se o limite inferior de peso é alterado para 5,550 g? Experimente com diferentes conjuntos de pesos e identifique limites que pareçam razoáveis, no sentido de que a máquina não aceite demasiadas moedas falsas, mas também aceite uma proporção razoável de moedas legítimas. Seria interessante construir um gráfico de distribuição normal razoavelmente preciso para as moedas boas, e outro para as moedas falsas; utilize a mesma escala horizontal e coloque um gráfico acima do outro, de modo que os números da escala horizontal fiquem alinhados. Com esses gráficos, queremos achar limites que incluam o máximo de moedas boas e o mínimo de moedas falsas. (Não há necessariamente uma resposta única; até certo ponto, as escolhas dependem de julgamentos subjetivos.) Redija um breve relatório que inclua os arranjos recomendados, juntamente com razões específicas para sua escolha.

ATIVIDADES DE GRUPO

1. *Atividade Extraclasse*: Utilize o cronometrador de reação seguindo as instruções a seguir. Colete tempos de reação para uma amostra de pelo menos 40 pessoas diferentes selecionadas de um grupo homogêneo, como, por exemplo, alunos destros de uma faculdade. Para cada pessoa, meça o tempo de reação para cada mão. Construa um histograma para os tempos da mão direita, outro para os tempos da mão esquerda. Com base nas formas dos histogramas, os dois conjuntos de tempo aparentam ter distribuição normal? Calcule a média e o desvio-padrão para cada um dos conjuntos e compare os dois conjuntos de resultados. Utilizando a média amostral $\bar{x}$ referente à mão direita como estimativa da média populacional μ, e o desvio-padrão amostral s referente à mão direita como estimativa do desvio-padrão populacional σ, determine os quartis Q_1, Q_2 e Q_3. Repita o processo para a mão esquerda. Se vamos utilizar os tempos de reação da mão direita para selecionar candidatos a emprego, qual é o tempo limite que separa os 5% mais rápidos dos 95% mais lentos?

ATIVIDADES DE GRUPO (Continuação)

Instruções para Uso do Cronometrador de Tempo de Reação

a. Corte o cronometrador ao longo da linha tracejada.
b. Peça à pessoa que mantenha em posição horizontal o polegar e o indicador; esses dedos devem ficar separados por uma distância igual à largura do cronometrador de tempo.
c. Segure o cronometrador de reação de modo que a margem inferior fique logo acima do polegar e do indicador da pessoa.
d. Peça à pessoa que agarre o cronometrador o mais rápido possível, soltando-o então após alguns segundos.
e. Registre o tempo de reação (em segundos) correspondente ao ponto em que a pessoa o apanha.

2. *Atividade em classe:* Use o quadro-negro para construir dois gráficos ramo-e-folhas, com ramos 0, 1, 2, ..., 9. Para o gráfico ramo-e-folhas à esquerda, cada estudante deve registrar os quatro últimos algarismos do número da carteira de identidade. Para o ramo-e-folhas à direita, cada estudante deve calcular a média $\bar{x}$ dos mesmos quatro algarismos, arredondar o resultado para o inteiro mais próximo, e introduzi-lo no gráfico ramo-e-folhas. Formam-se então grupo de três ou quatro para analisar os resultados e mostrar como eles ilustram o teorema central do limite.
3. *Atividade em classe:* Divida em grupos de três ou quatro estudantes. Para cada membro de um grupo, verifique o valor total das moedas em seu poder. Em seguida, ache a média do grupo. Partilhe os valores individuais e o valor médio com todos dos outros grupos. Qual é a distribuição dos valores individuais? Qual é a distribuição das médias dos grupos? Os resultados ilustram o teorema central do limite? Como?
4. *Atividade em classe:* Divida em grupos de três ou quatro estudantes. Utilizando uma moeda para simular nascimentos, cada membro de um grupo deve simular 25 nascimentos e registrar o número simulado de meninas. Combine todos os resultados no grupo e registre n = número total de nascimentos e x = número de meninas. Para grupos de n nascimentos, calcule a média e o desvio-padrão do número de meninas. O resultado simulado é usual ou raro? Por quê?

CRONOMETRADOR DE REAÇÃO

−0,204
−0,202
−0,199
−0,196
−0,194
−0,191
−0,188
−0,185
−0,183
−0,180
−0,177
−0,174
−0,171
−0,168
−0,165
−0,161
−0,158
−0,155
−0,151
−0,148
−0,144
−0,140
−0,137
−0,133
−0,129
−0,125
−0,121
−0,116
−0,112
−0,107
−0,102
−0,097
−0,091
−0,085
−0,079
−0,072
−0,065
−0,056
−0,046
−0,032

entrevista

David Burn

David Burn é o diretor do Statistical Services at Consumers Union (Serviço de Estatística da Sociedade dos Consumidores). A Sociedade dos Consumidores, que publica a revista *Consumer Reports*, é uma organização independente, não-lucrativa, de atendimento ao consumidor. David Burn foi, anteriormente, estatístico sênior na Minitab, Inc.

Nota do autor: O autor encontrou David Burn e os outros estatísticos na Sociedade dos Consumidores: M. Edna Derderian, Keith Newsom-Stewart, Martin Romm e Jed Schneider. O autor percorreu toda a instalação e constatou que a Sociedade dos Consumidores planeja, executa e analisa meticulosamente experimentos objetivos.

Qual o papel da estatística na avaliação de um produto?

O papel fundamental da estatística aqui na Sociedade dos Consumidores é garantir que os estudos que efetuamos sejam válidos, confiáveis e compreensíveis. Procuramos significância nos resultados, no sentido de que, se concluímos que o produto A é melhor do que o produto B, tal conclusão se baseia em uma teoria estatística boa e sólida. Mas temos sempre em mente que a significância estatística não ofusca a significância prática. Se há uma significância estatística de 0,01 na diferença entre dois alto-falantes, mas se há uma significância prática de 0,1 na maneira como o ouvido humano ouve esses alto-falantes, então a significância estatística não tem sentido. Esta é a maneira como o artigo seria escrito. Poderíamos dizer que a qualidade dos alto-falantes testados é tão boa que realmente não há diferença, e poderíamos comprar com base apenas no preço.

Quem trabalha em um projeto típico de avaliação?

Trabalhamos com diferentes departamentos que congregam funcionários com grande diversidade de formação. Tipicamente, o líder de um projeto provém do departamento ao qual o projeto pertence. Um engenheiro eletrônico pode liderar um projeto de estudo para CD players. Outros membros do conjunto tipicamente incluem um estatístico, um analista de mercado, um redator e um editor.

Que faz esse estatístico?

Para o estatístico, o trabalho começa realmente quando recebemos o relatório de pesquisa de mercado e vemos o que o consumidor realmente quer saber. A decisão sobre o que testar é tomada pelo grupo, com participação de todos os membros, inclusive o estatístico. Planejamos um experimento e um método de amostragem. Trabalhamos com o líder do projeto na fase de teste, na amostragem, e outros aspectos do planejamento experimental. O estatístico começa a análise dos dados de teste utilizando métodos exploratórios de dados com gráficos e mensurações numéricas. Exploramos os dados porque, na realidade, não sabemos o que vamos obter. Nesta altura, estamos utilizando instrumentos como diagramas de dispersão, *boxplots* e histogramas e busca de características como valores extremos (*outliers*). Passamos então a análises que podem mostrar que alguns produtos são diferentes dos outros. Finalmente, utilizamos métodos estatísticos para estabelecer escores globais que chamamos *ratings*. Um relatório estatístico é escrito e verificado por um dos outros estatísticos. Discutimos o tipo de análise para certificarmo-nos de que é a melhor escolha, verificamos se tudo foi feito corretamente e se os resultados são válidos.

Como assegurar objetividade?

Não aceitamos qualquer tipo de propaganda externa. Não representamos quaisquer organizações ou companhias: representamos o consumidor. Compramos todos os nossos produtos em um mercado aberto, tal como o faria um consumidor. Temos compradores anônimos em todos os Estados Unidos, nas principais áreas; o planejamento estatístico para o plano de amostragem pode exigir que compremos em San Francisco, Flórida, Nova York e Texas — o mesmo produto, simultaneamente. Dessa forma obtemos uma amostra representativa do produto que o fabricante está produzindo.

Usa computadores em seu trabalho?

Utilizamos computadores atualizados e uma diversidade de pacotes estatísticos e outros.

Recomendaria um curso de estatística para os alunos de faculdade?

Recomendaria vários cursos de estatística. Sem dúvida, os cursos de estatística elementar são importantes para uma visão geral dos métodos paramétricos e não-paramétricos. Além disso, os modelos lineares, a amostragem e o planejamento de experimentos proporcionam a profundidade e o conhecimento necessários para a revisão crítica de muitos tópicos que envolvem pesquisas estatísticas. Além disso, a capacidade de comunicar suas idéias é essencial. Como trabalhamos em vários projetos simultaneamente, a organização pessoal e o gerenciamento de projetos são fundamentais.

Rayane Andrade

6 Triola

Estimativas e Tamanho de Amostras

6-1 Aspectos Gerais

Identificam-se os objetivos do capítulo. Este capítulo e os próximos apresentam tópicos de *estatística inferencial*, que consiste em métodos de utilização de dados amostrais para tirar conclusões sobre parâmetros populacionais. O capítulo focaliza métodos de estimativa de médias, proporções ou variâncias populacionais; e método para determinar o tamanho da amostra necessário para aquelas estimativas.

6-2 Estimativa de uma Média Populacional: Grandes Amostras

Obtém-se uma aproximação de uma média populacional por um valor único (chamado estimativa pontual) e um intervalo de confiança, indicando-se um método para determinar o tamanho da amostra necessário para estimar uma média populacional. Nesta seção trabalhamos apenas com grandes amostras ($n > 30$).

6-3 Estimativa de uma Média Populacional: Pequenas Amostras

Nesta seção utilizam-se estatísticas extraídas de pequenas amostras ($n \leq 30$) para estimar valores de médias populacionais. Discutem-se as estimativas pontuais e os intervalos de confiança, e introduz-se a distribuição t de Student.

6-4 Estimativa de uma Proporção Populacional

Estima-se o valor de uma proporção populacional por meio de uma estimativa pontual e um intervalo de confiança. Descrevem-se processos para determinar o tamanho da amostra necessária para estimar uma proporção populacional. Esta seção apresenta o método comumente utilizado por pesquisadores para determinar quantas pessoas devem ser entrevistadas em várias pesquisas de opinião.

6-5 Estimativa de uma Variância Populacional

O valor de uma variância populacional é aproximado por uma estimativa pontual e um intervalo de confiança; apresenta-se um método para determinar o tamanho da amostra. Introduz-se a distribuição qui-quadrado.

Problema do Capítulo

A temperatura média do corpo humano é realmente 98,6°F?

A Tabela 6-1 relaciona 106 temperaturas do corpo humano (do Conjunto de Dados 2 do Apêndice B) obtidas por pesquisadores da Universidade de Maryland. Utilizando os métodos descritos no Capítulo 2, podemos obter as seguintes características importantes do conjunto de dados amostrais:

- Como se pode ver por um histograma, a distribuição dos dados tem aproximadamente a forma de um sino.
- A média é $\bar{x} = 98{,}20°F$
- O desvio-padrão é $s = 0{,}62°F$
- O tamanho da amostra é $n = 106$.

A maioria crê que a temperatura média do corpo humano seja de 98,6°F, mas os dados da Tabela 6-1 parecem sugerir uma média efetiva de 98,20°F. Sabemos que as amostras tendem a variar, de forma que talvez a verdadeira temperatura média seja 98,6°F e a média amostral $\bar{x} = 98{,}20°F$ seja o resultado de uma flutuação aleatória. Mas também a média amostral de 98,20°F pode ser correta, estando errado o valor comumente aceito de 98,6°F.

Com base em uma análise dos dados amostrais da Tabela 6-1, veremos se a temperatura média do corpo humano é ou não 98,6°F.

TABELA 6-1 Temperaturas do Corpo de 106 Adultos Sadios

98,6	98,6	98,0	98,0	99,0	98,4	98,4	98,4	98,4	98,6
98,6	98,8	98,6	97,0	97,0	98,8	97,6	97,7	98,8	98,0
98,0	98,3	98,5	97,3	98,7	97,4	98,9	98,6	99,5	97,5
97,3	97,6	98,2	99,6	98,7	99,4	98,2	98,0	98,6	98,6
97,2	98,4	98,6	98,2	98,0	97,8	98,0	98,4	98,6	98,6
97,8	99,0	96,5	97,6	98k,	96,9	97,6	97,1	97,9	98,4
97,3	98,0	97,5	97,6	98,2	98,5	98,8	98,7	97,8	98,0
97,1	97,4	99,4	98,4	98,6	98,4	98,5	98,6	983	98,7
98,8	99,1	98,6	97,9	98,8	98,0	98,7	98,5	98,9	98,4
98,6	97,1	97,9	98,8	98,7	97,6	98,2	99,2	97,8	98,0
98,4	97,8	98,4	97,4	98,0	97,0				

Temperaturas amostrais obtidas pelos Drs. Philip Mackowiak, Steven Wasserman e Myron Levine, pesquisadores da Universidade de Maryland.

6-1 Aspectos Gerais

No Capítulo 2 vimos que a *estatística descritiva* tem por objetivo resumir ou descrever características importantes de dados populacionais conhecidos; com a *inferência estatística*, entretanto, utilizamos os dados amostrais para fazer inferências (ou generalizações) sobre uma população. As duas principais aplicações da estatística inferencial envolvem a utilização de dados amostrais (1) para estimar o valor de um parâmetro populacional e (2) para formular uma conclusão sobre uma população. Este capítulo apresenta métodos de estimação de valores dos seguintes parâmetros populacionais: médias, proporções e variâncias. Dão-se também métodos para determinar o tamanho da amostra necessário para estimar esses parâmetros.

Ao prosseguirmos com os métodos de utilização de dados amostrais para fazer inferências sobre populações, convém recordar um ponto de extrema importância mencionado no Capítulo 1:

> **Dados coletados de forma imprecisa ou descuidada podem ser totalmente destituídos de valor, mesmo que a amostra seja suficientemente grande.**

Os métodos utilizados aqui e nos capítulos que seguem exigem processos bem fundamentados de amostragem. Por exemplo, um noticiário de TV poderia encorajar os espectadores a discar para um número grátis para registrar suas preferências entre candidatos presidenciais. Poderíamos então aplicar os métodos deste capítulo para fazer inferências com base nesses resultados amostrais que, entretanto, seriam totalmente sem valor (*hogwash*). Como as pessoas que responderam foram autoselecionadas, a amostra apresenta forte potencial de tendenciosidade, e qualquer inferência feita sobre a população carece de validade estatística.

6-2 Estimativa de uma Média Populacional: Grandes Amostras

Consideremos as 106 temperaturas dadas na Tabela 6-1 no início deste capítulo. Com base nesses valores amostrais, pretendemos estimar a média de *todas* as temperaturas. Poderíamos utilizar uma estatística como a mediana amostral, o ponto médio, ou a moda como estimativa dessa média populacional μ, mas a média amostral $\bar{x}$ em geral é a melhor estimativa de uma média populacional. A escolha de $\bar{x}$ se baseia em um estudo e uma análise cuidadosa das distribuições das diferentes estatísticas que poderiam ser usadas como estimadores.

DEFINIÇÕES

Um **estimador** é uma estatística amostral (como a média amostral $\bar{x}$) utilizada para obter uma aproximação de um parâmetro populacional. Uma **estimativa** é um valor específico, ou um intervalo de valores, usado para aproximar um parâmetro populacional.

Por exemplo, com base nos dados da Tabela 6-1, poderíamos utilizar o *estimador* $\bar{x}$ para concluir que a *estimativa* da temperatura média do corpo de todos os adultos sadios é 98,20°F.

Estimativa dos Tamanhos Populacionais na Vida Selvagem

O National Forest Management protege as espécies ameaçadas, inclusive a coruja pintada do norte, resultando daí que as madeireiras foram impedidas de derrubar árvores em vastas regiões do noroeste do Pacífico. Biólogos e estatísticos foram encarregados de analisar o problema e concluíram que as taxas de sobrevivência e os tamanhos populacionais estavam diminuindo no caso das fêmeas das corujas, que desempenham papel importante na sobrevivência da espécie. Aqueles cientistas estudaram também o salmão, nos rios Snake e Columbia no estado de Washington, e os pingüins da Nova Zelândia. No artigo "Sampling Wildlife Populations" (Amostragem de Populações da Vida Selvagem) (*Chance*, Vol.9, Nº 2), os autores Bryan Manly e Lyman McDonald comentam que, nesses estudos, "os biólogos ganham ao aperfeiçoarem sua habilidade em modelamento, que é uma das principais características da boa estatística. E os estatísticos ganham pelo fato de serem iniciados na realidade de problemas por biólogos, que sabem quais são os pontos cruciais."

Há duas razões importantes que explicam por que uma média amostral é um melhor estimador de uma média populacional μ do que quaisquer outros estimadores, como a mediana ou a moda.

1. Para muitas populações, a distribuição de médias amostrais $\bar{x}$ tende a ser mais consistente (apresentar *menor variação*) do que as distribuições de outras estatísticas amostrais. (Isto é, se utilizamos médias amostrais para estimar a média populacional μ, essas médias amostrais terão menor desvio-padrão do que outras estatísticas amostrais, tais como a mediana ou a moda.)
2. Para todas as populações, dizemos que a média amostral $\bar{x}$ é um **estimador não-tendencioso** da média populacional μ, o que significa que a distribuição de médias amostrais tende a centrar-se em torno da média populacional μ. (Isto é, as médias amostrais não tendem a sobreestimar nem a subestimar sistematicamente o valor de μ. Ao contrário, tendem para o valor-alvo que é o próprio valor de μ.)

Por estas razões, utilizaremos a média amostral $\bar{x}$ como a melhor estimativa da média populacional μ. Como a média amostral $\bar{x}$ é um valor único que corresponde a um ponto na escala numérica, chamamo-la uma *estimativa pontual*.

DEFINIÇÃO

Uma **estimativa pontual** é um valor (ou ponto) único usado para aproximar um parâmetro populacional.

A média amostral $\bar{x}$ é a melhor estimativa pontual da média populacional μ.

EXEMPLO Com a amostra de temperaturas do corpo humano dada na Tabela 6-1, determine a melhor estimativa pontual da média populacional μ das temperaturas de todos os corpos.

SOLUÇÃO A média amostral $\bar{x}$ é a melhor estimativa pontual da média populacional μ, e para os dados amostrais da Tabela 6-1, temos $\bar{x}$ = 98,20°F. Com base nesses valores amostrais particulares, a melhor estimativa pontual da média populacional μ de todas as temperaturas é, pois, 98,20°F.

Por que Necessitamos de Intervalos de Confiança? No exemplo precedente vimos que 98,20°F era nossa *melhor* estimativa pontual da média populacional μ, mas não tínhamos qualquer indicação sobre quão boa era essa estimativa. Se conhecêssemos apenas as quatro primeiras temperaturas de 98,6, 98,6, 98,0 e 98,0, nossa melhor estimativa pontual de μ seria sua média ($\bar{x}$ = 98,30°F), mas esta estimativa pontual não seria tão boa, porque se baseia em uma amostra demasiadamente pequena. Assim é que os estatísticos desenvolveram outro tipo de estimativa que, efetivamente, indica quão boa é uma estimativa pontual. Essa estimativa, chamada intervalo de confiança ou estimativa intervalar, consiste em uma amplitude (ou um intervalo) de valores, em lugar de um valor único.

DEFINIÇÃO

Um **intervalo de confiança** (ou **estimativa intervalar**) é uma amplitude (ou um intervalo) de valores que tem probabilidade de conter o verdadeiro valor da população.

Um *intervalo de confiança* está associado a um grau de confiança que é uma medida da nossa certeza de que o intervalo contém o parâmetro populacional. A definição de grau de confiança utiliza α para descrever uma probabilidade que corresponde a uma área. Veja a Figura 6-1, onde a probabilidade α está dividida igualmente entre duas regiões extremas sombreadas (geralmente chamadas caudas) na distribuição normal padronizada. (O papel de $z_{\alpha/2}$ será discutido mais adiante; por ora, notemos apenas que α está dividido igualmente entre as duas caudas.)

DEFINIÇÃO

O **grau de confiança** é a probabilidade $1 - \alpha$ (comumente expressa como o valor percentual equivalente) de o intervalo de confiança conter o verdadeiro valor do parâmetro populacional. (O grau de confiança é também chamado **nível de confiança**, ou **coeficiente de confiança**.)

São escolhas comuns para o *grau de confiança*: 90% (com α = 0,10), 95% (com α = 0,05) e 99% (com α = 0,01). A mais comum é a opção 95%, porque proporciona bom equilíbrio en-

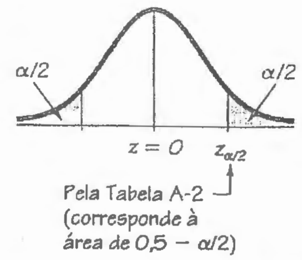


Fig. 6-1 A Distribuição normal padronizada: o valor crítico $z_{\alpha/2}$.

tre a precisão (refletida na amplitude do intervalo de confiança) e a confiabilidade (expressa pelo grau de confiança).

Eis um exemplo de intervalo de confiança baseado nos dados amostrais de 106 temperaturas da Tabela 6-1:

A estimativa intervalar, com grau de 95% (ou 0,95) de confiança da média populacional μ, é 98,08°F $< \mu <$ 98,32°F.

Note que esta estimativa consiste em um intervalo e está associada a um grau de confiança. Interpretamos este intervalo de confiança como segue: Se fôssemos selecionar muitas amostras diferentes de tamanho $n = 106$ da população de todos os cidadãos sadios, e construíssemos um intervalo de 95% de confiança análogo para cada amostra, a longo prazo, 95% desses intervalos conteriam efetivamente a média populacional μ. Devemos ter em mente que μ é um valor fixo, e não uma variável aleatória; portanto, é *errado* dizer que há 95% de chance de μ estar no intervalo. Qualquer intervalo de confiança contém, ou não contém μ, e como μ é fixo, não existe a *probabilidade* de μ estar em um intervalo.

Pelo Teorema Central do Limite, sabemos que as médias amostrais $\bar{x}$ tendem a distribuir-se normalmente, como na Figura 6-1. As médias amostrais apresentam uma chance relativamente pequena de estar em uma das caudas extremas da Figura 6-1. Denotando por $\alpha/2$ a área sombreada de cada cauda, vemos que há uma probabilidade total α de a média amostral estar em uma das duas caudas. Pela regra do complemento (do Capítulo 3), decorre que há uma probabilidade de $1 - \alpha$ de uma média amostral estar na região não-sombreada da Figura 6-1. O escore z que separa a região da cauda direita é denotado comumente por $z_{\alpha/2}$, e é chamado *valor crítico* porque está na fronteira que separa as médias amostrais passíveis de ocorrerem, das médias amostrais que provavelmente não ocorrerão.

Notação para o Valor Crítico

$z_{\alpha/2}$ é o valor positivo de z que está na fronteira vertical de uma área de $\alpha/2$ na cauda direita da distribuição normal padronizada. (O valor $-z_{\alpha/2}$ está na fronteira vertical da área $\alpha/2$ na cauda esquerda.)

DEFINIÇÃO

Um **valor crítico** é o número na fronteira que separa os valores das estatísticas amostrais prováveis de ocorrerem, dos valores que têm pouca chance de ocorrer. O número $z_{\alpha/2}$ é um valor crítico que é um escore z com a propriedade de separar uma área de $\alpha/2$ na cauda direita da distribuição normal padronizada. (Há uma área de $1 - \alpha$ entre as fronteiras verticais em $-z_{\alpha/2}$ e $z_{\alpha/2}$.)

EXEMPLO Ache o valor crítico $z_{\alpha/2}$ correspondente a um grau de confiança de 95%.

SOLUÇÃO Um grau de confiança de 95% corresponde a $\alpha = 0{,}05$. Veja a Figura 6-2, onde mostramos que a área em cada cauda sombreada é $\alpha/2 = 0{,}025$. Obtemos $z_{\alpha/2} = 1{,}96$, notando que a região à sua direita (e delimitada pela média $z = 0$) deve ser $0{,}5 - 0{,}025$, ou 0,475. Recorrendo então à Tabela A-2, vemos que a área de 0,4750 (encontrada no *corpo* da tabela) corresponde exatamente ao valor $z = 1{,}96$. O valor crítico para um grau de confiança de 95% é, pois, $z_{\alpha/2} = 1{,}96$.

O exemplo precedente mostrou que um grau de confiança de 95% resulta em um valor crítico $z_{\alpha/2} = 1{,}96$. Este é o valor crítico mais comum, e consta da tabela que segue, juntamente com outros dois valores também comuns.

Grau de Confiança	α	Valor Crítico $z_{\alpha/2}$
90%	0,10	1,645
95%	0,05	1,96
99%	0,01	2,575

Quando coletamos um conjunto de dados amostrais, como o conjunto de 106 temperaturas da Tabela 6-1, podemos calcular a média amostral $\bar{x}$; essa média amostral é tipicamente diferente da média populacional μ. A diferença entre a média amostral e a média populacional pode ser encarada como um erro. Na Seção 5-5 vimos que $\sigma/\sqrt{n}$ é o desvio-padrão das médias amostrais. Com $\sigma/\sqrt{n}$ e a notação $z_{\alpha/2}$ definimos como segue a margem de erro E.

DEFINIÇÃO

Quando utilizamos dados amostrais para estimar uma média populacional μ, a **margem de erro**, denotada por E, é a diferença máxima provável (com probabilidade $1 - \alpha$) entre a média amostral observada $\bar{x}$ e a verdadeira média populacional μ. A margem de erro E é chamada também **erro máximo da estimativa** e pode ser obtida multiplicando-se o valor crítico pelo desvio-padrão das médias amostrais, conforme Fórmula 6-1.

Fórmula 6-1 $$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Dada a maneira como se definiu a margem de erro E, há uma probabilidade de $1 - \alpha$ de uma média amostral conter um erro (diferir da média populacional μ) não superior a E, e há uma probabilidade de α de a média amostral conter um erro superior a E. O cálculo da margem de erro E, tal como dada pela Fórmula 6-1, exige o conhecimento do desvio-padrão populacional σ, mas, na realidade, é raro conhecermos σ quando a média populacional μ não é conhecida. É comum o método de cálculo seguinte

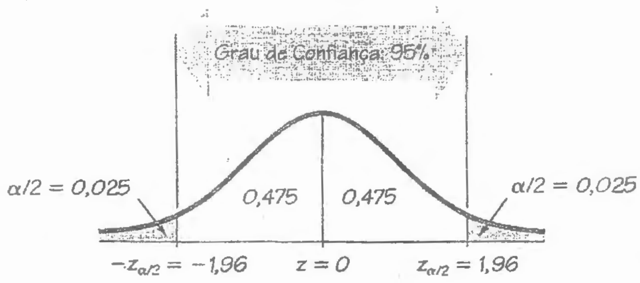


Fig. 6-2 Determinação de $z_{\alpha/2}$ para 95% de grau de confiança.

Cálculo de E Quando σ é Desconhecido

Se $n > 30$, podemos substituir σ na Fórmula 6-1 pelo desvio-padrão amostral s.
Se $n \leq 30$, a população deve ter distribuição normal, e devemos conhecer σ para aplicar a Fórmula 6-1. [Na próxima seção discutiremos um método alternativo para o cálculo da margem de erro E no caso de pequenas amostras ($n \leq 30$).]

Com base na definição da margem de erro E, podemos agora identificar o intervalo de confiança para a média populacional μ.

Intervalo de Confiança (ou Estimativa Intervalar) para a Média Populacional μ (Com Base em Grandes Amostras: $n > 30$)

$$\bar{x} - E < \mu < \bar{x} + E \quad \text{onde} \quad E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Outras formas equivalentes para o intervalo de confiança: $\mu = \bar{x} \pm E$ e $(\bar{x} - E, \bar{x} + E)$.

Os valores $\bar{x} - E$ e $\bar{x} + E$ são chamados **limites do intervalo de confiança.**

Passo a passo

Processo de Construção de um Intervalo de Confiança para μ (Com Base em uma Amostra Grande: $n > 30$)

1. Determinar o valor crítico $z_{\alpha/2}$ correspondente ao grau de confiança desejado. (Por exemplo, se o grau de confiança é 95%, o valor crítico é $z_{\alpha/2} = 1{,}96$.)
2. Calcular a margem de erro $E = z_{\alpha/2} \cdot \sigma/\sqrt{n}$. Se o desvio-padrão populacional σ não é conhecido, utilizar o desvio-padrão amostral s, desde que $n > 30$.
3. Com o valor calculado da margem de erro E e o valor da média amostral $\bar{x}$, calcular os valores de $\bar{x} - E$ e $\bar{x} + E$. Levar esses valores na expressão geral do intervalo de confiança:

$$\bar{x} - E < \mu < \bar{x} + E$$

ou

$$\mu = \bar{x} \pm E$$

ou

$$(\bar{x} - E, \quad \bar{x} + E)$$

$\bar{x} \pm 2 \cdot S_{\bar{x}}$

$S_{\bar{x}} = \frac{s}{\sqrt{n}}$

4. Arredondar os valores resultantes de acordo com a regra seguinte.

Regra de Arredondamento para Intervalos de Confiança Usados para Estimar μ

1. Quando utilizar o *conjunto original de dados* para construir um intervalo de confiança, arredondar os limites do intervalo para uma casa decimal além das usadas naquele conjunto.
2. Quando o conjunto original de dados é desconhecido, utilizando-se apenas as *estatísticas resumo* (n, $\bar{x}$, s), arredondar os limites do intervalo de confiança para o mesmo número de casas decimais utilizadas na média amostral.

O exemplo seguinte ilustra com clareza o processo relativamente simples (!) para construção de um intervalo de confiança. Os dados originais da Tabela 6-1 têm uma casa decimal, e as estatísticas resumo têm duas decimais; assim, os limites do intervalo de confiança serão arredondados para duas decimais.

EXEMPLO Para as temperaturas da Tabela 6-1, temos $n = 106$, $\bar{x} = 98{,}20$ e $s = 0{,}62$. Para um grau de confiança de 0,95, determine:

a. A margem de erro E
b. O intervalo de confiança para μ

SOLUÇÃO

a. O grau de confiança 0,95 implica que $\alpha = 0{,}05$, de forma que $z_{\alpha/2} = 1{,}96$, conforme exemplo precedente. Calcula-se como segue a margem de erro E com auxílio da Fórmula 6-1. (Note que, embora σ seja desconhecido, podemos utilizar $s = 0{,}62$ em lugar de σ porque $n > 30$.)

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 1{,}96 \cdot \frac{0{,}62}{\sqrt{106}} = 0{,}12$$

b. Com $\bar{x} = 98{,}20$ e $E = 0{,}12$, construímos o intervalo de confiança:

$$\bar{x} - E < \mu < \bar{x} + E$$
$$98{,}20 - 0{,}12 < \mu < 98{,}20 + 0{,}12$$
$$98{,}08 < \mu < 98{,}32$$

[Este resultado poderia expressar-se também como $\mu = 98{,}20 \pm 0{,}12$ ou como $(98{,}08; 98{,}32)$.] Para a amostra de 106 temperaturas da Tabela 6-1, o intervalo de confiança para a média populacional μ é $98{,}08°F < \mu < 98{,}32°F$, e este intervalo tem um grau de confiança de 0,95. Isto significa que, se escolhêssemos muitas amostras diferentes de tamanho 106 e construíssemos intervalos de confiança como fizemos aqui, 95% deles conteriam o valor da média populacional μ.

Note que os limites 98,08°F e 98,32°F do intervalo de confiança não contêm o valor 98,6°F, que se admite ser a temperatura média do corpo humano. Com base nos dados da Tabela 6-1, parece muito pouco provável que o valor correto de μ seja 98,6°F.

Utilização de Calculadoras e Computadores para Intervalos de Confiança

STATDISK, Minitab e muitos outros programas de computador são elaborados para gerar intervalos de confiança do tipo ilustrado no exemplo precedente. Com o STATDISK, selecionamos primeiro `Analysis` da barra de menu principal, em seguida selecionamos `Confidence Intervals`, e passamos a introduzir os dados solicitados.

O Minitab calcula intervalos de confiança, mas exige uma listagem original dos dados brutos. O Minitab não faz cálculos utilizando apenas as estatísticas resumo n, $\bar{x}$ e s.

Pode-se usar também a calculadora TI-83 para gerar intervalos de confiança. Introduza os dados em L1 ou tenha disponíveis as estatísticas resumo; acione então a tecla STAT. Selecione `TESTS` e escolha `ZInterval` para um intervalo de confiança baseado na distribuição normal (z). A TI-83 exige uma entrada para o desvio-padrão populacional; devemos, pois, estar preparados para introduzir o valor de s se σ não for conhecido. Se utilizarmos os dados da Tabela 6-1, o mostrador da calculadora apresentará o intervalo de confiança neste formato: (98.082, 98.318).

Interpretação de um Intervalo de Confiança Devemos ter cuidado para interpretar corretamente os intervalos de confiança. Desde que utilizemos dados amostrais para achar os limites específicos $\bar{x} - E$ e $\bar{x} + E$, esses limites incluirão, ou não incluirão, a média populacional μ; e isto não podemos determinar sem conhecer o verdadeiro valor de μ. É incorreto afirmar que μ tem 95% de chance de estar entre os limites específicos de 98,08 e 98,32, porque μ é uma constante, e não uma variável aleatória. Ou μ está entre esses limites, ou não está; não há qualquer probabilidade em jogo. É correto dizermos que, a longo prazo, esses métodos darão intervalos de confiança que conterão μ em 95% dos casos.

Suponhamos que, no exemplo precedente, as temperaturas realmente provenham de uma população cuja verdadeira média é 98,25°F. Então o intervalo de confiança obtido dos dados amostrais conteria a média populacional, porque 98,2 está entre 98,08 e 98,32. A Figura 6-3 ilustra este fato. (Veja o primeiro intervalo de confiança no gráfico.)

Rationale: A idéia básica na construção de intervalos de confiança está alicerçada no teorema central do limite, que indica que, com grandes ($n > 30$) amostras, a distribuição das médias amostrais é aproximadamente normal com média μ e desvio-padrão $\sigma/\sqrt{n}$. O formato do intervalo de confiança é, na realidade, uma variação da equação

$$z = \frac{\bar{x} - \mu}{\dfrac{\sigma}{\sqrt{n}}}$$

Resolvendo esta equação em relação a μ, obtemos $\mu = \bar{x} - z\dfrac{\sigma}{\sqrt{n}}$. Com os valores positivos e negativos de z obtemos os limites do intervalo de confiança que estamos utilizando.

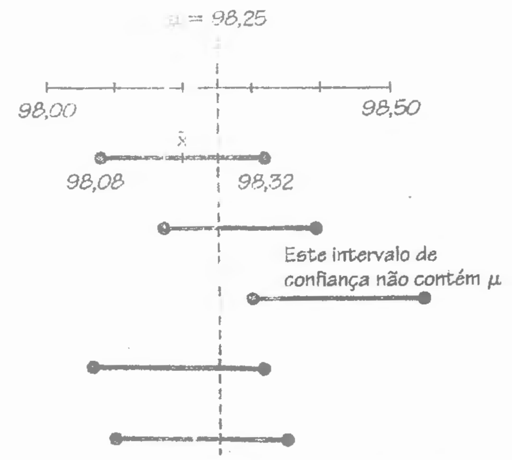


Fig. 6-3 Intervalos de confiança de diferentes amostras. O gráfico mostra vários intervalos de confiança; um deles não contém a média populacional μ. Ao nível de intervalos de confiança de 95%, espera-se que em 100 desses intervalos, 5 não contenham $\mu = 98,25$, enquanto os outros 95 contêm aquele valor.

Consideremos o caso específico de um grau de confiança de 95%, de forma que $\alpha = 0,05$ e $z_{\alpha/2} = 1,96$. Para este caso, há uma probabilidade de 0,05 de que uma média amostral diste mais de 1,96 desvios-padrão (ou $z_{\alpha/2}\sigma/\sqrt{n}$, que denotamos por E) da média populacional μ. Reciprocamente, há uma probabilidade de 0,95 de uma média amostral estar a menos de 1,96 desvios-padrão (ou $z_{\alpha/2}\sigma/\sqrt{n}$) de μ. (Veja a Figura 6-4.) Se a média amostral $\bar{x}$ está a menos de $z_{\alpha/2}\sigma/\sqrt{n}$ da média populacional μ, então μ deve estar entre $\bar{x} - z_{\alpha/2}\sigma/\sqrt{n}$ e $\bar{x} + z_{\alpha/2}\sigma/\sqrt{n}$; este fato se expressa no formato geral do intervalo de confiança (com $z_{\alpha/2}\sigma/\sqrt{n}$ denotado por E): $\bar{x} - E < \mu < \bar{x} + E$.

Determinação do Tamanho da Amostra

Até aqui, nesta seção, abordamos maneiras de achar estimativas pontuais e estimativas intervalares para uma média populacional μ. Tomamos por base dados amostrais conhecidos. Suponha, entretanto, que ainda não tenhamos coletado os dados. Como sabemos quantos elementos da população devem ser escolhidos? Suponha, por exemplo, que queiramos estimar a renda média de pessoas que concluíram um curso superior, no primeiro ano após a formatura. Quantas rendas devemos incluir em nossa amostra? A determinação do tamanho de uma amostra é um problema de grande importância, porque amostras desnecessariamente grandes acarretam desperdício de tempo e de dinheiro; e amostras demasiadamente pequenas podem levar a resultados não-confiáveis. Em muitos casos é possível determinar o tamanho mínimo de uma amostra para estimar determinado parâmetro, como a média populacional μ.

Partindo da expressão da margem de erro E (Fórmula 6-1) e resolvendo em relação ao tamanho n da amostra, obtemos:

Tamanho da Amostra para Estimar a Média μ	
Fórmula 6-2	$n = \left[\dfrac{z_{\alpha/2}\sigma}{E}\right]^2$

O tamanho da amostra deve ser um número inteiro, mas os cálculos para o tamanho amostral n quase nunca resultam em um número inteiro. Quando isto ocorre, devemos observar a seguin-

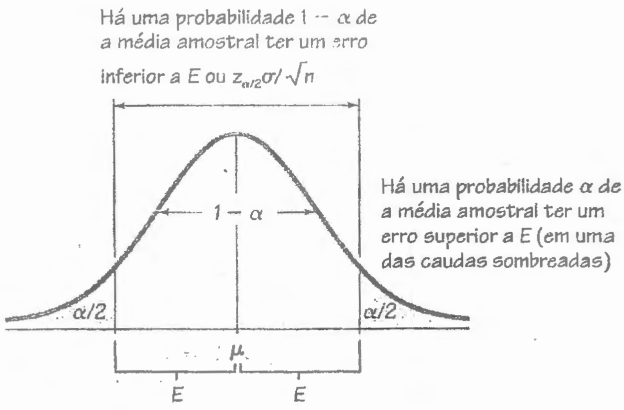


Fig. 6-4 Distribuição de médias amostrais.

te regra de arredondamento. (Baseia-se no princípio de que, quando o arredondamento se faz necessário, o tamanho da amostra deve ser arredondado *para cima*, a fim de ser ao menos adequadamente grande, em oposição a ligeiramente menor.)

Regra de Arredondamento para o Tamanho Amostral n

Ao determinar o tamanho da amostra n, se a aplicação da Fórmula 6-2 não conduz a um número inteiro, *aumente* sempre o valor de n para o próximo inteiro *maior*.

Com a Fórmula 6-2, pode-se determinar o tamanho da amostra necessária para dar resultados precisos, com um grau de confiança e uma margem de erro fixados. A fórmula deve ser usada quando conhecemos o valor de σ e queremos determinar o tamanho da amostra necessário para estabelecer, com um nível de confiança de $1 - \alpha$, o valor de μ a menos de $\pm E$. (Em geral não se conhece σ sem conhecer μ, mas σ pode ser conhecido de um estudo anterior, ou pode ser estimado com base em um estudo piloto ou uma regra empírica.) A existência de uma tal fórmula é algo notável, porque implica que o tamanho da amostra não depende do tamanho (N) da população; o tamanho da amostra depende do grau de confiança desejado, da margem de erro pretendida e do valor do desvio-padrão σ.

EXEMPLO Um economista deseja estimar a renda média para o primeiro ano de trabalho de um bacharel por uma faculdade, que teve a feliz idéia de fazer um curso de estatística. Quantos valores de renda devem ser tomados, se o economista deseja ter 95% de confiança em que a média amostral esteja a menos de \$500 da verdadeira média populacional? Suponha que saibamos, por um estudo prévio, que, para tais rendas, $\sigma = \$6250$.

SOLUÇÃO Queremos determinar o tamanho n da amostra, dado que $\alpha = 0{,}05$ (95% de confiança). Desejamos que a média amostral esteja a menos de \$500 da média populacional, de forma que $E = 500$. Supondo $\sigma = 6250$, aplicamos a Fórmula 6-2, obtendo

$$n = \left[\frac{z_{\alpha/2}\sigma}{E}\right]^2 = \left[\frac{1{,}96 \cdot 6250}{500}\right]^2 =$$

$$= 600{,}25 = 601 \quad \text{(arredondado para cima)}$$

Devemos, portanto, obter uma amostra de ao menos 601 rendas de primeiro ano, selecionadas aleatoriamente, de bacharéis de faculdades que tenham feito um curso de estatística. Com tal amostra teremos 95% de confiança em que a média amostral $\bar{x}$ difira em menos de \$500 da verdadeira média populacional μ.

Se nos contentamos com resultados menos precisos, utilizando uma margem de erro maior, como \$1000, o tamanho da amostra necessário cai para 150,0625, que arredondamos para 151. A duplicação da margem de erro faz com que o tamanho da amostra necessário se reduza a um quarto de seu valor original. Reciprocamente, reduzindo à metade a margem de erro, quadruplicamos o tamanho da amostra. O que tudo isto implica é que, se queremos resultados mais precisos, devemos aumentar substancialmente o tamanho da amostra. Como as grandes amostras em geral exigem tempo e dinheiro, devemos chegar a uma conciliação entre o tamanho da amostra e a margem de erro E.

Os Números de Série de Tanques Capturados Revelam o Tamanho da População

Durante a Segunda Guerra Mundial, os agentes do serviço secreto aliado tinham interesse em saber quantos tanques a Alemanha estava produzindo. As técnicas tradicionais de espionagem não deram resultados satisfatórios, mas os estatísticos obtiveram estimativas precisas analisando os números de série de tanques capturados. Um exemplo são os registros que mostram que a Alemanha produziu efetivamente 271 tanques em junho de 1941. A estimativa baseada em números de série foi de 244. Mas os métodos tradicionais de espionagem chegaram à cifra absurda de 1550 tanques. [Veja "An Empirical Approach to Economic Intelligence in World War II", por Ruggles and Brodie, *Journal of the American Statistical Association, Vol.* 42.]

E se σ não For Conhecido? A Fórmula 6-2 exige que se substitua por algum valor o desvio-padrão populacional σ, mas se este for desconhecido, devemos poder utilizar um valor preliminar obtido por processos como os que se seguem:

1. Utilizar a regra prática da Seção 2-5 para estimar o desvio-padrão da seguinte maneira: $\sigma \approx$ amplitude/4.
2. Realizar um estudo piloto, iniciando o processo de amostragem. Com base na primeira coleção de pelo menos 31 valores amostrais selecionados aleatoriamente, calcular o desvio-padrão amostral s e utilizá-lo em lugar de σ. Este valor pode ser refinado com a obtenção de mais dados amostrais.

Dos dois exemplos seguintes, o primeiro utiliza a regra prática para estimar σ, e o segundo emprega um estudo piloto que consiste nos dados amostrais encontrados no Apêndice B.

EXEMPLO Deseja-se estimar o preço médio de venda de um livro-texto para uma faculdade. Quantos exemplares devemos selecionar, para termos 95% de confiança de que a média amostral esteja a menos de \$2 da verdadeira média populacional μ?

SOLUÇÃO Procuramos o tamanho da amostra n, dado que $\alpha = 0{,}05$ (de 95% de confiança), de forma que $z_{\alpha/2} = 1{,}96$. Como queremos estar a menos de \$2, $E = 2$. Não conhecemos o desvio-padrão σ dos preços de venda de todos os livros, mas podemos estimar σ. Admitindo que os preços dos livros típicos de faculdades variem de \$10 a \$90, tem-se uma amplitude de \$80, de modo que

$$\sigma \approx \frac{\text{amplitude}}{4} = \frac{(90 - 10)}{4} = 20$$

Com $z_{\alpha/2} = 1{,}96$, $E = 2$ e $\sigma \approx 20$, aplicamos a Fórmula 6-2 como segue.

$$n = \left[\frac{z_{\alpha/2}\sigma}{E}\right]^2 = \left[\frac{1.96 \cdot 20}{2}\right]^2 = 384.16$$

$$= 385 \text{ (arredondado para cima)}$$

Devemos selecionar aleatoriamente 385 preços de venda de livros-texto e achar o valor da média amostral $\bar{x}$. Teremos 95% de confiança em que a média amostral está a menos de \$2 do verdadeiro preço médio de venda de todos os livros de faculdades.

Estimativa do Açúcar em Laranjas

No Flórida, a indústria cítrica utiliza amplamente métodos estatísticos. Uma aplicação específica envolve a maneira como os plantadores são pagos pelas laranjas usadas para fazer suco. Inicialmente, pesa-se um carregamento recebido na fábrica; em seguida, seleciona-se uma amostra aleatória de cerca de uma dúzia de laranjas. Essa amostra é pesada e espremida, medindo-se a quantidade de açúcar no suco. Com base nos resultados amostrais, estima-se a quantidade total de açúcar presente em todo o carregamento.

O pagamento do carregamento de laranjas se baseia na estimativa da quantidade de açúcar, porque as laranjas mais doces têm mais valor do que as menos doces, mesmo que as quantidades de suco sejam iguais.

EXEMPLO Se queremos estimar o peso médio do plástico descartado por residências em uma semana, quantas residências devemos selecionar aleatoriamente para termos 99% de confiança em que a média amostral esteja a menos de 0,250 lb da verdadeira média populacional?

SOLUÇÃO Procuremos o tamanho da amostra n, dado que $\alpha = 0{,}01$ (de 99% de confiança), de forma que $z_{\alpha/2} = 2{,}575$. Como desejamos ficar a menos de 0,250 lb da verdadeira média, $E = 0{,}250$. O valor do desvio-padrão populacional σ é desconhecido, mas podemos recorrer ao Conjunto de Dados 1 do Apêndice B, que compreende os pesos do plástico descartado por 62 residências. Utilizando essa amostra como estudo piloto, calculamos o valor do desvio-padrão, obtendo $s = 1{,}065$ lb; como a amostra é grande ($n > 30$), tomamos $s = 1{,}065$ como estimativa do desvio-padrão populacional σ, conforme mostrado a seguir.

$$n = \left[\frac{z_{\alpha/2}\sigma}{E}\right]^2 = \left[\frac{2{,}575 \cdot 1{,}065}{0{,}250}\right]^2 =$$

$$= 120{,}3 = 121 \text{ (arredondado para cima)}$$

Com base no desvio-padrão populacional estimado pelo Conjunto de Dados 1 do Apêndice B, devemos tomar uma amostra de pelo menos 121 residências, selecionadas aleatoriamente, para termos 99% de confiança em que a média amostral esteja a menos de 0,250 lb da verdadeira média populacional.

Esta seção abordou a construção de estimativas pontuais e estimativas intervalares para médias populacionais, e apresentou um método para determinar tamanhos de amostras necessários para estimar médias populacionais com o grau desejado de precisão. Todos os exemplos e exercícios desta seção envolvem grandes amostras ($n > 30$). Na seção seguinte iremos descrever processos a serem usados quando as amostras são pequenas ($n \leq 30$).

6-2 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-4, determine o valor crítico $z_{\alpha/2}$ que corresponde ao grau de confiança indicado.

1. 99% 2. 94% 3. 98% 4. 92%

Nos Exercícios 5-8, use o grau de confiança e os valores amostrais dados para achar (a) a margem de erro e (b) o intervalo de confiança para a média populacional μ.

5. Alturas de mulheres: 95% de confiança; $n = 50$, $\bar{x} = 63{,}4$ in., $s = 2{,}4$ in.
6. Médias de notas: 99% de confiança; $n = 75$, $\bar{x} = 2{,}76$, $s = 0{,}88$.
7. Notas de um teste: 90% de confiança; $n = 150$, $\bar{x} = 77{,}6$, $s = 14{,}2$
8. Salários da polícia: 92% de confiança; $n = 64$, $\bar{x} = \$23.228$, $s = \$8779$
9. Obtém-se uma amostra de 35 crânios de homens egípcios que viveram por volta de 1850 AC. Mede-se a largura máxima de cada crânio, obtendo-se $\bar{x} = 134{,}5$ mm e $s = 3{,}48$ mm (com base em dados de *Ancient Races of the Thebaid*, por Thomson e Randall-Maciver). Com esses dados amostrais, construa um intervalo de 95% de confiança para a média populacional μ.
10. Uma amostra consiste em 75 aparelhos de TV adquiridos há vários anos. Os tempos de substituição desses aparelhos têm média de 8,2 anos e desvio-padrão de 1,1 ano (com base em dados de "Getting Things Fixed," *Consumer Reports*). Construa um intervalo de 90% de confiança para o tempo médio de substituição de todos os aparelhos de TV daquela época. O resultado se aplica aos aparelhos de TV que estão sendo vendidos atualmente?
11. O National Center for Education Statistics pesquisou 4400 bacharéis de faculdades sobre o tempo que levaram para obter seus diplomas. A média é 5,5 anos e o desvio-padrão é 1,68 ano. Com base nesses dados amostrais, construa um intervalo de 99% de confiança para o tempo médio gasto por todos os bacharéis.
12. O Corpo de Fuzileiros Navais dos EUA está revendo seus pedidos de uniformes, porque tem sobra de uniformes para recrutas de porte elevado e falta de uniformes para recrutas mais baixos. Analisam-se então as alturas de 772 homens entre as idades de 18 e 24 anos. Esse grupo amostral acusa altura média de 69,7 in., com desvio-padrão de 2,8 in. (veja a publicação USDHEW 79-1659). Com esses dados amostrais, construa um intervalo de 95% de confiança para a altura média de todos os homens entre 18 e 24 anos.
13. O teste de QI padrão é planejado de modo que a média seja 100 e o desvio-padrão para adultos normais seja 15. Ache o tamanho da amostra necessário para estimar o QI médio dos instrutores de estatística. Queremos ter 98% de confiança em que nossa média amostral esteja a menos de 1,5 pontos de QI da verdadeira média. A média para esta população é obviamente superior a 100, e o desvio-padrão é provavelmente inferior a 15, porque se trata de um grupo com menor variação do que um grupo selecionado aleatoriamente da população geral; portanto, se tomamos $\sigma = 15$, estaremos sendo conservadores, por utilizarmos um valor que dará um tamanho de amostra no mínimo tão grande quanto necessário. Suponha $\sigma = 15$ e determine o tamanho da amostra necessário.
14. A Washington Vending Machine Company deve ajustar suas máquinas para que aceitem apenas moedas com pesos especificados. Vamos obter uma amostra de moedas de 25 centavos e pesá-las, a fim de determinar sua média. Quantas moedas devem ser selecionadas aleatoriamente e pesadas para termos 99% de confiança em que a média amostral esteja a menos de 0,025 g da verdadeira média populacional de todas as moedas de 25 centavos? Se utilizamos a amostra de moedas de 25 centavos do Conjunto de Dados 13 do Apêndice B, podemos estimar em 0,068 g o desvio-padrão da população.
15. Referindo-nos aos pesos (em gramas) de moedas de 25 centavos relacionados no Conjunto de Dados 13 do Apêndice B, encontramos 50 pesos com média de 5,622 g e desvio-padrão de 0,068 g.

Com base nessa amostra de moedas em circulação, construa um intervalo de 98% de confiança para a estimativa da média populacional de todas as moedas de 25 centavos em circulação. O Ministério da Fazenda dos EUA alega que as moedas de 25 centavos cunhadas devem pesar em média 5,670 g. Esta alegação é compatível com o intervalo de confiança? Caso contrário, qual é a explicação plausível para a discrepância?

16. O Conjunto de Dados 1 do Apêndice B inclui os pesos *totais* do lixo descartado por 62 residências durante uma semana (com base em dados coletados como parte do Projeto do Lixo da Universidade do Arizona). Para essa amostra, a média é de 27,44 lb e o desvio-padrão é de 12,46 lb. Construa uma estimativa intervalar de 97% de confiança para o peso médio do lixo descartado por todas as residências. Se a cidade de Providence comporta o lixo até um peso médio de 35 lb, há alguma causa de preocupação quanto a um eventual excesso de lixo?

17. Uma psicóloga elaborou um novo teste de percepção espacial e deseja estimar o escore médio alcançado por pilotos do sexo masculino. Quantas pessoas ela deve testar para que o erro da média amostral não exceda 2,0 pontos, com 95% de confiança? Estudo anterior sugere $\sigma = 21,2$.

18. Para planejar o manuseio adequado do lixo doméstico, a cidade de Providence deve estimar o peso médio do lixo descartado pelas residências em uma semana. Determine o tamanho da amostra necessário para estimar essa média, para que tenhamos 96% de confiança em que a média amostral esteja a menos de 2 lb da verdadeira média populacional. Para o desvio-padrão populacional σ, use o valor 12,46 lb, que é o desvio-padrão da amostra de 62 residências incluídas no estudo do Projeto do Lixo na Universidade do Arizona.

19. O Department of Health, Education and Welfare — USDHEW coletou dados amostrais referentes a 1525 mulheres com idades de 18 a 24 anos. Esse grupo amostral tem nível médio de colesterol sérico (medido em mg/100 ml) de 191,7, com desvio-padrão de 41,0 (ver publicação 78-1652 do USDHEW). Com esses dados amostrais, determine o intervalo de 90% de confiança para o nível médio de colesterol sérico de todas as mulheres na faixa etária 18-24. Se um médico afirma que o nível médio de colesterol sérico para as mulheres naquela faixa etária é 200, tal afirmação se afigura compatível com o intervalo de confiança?

20. Uma amostra aleatória de 200 possuidores de cartão de crédito mostra que o débito médio anual nesses cartões, para contas individuais, é \$1592, com desvio-padrão de \$997 (com base em dados do *USA Today*). Com essas estatísticas, construa um intervalo de [illegible] de confiança para o débito médio anual em cartões de crédito para a população de todas as contas.

21. A Nielsen Media Research deseja estimar o tempo médio (em horas) que os estudantes universitários de tempo integral passam vendo televisão em cada dia da semana. Determine o tamanho da amostra necessário para estimar essa média com uma margem de erro de 0,25 h (ou 15 minutos). Suponha que se exija um grau de 96% de confiança. Suponha também que um estudo piloto tenha indicado que o desvio-padrão é estimado em 1,87 horas.

22. Ao decidirem entrar para uma faculdade, muitos estudantes se deixam influenciar pelo crescente potencial de ganho que um diploma pode proporcionar. Dados recentes do Ministério do Trabalho (dos EUA) mostram que a renda média anual dos que concluíram o curso secundário é de \$21.652, enquanto a renda média anual dos bacharéis por uma faculdade é de \$40.202. Determine o tamanho da amostra necessária para estimar a renda média anual dos bacharéis no próximo ano. Suponha um grau de 94% de confiança em que a média amostral esteja a menos de \$1000 da verdadeira média populacional, e admita que a estimativa do desvio-padrão seja \$32.896.

23. O leitor acaba de ser contratado pela Boston Marketing Company para realizar uma pesquisa a fim de estimar o gasto médio (por filme) de espectadores de filmes em Massachusetts. Aplique primeiro a regra prática para fazer uma estimativa grosseira do desvio-padrão das quantias gastas. É razoável admitir que a despesa típica varie de \$3 a cerca de \$15. Use então o desvio-padrão estimado para determinar o tamanho da amostra correspondente a 98% de confiança e a uma margem de erro de 25 centavos.

24. Estime as idades mínima e máxima dos livros-texto típicos normalmente usados em cursos de faculdades e aplique a regra prática para estimar o desvio-padrão. Em seguida, determine o tamanho da amostra necessário para estimar a idade média (em anos) dos livros-texto normalmente utilizados em cursos de faculdades. Admita um grau de 96% de confiança em que o erro da média amostral não supere 0,25 ano.

25. Recorra ao Conjunto de Dados 1 do Apêndice B para os 62 pesos (em libras) de papel descartado por residências [com base em dados do Garbage Project da Universidade do Arizona]. Utilizando essa amostra, construa uma estimativa intervalar de 95% de confiança para o peso médio do papel descartado por todas as residências.

26. Recorrendo ao Conjunto de Dados 7 do Apêndice B, construa um intervalo de 99% de confiança para a média da precipitação pluviométrica total anual em Iowa.

27. Recorrendo ao Conjunto de Dados 11 do Apêndice B, construa um intervalo de 97% de confiança para o peso médio dos bombons M&M marrons. Podemos utilizar os métodos desta seção para construir um intervalo de 97% de confiança para o peso médio dos bombons M&M azuis? Por que ou por que não?

28. Com base no Conjunto de Dados 8 do Apêndice B, construa um intervalo de 98% de confiança para o valor médio das moedas em poder de estudantes de estatística. Há razão para crer que esse valor seja diferente do valor médio das moedas em poder de pessoas selecionadas aleatoriamente da população de adultos americanos?

6-2 Exercícios B: Além do Básico

29. Na Fórmula 6-1, admitimos que a população seja infinita, que a amostragem seja feita com reposição, ou que a população seja muito grande. Se tivermos uma população relativamente pequena e a amostragem se processar sem reposição, é necessário modificar E, com a inclusão de um fator de *correção para população finita*, como segue:

$$E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

onde N é o tamanho da população. Mostre que a expressão precedente pode ser resolvida em relação a n, dando

$$n = \frac{N\sigma^2 (z_{\alpha/2})^2}{(N-1)E^2 + \sigma^2(z_{\alpha/2})^2}$$

Refaça o Exercício 13, admitindo que os professores de estatística sejam selecionados aleatoriamente, sem reposição, de uma população de $N = 200$ professores.

30. O erro-padrão da média é $\sigma/\sqrt{n}$, desde que a população seja infinita. Se o tamanho da população é finito, e denotado por N, deve-se aplicar o fator de correção

$$\sqrt{\frac{N-n}{N-1}}$$

sempre que $n > 0{,}05N$. Este fator multiplica o erro-padrão da média, conforme Exercício 29. Determine um intervalo de confiança de 95% para a média de 100 valores de QI, se uma amostra de 31 desses valores acusa média de 132 e desvio-padrão de 10.

31. Constatou-se a necessidade de uma amostra de tamanho 810 para estimar o peso médio (em mg) de comprimidos de Bufferin. Esse tamanho de amostra se baseia em um grau de 95% de confiança e um desvio-padrão populacional estimado pelo desvio-padrão amostral do Conjunto de Dados 14 do Apêndice B. Determine a margem de erro E.

32. $430 < \mu < 470$ é um intervalo de 95% de confiança para as vidas (em minutos) de pilhas Kodak AA. [Veja Programa 1 de *Against All Odds: Inside Statistics*.] Suponha que este resultado se baseie em uma amostra de tamanho 100.

    a. Construa o intervalo de 99% de confiança.
    b. Qual é o valor da média amostral?
    c. Qual é o valor do desvio-padrão amostral?
    d. Se se obtém com os mesmos dados o intervalo de confiança $432 < \mu < 468$, qual é o grau de confiança?

6-3 Estimativa de uma Média Populacional: Pequenas Amostras

Na Seção 6-2 iniciamos nosso estudo de estatística inferencial considerando estimativas pontuais e intervalos de confiança para estimar o valor de uma média populacional μ. Todos os exemplos e exercícios da Seção 6-2 envolviam amostras grandes, com tamanho $n > 30$. Fatores como custo e tempo, entretanto, não raro impõem sérios limites ao tamanho de uma amostra, de tal modo que a distribuição normal pode não ser uma aproximação adequada da distribuição de médias de pequenas amostras. Nesta seção vamos estudar métodos para estimar uma média populacional μ quando o tamanho n da amostra é pequeno, entendendo-se como *pequeno* 30 ou menos.

Notemos, de início, que, no caso de pequenas amostras, a média amostral $\bar{x}$ é, em geral, a melhor *estimativa pontual* da média populacional μ. Em segundo lugar, podem-se construir intervalos de confiança para pequenas amostras utilizando-se a distribuição normal com a mesma margem de erro da seção precedente, desde que a população original tenha distribuição normal e que se conheça o desvio-padrão populacional σ (uma condição não muito comum nas aplicações reais). Mais adiante, nesta seção, elaboraremos e resumiremos essas condições; mas vamos considerar primeiro os casos de pequenas amostras em que a população tem distribuição normal, mas o desvio-padrão populacional σ não é conhecido

Se temos uma amostra pequena ($n \leq 30$) e pretendemos construir um intervalo de confiança mas não conhecemos σ, podemos eventualmente utilizar a distribuição t de Student, desenvolvida por William Gosset (1876-1937). Gosset era empregado da Cervejaria Guinness e precisava de uma distribuição que pudesse ser utilizada com pequenas amostras. Como a cervejaria irlandesa para a qual ele trabalhava não permitia a publicação de resultados de pesquisas, Gosset publicou-os com o pseudônimo de *Student*. Graças a seus experimentos e estudos iniciais, podemos hoje trabalhar com a distribuição t de Student.

A Distribuição t de Student

Se a distribuição de uma população é essencialmente normal (com a forma aproximadamente de um sino), então a distribuição de

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

é essencialmente uma **distribuição t de Student** para todas as amostras de tamanho n. A distribuição t de Student, geralmente conhecida como **distribuição t**, é utilizada na determinação de valores críticos denotados por $t_{\alpha/2}$.

A Tabela A-3 relaciona valores da distribuição t juntamente com áreas denotadas por α. Obtêm-se valores de $t_{\alpha/2}$ na Tabela A-3 localizando o número adequado de graus de liberdade na coluna à esquerda e percorrendo a linha correspondente até atingir o número diretamente abaixo do valor aplicável (bilateral) de α.

DEFINIÇÃO

O número de **graus de liberdade** para um conjunto de dados corresponde ao número de valores que podem variar após terem sido impostas certas restrições a todos os valores.

Por exemplo, se 10 estudantes têm em um teste notas com média 80, podemos atribuir valores arbitrários a 9 delas, mas a décima fica determinada univocamente. A soma das 10 notas deve ser 800, de modo que a décima nota deve ser igual a 800 menos a soma das 9 primeiras. Como as 9 primeiras notas podem ser escolhidas arbitrariamente, dizemos que há nove graus de liberdade. Para as aplicações desta seção, o número de graus de liberdade é simplesmente o tamanho da amostra menos 1.

$$\textbf{graus de liberdade} = n - 1$$

Condição: Para que a distribuição t de Student seja aplicável, a distribuição da população básica deve ser essencialmene normal; não precisa ser exatamente normal, mas se tem apenas uma moda e é basicamente simétrica, obtemos bons resultados em geral (tais como intervalos de confiança precisos). Se há forte evidência de que a população tem uma distribuição bastante não-normal, então devemos utilizar métodos não-paramétricos (veja Capítulo 13) ou métodos de reamostragem.

Pode parecer um pouco estranho que, com uma população distribuída normalmente, venhamos eventualmente a utilizar a distribuição t para achar valores críticos; mas quando σ não é conhecido, a utilização de s de uma amostra pequena incorpora outra fonte de erro. Para manter o grau desejado de confiança, compensamos a variabilidade adicional ampliando o intervalo de confiança por um processo que substitui o valor crítico $z_{\alpha/2}$ (obtido na Tabela A-2 de valores da distribuição normal padronizada) por um valor crítico maior $t_{\alpha/2}$ (obtido na Tabela A-3 de valores da distribuição t).

Hispânicos Maltratados em Pesquisas

O Bureau of Census dos EUA define povos de *origem hispânica* como aqueles de origem ou ascendência espanhola/hispânica, incluindo mexicanos, cubanos, porto-riquenhos, povos de países de língua espanhola e outros. A definição inclui o comentário de que "pessoas de origem espanhola/hispânica podem ser de qualquer raça". Entretanto, ao listarem dados, as estatísticas e os resumos governamentais às vezes combinam raça e etnia. Em um caso, os dados do censo são relacionados sob o título "Raça e Origem Hispânica", com categorias de brancos, pretos, povos de origem hispânica e outros. Mas como classificar um hispânico branco? As categorias são confusas ou se superpõem. Um estudo realizado pelo Office of Management and Budget (Departamento de Administração e Orçamento) mostrou que o formato combinado de "Raça e Origem Hispânica" resulta em uma subavaliação do número de hispânicos da ordem de 20 a 30 por cento. [O autor agradece a Joseph Diaz-Calderón, um estudante que utilizou a edição anterior deste livro, por ter suscitado a questão.]

Propriedades Importantes da Distribuição *t* de Student

1. A distribuição *t* de Student é diferente, conforme o tamanho da amostra. (Ver Figura 6-5 para os casos $n = 3$ e $n = 12$.)
2. A distribuição *t* de Student tem a mesma forma geral simétrica (forma de sino) que a distribuição normal, mas reflete a maior variabilidade (com distribuições mais amplas) que é esperada em pequenas amostras.
3. A distribuição *t* de Student tem média $t = 0$ (tal como a distribuição normal padronizada, com média $z = 0$).
4. O desvio-padrão da distribuição *t* de Student varia com o tamanho da amostra, mas é superior a 1 (ao contrário da distribuição normal padronizada, em que $\sigma = 1$).
5. Na medida em que aumenta o tamanho n da amostra, a distribuição *t* de Student se aproxima mais e mais da distribuição normal padronizada. Para valores $n > 30$, as diferenças são tão pequenas que podemos utilizar os valores críticos z em lugar de elaborar uma tabela muito maior de valores críticos t. (Os valores da última linha da Tabela A-3 são iguais aos valores críticos z correspondentes da distribuição normal padronizada.)

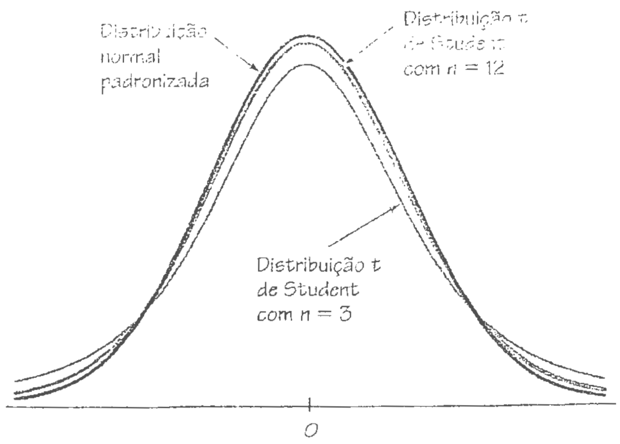


Fig. 6-5 Distribuições *t* de Student para n = 3 e n = 12. A distribuição *t* de Student tem a mesma forma geral e simetria que a distribuição normal padronizada, mas reflete a maior variabilidade esperada com pequenas amostras.

Segue um resumo das condições que indicam o uso de uma distribuição *t* em lugar da distribuição normal padronizada. (Essas mesmas condições também se aplicam no Capítulo 7.)

Condições para Utilização da Distribuição *t* de Student

1. O tamanho da amostra é pequeno ($n \leq 30$); e
2. σ é desconhecido; e
3. A população original tem distribuição essencialmente normal. (Como a distribuição da população original em geral é desconhecida, estimamo-la construindo um histograma de dados amostrais.)

Podemos agora determinar valores da margem de erro E ao estimar μ quando se aplica uma distribuição t. Pode-se utilizar essa margem de erro para construir intervalos de confiança.

Margem de Erro para a Estimativa de μ (Com base em uma Amostra Pequena ($n \leq 30$) e σ Desconhecido)

Fórmula 6-3 $E = t_{\alpha/2}\dfrac{s}{\sqrt{n}}$ onde $t_{\alpha/2}$ tem $n - 1$ graus de liberdade

Intervalo de Confiança para a Estimativa de μ (Com base em uma Amostra Pequena ($n \leq 30$) e σ Desconhecido)

$$\bar{x} - E < \mu < \bar{x} + E$$

onde

$$E = t_{\alpha/2}\frac{s}{\sqrt{n}}$$

EXEMPLO Com um *teste destrutivo*, as amostras são destruídas no processo de teste. O teste de colisão de carros é um exemplo muito dispendioso de teste destrutivo. Se o leitor fosse responsável por tais testes de colisão, dificilmente convenceria seu chefe da necessidade de fazer colidir e destruir mais de 30 carros, a fim de poder utilizar uma distribuição normal. Suponha que tenhamos feito teste de colisão em 12 carros esporte Dodge Viper (preço de venda atual: \$59.300) sob uma diversidade de condições que simulam colisões típicas. A análise dos 12 carros danificados resulta em custos de conserto que parecem ter distribuição em forma de sino com média $\bar{x} = \$26.227$ e desvio-padrão $s = \$15.873$ (com base em dados do Highway Loss Data Institute). Determine:

a. A melhor estimativa pontual de μ, o custo médio de conserto de todos os Dodge Vipers envolvidos em colisões.
b. A estimativa intervalar de 95% de μ.

SOLUÇÃO

a. A melhor estimativa pontual da média populacional μ é o valor da média amostral $\bar{x}$. Neste caso, a melhor estimativa pontual de μ é \$26.227.
b. Passaremos à construção de um intervalo de 95% de confiança utilizando a distribuição t, porque são verificadas

as condições seguintes: (1) A amostra é pequena ($n \leq 30$), (2) o desvio-padrão populacional σ é desconhecido e (3) a população parece ter distribuição normal, porque os dados amostrais têm distribuição em forma de sino.

Começamos determinando o valor da margem de erro conforme a seguir. Note que o valor crítico $t_{\alpha/2} = 2{,}201$ é obtido na Tabela A-3 na interseção da coluna rotulada "0,05 bilateral" (de 95% de confiança) com a linha correspondente a 11 graus de liberdade ($n - 1 = 11$).

$$E = t_{\alpha/2}\frac{s}{\sqrt{n}} = 2{,}201\,\frac{15.873}{\sqrt{12}} = 10.085{,}29$$

Podemos agora construir a estimativa intervalar de 95% para μ utilizando $E = 10.085{,}29$ e $\bar{x} = 26.227$. Como as estatísticas resumo são arredondadas para o dólar mais próximo, os limites do intervalo também serão arredondados para o dólar mais próximo.

$$\bar{x} - E < \mu < \bar{x} + E$$
$$26.227 - 10.085{,}29 < \mu < 26.227 + 10.085{,}29$$
$$\$16.142 < \mu < \$36.312$$

[Este resultado poderia expressar-se também como $\mu = \$26.227 \pm \$10.085$ ou como (\$16.142, \$36.312).]

Com base nos resultados apresentados, temos 95% de confiança de que os limites de \$16.142 e \$36.312 efetivamente contêm o valor da média populacional μ. Esses custos de conserto parecem ser bastante elevados. Na verdade, o Dodge Viper é o carro mais caro para consertar após uma colisão. Esta informação é de grande importância para companhias que aceitam seguro de Dodge Vipers contra colisão.

Trechos de uma Circular do Ministério dos Transportes

Os trechos seguintes de uma circular do Ministério dos Transportes dos EUA se referem a algumas exigências quanto à precisão de instrumentos de navegação usados em aviões. Observe o uso do intervalo de confiança.

"O total das contribuições de erros de equipamento aerotransportado, quando combinado com o total adequado de erros técnicos de vôo relacionados, não deve exceder o que segue, com 95% de confiança (2 sigma) em um período de tempo igual ao ciclo de atualização."

"O sistema de vias e rotas aéreas nos EUA tem faixas de proteção de vôo usadas em um sistema VOR, com uma precisão de ±4,5 graus, com uma base probabilística de 95%."

Utilização de Calculadoras e Computadores para Intervalos de Confiança

Uma vez dominada a teoria básica fundamental do uso de intervalos de confiança na estimativa de parâmetros, e entendido como construí-los e interpretá-los, podemos utilizar programas que simplifiquem os cálculos manuais. Com o STATDISK, selecione `Analysis` do menu principal, em seguida selecione `Confidence Intervals`, e introduza os itens solicitados. Após ter *clicado* `Evaluate`, aparecerá o intervalo de confiança.

O Minitab exige uma listagem original dos dados brutos e não funciona com estatísticas resumo. Para contornar essa limitação, e para maiores detalhes sobre a utilização do Minitab, consulte o *Minitab Student Laboratory Manual and Workbook*.

A Calculadora TI-83 pode gerar intervalos de confiança do tipo abordado nesta seção. Introduza os dados em L1 ou obtenha as estatísticas resumo; acione então `TInterval` (após acionar a tecla STAT e selecionar `Tests`). Utilizando os dados do exemplo anterior, TI-83 irá exibir um resultado que inclui este intervalo de confiança: (16142, 36312).

Escolha da Distribuição Apropriada

Às vezes é difícil decidir se devemos utilizar a Fórmula 6-1 (e a distribuição normal padronizada) ou a Fórmula 6-3 (e a distribuição t de Student). O fluxograma da Figura 6-6 resume os pontos-chave a serem considerados na construção de intervalos de confiança para estimar a média populacional μ. Na Figura 6-6, note que, no caso de uma amostra pequena ($n \leq 30$) extraída de uma distribuição essencialmente não-normal, não é possível utilizarmos os métodos descritos neste capítulo. Uma alternativa consiste no emprego de métodos não-paramétricos (veja Capítulo 13), e outra alternativa é a utilização do método de reamostragem, que não faz hipótese alguma sobre a população original. Descreve-se este método no Projeto para Computador no final deste capítulo.

EXEMPLO Suponha que o leitor deseja construir um intervalo de confiança para a média populacional μ. Utilize os dados a seguir para determinar se a margem de erro E deve ser calculada com auxílio da distribuição normal, da distribuição t, ou nenhuma das duas (quando então não podemos usar os métodos deste capítulo).

a. $n = 50$, $\bar{x} = 77{,}6$, $s = 14{,}2$, e a forma da distribuição é assimétrica.
b. $n = 25$, $\bar{x} = 77{,}6$, $s = 14{,}2$ e a distribuição tem forma de sino.
c. $n = 25$, $\bar{x} = 77{,}6$, $\sigma = 14{,}2$ e a distribuição tem forma de sino.
d. $n = 25$, $\bar{x} = 77{,}6$, $\sigma = 14{,}2$ e a distribuição é extremamente assimétrica.

SOLUÇÃO
Recorra à Figura 6-6 e use o fluxograma para determinar:

a. Como a amostra é grande ($n > 30$), aplique a distribuição normal. Ao calcular a margem de erro E, use a Fórmula 6-1, onde o desvio-padrão σ é substituído por s.
b. Use a distribuição t porque (1) a amostra é pequena, (2) a população parece ter distribuição normal, e (3) σ é desconhecido.
c. Aplique a distribuição normal porque a população parece ter distribuição normal e σ é conhecido.
d. Como a amostra é pequena e a população tem distribuição essencialmente não-normal, não podemos aplicar os métodos deste capítulo. Não podemos utilizar nem a distribuição t nem a distribuição normal.

Amostra Pequena

O Children's Defense Fund (Fundo de Amparo à Criança) foi organizado para promover o bem-estar da criança. O grupo publicou *Children Out Of School In America*, onde relata que, em determinada região, 37,5% dos adolescentes de 16 e 17 anos não freqüentavam escola. Esta estatística, que mereceu

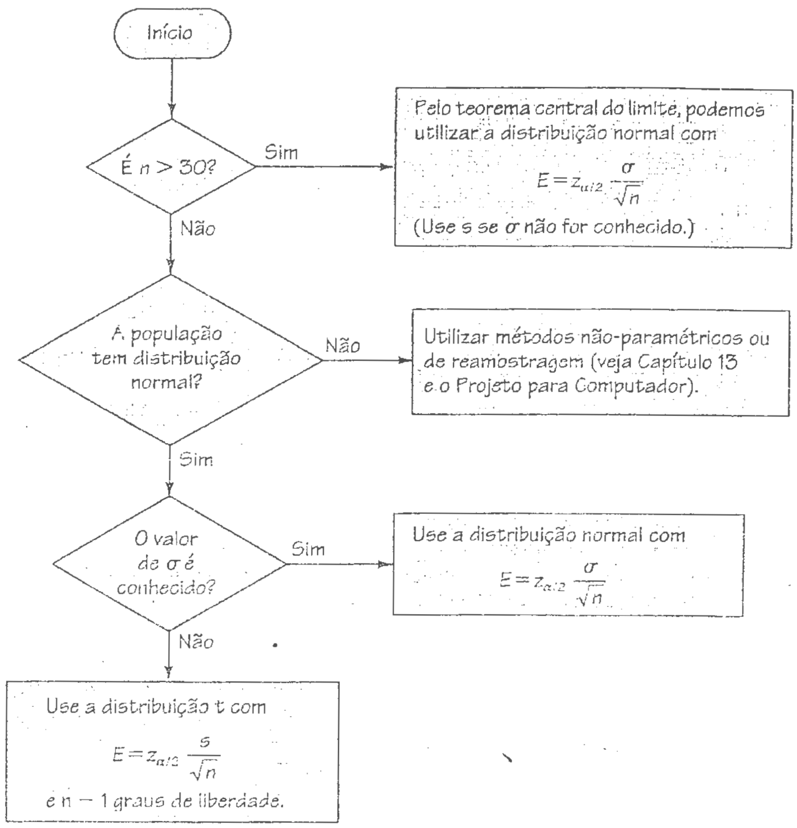


Fig. 6-6 Escolha entre as distribuições normal (z) e t.

ampla cobertura da imprensa, baseou-se em uma amostra de apenas 16 crianças. Outra estatística teve como base apenas três estudantes. [Veja: "Firsthand Report: How Flawed Statistics Can Make an Ugly Picture Look Even Worse"; *American School Board Journal*, Vol. 162.)]

Na Seção 6-2 apresentamos os três conceitos mais importantes de estimativa pontual, intervalo de confiança e determinação do tamanho da amostra. Nesta seção ampliam-se os conceitos de estimativa pontual e intervalo de confiança ao caso de pequenas amostras. Esta seção não inclui a determinação do tamanho da amostra porque, com σ conhecido, aplica-se a distribuição normal, e com σ desconhecido necessitamos de uma amostra maior para justificar a estimação de σ por s; mas a amostra grande permite-nos novamente utilizar a distribuição normal. Baseamos, pois, nossa determinação do tamanho da amostra somente na Fórmula 6-2, e não consideramos qualquer circunstância em que a distribuição t deva ser usada.

A seção a seguir aborda novamente os conceitos de estimativa pontual, intervalo de confiança e determinação do tamanho da amostra, mas vamos focalizar a proporção populacional, em lugar da média populacional.

6-3 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-4, ache o valor crítico $t_{\alpha/2}$ *que corresponde ao grau de confiança e ao tamanho* n *da amostra de dados.*

1. 99%; $n = 10$
2. 95%; $n = 16$
3. 98%; $n = 21$
4. 90%; $n = 8$

Nos Exercícios 5-8, dados o grau de confiança e os elementos amostrais, determine (a) a margem de erro e (b) o intervalo de confiança para a média populacional μ. *Em ambos os casos, admita que a população tenha distribuição normal.*

5. Alturas de mulheres: 95% de confiança; $n = 10$, $\bar{x} = 63{,}4$ in., $s = 2{,}4$ in.
6. Médias de notas: 99% de confiança; $n = 15$, $\bar{x} = 2{,}76$, $s = 0{,}88$
7. Notas de testes: 90% de confiança; $n = 16$, $\bar{x} = 77{,}6$, $s = 14{,}2$
8. Salários da polícia: 98% de confiança; $n = 19$, $\bar{x} = \$23.228$, $s = \$8779$

Nos Exercícios 9-20, determine corretamente se os intervalos de confiança são calculados com a distribuição normal ou com a distribuição t.

9. Em testes de colisão feitos em 15 minivans Honda Odyssey, os custos de conserto apresentam uma distribuição aproximadamente em forma de sino, com média de \$1786 e desvio-padrão de \$937 (com base em dados do Highway Loss Data Institute). Construa um intervalo de confiança de 99% para o custo médio de conserto para as colisões de todos os veículos desse tipo.
10. O Conjunto de Dados 2 do Apêndice B apresenta 106 temperaturas do corpo humano tomadas às 12:00 h do dia 2. Suponha que tenhamos apenas as 10 temperaturas dadas a seguir. Para esses valores,

$\bar{x} = 98{,}44$ e $s = 0{,}30$. Construa um intervalo de confiança de 95% para a média de todas as temperaturas do corpo humano. (Estudo anterior mostra que essas temperaturas têm distribuição normal.)

98,6 98,6 98,0 98,0 99,0 98,4 98,4 98,4 98,4 98,6

11. Em um estudo sobre aplicação do tempo, constatou-se que 20 administradores selecionados aleatoriamente gastam uma média de 2,40 horas por dia com trabalho burocrático. O desvio-padrão das 20 horas é 1,30 h (com base em dados da Adia Personnel Services). Os dados aparentam ter uma distribuição normal. Construa o intervalo de confiança de 95% para o tempo médio gasto em trabalho burocrático por todos os administradores.
12. Em um estudo do tempo necessário para o serviço de entrega no apartamento em um Hotel Radisson recentemente aberto, 20 entregas acusaram média de 24,2 minutos, com desvio-padrão de 8,7 minutos. Os dados amostrais parecem ter distribuição em forma de sino. Construa um intervalo de confiança de 90% para a média de todas as entregas.
13. Em um estudo de atração física e distúrbios mentais, 231 pessoas foram classificadas quanto à atração; a média amostral e o desvio-padrão amostral resultantes foram 3,94 e 0,75, respectivamente. (Veja o artigo "Physical Attractiveness and Self-Perception of Mental Disorder", por Burns e Farina, *Journal of Abnormal Psychology*, Vol. 96, No. 2.) Com esses dados amostrais, construa um intervalo de confiança de 95% para a média populacional.
14. Fez-se um estudo para estimar o custo hospitalar de vítimas de acidentes que usavam cinto de segurança. Vinte casos selecionados aleatoriamente têm distribuição aparentemente em forma de sino, com média de \$9004 e desvio-padrão de \$5629 (com base em dados do Ministério dos Transportes dos EUA). Construa um intervalo de confiança de 99% para a média de todos esses custos.
15. Construa um intervalo de 98% de confiança para a renda média de todos os empregados de tempo integral que têm grau de bacharel. Uma amostra de 25 desses empregados revelou que a distribuição das rendas é aproximadamente normal, com média de \$39.271 e desvio-padrão de \$18.933 (com base em dados do Ministério do Trabalho dos EUA).
16. Uma amostra aleatória de 19 mulheres acusou altura média de 63,85 in. Outro estudo revelou que as alturas das mulheres têm distribuição normal com desvio-padrão $\sigma = 2{,}5$ in. Construa o intervalo de confiança de 98% para a altura média de todas as mulheres.
17. Em um estudo da utilização da hipnose para aliviar a dor, obtiveram-se as taxas sensoriais para 16 indivíduos, com os resultados dados a seguir [com base em dados de "An Analysis of Factors That Contribute to the Efficacy of Hypnotic Analgesia", por Price e Barber, *Journal of Abnormal Psychology*, Vol. 96, No. 1]. Com esses dados amostrais, construa o intervalo de confiança de 95% para a taxa sensorial média da população da qual se extraiu a amostra.

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

18. Com base no Conjunto de Dados 11 do Apêndice B, use apenas a média amostral dos bombons M&M *vermelhos* para construir um intervalo de confiança de 95% para o peso médio de todos os M&Ms.
19. Com base no Conjunto de Dados 11 do Apêndice B, use toda a amostra de 100 bombons M&M para construir um intervalo de confiança de 95% para o peso médio de todos os M&Ms. Compare o resultado com o intervalo de confiança obtido no Exercício 18.
20. Com referência ao Conjunto de Dados 10 no Apêndice B.

    a. Construa um intervalo de confiança de 95% para o comprimento médio de filmes classificados como R.
    b. Construa um intervalo de confiança de 95% para o comprimento médio de filmes classificados como PG ou PG-13.
    c. Compare os métodos e os resultados das partes (a) e (b).

6-3 Exercícios B: Além do Básico

21. Suponha que uma amostra pequena ($n \leq 30$) seja selecionada aleatoriamente de uma população distribuída normalmente com σ desconhecido. Na construção de um intervalo de confiança deve-se usar a distribuição t, mas se se utiliza incorretamente a distribuição normal, como são afetados os limites do intervalo de confiança?
22. Constrói-se um intervalo de confiança para uma pequena amostra de temperaturas (em graus Fahrenheit) selecionada aleatoriamente de uma população normalmente distribuída com σ desconhecido (como o conjunto de dados do Exercício 10).

    a. Como é afetada a margem E de erro, se cada temperatura é convertida para a escala Celsius? $[C = \frac{5}{9}(F - 32)]$
    b. Denotando por a e b os limites do intervalo de confiança, ache as expressões para esses limites após converter as temperaturas originais para a escala Celsius.
    c. Com base nos resultados da parte (b), podemos obter os limites do intervalo de confiança para as temperaturas Celsius, simplesmente convertendo os limites da escala Fahrenheit para a escala Celsius?

6-4 Estimativa de uma Proporção Populacional

Nesta seção vamos abordar os mesmos três conceitos estudados nas Seções 6-2 e 6-3: (1) estimativa pontual, (2) intervalo de confiança, e (3) determinação do tamanho n da amostra. Nas Seções 6-2 e 6-3 aplicamos esses conceitos a estimativas de uma média populacional μ; nesta seção vamos aplicá-los à proporção populacional p. Por exemplo, a Nielsen Media Research poderia querer estimar a proporção de residências que sintonizam o programa *Super Bowl*, e à Companhia de Seguros Hartford poderia interessar a estimativa da proporção de motoristas embriagados.

Embora, nesta seção, enfatizemos a proporção populacional p, podemos também considerar uma probabilidade ou uma percentagem. Tanto as proporções como as probabilidades são expressas em forma decimal ou fracionária. Ao trabalharmos com percentagens, transformamo-las em proporções eliminando o sinal % e dividindo por 100. Por exemplo, a taxa de 48,7% de pessoas que não compram livros pode expressar-se em forma decimal como 0,487. O símbolo p pode, pois, representar uma proporção, uma probabilidade ou o equivalente decimal de uma percentagem. Vamos introduzir a notação $\hat{p}$ (lê-se "p chapéu") para a proporção amostral.

Notação para Proporção

p = *proporção populacional*

$\hat{p} = \frac{x}{n}$ proporção amostral de x sucessos em uma amostra de tamanho n

Em capítulos anteriores definimos $q = 1 - p$; é natural, pois, fazermos aqui $\hat{q} = 1 - \hat{p}$. Por exemplo, se pesquisarmos 1068 americanos e constatarmos que 673 deles têm secretária eletrônica, a proporção amostral é $\hat{p} = x/n = 673/1068 = 0{,}630$, e $\hat{q} = 0{,}370$ $(1 - 0{,}630)$. Em alguns casos, o valor de $\hat{p}$ pode ser conhecido porque a proporção ou percentagem amostral é dada diretamente. Se, em 1068 espectadores de TV americanos, 25% são bacharéis, então $\hat{p} = 0{,}25$ e $\hat{q} = 0{,}75$.

Campanhas Políticas Disfarçadas

Uma forma da campanha política disfarçada tem como objetivo desviar eleitores dos candidatos da oposição, formulando perguntas tendenciosas para desacreditá-los. Eis um exemplo de uma dessas questões: "Queira dizer se estaria mais inclinado ou menos inclinado a votar em Roy Romer se soubesse que o Gov. Romer nomeou uma comissão que decidiu pela liberdade de uma média de quatro criminosos por dia desde que tomou posse." O National Council of Public Polls caracteriza tais pesquisas como antiéticas, mas alguns pesquisadores profissionais não condenam a prática, desde que as questões não envolvam mentiras cabais.

Estimativa Pontual

A proporção amostral $\hat{p}$ é a melhor estimativa pontual da proporção populacional p.

Margem de Erro da Estimativa de p:

Fórmula 6-4 $\quad E = z_{\alpha/2}\sqrt{\dfrac{\hat{p}\hat{q}}{n}}$ $\rightarrow S_{\hat{p}} = \sqrt{\dfrac{\hat{p}.\hat{q}}{n}}$

Intervalo de Confiança (ou Estimativa Intervalar) para a Proporção Populacional p:

$$\hat{p} - E < p < \hat{p} + E \quad \text{onde} \quad E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$\hookrightarrow \hat{p} \pm z.(S_{\hat{p}})$

Eventualmente, expressa-se o intervalo de confiança como

$$p = \hat{p} \pm E$$

ou

$$(\hat{p} - E, \hat{p} + E)$$

No Capítulo 3 arredondamos para três algarismos significativos as probabilidades expressas em forma decimal. Vamos usar aqui a mesma regra de arredondamento.

Regra de Arredondamento para Estimativas Intervalares de p

Arredondar para três algarismos significativos os limites do intervalo de confiança.

EXEMPLO Os pesquisadores de opinião são atormentados por uma diversidade de fatores de confusão, como secretárias eletrônicas. Em uma pesquisa junto a 1068 americanos, 673 informaram ter secretária eletrônica (com base em dados da International Mass Retail Association, relatado em *USA Today*). Com esses resultados amostrais, determine:

a. a estimativa pontual da proporção populacional de todos os americanos que têm secretária eletrônica.
b. a estimativa intervalar de 95% da proporção populacional de todos os americanos que têm secretária eletrônica.

SOLUÇÃO
a. A estimativa pontual de p é

$$\hat{p} = \frac{x}{n} = \frac{673}{1068} = 0{,}630$$

b. A construção do intervalo de confiança exige primeiro o cálculo da margem de erro E, que pode ser obtido pela Fórmula 6-4. Temos $\hat{p} = 0{,}630$ (da parte a), $\hat{q} = 0{,}370$ (de $\hat{q} = 1 - \hat{p}$) e $z_{\alpha/2} = 1{,}96$ (da Tabela A-2, onde 95% é convertido para $\alpha = 0{,}05$, que é dividido igualmente entre as duas caudas, de forma que $z = 1{,}96$ corresponde a uma área de 0,4750).

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 1{,}96\sqrt{\frac{(0{,}630)(0{,}370)}{1068}} = 0{,}0290$$

Podemos agora achar o intervalo de confiança, utilizando $\hat{p} = 0{,}630$ e $E = 0{,}0290$.

$$\hat{p} - E < p < \hat{p} + E$$
$$0{,}630 - 0{,}0290 < p < 0{,}630 + 0{,}0290$$
$$0{,}601 < p < 0{,}659$$

Se quiséssemos o intervalo de 95% de confiança para a verdadeira *percentagem* populacional, poderíamos expressar este resultado como $60{,}1\% < p < 65{,}9\%$. Este resultado costuma ser apresentado da seguinte forma: "Entre os americanos, a percentagem dos que têm secretária eletrônica é estimada em 63%, com uma margem de erro de ±2,9 pontos percentuais." Trata-se de uma formulação verbal da expressão do intervalo de confiança: $p = 63\% \pm 2{,}9\%$. (Dever-se-ia indicar também o nível de confiança, mas isto raramente acontece. As publicações utilizam tipicamente um grau de confiança de 95%, mas não fazem qualquer referência a isso.) Pode-se expressar o intervalo de confiança também como: (0,601; 0,659).

Utilização de Calculadoras e Computadores para Intervalos de Confiança

O STATDISK constrói intervalos de confiança para uma proporção populacional. Para isto, selecione `Analysis`, em seguida `Confidence Intervals` e introduza os itens solicitados. Pode-se usar também a calculadora TI-83, selecionando `1-PropZInt` (após acionar a tecla STAT e selecionar `TESTS`). Se estiver usando os dados do exemplo precedente, a TI-83 apresentará um quadro que inclui o intervalo de confiança na forma (.6012, .6591).

Rationale: Utilizamos $\hat{p}$ como estimativa pontual de p (assim como usamos $\bar{x}$ como estimativa pontual de μ), porque é não-tendencioso e, entre os estimadores que podem ser usados, é o mais consistente. É não-tendencioso no sentido de que a distribuição de proporções amostrais tende a centrar-se em torno do valor de p; isto é, as proporções amostrais $\hat{p}$ não tendem nem a subestimar, nem a sobreestimar sistematicamente p. A proporção amostral $\hat{p}$ é o estimador mais consistente no sentido de que o desvio-padrão das proporções amostrais tende a ser menor do que o desvio-padrão de quaisquer outros estimadores não-tendenciosos.

Suposições: Como estamos lidando com proporções, podemos utilizar a distribuição binomial apresentada na Seção 4-3. Admitimos nesta seção que sejam essencialmente satisfeitas as quatro condições da Seção 4-3 para a distribuição binomial: (1) Há um número fixo de provas, (2) as provas são independentes, (3) há duas categorias de resultados, e (4) as probabilidades permanecem constantes em todas as provas. Admitimos também que se possa utilizar a distribuição normal como aproximação da distribuição de proporções amostrais (porque as condições $np \geq 5$ e $nq \geq 5$ são ambas satisfeitas). Podemos, assim, utilizar os resultados da Seção 5-6 para concluir que μ e σ são dados por $\mu = np$ e $\sigma = \sqrt{npq}$. Ambos os parâmetros se referem a n provas, mas convertemo-los para uma base *por prova* dividindo por n:

$$\text{Média de proporções amostrais} = \frac{np}{n} = p$$

$$\text{Desvio-padrão de proporções amostrais} = \frac{\sqrt{npq}}{n} = \sqrt{\frac{npq}{n^2}} = \sqrt{\frac{pq}{n}}$$

O primeiro resultado pode parecer trivial porque já estipulamos que a verdadeira proporção populacional é p. Já o segundo resultado é não-trivial e serve para descrever a margem de erro E, mas devemos substituir o produto pq por $\hat{p}\,\hat{q}$ porque ainda não conhecemos o valor de p (é o valor que estamos procurando estimar). A Fórmula 6-4 para a margem de erro reflete o fato de que $\hat{p}$ tem uma probabilidade $1 - \alpha$ de estar a menos de $z_{\alpha/2}\sqrt{pq/n}$ de p. O intervalo de confiança para p, dado anteriormente, reflete o fato de que há uma probabilidade $1 - \alpha$ de $\hat{p}$ diferir de p por menos da margem de erro E.

Tamanho de Amostra em TV

Um artigo de *Newsweek* menciona a utilização de pesquisadores como uma forma de determinar quantas pessoas estão assistindo a diferentes programas de televisão. O artigo afirmava que "os estatísticos vêm afirmando, de longa data, que as amostras de residências usadas são demasiadamente pequenas para determinar com precisão o que a América está assistindo. Diante disso, é interessante notar que as 4000 residências visitadas pelos pesquisadores constituem exatamente 0,0045% da população ligada à rede de TV."

A alegação de que o tamanho da amostra de 4000 residências é por demais pequeno não é válida. Recorrendo a métodos deste capítulo, mostra-se que uma amostra aleatória de tamanho 4000 pode dar resultados assaz bons, mesmo que represente apenas 0,0045% da população.

Determinação do Tamanho da Amostra

Tendo discutido as estimativas pontuais e os intervalos de confiança para p, passamos a descrever um processo para determinar o tamanho necessário da amostra, quando desejamos achar o valor aproximado de uma proporção populacional. Na seção anterior, partimos da expressão da margem de erro E e resolvemos em relação a n. Utilizando processo análogo, começamos com

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

e resolvemos em relação a n para obter o tamanho da amostra, tal como dado pela Fórmula 6-5. Essa fórmula exige $\hat{p}$ como estimativa da proporção populacional p, mas se não se conhece tal estimativa, substituímos $\hat{p}$ por 0,5 e $\hat{q}$ por 0,5, com o resultado dado pela Fórmula 6-6.

Tamanho da Amostra para Estimar a Proporção p

Quando se Conhece uma Estimativa $\hat{p}$:

Fórmula 6-5 $\quad n = \dfrac{[z_{\alpha/2}]^2\hat{p}\hat{q}}{E^2}$

Quando Não se Conhece uma Estimativa $\hat{p}$:

Fórmula 6-6 $\quad n = \dfrac{[z_{\alpha/2}]^2 \cdot 0{,}25}{E^2}$

Regra de Arredondamento na Determinação do Tamanho da Amostra

Se o tamanho da amostra calculado não é um número inteiro, devemos arredondá-lo para o próximo inteiro *mais elevado*.

$\hat{p}$	$\hat{q}$	$\hat{p} \cdot \hat{q}$
0,1	0,9	0,09
0,2	0,8	0,16
0,3	0,7	0,21
0,4	0,6	0,24
0,5	0,5	0,25
0,6	0,4	0,24
0,7	0,3	0,21
0,8	0,2	0,16
0,9	0,1	0,09

Utilize a Fórmula 6-5 quando puder fazer estimativas razoáveis de $\hat{p}$ com base em amostras prévias, um estudo piloto ou o conhecimento de um perito. No caso de não dispormos de tais recursos, podemos atribuir o valor 0,5 tanto a $\hat{p}$ como a $\hat{q}$, de forma que o tamanho da amostra resultante seja no mínimo tão grande quanto devia ser. A razão subjacente para atribuir o valor 0,5 é a seguinte: O valor máximo possível do produto $\hat{p}\,\hat{q}$ é 0,25, que ocorre quando $\hat{p} = 0{,}5$ e $\hat{q} = 0{,}5$. (Ver a tabela a seguir, que relaciona alguns valores de $\hat{p}$ e $\hat{q}$.) Note que as Fórmulas 6-5 e 6-6 não incluem o tamanho N da população; N é, assim, irrelevante. (*Exceção*: Amostragem sem reposição de uma população finita relativamente pequena. Ver Exercício 29.)

EXEMPLO As companhias de seguro estão ficando preocupadas com o fato de que o número crescente de telefones celulares resulte em maior número de colisões de carros; estão, por isso, pensando em cobrar prêmios mais elevados para os motoristas que utilizam celulares. Desejamos estimar, com uma margem de erro de três pontos percentuais, a percentagem de motoristas que falam ao celular enquanto estão dirigindo. Supondo que se pretende um nível de confiança de 95% nos resultados, quantos motoristas devem ser pesquisados?

a. Suponha que tenhamos uma estimativa $\hat{p}$ com base em estudo anterior, que mostrou que 18% dos motoristas falam ao celular (com base em dados da revista *Prevention*).
b. Suponha que não tenhamos qualquer informação que possa sugerir um valor de $\hat{p}$.

SOLUÇÃO

a. O estudo anterior sugere $\hat{p} = 0{,}18$, de forma que $\hat{q} = 0{,}82$ (de $\hat{q} = 1 - \hat{p}$). Ao nível de 95% de confiança, temos $\alpha = 0{,}05$, donde $z_{\alpha/2} = 1{,}96$. Outrossim, a margem de erro é $E = 0{,}03$ (o equivalente decimal de "três pontos percentuais"). Como temos um valor estimado de $\hat{p}$, utilizamos a Fórmula 6-5 como segue.

$$n = \frac{[z_{\alpha/2}]^2\hat{p}\hat{q}}{E^2} = \frac{[1{,}96]^2(0{,}18)(0{,}82)}{0{,}03^2}$$
$$= 630{,}0224 = 631 \text{ (arredondado para cima)}$$

Devemos pesquisar ao menos 631 motoristas selecionados aleatoriamente.

b. Tal como na parte (a), utilizamos novamente $z_{\alpha/2} = 1{,}96$ e $E = 0{,}03$, mas sem qualquer conhecimento prévio de $\hat{p}$ (ou $\hat{q}$), aplicamos a Fórmula 6-6:

$$n = \frac{[z_{\alpha/2}]^2 \cdot 0{,}25}{E^2} = \frac{[1{,}96]^2 \cdot (0{,}25)}{0{,}03^2}$$
$$= 1067{,}1111 = 1068 \text{ (arredondado para cima)}$$

Para termos 95% de confiança de que nossa percentagem amostral está a menos de três pontos percentuais da verdadeira percentagem de todos os motoristas, devemos selecionar aleatoriamente e pesquisar 1068 motoristas. Comparando este resultado com o tamanho amostral de 631 obtido na parte (a), podemos ver que, na ausência de conhecimento de um estudo prévio, é necessário uma amostra maior para obtermos os mesmos resultados que obteríamos se pudéssemos estimar o valor de $\hat{p}$.

A parte (b) do exemplo precedente envolve a aplicação da Fórmula 6-6, a mesma fórmula freqüentemente usada por Nielsen, Gallup e outros pesquisadores profissionais. Muitos acham, incorretamente, que devemos pesquisar uma percentagem da população, mas a Fórmula 6-6 mostra que o tamanho da população é irrelevante. (Na verdade, às vezes utilizamos o tamanho da população, mas somente em casos de amostragem sem reposição de uma população relativamente pequena. Veja Exercício 29.) A maioria das pesquisas apresentadas em jornais, revistas e TV envolve amostras com tamanho de 1000 a 2000. Mesmo que tais pesquisas abranjam uma percentagem muito pequena da população, podem dar resultados bastante satisfatórios. Quando Nielsen pesquisa 1068 residências com TV em uma população de 93 milhões, apenas 0,001% das residências é pesquisado; entretanto, mesmo assim podemos ter 95% de confiança de que a percentagem amostral estará a menos de três pontos percentuais da percentagem populacional.

As pesquisas de opinião adquiriram grande importância e predominância nos Estados Unidos. Elas afetam os *shows* de TV a que assistimos, os governantes que elegemos, a legislação que nos rege e os produtos que consumimos. A compreensão dos conceitos desta seção elimina grande parte do mistério e incompreensão que cercam as pesquisas.

6-4 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-4, suponha que utilizemos uma amostra para estimar uma proporção populacional p. *Determine a margem de erro que corresponde aos valores dados de* n *e* x *e ao grau de confiança.*

1. $n = 800, x = 200$, 95% **2.** $n = 1400, x = 420$, 99%
3. $n = 4275, x = 2576$, 98% **3.** $n = 887, x = 209$, 90%

Nos Exercícios 5-8, utilize os dados amostrais e o grau de confiança para construir uma estimativa intervalar para a proporção populacional p.

5. $n = 800, x = 600$, 95% de confiança
6. $n = 2000, x = 300$, 99% de confiança
7. $n = 2475, x = 992$, 90% de confiança
8. $n = 5200, x = 1024$, 98% de confiança

Nos Exercícios 9-12, utilize os dados para achar o menor tamanho de amostra necessário para estimar uma proporção ou percentagem populacional.

9. Margem de erro: 0,02; nível de confiança: 95%; $\hat{p}$ e $\hat{q}$ desconhecidos
10. Margem de erro: 0,01; nível de confiança: 90%; $\hat{p}$ e $\hat{q}$ desconhecidos
11. Margem de erro: quatro pontos percentuais; nível de confiança: 99%; $\hat{p}$ é estimado em 0,20 com base em estudo anterior
12. Margem de erro, dois pontos percentuais; nível de confiança: 97%; $\hat{p}$ é estimado em 0,85 com base em estudo anterior
13. A Hartford Insurance Company deseja estimar a percentagem dos motoristas que trocam fita ou CD enquanto dirigem. Uma amostra de 850 motoristas acusou 544 que trocam fita ou CD quando na direção (com base em dados da revista *Prevention*).

    a. Determine a estimativa pontual da *percentagem* de todos os motoristas que trocam fita ou CD quando dirigindo.
    b. Determine uma estimativa intervalar de 90% da *percentagem* de todos os motoristas que trocam fita ou CD quando dirigindo.

14. Selecionados aleatoriamente e pesquisados 500 universitários, verificou-se que 135 deles têm computadores pessoais (com base em dados da America Passage Media Corporation).

    a. Determine a estimativa pontual da verdadeira proporção de todos os universitários que têm computador pessoal.
    b. Determine um intervalo de 95% de confiança para a verdadeira proporção de todos os universitários que têm computador pessoal.

15. Uma repórter da revista *Byte* deseja fazer uma pesquisa para estimar a verdadeira proporção de todos os universitários que têm computador pessoal, e quer ter 95% de confiança de que seus resultados tenham uma margem de erro de 0,04. Quantos universitários devem ser pesquisados?

    a. Suponha que tenhamos uma estimativa de $\hat{p}$, obtida em estudo anterior, que revele uma percentagem de 27% (com base em dados de America Passage Media Corporation).
    b. Suponha que não temos qualquer informação anterior que sugira um possível valor de $\hat{p}$.

16. Como fabricantes de equipamento para golfe, a Spalding Corporation deseja estimar a proporção de golfistas canhotos. (Com esta informação, a companhia pode planejar o número de equipamentos para destros e para canhotos que deve produzir.) Quantos golfistas devem ser pesquisados, para que a companhia tenha 98% de confiança de que a proporção amostral apresenta uma margem de 0,025 de erro?

    a. Suponha que não se dispõe de qualquer informação que permita estimar $\hat{p}$.
    b. Suponha que temos uma estimativa de $\hat{p}$, obtida em um estudo prévio, que sugere que 15% dos golfistas são canhotos (com base em um artigo de *USA Today*).

17. Uma pesquisa de mercado para a Ford Motor Company revela que uma amostra de 1220 residências selecionadas aleatoriamente inclui 1054 que possuem um veículo (dados do Bureau of Census). Com base nesses resultados, construa um intervalo de 98% de confiança para a percentagem de todas as residências que possuem um veículo.
18. Um relatório do repórter de TV do programa *Prime Time Live* abordou o problema do prejuízo do consumidor num supermercado de Connecticut. Falando pela Connecticut Consumer Protection Agency, Gloria Shaffer afirmou que, de 1527 pacotes conferidos, 706 acusavem peso a menos. Segundo ela, trata-se de uma percentagem muito elevada. Com os dados amostrais, construa um intervalo de confiança de 99% para a percentagem de todos os pacotes que acusam peso a menos. Com base no resultado, concorda com a afirmação de que se trata de uma percentagem muito alta?
19. A Greybar Tax Company acha que seus clientes são selecionados para sofrerem fiscalização numa proporção bem mais alta do que a da população em geral. O Imposto de Renda afirma que verifica 4,3% das declarações dos que ganham mais de \$100.000, mas em uma verificação de 400 clientes da Greybar com renda superior a \$100.000, selecionados aleatoriamente, constatou-se que 56 deles foram fiscalizados. Com um nível de 99%, construa um intervalo de confiança para a percentagem de declarações de clientes da Greybar, com renda superior a \$100.000, que são fiscalizados. Com base no resultado, parece que os clientes de renda elevada da Greybar são fiscalizados com freqüência substancialmente maior do que a da população em geral?
20. O Locust Tree Restaurant mantém registros das reservas e de não-comparecimentos. Selecionadas aleatoriamente 150 reservas de um sábado, verificou-se que houve 70 não-comparecimentos (com base em dados do American Express). Determine um intervalo de confiança para a proporção dos não-comparecimentos aos sábados, utilizando um grau de 90% de confiança.
21. Quantas residências com TV a Nielsen deve pesquisar para estimar a percentagem das que estão sintonizadas no programa *The Late Show with David Letterman*? Adote a margem de 97% de confiança em que sua percentagem amostral tenha uma margem de erro de dois pontos percentuais. Admita também que nada se sabe sobre a percentagem de residências sintonizadas para qualquer *show* de TV após 11 horas da noite.
22. A West American Communications Company está cogitando de fazer uma concorrência para o serviço telefônico interurbano. Deseja-se fazer uma pesquisa para estimar a percentagem de assinantes que estão satisfeitos com o atual serviço de interurbanos. Queremos ter 90% de confiança em que a percentagem amostral esteja a menos de 2,5 pontos percentuais do verdadeiro valor populacional, e uma pesquisa da Roper sugere que essa percentagem deve girar em torno de 85%. Qual deve ser o tamanho da amostra?
23. A cadeia de hotéis American Resort dá um teste de aptidão aos candidatos a emprego, e considera fácil uma questão do tipo múltipla escolha se ao menos 80% das respostas são corretas. Uma amostra aleatória de 6503 respostas a determinada questão apresenta 84% de respostas corretas. Construa o intervalo de confiança de 99% para a verdadeira percentagem de respostas corretas. É admissível que a questão seja realmente fácil?
24. A indústria do tabaco fiscaliza severamente todas as pesquisas que envolvem o fumo. Uma pesquisa revelou que, em 785 indivíduos com quatro anos de faculdade selecionados aleatoriamente, 18,3% fumam (com base em dados da American Medical Association). Construa o intervalo de 98% de confiança para a verdadeira percentagem dos fumantes entre todos os que completaram quatro anos de faculdade. Com base no resultado, a taxa de fumantes entre bacharéis parece ser substancialmente diferente da taxa geral de 27%?
25. Numa pesquisa feita em um supermercado, verificaram-se 1234 itens, constatando-se 20 deles com preço excessivo (com base em dados de "UPC Scanner Pricing Systems: Are They Accurate?" por Goodstein, *Journal of Marketing*, Vol. 58).

    a. Com os dados amostrais, construa um intervalo de confiança de 95% para a proporção de todos os artigos que acusam preço excessivo.
    b. Utilizando os dados amostrais como estudo piloto, determine o tamanho da amostra necessário para estimar a proporção de itens que acusam preço excessivo. Admita um nível de confiança de 99% em que a estimativa não apresente erro superior a 0,005.

26. Um estudo de saúde envolve 1000 mortes selecionadas aleatoriamente, dentre as quais 331 causadas por doenças cardíacas (com base em dados do Center for Disease Control).

    a. Com os dados amostrais, construa um intervalo de confiança de 99% para a proporção de todas as mortes causadas por doenças cardíacas.
    b. Utilizando os dados amostrais como estudo piloto, determine o tamanho da amostra necessário para estimar a proporção de todas as mortes causadas por doenças cardíacas. Admita um nível de confiança de 98%, em que o erro da estimativa não supere 0,01.

27. O presidente da Columbia Pictures, Mark Canton, afirma que 58% dos filmes feitos são classificados como R. Com base nos dados do Conjunto 10 do Apêndice B, construa o intervalo de confiança de 95% para a percentagem dos filmes classificados como R. O intervalo de confiança resultante é compatível com a afirmação de Canton?
28. Com base no Conjunto de Dados 11 do Apêndice B, ache a proporção amostral dos M&Ms vermelhos. Utilize o resultado para construir um intervalo de 95% de confiança para estimar a população dos M&Ms vermelhos. O resultado é compatível com a taxa de 20% reportada pelo fabricante Mars?

6-4 Exercícios B: Além do Básico

29. Nesta seção apresentamos as Fórmulas 6-5 e 6-6, utilizadas para determinar o tamanho da amostra. Em ambos os casos admitimos

que a população seja infinita ou muito grande, ou que a amostragem seja feita com reposição. Quando temos uma população relativamente pequena, de tamanho N, e a amostragem é sem reposição, modificamos E para incluir o *fator de correção para população finita* dado a seguir, e resolvemos em relação a n, obtendo o resultado à direita. Com este resultado, refaça o Exercício 21, supondo que limitamos nossa população a uma cidade de 5000 habitantes.

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}\sqrt{\frac{N-n}{N-1}} \qquad n = \frac{N\hat{p}\hat{q}[z_{\alpha/2}]^2}{\hat{p}\hat{q}\,[z_{\alpha/2}]^2 + (N-1)E^2}$$

30. Um artigo no *New York Times* sobre resultados de pesquisas afirma: "Teoricamente, de 19 casos em 20, os resultados de tal pesquisa não devem diferir em mais de um ponto percentual, em uma ou outra direção, do valor que obteríamos entrevistando todos os eleitores nos EUA." Determine o tamanho da amostra sugerido por esta afirmação.
31. Segundo um artigo de jornal, uma estimativa da taxa de desemprego envolve uma pesquisa de 47.000 indivíduos (tamanho típico de amostra em pesquisas do Departamento de Estatísica do Trabalho dos EUA). Se a taxa de desemprego reportada não deve ter erro superior a 0,2 ponto percentual, e se sabemos que a taxa é de cerca de 8%, determine o nível de confiança correspondente.
32. As alturas das mulheres têm distribuição normal com média de 63,6 in. e desvio-padrão de 2,5 in. Quantas mulheres devem ser pesquisadas, se queremos estimar a percentagem das que têm mais de 5 ft de altura? Admita um nível de confiança de 98%, em que o erro não supere 2,5 pontos percentuais. (*Sug.:* A resposta é substancialmente inferior a 2172.)
33. Um *intervalo de confiança unilateral* para p pode ser representado como $p < \hat{p} + E$ ou $p > \hat{p} - E$, com a margem de erro modificada substituindo-se $z_{\alpha/2}$ por z_α. Se a Air America deseja manter uma taxa de pontualidade de pelo menos x por cento com 95% de confiança, construa o intervalo de confiança unilateral adequado e determine a percentagem em questão. Suponha que, em 750 vôos, 630 estejam no horário.
34. Existem tabelas especiais para achar intervalos de confiança para proporções que envolvem números pequenos de casos em que não se pode usar a aproximação normal. Por exemplo, para três sucessos em oito provas, o intervalo de confiança dado em *Standard Probability and Statistics Tables and Formulae* (CRC Press) é $0{,}085 < p < 0{,}755$. Ache o intervalo de confiança resultante se utilizássemos incorretamente a distribuição normal como aproximação da binomial. Os resultados estão razoavelmente próximos?

6-5 Estimativa de uma Variância Populacional

Nesta seção vamos considerar os mesmos três conceitos abordados nas Seções 6-2, 6-3 e 6-4: (1) estimativa pontual, (2) intervalo de confiança e (3) determinação do tamanho da amostra. As Seções 6-2, 6-3 e 6-4 aplicam estes conceitos a estimativas de médias e de proporções; nesta seção, vamos aplicá-los à variância populacional σ^2, ou ao desvio-padrão σ. Em muitas situações reais, como o controle de qualidade em processos de fabricação, devemos estimar valores de variâncias ou desvios-padrão populacionais. Além de medidas que apresentem uma média desejada, o fabricante deve produzir artigos de qualidade *consistente*, que não variem de extremamente bons a extremamente maus. Esta consistência pode ser geralmente avaliada pela variância ou pelo desvio-padrão, que são, assim, estatísticas vitais para a manutenção da qualidade de produtos.

Suposição: Nesta seção admitimos que a população tenha valores distribuídos normalmente. Esta suposição já foi feita em seções anteriores, mas aqui ela é mais crítica. Ao aplicar a distribuição t de Student na Seção 6-3, por exemplo, exigimos que a distribuição de valores fosse aproximadamente normal, embora aceitando desvios da normalidade que não fossem muito sérios. Ao calcular variâncias com os métodos desta seção, entretanto, a utilização de populações com distribuições muito não-normais pode levar a erros sérios. Conseqüentemente, a exigência de que tenhamos uma distribuição normal é muito mais estrita, tornando-se necessário verificar a distribuição dos dados com a construção de um histograma, para ver se é simétrica e se tem a forma de um sino. Descrevemos esta sensitividade a uma distribuição normal dizendo que as inferências sobre a variância populacional σ^2 (ou sobre o desvio-padrão populacional σ), feitas com base na distribuição qui-quadrado (a ser definida mais adiante), não são *robustas*, o que significa que as inferências podem ser enganosas se a população não tem uma distribuição normal. Em contrapartida, as inferências feitas sobre a média populacional μ, com base na distribuição t de Student, são razoavelmene robustas, porque os desvios da normalidade que não são muito extremos não conduzem a erro sério.

Quando abordamos estimativas de médias e de proporções nas Seções 6-2, 6-3 e 6-4, utilizamos as distribuições normal e t de Student. Já ao estabelecermos estimativas de variâncias ou desvios-padrão, lançamos mão de outra distribuição, chamada distribuição qui-quadrado. Vamos examinar características importantes dessa distribuição antes de passarmos ao estabelecimento de intervalos de confiança.

A Distribuição Qui-Quadrado

Em uma população distribuída normalmente com variância σ^2, escolhemos aleatoriamente amostras independentes de tamanho n e calculamos a variância amostral s^2 (ver Fórmula 2-5) para cada amostra. A estatística amostral $\chi^2 = (n-1)s^2/\sigma^2$ tem uma distribuição chamada **distribuição qui-quadrado**.

A Distribuição Qui-Quadrado	
Fórmula 6-7	$\chi^2 = \dfrac{(n-1)s^2}{\sigma^2}$
onde	n = tamanho da amostra
	s^2 = variância amostral
	σ^2 = variância populacional

Denotamos qui-quadrado por χ^2. (As equações matemáticas que definem esta distribuição não são dadas aqui, porque ultrapassam o âmbito deste livro.) Para achar valores críticos da distribuição qui-quadrado, recorremos à Tabela A-4. A distribuição qui-quadrado é determinada pelo número de graus de liberdade; neste capítulo utilizamos $n - 1$ graus de liberdade.

$$\textbf{graus de liberdade} = n - 1$$

Em capítulos posteriores encontraremos situações em que o número de graus de liberdade não é $n - 1$; não devemos, assim, tomar sempre esse número como $n - 1$.

Propriedades da Distribuição da Estatística Qui-Quadrado

1. A distribuição qui-quadrado não é simétrica, ao contrário das distribuições normal e t de Student (ver Figura 6-7). (Na medida em que o número de graus de liberdade aumenta, a distribuição vai se tornando menos assimétrica, conforme ilustrado na Figura 6-8.)
2. Os valores de qui-quadrado podem ser zero ou positivos; nunca podem ser negativos (ver Figura 6-7).
3. Há uma distribuição qui-quadrado diferente para cada número de graus de liberdade (ver Figura 6-8) que é gl $= n - 1$ nesta seção. À medida que o número de graus de liberdade aumenta, a distribuição qui-quadrado tende para uma distribuição normal.

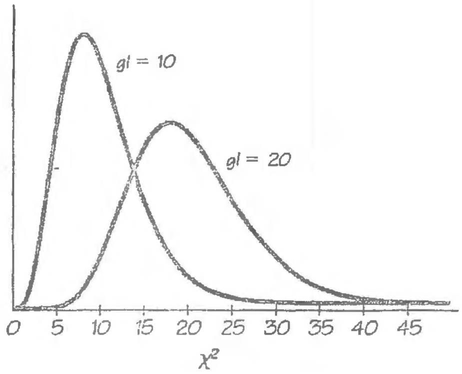


Fig. 6-8 Distribuição Qui-Quadrado para gl = 10 e gl = 20.

Em razão da natureza da distribuição χ^2, os métodos desta seção são muito diferentes dos das seções precedentes. Pode-se ver uma diferença no processo de determinação dos pontos críticos ilustrado no exemplo que se segue. Atente para a seguinte característica essencial da Tabela A-4:

Na Tabela A-4, cada valor crítico de χ^2 corresponde a uma área dada na linha superior da tabela, e essa área representa a *região total localizada à direita* do valor crítico.

EXEMPLO Determine os valores críticos de χ^2 que definem regiões críticas contendo uma área de 0,025 em cada cauda. Suponha que o tamanho da amostra seja 10, de modo que o número de graus de liberdade é $10 - 1 = 9$.

SOLUÇÃO Veja a Figura 6-9 e consulte a Tabela A-4. Obtém-se o valor crítico à direita ($\chi^2 = 19{,}023$) diretamente, localizando 9 na coluna de graus de liberdade à esquerda e 0,025 na parte superior. O valor crítico $\chi^2 = 2{,}700$ à esquerda mais uma vez corresponde a 9 na coluna de graus de liberdade, mas devemos localizar 0,975 ($1 - 0{,}025$) na parte superior, porque os valores no topo são sempre *áreas à direita* do valor crítico. Verifique, na Figura 6-9, que a área total à direita de $\chi^2 = 2{,}700$ é 0,975.

A Figura 6-9 mostra que, para uma amostra de 10 valores extraídos de uma população distribuída normalmente, a estatística qui quadrado $(n-1)s^2/\sigma^2$ tem 0,95 de probabilidade de estar entre os valores críticos qui-quadrado 2,700 e 19,023.

Ao obter valores críticos da distribuição qui-quadrado na Tabela A-4, observe que os números de graus de liberdade são inteiros consecutivos de 1 a 30, seguidos por 40, 50, 60, 70, 80, 90 e 100. Quando um número de graus de liberdade (como 52) não se encontra na tabela, podemos em geral tomar o valor crítico mais próximo. Por exemplo, se o número de graus de liberdade é 52, recorra à Tabela A-4 e use 50 graus de liberdade. (Se o número de graus de liberdade está exatamente a meio caminho entre valores da tabela, como 55, basta acharmos a média dos dois valores χ^2.) Para números de graus de liberdade superiores a 100, utilize a equação dada no Exercício 22, uma tabela mais detalhada, ou um pacote estatístico.

Fig. 6-7 Distribuição Qui-Quadrado.

Estimadores de σ^2

Como as variâncias amostrais s^2 (obtidas com a Fórmula 2-5) tendem a centrar-se no valor da variância populacional σ^2, dizemos que s^2 é um *estimador não-tendencioso* de σ^2. Ou seja, as variâncias amostrais s^2 não tendem sistematicamente nem a sobreestimar nem a subestimar o valor de σ^2. Ao contrário, tendem para o próprio valor alvo de σ^2. Outrossim, os valores de s^2 tendem a produzir erros menores, pelo fato de estarem mais próximos de σ^2 do que quaisquer outras medidas de variação. Por estas razões, o valor de s^2 é, em geral, o melhor valor isolado (ou estimativa pontual) dentre as várias estatísticas que poderíamos usar para estimar σ^2.

A variância amostral s^2 é a melhor estimativa pontual da variância populacional σ^2.

Como s^2 é a melhor estimativa pontual de σ^2, seria natural esperarmos que s fosse a melhor estimativa pontual de σ, mas isso não ocorre, porque s é um estimador tendencioso de σ. Entretanto, se o tamanho da amostra é grande, a tendenciosidade é tão pequena que podemos utilizar s como uma estimativa razoavelmente boa de σ.

Embora s^2 seja a melhor estimativa pontual de σ^2, não há indicação de quão bom realmente seja. Para compensar esta deficiência, estabelecemos uma estimativa intervalar (ou intervalo de confiança) mais reveladora.

Intervalo de Confiança (ou Estimativa Intervalar) para a Variância Populacional σ^2

$$\frac{(n-1)s^2}{\chi^2_R} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_L}$$

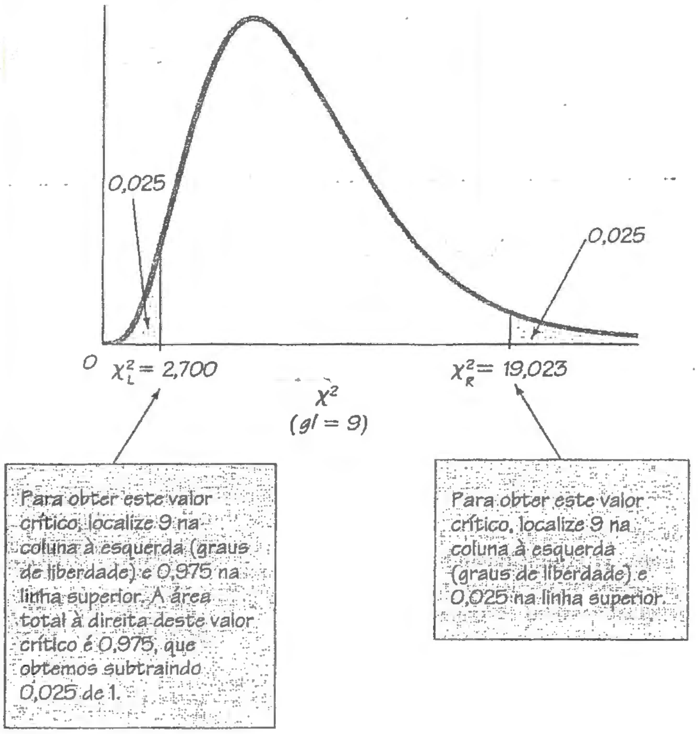


Fig. 6-9 Valores críticos da distribuição Qui-Quadrado.

Com esta expressão, obtemos um intervalo de confiança para σ^2; e o intervalo de confiança (ou estimativa intervalar) para o desvio-padrão é dado pela raiz quadrada de cada componente, conforme se segue:

$$\sqrt{\frac{(n-1)s^2}{\chi_R^2}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi_L^2}}$$

As notações χ_R^2 e χ_L^2* nas expressões precedentes são definidas como se segue. (Note que alguns textos usam $\chi^2_{\alpha/2}$ em lugar de χ_R^2 e $\chi^2_{1-\alpha/2}$ em lugar de χ_L^2.)

Notação

Com uma área total α dividida igualmente entre as duas extremidades de uma distribuição qui-quadrado, χ_L^2 denota o valor crítico da extrema esquerda e χ_R^2 denota o valor crítico da extrema direita. (Ver Figura 6-10.)

Os limites do intervalo de confiança para σ^2 e σ devem ser arredondados de acordo com a regra seguinte, que é, na verdade, a mesma regra básica dada na Seção 6-2.

Regra de Arredondamento para Estimativas de Intervalos de Confiança de σ e σ^2

1. Quando utilizar o conjunto original de dados para construir um intervalo de confiança, arredonde os limites do intervalo com uma decimal a mais além das que aparecem nos dados originais.
2. Quando o conjunto original de dados não é conhecido, utilizando-se apenas estatísticas resumo (n, s), arredonde os limites do intervalo de confiança para o mesmo número de decimais que aparecem no desvio-padrão ou na variância.

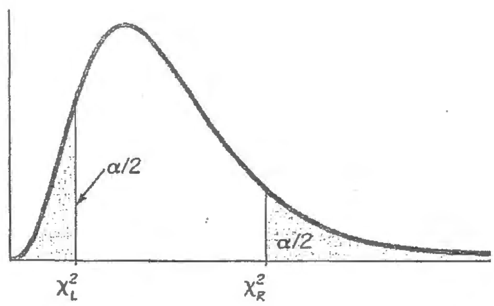


Fig. 6-10 Distribuição Qui-Quadrado com valores críticos χ_L^2 e χ_R^2. Esses valores críticos separam as áreas extremas correspondentes às variâncias amostrais improváveis (com probabilidade α).

*R de *right*, L de *left*. (N. do T.)

EXEMPLO A confeitaria Hudson Valley fabrica sonhos que são embalados em pacotes com a indicação de que há 12 sonhos pesando um total de 42 oz. Se a variação entre os sonhos é muito grande, algumas caixas terão peso a menos (prejudicando o consumidor) e outras terão peso a mais (diminuindo o lucro). É claro que o consumidor não ficaria satisfeito com um sonho tão pequeno que só pudesse ser visto ao microscópio, nem com um sonho tão grande que se assemelhasse a um pneu de trator. O supervisor de controle de qualidade constatou que esses problemas podem ser evitados se os sonhos tiverem um peso médio de 3,50 oz e um desvio-padrão de 0,06 oz ou menos. Selecionam-se aleatoriamente, na linha de produção, doze sonhos, que são pesados, dando os resultados a seguir. Construa dois intervalos de confiança de 95%, um para σ^2 e outro para σ, e determine se o supervisor de controle está com problemas.

3,58 3,50 3,68 3,61 3,42 3,52 3,66 3,50 3,36 3,42

SOLUÇÃO Com base nos dados amostrais, a média $\bar{x} = 3{,}504$ parece satisfatória, pois está muito próxima do valor desejado de 3,50 oz. Os valores dados acusam um desvio-padrão $s = 0{,}109$, superior ao valor desejado de 0,06 ou menos. Passemos à construção do intervalo de confiança para σ^2.

Com uma amostra de 12 valores, temos 11 graus de liberdade. Com um grau de confiança de 95%, dividimos $\alpha = 0{,}05$ igualmente entre as duas caudas da distribuição χ^2 e localizamos os valores 0,975 e 0,025 na linha superior. Os valores críticos de χ^2 são $\chi^2_L = 3{,}816$ e $\chi^2_R = 21{,}920$. Com esses valores críticos, o desvio-padrão amostral $s = 0{,}109$ e o tamanho 12 da amostra, construímos como segue o intervalo de 95% de confiança:

$$\frac{(12-1)(0{,}109)^2}{21{,}920} < \sigma^2 < \frac{(12-1)(0{,}109)^2}{3{,}816}$$

Esta expressão se reduz a $0{,}006 < \sigma^2 < 0{,}034$. Tomando a raiz quadrada de cada membro (antes de arredondar), vem $0{,}077 < \sigma < 0{,}185$. Com base no intervalo de confiança de 95%, parece que o desvio-padrão é superior ao valor desejado de 0,06 oz; surge assim um problema para o supervisor de controle que deve tomar providências corretivas para que os pesos dos sonhos sejam mais consistentes.

O intervalo de confiança $0{,}077 < \sigma < 0{,}185$ pode também expressar-se como (0,077; 0,185), mas *não podemos* utilizar a expressão $\sigma = s \pm E$, porque s não está no centro do intervalo de confiança.

Amostragem em Companhias Aéreas

As companhias de aviação costumavam utilizar um sistema contábil dispendioso para dividir a receita de bilhetes que envolviam duas ou mais companhias. Agora, as companhias aplicam um método de amostragem em que se seleciona aleatoriamente uma pequena percentagem desses bilhetes "fracionados"; essa amostra é usada para se fazer a divisão de toda a receita. O erro causado por este processo pode fazer com que algumas companhias recebam um pouco menos do que a sua quota exata, mas essas perdas são amplamente compensadas pela economia de mão-de-obra. Com o novo sistema, as companhias economizam milhões de dólares a cada ano.

Utilização de Programas de Computador para Intervalos de Confiança para σ ou σ^2

O STATDISK pode ser utilizado facilmente para construir intervalos de confiança para desvios-padrão ou variâncias. Selecione `Analysis` no menu principal, em seguida selecione `Confidence Intervals` e passe a introduzir os dados solicitados.

Com o Minitab, introduza os dados na coluna C1, selecione `Editor`, em seguida selecione `Enable Command Language` e introduza o comando `%DESCRIBE C1` para obter resultados que incluem intervalos de 95% para μ e σ. O grau de confiança pode ser mudado (95% é o *default*.)

As calculadoras TI-83 não dão intervalos de confiança para σ^2 ou σ.

Rationale: Consideremos agora por que os intervalos de confiança para σ^2 e σ têm as formas que acabamos de dar. Se extrairmos amostras de tamanho n de uma população com variância σ^2, a distribuição dos valores $(n-1)s^2/\sigma^2$ será a exibida na Figura 6-10. Para uma amostra aleatória, há uma probabilidade $1-\alpha$ de a estatística $(n-1)s^2/\sigma^2$ estar entre os valores críticos χ^2_L e χ^2_R. Em outras palavras (e símbolos), há uma probabilidade de $1-\alpha$ de se verificarem ambas as condições

$$\frac{(n-1)s^2}{\sigma^2} < \chi^2_R \quad \text{e} \quad \frac{(n-1)s^2}{\sigma^2} > \chi^2_L$$

Multiplicando por σ^2 ambas as desigualdades precedentes e dividindo cada uma delas pelo valor adequado de χ^2, vemos que as duas desigualdades podem expressar-se nas formas equivalentes

$$\frac{(n-1)s^2}{\chi^2_R} < \sigma^2 \quad \text{e} \quad \frac{(n-1)s^2}{\chi^2_L} > \sigma^2$$

As duas últimas desigualdades podem ser combinadas em uma única desigualdade (dupla):

$$\frac{(n-1)s^2}{\chi^2_R} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_L}$$

Há uma probabilidade de $1-\alpha$ de esses limites de intervalo de confiança conterem a variância populacional σ^2.

Determinação do Tamanho da Amostra

Os processos para achar o tamanho da amostra necessário para estimar σ^2 são muito mais complexos do que os processos estudados anteriormente para médias e proporções. Em lugar de processos bastante complicados, vamos utilizar a Tabela 6-2.

EXEMPLO Com 95% de confiança, queremos estimar σ a menos de 10%. Qual deve ser o tamanho da amostra? Admita que a população tenha distribuição normal.

SOLUÇÃO Pela Tabela 6-2, vemos que 95% de confiança e um erro de 10% para σ correspondem a um tamanho amostral de 191. Devemos, pois, selecionar aleatoriamente 191 valores da população.

TABELA 6-2 Tamanho de amostra para σ^2		Tamanho de amostra para σ	
Para estarmos 95% confiantes de que s^2 esteja a menos do valor de	do valor de σ^2, o tamanho n da amostra deve ser no mínimo	Para estarmos 95% confiantes de que o valor de s esteja a menos de	do valor de σ, o tamanho n da amostra deve ser pelo menos
1%	77.207	1%	19.204
5%	3.148	5%	767
10%	805	10%	191
20%	210	20%	47
30%	97	30%	20
40%	56	40%	11
50%	37	50%	7
Para estarmos 99% confiantes de que s^2 esteja a menos do valor de	do valor de σ^2, o tamanho n da amostra deve ser no mínimo	Para estarmos 99% confiantes de que o valor de s esteja a menos de	do valor de σ, o tamanho n da amostra deve ser pelo menos
1%	133.448	1%	33.218
5%	5.457	5%	1.335
10%	1.401	10%	335
20%	368	20%	84
30%	171	30%	37
40%	100	40%	21
50%	67	50%	13

6-5 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-4, ache os valores críticos χ^2_L e χ^2_R que correspondem ao grau de confiança e ao tamanho da amostra dados.

1. 95%; $n = 26$
2. 99%; $n = 17$
3. 90%; $n = 60$
4. 95%; $n = 50$

Nos Exercícios 5-8, use o grau de confiança e os dados amostrais indicados para achar um intervalo de confiança para o desvio-padrão populacional σ. Em cada caso, admita que a população tenha distribuição normal.

5. Alturas de mulheres: 95% de confiança; $n = 10$, $\bar{x} = 63{,}4$ in., $s = 2{,}4$ in.
6. Médias de notas: 99% de confiança; $n = 15$, $\bar{x} = 2{,}76$, $s = 0.88$
7. Notas de testes: 90% de confiança; $n = 16$, $\bar{x} = 77{,}6$, $s = 14{,}2$
8. Salários de policiais: 95% de confiança; $n = 19$, $\bar{x} = \$23.228$, $s = \$8.779$

Nos Exercícios 9-12, suponha que cada amostra seja obtida mediante seleção aleatória de uma população normalmente distribuída.

9. Determine o tamanho mínimo da amostra necessária para termos 95% de confiança em que o desvio-padrão amostral s esteja a menos de 30% de σ.
10. Determine o tamanho mínimo da amostra necessária para termos 99% de confiança em que o desvio-padrão amostral s esteja a menos de 20% de σ.
11. Determine o tamanho mínimo da amostra necessária para termos 99% de confiança em que a variância amostral esteja a menos de 30% da variância populacional.
12. Determine o tamanho mínimo da amostra necessária para termos 95% de confiança em que a variância amostral esteja a menos de 40% da variância populacional.

Nos Exercícios 13-20, admita que cada amostra seja obtida mediante seleção aleatória de valores de uma população com distribuição normal.

13. Supõe-se que um contêiner de anticongelante para carros comporte 3785 mL do líquido. Como as variações são inevitáveis, o controlador de qualidade quer ter a certeza de que o desvio-padrão seja inferior a 30 mL. Em caso contrário, alguns contêineres transbordariam, enquanto outros não teriam a quantidade suficiente de anticongelante. O controlador seleciona uma amostra, obtendo os resultados dados a seguir. Com esses dados amostrais, construa um intervalo de confiança de 99% de confiança para o verdadeiro valor de σ. Este intervalo de confiança sugere que o nível das flutuações seja aceitável?

3761	3861	3769	3772	3675	3861	$n = 18$
3888	3819	3788	3800	3720	3748	$\bar{x} = 3787{,}0$
3753	3821	3811	3740	3740	3839	$s = 55{,}4$

14. O National Center for Educational Statistics fez uma pesquisa junto a bacharéis por faculdades, sobre o tempo gasto para obterem seus títulos de bacharel. A média foi 5,15 anos e o desvio-padrão 1,68

ano. Suponha que o tamanho da amostra seja 101. Com base nesses dados amostrais, construa o intervalo de 99% de confiança para o desvio-padrão dos tempos gastos por todos os bacharéis.

15. Obtém-se uma amostra de 15 crânios de homens egípcios que viveram por volta de 1850 A.C. Mede-se a largura máxima de cada crânio, com o resultado $\bar{x} = 134{,}5$ mm e $s = 3{,}5$ mm (com base em dados de *Ancient Races of the Thebaid*, por Thomson e Randall-Maciver). Com esses dados amostrais, construa um intervalo de 95% de confiança para o desvio-padrão populacional σ.

16. Uma amostra consiste em 75 aparelhos de TV adquiridos há vários anos. Os tempos de substituição desses aparelhos acusam média de 8,2 anos com desvio-padrão de 1,1 ano (com base em dados de "Getting Things Fixed," *Consumer Reports*). Construa um intervalo de 90% de confiança para o desvio-padrão dos tempos de substituição de todos os aparelhos de TV daquela época. Os resultados se aplicam aos aparelhos que estão sendo vendidos atualmente?

17. Os valores relacionados são tempos de espera (em minutos) de clientes no Jefferson Valley Bank, onde os clientes entram em uma fila única que é atendida por três guichês. Construa um intervalo de 95% de confiança para o desvio-padrão populacional σ.

6,5 6,6 6,7 6,8 7,1 7,3 7,4 7,7 7,7 7,7

18. Os valores relacionados são tempos de espera (em minutos) de clientes do Bank of Providence, onde os clientes podem entrar em qualquer uma de três filas atendidas por três guichês diferentes. Construa um intervalo de 95% de confiança para σ e compare com o resultado do intervalo obtido para os dados do Exercício 17. Os intervalos de confiança sugerem uma diferença na variação entre os tempos de espera? Que sistema parece melhor: o de fila única ou o de filas múltiplas?

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

19. Veja o Conjunto de Dados 7 do Apêndice B.
   a. Aplique a regra prática (veja Seção 2-5) para estimar σ, o desvio-padrão da precipitação pluviométrica anual total em Iowa.
   b. Construa um intervalo de confiança de 99% para σ.
   c. O intervalo de confiança contém seu valor estimado de σ?

20. Veja o Conjunto de Dados 11 do Apêndice B.
   a. Aplique a regra prática (veja Seção 2-5) para estimar σ, o desvio-padrão dos pesos dos bombons M&M marrons.
   b. Construa um intervalo de 98% de confiança para σ.
   c. O intervalo de confiança contém seu valor estimado de σ?

6-5 Exercícios B: Além do Básico

21. Um artigo de jornal inclui um gráfico mostrando que certos dados amostrais são distribuídos normalmente.
   a. Inadvertidamente, omitiu-se o grau de confiança quando foi dado o intervalo de confiança $2{,}8 < \sigma < 6{,}0$. Determine o grau de confiança, com as seguintes estatísticas amostrais: $n = 20$, $\bar{x} = 45{,}2$ e $s = 3{,}8$.
   b. Dá-se o seguinte intervalo de confiança: $19{,}1 < \sigma < 45{,}8$. Determine o valor do desvio-padrão s, que foi omitido.

22. Ao construirmos intervalos de confiança para σ e σ^2, utilizamos a Tabela A-4 para achar os valores críticos χ^2_L e χ^2_R, mas essa tabela só se aplica a casos em que $n \leq 101$, de modo que o número máximo de graus de liberdade é 100. Para números maiores de graus de liberdade, podemos aproximar χ^2_L e χ^2_R com a expressão

$$\chi^2 = \frac{1}{2}[\pm z_{\alpha/2} + \sqrt{2k-1}]^2$$

onde k é o número de graus de liberdade. Construa o intervalo de confiança de 95% para σ utilizando os seguintes dados amostrais: As alturas de 772 homens entre 18 e 24 anos de idade têm desvio-padrão de 2,8 in. (Com base em dados do National Health Survey.)

Vocabulário

estimador
estimativa
estimativa pontual
intervalo de confiança
estimativa intervalar
grau de confiança
nível de confiança
coeficiente de confiança
valor crítico
margem de erro (E)
limites do intervalo de confiança
distribuição t de Student
distribuição t
graus de liberdade
ditribuição qui-quadrado

Revisão

Este capítulo e o seguinte introduzem conceitos fundamentais e importantes de inferência estatística. Aqui focalizamos *estimativas* de parâmetros ao considerar médias, proporções e variâncias populacionais, para estabelecer processos para:

- Identificar uma estimativa pontual
- Construir um intervalo de confiança
- Determinar o tamanho da amostra necessário

Abordamos a estimativa pontual (ou estimativa de valor único) e formulamos as seguintes conclusões:

- A melhor estimativa pontual de μ é $\bar{x}$
- A melhor estimativa pontual de p é $\hat{p}$
- A melhor estimativa pontual de σ^2 é s^2.

Como valores únicos, as estimativas pontuais não dizem muito sobre sua confiabilidade; por isso, introduzimos intervalos de confiança (ou estimativas intervalares) como estimativas mais informativas. Abordamos também maneiras de determinar tamanhos amostrais necessários para estimar parâmetros dentro de um fator de tolerância. Este capítulo introduziu também as distribuições t de Student e qui-quadrado. Devemos ter o cuidado de utilizar a distribuição correta para cada conjunto de circunstâncias.

É importante ter em mente que todos os processos para intervalos de confiança e tamanhos de amostra neste capítulo exigem que tenhamos uma população com distribuição aproximadamente normal. Se a distribuição se afasta muito da normal, devemos utilizar outros métodos, como o método descrito no Projeto para Computador no final deste capítulo.

Exercícios de Revisão

1. No Conjunto de Dados 3 do Apêndice B, temos 54 valores de comprimento de cabeças de urso, com média de 12,95 in., desvio-padrão de 2,14 in. e distribuição aproximadamente normal.
   a. Com base nessa amostra, construa uma estimativa intervalar de 98% de confiança para a média populacional de todas os comprimentos de cabeças de urso.
   b. Use os dados amostrais para construir uma estimativa intervalar de 98% de confiança para o desvio-padrão de todos os comprimentos de cabeças de urso.
   c. Utilizando o desvio-padrão amostral como estimativa do desvio-padrão populacional, determine o tamanho da amostra necessária para estimar o comprimento médio das cabeças de todos os ursos. Trabalhe com 98% de confiança em que a margem de erro seja de 0,25 in.

2. O leitor acaba de ser contratado pela General Motors para viajar pelos EUA fazendo testes de direção em um novo Corvette com motoristas selecionados aleatoriamente. Após o teste, deverá perguntar ao motorista se ele compraria um Corvette. Quantos motoristas devem ser pesquisados, para termos 97% de confiança em que a percentagem amostral não apresente erro superior a dois pontos percentuais?
3. Um fornecedor da NAPA Auto Parts deseja obter informação sobre o tempo durante o qual os proprietários de automóveis desejam conservá-los. Uma amostra aleatória de 25 proprietários de carros acusou $\bar{x}$ = 7,01 anos e s = 3,74 anos, respectivamente (com base em dados de uma pesquisa Roper). Admitindo que a amostra provenha de uma população distribuída normalmente, determine um intervalo de 95% de confiança para a média populacional.
4. Com os mesmos dados amostrais do Exercício 3, determine um intervalo de confiança de 95% para o desvio-padrão populacional.
5. De 1475 empregados em transporte, aleatoriamente selecionados, 32,0% pertencem a sindicatos (com base em dados do U.S. Bureau of Labor Statistics). Construa um intervalo de 95% de confiança para a verdadeira proporção de todos os empregados em transporte sindicalizados. Que margem de erro (em pontos percentuais) deve ser reportada juntamente com a percentagem amostral de 32,0%?
6. Em uma pesquisa Gallup junto a 1004 adultos, 93% afirmaram que os restaurantes e bares devem recusar o atendimento a clientes que se apresentem embriagados. Se o pesquisador planeja uma nova pesquisa para confirmar que aquela percentagem continua a ser correta, quantos adultos selecionados aleatoriamente devem ser pesquisados, a fim de termos 98% de confiança de que a margem de erro é de quatro pontos percentuais?
7. Ao planejar uma nova máquina a ser utilizada em uma linha de montagem da General Motors, um engenheiro obtém as medidas dos comprimentos de braços de uma amostra aleatória de operários homens, obtendo os seguintes valores (em centímetros). Construa um intervalo de confiança de 95% para o comprimento médio dos braços desses operários.

76,8	75,6	69,3	75,7	75,5	71,2	72,5	71,9
70,9	69,4	71,7	72,5	72,2	68.5	75,9	73,0

8. As notas de um teste verbal PSAT de uma amostra aleatória de 40 candidatos a uma faculdade têm média de 40,7, desvio-padrão de 10,2 e distribuição normal (com base em dados do Educational Testing Service). Determine uma estimativa intervalar de 99% de confiança para a média populacional.
9. Um pesquisador médico deseja estimar o nível de colesterol (em mg/100mL) de todas as mulheres na faixa etária de 18 a 24. Há forte evidência de que σ = 41,0 mg/100mL (com base em dados de uma pesquisa feita com 1524 mulheres com 18 a 24 anos de idade, como parte da National Health Survey). Se o pesquisador deseja ter 95% de confiança em obter uma amostra que não apresente erro superior a quatro unidades, qual deve ser o tamanho da amostra?
10. Em uma pesquisa Roper junto a 1.998 adultos selecionados aleatoriamente, 24% incluíam anúncios comerciais entre os aspectos desagradáveis da televisão. Construa um intervalo de confiança de 99% para a percentagem de todos os adultos a quem os comerciais da TV desagradam.

Exercícios Cumulativos de Revisão

1. O Corpo de Fuzileiros Navais dos EUA está revisando seus pedidos referentes a uniformes por ter constatado uma sobra de uniformes para recrutas de porte mais elevado e uma falta de uniformes para recrutas mais baixos. Essa revisão envolve a amostra aleatória de alturas de recrutas homens entre 18 e 24 anos, conforme relacionadas (em polegadas):

69,9	69,4	72,6	70,0	70,2	71,8	70,6	72,8
69,0	68,4	68,3	69,6	71,7	69,2	70.8	71.0
70,4	66,8	69,9	69,2	70,5	70,2	70,8	70,0

Determine:
a. média b. mediana c. moda d. ponto médio
e. amplitude f. variância g. desvio-padrão
h. Q_1 i. Q_2 j. Q_3
k. Qual é o nível de mensuração desses dados (nominal, ordinal, intervalar, razão)
l. Construa um diagrama em caixas para os dados.
m. Construa um histograma e identifique sua forma geral.
n. Construa um intervalo de 99% de confiança para a média populacional.
o. Construa um intervalo de 99% de confiança para o desvio-padrão σ.
p. Determine o tamanho da amostra necessária para estimar a altura média de modo que haja 99% de confiança de que a média amostral não apresente erro superior a 0,2 in. Use o desvio-padrão s da parte (g) como estimativa do desvio-padrão populacional σ.
q. Selecionados da população geral homens na faixa etária 18-24, suas alturas acusam distribuição normal com média de 69,7 in. e desvio-padrão de 2,8 in. (conforme dados do National Health Survey). Com base nos resultados precedentes, as alturas da amostra de recrutas (relacionados anteriormente) concordam com os parâmetros populacionais? Explique.

2. Um geneticista constatou que, para certos casais, há uma probabilidade de 0,25 de um filho apresentar um certo distúrbio recessivo ligado ao cromossoma X.
a. Ache a probabilidade de que, em 200 dessas crianças, pelo menos 65 tenham aquele distúrbio.
b. Estudo subseqüente de 200 nascimentos revelou que 65 das crianças são portadoras daquele distúrbio. Com base nesses resultados amostrais, construa um intervalo de confiança de 95% para a proporção de todas as crianças que apresentam o distúrbio.
c. Com base nas partes (a) e (b), parece correta a avaliação do geneticista, de uma probabilidade de 0,25? Explique.

Projeto de Computador

O *método de reamostragem* pode ser utilizado para construir intervalos de confiança para situações em que os métodos tradicionais não podem (ou não devem) ser usados. A seguinte amostra aleatória de 10 valores, por exemplo, foi extraída de uma distribuição extremamente não-normal, de modo que os métodos previamente estudados não podem ser aplicados.

2,9 564,2 1,4 4.7 67,6 4,8 51,3 3,6 18,0 3,6

Os métodos deste capítulo exigem que a população tenha uma distribuição pelo menos aproximadamente normal. O método de reamostragem descrito a seguir não faz qualquer suposição sobre a população original e exige, tipicamente, um computador para construir uma nova população replicando (duplicando-a) uma amostra muitas vezes. Podemos fazer extrações com reposição, criando assim uma aproximação da população original. Com os dados amostrais anteriores, construa uma estimativa intervalar de 95% de confiança para a média populacional m, seguindo os passos descritos a seguir.

a. Crie 500 novas amostras, cada uma de tamanho 10, selecionando 10 notas, com reposição das 10 relacionadas. Com o Minitab, primeiro introduza as notas na coluna C1; introduza a seguir as probabilidades 0,1, 0,1, ..., 0,1 (dez vezes) na Coluna C2. Selecione `Calc` da barra principal de menu; selecione então `Random Data` e, em seguida, `Discrete`. Passe a gerar 500 linhas de dados, armazenados nas colunas C11-C20, com valores em C1 e probabilidades em C2. Clique OK.
b. Ache as médias das 500 amostras geradas na parte (a). Selecione `Calc, Row Statistics, Mean` e introduza as variáveis de C11-C20 com resultados a serem armazenados em C21; clique OK.
c. Ordene as 500 médias. Selecione `Manip` da barra de menu principal, escolha a opção de `Sort` e passe a ordenar a coluna C21; armazene a coluna ordenada em C21 e ordene pela coluna C21. Clique OK.
d. Ache os percentis $P_{2,5}$ e $P_{97,5}$ para as médias ordenadas resultantes do passo precedente ($P_{2,5}$ é a média dos 12° e 13° valores na lista ordenada da coluna C21; $P_{97,5}$ é a média dos 487° e 488° valores da

coluna C21). Identifique o intervalo de confiança resultante levando os valores de $P2_{,5}$ e $P_{97,5}$ em $P_{2,5} < \mu < P_{97,5}$. Este intervalo de confiança contém o verdadeiro valor de μ, que é 148?

Utilize agora o mesmo método para achar um intervalo de 95% de confiança para o desvio-padrão populacional σ. (Siga os mesmos passos relacionados antes, mas especifique *desvio-padrão* em lugar de *média* na parte b.) Compare seu resultado com o intervalo $318{,}4 < \sigma < 1079{,}6$, obtido com a aplicação incorreta dos métodos descritos na Seção 6-5. (Sua utilização é incorreta porque a distribuição da população é bastante diferente da normal.) Este intervalo de confiança incorreto para σ não contém o verdadeiro valor de σ, que é 232,1. O método de reamostragem dá um intervalo de confiança para σ que contenha 232,1, confirmando assim que esse método é eficiente?

Uma alternativa para o uso do Minitab consiste em usar programas especiais elaborados especificamente para métodos de carga de reamostragem. O autor recomenda Resampling Stats, que pode ser encontrado em Resampling Stats, Inc., 612 N. Jackson St., Arlington, VA 22201.

DOS DADOS À DECISÃO

Ele Está Zangado, Mas Tem Razão?

Dá-se a seguir um extrato de uma carta dirigida pelo presidente de uma empresa à Associated Press.

> **Quando você ou qualquer outra pessoa tenta dizer a mim e a meus associados que 1223 pessoas representam as opiniões e preferências na América, quase enlouqueço. Como ousam? Dizer que 1223 pessoas representam a América é algo assombroso e injusto, e quem o diz deveria ser banido.**

O missivista alega então que, como o tamanho da amostra de 1223 pessoas representa 120 milhões de pessoas, sua carta representa 98.000 pessoas (120 milhões divididos por 1223) que compartilham do mesmo ponto de vista.

a. Como o tamanho da amostra é 1223 e o grau de confiança é 95%, determine a margem de erro para a proporção. Suponha que não haja qualquer conhecimento prévio do valor dessa proporção.

b. O autor da carta adota a posição de que uma amostra de 1223 extraída de uma população de 120 milhões é demasiadamente pequena para significar alguma coisa. O leitor concorda ou discorda? Escreva uma resposta que apóie ou refute a posição do autor da carta de que a amostra é demasiadamente pequena.

c. O autor da carta alega também que, como a pesquisa de 1223 pessoas foi projetada de modo a refletir a opinião de 120 milhões, 1 pessoa qualquer representa 98.000 outras pessoas. Como o autor é 1 pessoa, pretende representar 98.000 outras pessoas. Essa alegação é correta? Explique.

ATIVIDADES EM GRUPO

1. *Atividades Extraclasse*: Estime o erro médio (em segundos) em relógios de pulso. Primeiro, utilizando um relógio de pulso razoavelmente preciso, acerte a hora para que ela fique exata. Use uma estação de rádio ou um serviço telefônico da hora que diz "quando soar o sinal serão ... horas etc." Se você não puder acertar o relógio para menos de um segundo, registre o erro no relógio que estiver usando. Compare então a hora em seu relógio com a hora em outros relógios. Registre as horas com sinal positivo para os relógios que estiverem adiantados, e com sinal negativo para os que estiverem atrasados. Determine estimativas pontuais e intervalos de confiança para o erro médio e o desvio-padrão dos erros. O intervalo de confiança para o erro médio contém o zero? Com base nos resultados, o que podemos concluir sobre a precisão dos relógios da população? Os desvios em relação à hora certa são simples flutuações, ou há outros fatores a serem considerados?
2. *Atividade em Classe*: Forme grupos com aproximadamente 10 alunos cada um. Tome o cronômetro de reação da Primeira Atividade em Grupo do Capítulo 5 e meça o tempo de reação de cada elemento do grupo. (Os estudantes destros devem utilizar sua mão direita, e os canhotos, a mão esquerda.) Utilize os métodos deste capítulo para estimar o tempo médio de reação para todos os estudantes de faculdades. Construa uma estimativa intervalar de 90% de confiança para essa média. Compare os resultados com os de outros grupos.
3. *Atividade em Classe*: Divida em grupos de três ou quatro. Suponha que deva ser feita uma pesquisa com estudantes de tempo integral em sua faculdade, com o objetivo de identificar a percentagem dos estudantes que são eleitores.* Identifique uma margem razoável de erro, escolha um grau de confiança adequado e ache o tamanho mínimo da amostra. Descreva um plano de amostragem capaz de dar bons resultados.
4. *Atividade em Classe*: Forme grupos de aproximadamente 10 estudantes cada um. Primeiro, cada elemento do grupo deve escrever uma estimativa da quantia média que os estudantes do grupo têm no bolso. Em seguida, cada membro do grupo deve indicar a quantia que realmente tem. (As importâncias devem ser escritas anonimamente em folhas de papel separadas, que são então misturadas, não comprometendo a privacidade de ninguém.) Calcule $\bar{x}$ e s e construa uma estimativa intervalar de 95% de confiança para a média μ. Descreva com precisão a população que está sendo estimada. Qual o membro do grupo que está mais próximo do valor de $\bar{x}$? Compare os resultados com os de outros grupos. Os resultados são consistentes, ou há grandes variações entre os resultados dos grupos?

*Nos EUA não se é obrigado a votar.

entrevista

Barry Cook

Vice-Presidente Sênior da Nielsen Media Research

Barry Cook é Vice-Presidente Sênior e Chefe de Pesquisas da Nielsen Media Research. Ensinou em Yale e no Hunter College e trabalhou para a NBC e a USA Cable Network. Tem atualmente a seu cargo o sistema de avaliação da Nielsen, efetuando pesquisas para melhor compreender como funcionam as mensurações e para desenvolver novos sistemas de mensurações.

Como o senhor se envolveu em seu trabalho atual com pesquisas para os meios de comunicação?

Eu costumava ensinar estatística. Antes de entrar no aspecto comercial do negócio de pesquisas, estive no departamento de psicologia de Yale e no departamento de sociologia do Hunter College. Lecionava um curso introdutório de estatística a cada semestre. Após lecionar por vários anos, trabalhei em um programa de pesquisa social aplicada, em que treinava candidatos à pesquisa de mercado. Observei para onde meus alunos estavam indo e, como me parecia uma boa opção, segui-os. Entrei na pesquisa de meios de comunicação na NBC fazendo pesquisa de mercado para seus programas de notícias. Posteriormente, fiz trabalhos de avaliação na NBC. Em seguida, envolvi-me com uma rede de televisão a cabo e com a Nielsen.

Quais as tendências principais que observa na maneira como os americanos assistem à TV?

Em 1985, a residência média tinha acesso a 18,8 canais; em 1990 esse número subiu para 33,2 canais. Isto obviamente influi no que o espectador escolhe para assistir.

Qual é o seu tamanho de amostra?

Para pesquisas nacionais, utilizamos pesquisadores em 4000 residências com cerca de 11.000 moradores. Aumentamos o tamanho da amostra porque o uso da televisão se modificou. Em lugar de apenas três fontes principais de TV (ABC, NBC, CBS), há atualmente dúzias de fontes de programação, muitas das quais têm apenas uma pequena parcela da audiência. A fim de avaliar essas parcelas menores com precisão, torna-se necessário uma amostra maior. Além disso, também temos contadores em 25 mercados; os aparelhos de televisão (mas não as pessoas) são contados em 250 a 500 residências por mercado. A Nielsen ainda é muito grande no negócio diário; não para assistência em âmbito nacional, mas para avaliar assistência em cerca de 200 mercados separados através do país. Esses diários atingem um tamanho amostral combinado de 100.000, quatro vezes por ano.

Tem experimentado maior resistência a consultas e pesquisas?

Não resta dúvida de que temos assistido a um declínio na cooperação tanto em pesquisas telefônicas como em contactos pessoais — em toda a indústria de pesquisas. Há preocupação com a privacidade. Confundem-se os esforços para coleta de dados com os esforços para venda, e isto provavelmente contribui para um declínio na taxa de cooperação. Por outro lado, as secretárias eletrônicas tornam mais difíceis os contactos pessoais.

Costuma ponderar os resultados amostrais de modo a refletirem melhor os parâmetros populacionais?

Temos uma política contra isto. Procuramos representar os parâmetros populacionais fazendo a amostragem corretamente, em primeiro lugar —, de tal maneira que haja a mesma probabilidade de escolha de todas as residências nos 50 estados. Como resultado, há um volume conhecido de erro amostral e também um volume desconhecido de erro não-amostral, mas fizemos uma pesquisa de validação para estimar a aderência das amostras às medidas da população.

Quais são alguns dos métodos estatísticos específicos que utiliza?

Utilizamos vários métodos, para que não só nós, mas também nossos clientes, compreendamos a amostragem e as tendências. Entramos no teste de hipóteses quando procuramos compreender por que as coisas se modificam. Os intervalos de confiança têm grande importância na interpretação de estimativas da população.

Poderia citar uma estratégia de programação de TV baseada em resultados de pesquisas?

A estratégia mais importante é chamada "tempo principal" (*prime time*). A maior utilização da TV ocorre naquelas horas da noite em que quase todos estão em casa. A estratégia mais geral de programação consiste em apresentar os melhores programas quando há maior número de espectadores. Com minisséries, o que se obtém no primeiro episódio serve quase como indicador do que se pode obter em seguida. É conveniente obter a audiência potencial máxima possível para a primeira parte, e para isto serve o domingo à noite.

Rayane Andrade
MQ - Introd. à Estatística

7 Triola

Teste de Hipóteses

7-1 Aspectos Gerais

Definem-se os objetivos do capítulo. Este capítulo introduz os conceitos e processos básicos para testar afirmações sobre parâmetros populacionais. Os processos de teste de hipóteses deste capítulo e os processos de estimação do Capítulo 6 são dois tópicos fundamentais da inferência estatística.

7-2 Fundamentos do Teste de Hipóteses

Apresenta-se um exemplo informal de um teste de hipótese, descrevendo-se suas componentes importantes. Abordam-se também os tipos de erros que podem ocorrer.

7-3 Teste de uma Afirmação sobre uma Média: Grandes Amostras

Apresenta-se a abordagem tradicional do teste de hipóteses, juntamente com o método P e um terceiro processo baseado em intervalos de confiança. Nesta seção limitamo-nos apenas aos casos que envolvem grandes amostras ($n > 30$).

7-4 Teste de uma Afirmação sobre uma Média: Pequenas Amostras

Apresentamos aqui o processo de teste de uma afirmação sobre a média de uma população, no caso de a amostra ser pequena (30 valores ou menos) e de o desvio-padrão populacional não ser conhecido. Em tais casos, utiliza-se a distribuição t.

7-5 Teste de uma Afirmação sobre uma Proporção

Descreve-se o método para testar uma afirmação sobre uma proporção ou percentagem populacional. Utiliza-se a distribuição normal nos exemplos e exercícios desta seção.

7-6 Teste de uma Afirmação sobre um Desvio-Padrão ou uma Variância

Indica-se o método de teste de uma afirmação sobre o desvio-padrão ou a variância de uma população. Em tais testes emprega-se a distribuição qui-quadrado, descrevendo-se seu papel.

Problema do Capítulo

A temperatura média do corpo humano é de 98,6°F?

Quando perguntado sobre a temperatura média de um adulto sadio, quase todo mundo responde que é de 98,6°F. A Tabela 7-1 relaciona 106 temperaturas constantes do Conjunto de Dados 2 do Apêndice B (para o meio-dia do dia 2). Essas 106 temperaturas, obtidas por pesquisadores da Universidade de Maryland, acusam média $\bar{x} = 98{,}20°F$ e desvio-padrão $s = 0{,}62°F$. No Capítulo 6, utilizamos o mesmo conjunto de temperaturas para estimar μ, a temperatura média do corpo humano, e obtivemos o intervalo de confiança de 95%: $98{,}08°F < \mu < 98{,}32°F$. Eis o problema: os limites do intervalo de confiança, 98,08°F e 98,32°F, não contêm 98,6°F, o valor geralmente aceito como temperatura média do corpo humano. O intervalo de confiança é uma estimativa da temperatura média do corpo, mas os pesquisadores foram além, alegando que 98,6°F "deve ser desprezado como um conceito desprovido de qualquer significado para a temperatura normal do corpo." Devemos rejeitar a suposição comum, de que a temperatura média do corpo de adultos sadios é 98,6°F? Há um processo-padrão para testar tais afirmações, que será estudado neste capítulo.

Tabela 7-1 Temperaturas de 106 Adultos Sadios

98,6	98,6	98,0	98,0	99,0	98,4	98,4	98,4	98,4	98,6	98,6
98,8	98,6	97,0	97,0	98,8	97,6	97,7	98,8	98,0	98,0	98,3
98,5	97,3	98,7	97,4	98,9	98,6	99,5	97,5	97,3	97,6	98,2
99,6	98,7	99,4	98,2	98,0	98,6	98,6	97,2	98,4	98,6	98,2
98,0	97,8	98,0	98,4	98,6	98,6	97,8	99,0	96,5	97,6	98,0
96,9	97,6	97,1	97,9	98,4	97,3	98,0	97,5	97,6	98,2	98,5
98,8	98,7	97,8	98,0	97,1	97,4	99,4	98,4	98,6	98,4	98,5
98,6	98,3	98,7	98,8	99,1	98,6	97,9	98,5	98,0	98,7	98,5
98,9	98,4	98,6	97,1	97,9	98,8	98,7	97,6	98,2	99,2	97,8
98,0	98,4	97,8	98,4	97,4	98,0	97,0				

7-1 Aspectos Gerais

No Capítulo 6 introduzimos um tópico de grande importância para a inferência estatística: a estimação de parâmetros populacionais com base em estatísticas amostrais. Neste capítulo apresentamos outro tópico relevante da inferência estatística: o teste de afirmações (ou *hipóteses*) sobre parâmetros de uma população.

DEFINIÇÃO

Em estatística, uma **hipótese** é uma alegação, ou afirmação, sobre uma propriedade de uma população.

As afirmações seguintes são exemplos de hipóteses a serem testadas pelos processos apresentados neste capítulo:

- Pesquisadores médicos afirmam que a temperatura média do corpo humano não é igual a 98,6°F.
- A percentagem de motoristas hospitalizados em conseqüência de acidentes é menor no caso de carros equipados com *airbag* do que no caso de carros sem esse equipamento.
- Quando se utiliza equipamento novo na fabricação de altímetros para aviões, a variação nos erros é reduzida, de modo que os resultados apresentados são mais consistentes.

Antes de iniciar o estudo deste capítulo, deve-se ter em mente esta diretriz geral para o raciocínio estatístico:

Analisar uma amostra para distinguir entre resultados que podem ocorrer facilmente e os que dificilmente ocorrem.

Podemos explicar a ocorrência de resultados altamente improváveis dizendo que ou ocorreu efetivamente um evento raro, ou as coisas não são como supúnhamos. Vamos aplicar este raciocínio ao exemplo seguinte.

EXEMPLO A ProCare Industries, Ltd., lançou, certa vez, um produto chamado "Gender Choice" (Escolha do Sexo) que, de acordo com a propaganda, permitia que os casais aumentassem em 85% a chance de terem um filho, e em 80% a chance de terem uma filha. O Gender Choice era vendido em embalagem azul para os casais que queriam um menino, e em embalagem rosa para os que queriam menina. Suponha que se faz um experimento que inclui 100 casais que querem ter menina, e que todos eles sigam as instruções indicadas na embalagem rosa. Utilizando o bom senso, e sem qualquer método estatístico formal, que se poderia concluir sobre a eficácia do Gender Choice se as 100 crianças compreendem

a. 52 meninas
b. 97 meninas?

SOLUÇÃO

a. Normalmente, esperamos 50 meninas em 100 nascimentos. O resultado 52 meninas está próximo de 50, e assim não devemos concluir que o Gender Choice seja eficaz. Se os 100 casais não usassem métodos específicos de escolha de sexo, o resultado de 52 meninas poderia ocorrer por mera chance.
b. A ocorrência de 97 meninas em 100 nascimentos é extremamente improvável, e poderia ser explicada de duas maneiras: Ou ocorreu, por puro acaso, um evento *extremamente* raro, ou o Gender Choice é eficaz. Diante da probabilidade extremamente baixa de obter, por puro acaso, 97 meninas, a explicação mais provável é que o produto é eficaz.

O ponto-chave do exemplo precedente é que devemos concluir que o produto é eficaz somente se obtemos um número significativamente maior de meninas do que poderíamos esperar em circunstâncias normais. Embora ambos os resultados, 52 meninas e 97 meninas, estejam acima da média, o resultado 52 não é significativo, enquanto 97 meninas constitui um resultado realmente significativo.

Este rápido exemplo ilustra a abordagem básica usada no teste de hipóteses. O método formal envolve uma diversidade de termos e condições-padrão incorporadas em um processo organizado. Recomendamos começar o estudo deste capítulo lendo primeiro as Seções 7-2 e 7-3 para obter uma idéia geral de seus conceitos, e em seguida reler mais detalhadamente a Seção 7-2 a fim de familiarizar-se com a terminologia.

Uma Grande Amostra Nem Sempre É Suficientemente Boa

Não devemos utilizar dados amostrais tendenciosos em inferências, por maior que seja o tamanho da amostra. Por exemplo, em *Women and Love: A Cultural Revolution in Progress*, Shere Hite fundamenta suas conclusões em 4500 respostas recebidas através de 100.000 questionários enviados a vários grupos de mulheres. Uma amostra *aleatória* de 4500 indivíduos em geral daria bons resultados, mas a amostra de Hite é tendenciosa. É criticada por supervalorizar as mulheres que aderem a grupos e as que se sentem seguras sobre os tópicos abordados. Como a amostra de Hite é tendenciosa, suas inferências não são válidas, mesmo que 4500 possa parecer um tamanho de amostra suficientemente grande.

7-2 Fundamentos do Teste de Hipóteses

Iniciamos esta seção com um exemplo informal, e em seguida identificamos os componentes formais do método-padrão de teste de hipóteses. Na Seção 7-3 incluiremos essas componentes em um processo formal. Como esse processo envolve passos específicos, poderíamos cogitar de memorizá-los e aplicá-los mecanicamente, mas se procurarmos *entendê-los*, estaremos mais bem preparados para aplicá-los a diferentes circunstâncias, evitando erros sérios. O leitor deve estudar o exemplo que segue até compreendê-lo bem, dominando, assim, um conceito básico da estatística.

EXEMPLO No Problema do Capítulo vimos que os pesquisadores afirmam que a temperatura média do corpo humano não é igual a 98,6°F. Os pesquisadores de Maryland coletaram dados amostrais com as características: $n = 106$, $\bar{x} = 98{,}20$, $s = 0{,}62$, e distribuição de forma aproximadamente normal. Eis a questão-chave: *os dados amostrais (com $\bar{x} = 98{,}20$), constituem evidência suficiente para rejeitar a crença comum de que $\mu = 98{,}6$?*

Concluímos que há evidência suficiente para rejeitar a hipótese comumente aceita de que $\mu = 98{,}6$ porque, se a média é realmente 98,6, a probabilidade de obter a média amostral de 98,20 é aproximadamente 0,0002, isto é, demasiadamente pequena. (Mais adiante mostraremos como determinar o valor 0,0002.)

Se admitirmos que $\mu = 98{,}6$ e aplicarmos o teorema central do limite (Seção 5.5), sabemos que as médias amostrais tendem a distribuir-se normalmente com os parâmetros

$$\mu_{\bar{x}} = \mu = 98{,}6 \quad \text{(por hipótese)}$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \approx \frac{s}{\sqrt{n}} = \frac{0{,}62}{\sqrt{106}} = 0{,}06$$

Construímos a Figura 7-1 admitindo $\mu = 98{,}6$ e utilizando os parâmetros apresentados. A Figura 7-1 também mostra que se μ é realmente 98,6, então 95% de todas as médias amostrais deveriam estar entre 98,48 e 98,72. (Obtivemos os valores 98,48 e 98,72 com os métodos da Seção 5-5. Especificamente, 95% das médias amostrais deveriam estar a menos de 1,96 desvios-padrão de μ. Com $\sigma_{\bar{x}} = 0{,}06$, 95% das médias amostrais devem estar a menos de $1{,}96 \times 0{,}06 \approx 0{,}12$ de 98,6. Estar a menos de 0,12 de 98,6 equivale a estar entre 98,48 e 98,72.)

Eis os pontos-chave:

- A suposição comum é que $\mu = 98{,}6$.
- A amostra acusou $\bar{x} = 98{,}20$.
- Considerando a distribuição de médias amostrais, o tamanho da amostra e a magnitude da discrepância entre 98,6 e 98,20, vemos que a ocorrência de uma média amostral de 98,20 (com menos de 5% de chance) é bastante improvável, se μ é realmente 98,6.
- Há duas explicações razoáveis para a média amostral de 98,20: ou ocorreu um evento assaz raro, ou μ não é realmente 98,6. Como a probabilidade de obter uma média amostral 98,20 (quando $\mu = 98{,}6$) é tão pequena, ficamos com a explicação mais razoável: o valor de μ não é 98,6, como se supõe em geral.

A Pena de Morte Como Repressão

Um argumento comum em favor da pena de morte é que ela desencoraja a prática de crimes. Jeffrey Grogger, da Universidade da Califórnia, analisou dados diários de

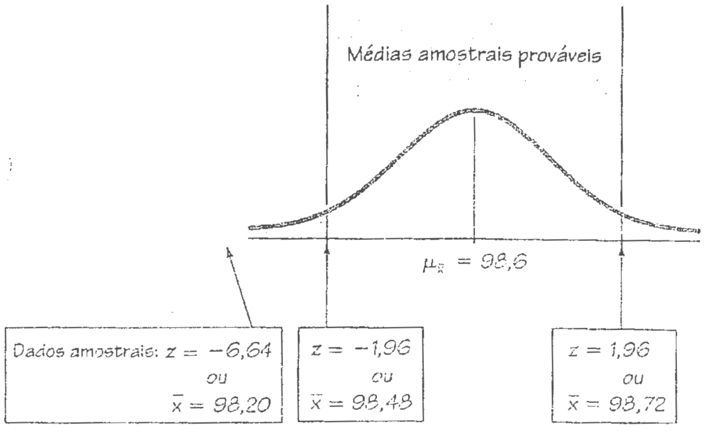


Fig. 7-1 Teorema Central do Limite.
Distribuição esperada de médias amostrais, supondo $\mu = 98{,}6$.

homicídios em um período de quatro anos, durante o qual as execuções foram freqüentes. Dentre suas conclusões publicadas no *Journal of the American Statistical Association* (Vol. 85, No. 410): "Análises efetuadas consistentemente indicam que esses dados não apóiam a hipótese de que as execuções refreiem o crime a curto prazo." Trata-se de um problema social relevante, e os esforços de pessoas como o Professor Grogger contribuem para dissipar concepções errôneas, de modo que tenhamos informações precisas para abordar tais problemas.

O exemplo precedente ilustra claramente a linha básica de raciocínio. Leia-o cuidadosamente várias vezes até compreendê-lo bem. Procure não deter-se em detalhes de cálculo, fixando-se na idéia-chave de que embora seja comumente aceito que $\mu = 98{,}6$, a média amostral é $\bar{x} = 98{,}20$. Com base no teorema central do limite, determinamos que, se a média é realmente 98,6, então a probabilidade de obter uma amostra com média de 98,20 é muito pequena, sugerindo que deve ser rejeitada a suposição de que $\mu = 98{,}6$. Na Seção 7-3 detalharemos o processo específico usado ao testar uma hipótese, mas primeiro vamos descrever as componentes de um **teste de hipótese**, ou **teste de significância** formal.

Componentes de um Teste de Hipótese Formal

- A **hipótese nula** (denotada por H_0) é uma afirmação sobre o valor de um parâmetro populacional (como a média), deve conter a condição de igualdade e deve escrever-se como =, ≤ ou ≥. (Ao fazermos efetivamente o teste, trabalhamos com a hipótese de que o parâmetro é *igual* a um valor específico.) Para a média, temos as três formas possíveis para a hipótese nula:

  H_0: μ = algum valor H_0: $\mu \leq$ algum valor
  H_0: $\mu \geq$ algum valor

  Por exemplo, a hipótese nula correspondente à suposição geral de que a temperatura média do corpo humano é 98,6°F se expressa como H_0: $\mu = 98{,}6$. Testamos a hipótese nula diretamente no sentido de que, supondo-a verdadeira, procuramos chegar a uma conclusão que nos leve a rejeitar H_0, ou não rejeitar H_0.

- A **hipótese alternativa** (denotada por H_1) é a afirmação que deve ser verdadeira se a hipótese nula é falsa. Para a média, a hipótese alternativa comporta apenas uma das três formas:

  H_1: $\mu \neq$ algum valor H_1: $\mu >$ algum valor
  H_1: $\mu <$ algum valor

  Note que H_1 é o oposto de H_0. Por exemplo, se H_0 é dada como $\mu = 98.6$, então a hipótese alternativa é H_1: $\mu \neq 98.6$.

Nota sobre o Uso de ≤ ou ≥ em H_0: Mesmo que por vezes expressemos H_0 com o símbolo ≤ ou ≥, como em H_0: $\mu \leq 98{,}6$ ou H_0: $\mu \geq 98{,}6$, fazemos o teste supondo que $\mu = 98{,}6$ seja verdadeira. Devemos ter um valor fixo único para μ, de modo que possamos trabalhar com uma única distribuição com média específica. (Alguns livros e programas de computador usam uma notação em que H_0 *sempre* contém apenas o símbolo de igualdade. Onde este livro e muitos outros usam $\mu \leq 98{,}6$ e $\mu > 98{,}6$ para H_0 e H_1, respectivamente, outros podem usar $\mu = 98{,}6$ e $\mu > 98{,}6$.)

Nota sobre a Indicação de Suas Próprias Hipóteses: Se o leitor está fazendo uma pesquisa e deseja usar um teste de hipótese para *apoiar* sua afirmação, essa afirmação deve ser formulada de maneira que se torne a hipótese alternativa, não podendo, assim, conter a condição de igualdade. Por exemplo, se o leitor acha que a marca do seu refrigerador dura mais do que a média de 14 anos das outras marcas, formule a afirmação como $\mu > 14$, onde μ é a vida média de seus refrigeradores. (Neste contexto de procurar apoiar o objetivo da pesquisa, a hipótese alternativa é às vezes designada como *hipótese de pesquisa*. Ainda neste contexto, supõe-se verdadeira a hipótese nula para se fazer o teste da hipótese, mas espera-se que a conclusão seja pela rejeição da hipótese nula, de forma a apoiar a hipótese de pesquisa.)

Nota sobre o Teste da Validade de uma Afirmação Alheia: Por vezes testamos a validade de uma afirmação alheia, como a afirmação da Coca Cola Bottling Company de que "a quantidade média de Coca em latas é no mínimo 12 oz", que passa a ser a hipótese nula H_0: $\mu \geq 12$. Neste contexto de teste da validade de uma afirmação alheia, a afirmação original às vezes se torna a hipótese nula (porque contém a igualdade) e por vezes passa a ser a hipótese alternativa (porque não contém a igualdade).

TABELA 7-2 Erros Tipo I e Tipo II

		O Verdadeiro Estado da Natureza	
		A hipótese nula é verdadeira.	A hipótese nula é falsa.
Decisão	Decidimos rejeitar a hipótese nula.	**Erro tipo I** (rejeição de uma hipótese nula verdadeira)	Decisão correta
	Não rejeitamos a hipótese nula.	Decisão correta	**Erro tipo II** (Não rejeição de uma hipótese nula falsa)

Ao testarmos uma hipótese nula, chegamos a uma conclusão: rejeitá-la ou não rejeitá-la. Tais conclusões ora são corretas, ora são incorretas (mesmo quando fazemos tudo corretamente). Há dois tipos diferentes de erro que podemos cometer. A Tabela 7-2 resume as diferentes possibilidades e mostra que tomamos uma decisão correta quando, ou rejeitamos uma hipótese nula que é falsa, ou deixamos de rejeitar uma hipótese nula que é verdadeira. Todavia, cometemos um erro quando rejeitamos uma hipótese nula verdadeira, ou deixamos de rejeitar uma hipótese nula falsa. Definem-se como se segue os erros tipo I e tipo II.

- **Erro tipo I**: Consiste em *rejeitar a hipótese nula quando ela é verdadeira.* Para o exemplo informal precedente, um erro tipo I consistiria em rejeitar a hipótese nula de que a temperatura média é 98,6, quando aquela média é, de fato, 98,6. O erro tipo I não é um cálculo malfeito ou uma fase de processo mal desempenhada; é um erro que pode ocorrer como conseqüência casual de um evento raro. *A probabilidade* de rejeitar a hipótese nula quando ela é verdadeira é chamada **nível de significância** e se denota por α (alfa). O valor de α é tipicamente predeterminado; são comuns as escolhas $\alpha = 0{,}05$ e $\alpha = 0{,}01$.
- **Erro tipo II**: Consiste em *não rejeitar a hipótese nula quando ela é falsa.* No exemplo informal precedente, um erro tipo II consistiria em não rejeitar a hipótese nula ($\mu = 98{,}6$), quando ela é, de fato, falsa (isto é, a média não é 98,6). Usa-se o símbolo β (beta) para representar a probabilidade de um erro tipo II.

Associam-se os seguintes termos às componentes principais em um processo de teste de hipótese.

- **Estatística de teste:** É uma *estatística amostral*, ou um *valor baseado nos dados amostrais*. Utiliza-se uma estatística de teste para tomar uma decisão sobre a rejeição da hipótese nula. Para os dados do exemplo informal precedente, podemos aplicar o teorema central do limite para obter a estatística de teste $z = -6{,}64$ da seguinte forma:

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}} = \frac{98{,}20 - 98{,}6}{\frac{0{,}62}{\sqrt{106}}} = -6{,}64$$

- **Região crítica:** É o *conjunto de todos os valores da estatística de teste que levam à rejeição da hipótese nula.* No exemplo precedente, a região crítica é representada pela parte sombreada da Figura 7-1 e consiste nos valores da estatística de teste inferiores a $z = -1{,}96$ ou superiores a $z = 1{,}96$.
- **Valor crítico:** *É o valor, ou valores, que separa(m) a região crítica dos valores da estatística de teste que não levam à rejeição da hipótese nula.* Os valores críticos dependem da natureza da hipótese nula, da distribuição amostral principal, e do nível de significância α. No exemplo precedente, os valores críticos $z = -1{,}96$ e $z = 1{,}96$ separam as regiões críticas sombreadas. Veja Figura 7-1.

> **Estatística: Empregos e Empregadores**
>
> Eis uma pequena amostra de empregos anunciados no campo da estatística: previsor, analista de dados, cientista de marketing, gerente de risco de crédito, pesquisador e avaliador de câncer, analista de risco de seguro, pesquisador de testes educacionais, bioestatístico, estatístico para produtos farmacêuticos, criptologista, programador estatístico.
>
> Uma pequena amostra de firmas ou entidades que oferecem emprego no campo da estatística (nos EUA): Centros de Controle e Prevenção de Doenças, Cardiac Pacemakers (marcapassos), National Institutes of Health (Institutos Nacionais da Saúde), National Cancer Institute (Instituto Nacional do Câncer), CNA Insurance Companies (companhias de seguro CNA), Educational Testing Services (Serviços de Testes Educacionais), Rosewell Park Cancer Institute, Cleveland Clinic Foundation, National Security Agency (Agência de Segurança Nacional), Quontiles, 3M, IBM, Nielsen Media Research, AT&T Labs, Bell Labs, Hewlett Packard, Johnson & Johnson, Smith Hanley.

Abordaremos a seguir alguns aspectos importantes do teste de hipóteses.

Controle dos Erros Tipo I e Tipo II Vimos que α é a probabilidade de um erro tipo I (rejeitar uma hipótese nula verdadeira) e β é a probabilidade de um erro tipo II (não rejeitar uma hipótese nula falsa). Uma das etapas de nosso processo de teste de hipóteses envolve a escolha do nível de significância α, que é a probabilidade de um erro tipo I. Entretanto, não selecionamos β (a probabilidade de um erro tipo II). Seria ótimo se pudéssemos ter sempre $\alpha = 0$ e $\beta = 0$, mas isto não é possível; devemos, pois, procurar controlar as probabilidades de erro α e β. Pode-se mostrar, matematicamente, que α, β e o tamanho da amostra n estão todos inter-relacionados, de forma que, escolhidos quaisquer dois deles, o terceiro está automaticamente determinado. Poderíamos, pois, escolher α e β (e o tamanho n da amostra estaria determinado), mas a prática comum na pesquisa e na indústria consiste em determinar previamente os valores de α e n, de modo que o valor de β fica determinado. Dependendo da seriedade de um erro tipo I, devemos tentar utilizar o maior valor tolerável de α. Para erros tipo I com conseqüências mais sérias, devemos escolher valores menores de α. Escolhe-se então um tamanho n de amostra tão grande quanto razoável em face do tempo, custo e outros fatores relevantes. (A determinação do tamanho da amostra foi estudada na Seção 6-2.) Valem as seguintes considerações de ordem prática:

1. Para α fixo, um aumento do tamanho n da amostra ocasiona uma redução de β; isto é, uma amostra maior reduz a chance de cometermos o erro de não rejeitar a hipótese nula quando ela é falsa.
2. Para um tamanho n, fixo, de amostra, uma diminuição de α acarreta um aumento de β; reciprocamente, um aumento de α acarreta uma diminuição de β.
3. Para reduzir α e β, devemos aumentar o tamanho da amostra.

Para dar sentido a essas idéias abstratas, consideremos os bombons M&M (produzidos pela Mars, Inc.) e os tabletes de aspirina Bufferin (produzidos pela Bristol-Myers Products). O pacote de bombons M&M contém 1498 unidades. O peso médio de cada bombom deve ser pelo menos 0,9085 g, porque o pacote anuncia um peso total de 1361 g. O pacote de tabletes de Bufferin contém 30 tabletes, cada um dos quais com 325 mg de aspirina. (Veja os dados amostrais dos Conjuntos 11 e 14 do Apêndice B.) Como os M&Ms são bombons, destinados ao prazer, enquanto os tabletes de Bufferin são remédios destinados ao tratamento da saúde, temos diante de nós dois níveis de seriedade completamente diferentes. Se os bombons M&M não têm um peso médio de 0,9085 g, as conseqüências são irrelevantes; mas se os tabletes de Bufferin não têm uma média de 325 mg de aspirina, as conseqüências podem ser muito sérias. Se os confeitos M&M têm uma média demasiadamente grande, a Mars terá um certo prejuízo, mas os consumidores não se queixarão. Em contraste, se os tabletes de Bufferin contêm muita aspirina, a Bristol-Myers poderá ser acionada por consumidores, enfrentando um processo da FDA (Administração Federal de Remédios). Conseqüentemente, ao testar a afirmação que $\mu = 0{,}9085$ g para os bombons M&M, poderemos escolher $\alpha = 0{,}05$, e um tamanho de amostra $n = 100$; já para testar a afirmação que $\mu = 325$ mg para os tabletes de Bufferin, podemos escolher $\alpha = 0{,}01$ e uma amostra de tamanho $n = 500$. Escolhem-se um menor nível de significância α e um maior tamanho n de amostra em razão das conseqüências mais sérias associadas a um remédio.

Conclusões no Teste de Hipóteses

Já vimos que a afirmação original, ou básica, ora se torna a hipótese nula, ora se transforma na hipótese alternativa. Todavia, nosso processo exige que sempre testemos a hipótese nula. Na Seção 7-3 veremos que nossa conclusão inicial será sempre uma das seguintes:

1. Não rejeitar a hipótese nula H_0.
2. Rejeitar a hipótese nula H_0.

A conclusão de rejeitar a hipótese nula, ou não rejeitá-la, é ótima para os que tomaram a sábia decisão de fazer um curso de estatística, mas devemos mostrar em termos simples, não-técnicos, o que a conclusão sugere. Os estudantes em geral têm dificuldade em formular esta afirmação final não-técnica, que descreve a conseqüência prática dos dados e dos cálculos. A linguagem utilizada deve ser precisa; as implicações de palavras ou expressões como "apoiar" e "não rejeitar" são muito diferentes. A Figura 7-2 mostra como formular corretamente a conclusão final. Note que apenas um caso conduz a uma frase que indica que os dados amostrais efetivamente *apóiam* a conclusão. Se

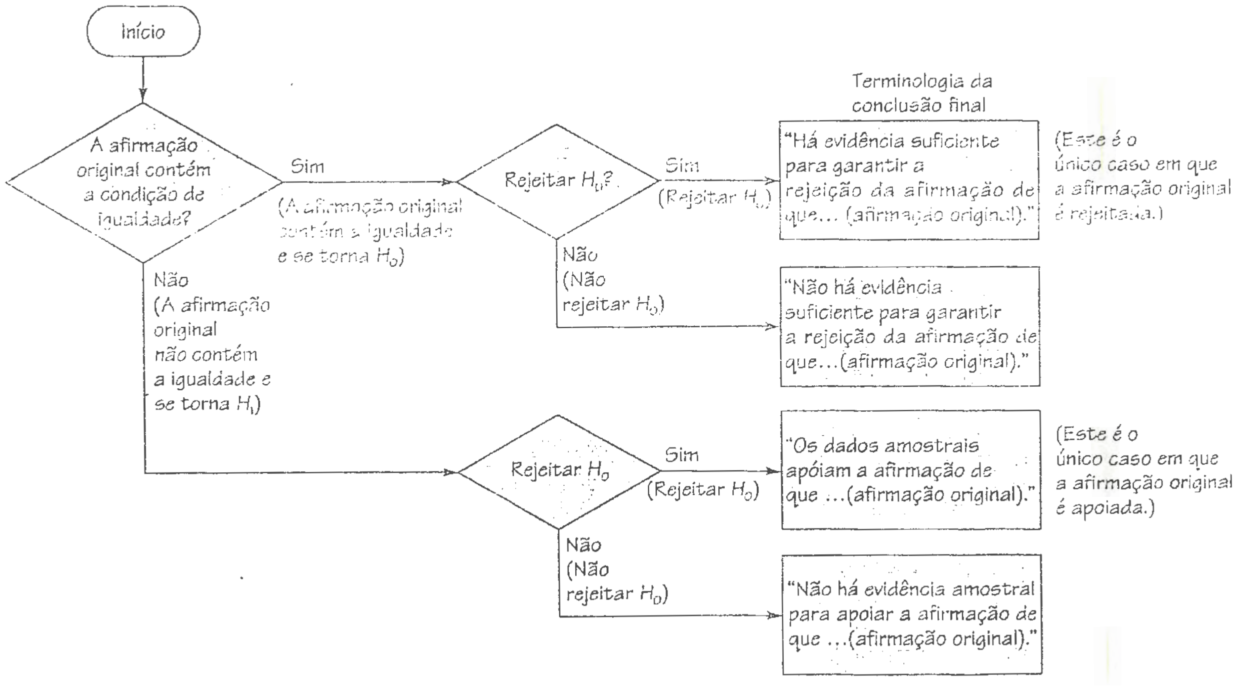


Fig. 7-2 Terminologia das conclusões no teste de hipóteses.

queremos justificar uma afirmação, devemos formulá-la de maneira que ela se torne a hipótese alternativa; esperamos então que a hipótese nula seja rejeitada. Por exemplo, se queremos justificar a afirmação de que a temperatura média do corpo humano é diferente de 98,6°F, fazemos a afirmação $\mu \neq 98{,}6$. Esta afirmação será uma hipótese alternativa que será apoiada se rejeitarmos a hipótese nula H_0: $\mu = 98{,}6$. Se, por outro lado, alegamos que $\mu = 98{,}6$, ou rejeitamos, ou não rejeitamos a afirmação; em qualquer caso, não *apoiaremos* a afirmação original.

Alguns textos dizem "aceitar a hipótese nula" em vez de "não rejeitar a hipótese nula". Quer empreguemos *aceitar* ou *não rejeitar*, devemos reconhecer que *não estamos provando a hipótese nula*; estamos apenas dizendo que a evidência amostral não é suficientemente forte para recomendar a rejeição da hipótese nula. É como um júri decidir que não há evidência suficiente para condenar um acusado. O termo *aceitar* pode ser enganoso, porque parece implicar incorretamente que a hipótese nula foi provada. A expressão *não rejeitar* diz mais corretamente que a evidência de que dispomos não é bastante forte para recomendar a rejeição da hipótese nula. Neste texto usaremos a conclusão *não rejeitar a hipótese nula* em lugar de *aceitar a hipótese nula.*

inferior superior

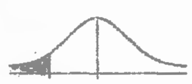
Bilateral, Unilateral Esquerdo, Unilateral Direito

As *caudas* em uma distribuição são as regiões extremas delimitadas por valores críticos. Nosso exemplo informal de teste de hipótese envolveu um **teste bilateral** no sentido de que a região crítica da Figura 7-1 está situada nas duas regiões extremas (caudas) sob a curva. Rejeitamos a hipótese nula H_0 se nossa estatística de teste está na região crítica, porque isto indica uma discrepância significativa entre a hipótese nula e os dados amostrais. Alguns testes são **unilaterais esquerdos**, com a região crítica localizada na região extrema esquerda sob a curva. Outros podem ser **unilaterais direitos**, com a região crítica localizada na região extrema direita sob a curva.

Nos testes bilaterais, o nível de significância α é dividido igualmente entre as duas caudas que constituem a região crítica. Por exemplo, em um teste bilateral, com nível de significância $\alpha = 0{,}05$, há uma área de 0,025 em cada uma das duas caudas. Em testes unilaterais direitos ou esquerdos, a área da região crítica é α.

Examinando a hipótese nula H_0, devemos poder deduzir se se trata de um teste unilateral direito, unilateral esquerdo ou bilateral. A cauda corresponderá à região crítica que contém os valores significativamente conflitantes com a hipótese nula. As figuras resumem uma verificação conveniente e mostram como o sinal de desigualdade em H_1 aponta na direção da região crítica. Em linguagem de programação, o símbolo $\neq$ é freqüentemente expresso como <>, e isto nos faz lembrar que uma hipótese alternativa como $\mu \neq 98{,}6$ corresponde a um teste bilateral.

Sinal de H_1: >
Teste unilateral direito



Sinal de H_1: <
Teste unilateral esquerdo

Sinal de H_1: $\neq$
Teste bilateral

Na Seção 7-3 abordaremos processos formais de teste de hipóteses. Por ora, examinaremos as componentes incluídas nos seguintes exemplos.

EXEMPLO *Identificar as componentes de um teste de hipótese:* Após obter amostras nas bombas de gasolina no Premium Auto Service Center, a Connecticut Consumer Protection Agency afirma que os consumidores estão sendo prejudicados em virtude da condição seguinte: Quando o marcador indica 1 galão, a quantidade média de combustível fornecida é realmente inferior a 1 galão.

a. Expresse, de forma simbólica, a afirmação de que a Premium Auto Service está prejudicando os consumidores. $\mu < 1$
b. Identifique a hipótese nula H_0. $H_0: \mu \geq 1$
c. Identifique a hipótese alternativa H_1. H_1 $\mu < 1$
d. Identifique este teste como bilateral, unilateral direito ou unilateral esquerdo. unil. esq
e. Identifique o erro tipo I para este teste.
f. Identifique o erro tipo II para este teste.
g. Suponha que a conclusão seja rejeitar a hipótese nula. Enuncie a conclusão em termos não-técnicos; certifique-se de que está abordando a afirmação original.
h. Suponha que a conclusão seja não rejeitar a hipótese nula. Enuncie a conclusão em termos não-técnicos; certifique-se de que está abordando a afirmação original.

SOLUÇÃO

a. A afirmação de que os consumidores estão sendo prejudicados é equivalente a afirmar que a média é inferior a 1 gal, o que, em forma simbólica, se expressa como $\mu < 1$ gal.
b. A afirmação original $\mu < 1$ gal não contém a igualdade, conforme exigido pela hipótese nula. A afirmação original é, pois, a hipótese alternativa; a hipótese nula é H_0: $\mu \geq 1$.
c. Veja a parte (b). A hipótese alternativa é H_1: $\mu < 1$.
d. Este teste é unilateral esquerdo, porque a hipótese nula é rejeitada se a média amostral é significativamente inferior a 1 (está à esquerda de 1). (Como uma **dupla verificação**, note que a hipótese alternativa $\mu < 1$ contém o sinal <, que aponta para a esquerda.)
e. O erro tipo I (rejeição de uma hipótese nula verdadeira) consiste em rejeitar a hipótese nula $\mu \geq 1$ quando a média populacional é realmente igual ou superior a 1. (Trata-se de um erro sério, porque o negócio será acusado de prejudicar os consumidores quando, na realidade, não há tal prejuízo.)
f. O erro tipo II (não rejeitar uma hipótese nula falsa) consiste em não rejeitar a hipótese nula $\mu \geq 1$ quando a média populacional é realmente inferior a 1. (Isto é, concluímos que não há evidência suficiente para comprovar o prejuízo, quando esse prejuízo está efetivamente ocorrendo.)

g. Veja a Figura 7-2 para o caso em que a afirmação original não contém a igualdade, mas H_0 é rejeitada. Concluir que há evidência suficiente para apoiar a afirmação de que a média é inferior a 1.

h. Veja a Figura 7-2 para o caso em que a afirmação original não contém a igualdade, e não rejeitamos H_0. Concluir que não há evidência suficiente para apoiar a afirmação de que a média é inferior a 1.

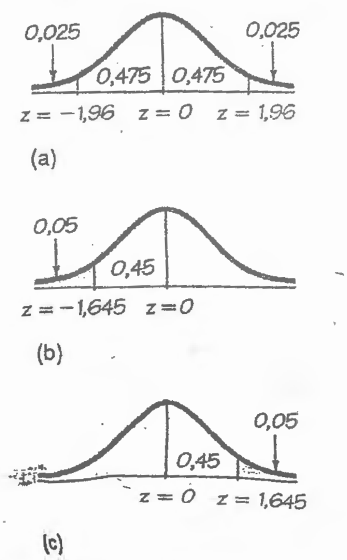
EXEMPLO *Determinação de Valores Críticos:* Muitos passageiros de navios de cruzeiro utilizam adesivos que fornecem dramamina ao corpo a fim de evitar o enjôo. Testa-se uma afirmação sobre a quantidade da dosagem média, ao nível de significância $\alpha = 0{,}05$. As condições são tais que permitem a utilização da distribuição normal (porque o teorema central do limite é aplicável). Determine o(s) valor(es) crítico(s) de z se o teste é (a) bilateral, (b) unilateral esquerdo e (c) unilateral direito.

SOLUÇÃO

a. Em um teste bilateral, o nível de significância de $\alpha = 0{,}05$ é dividido igualmente entre as duas caudas, o que determina uma área de 0,025 em cada cauda. Podemos achar os valores críticos na Tabela A-2 como os valores que correspondem a áreas de 0,4750 (0,5 - 0,025) à direita e à esquerda da média. Obtemos os valores críticos $z = -1{,}96$ e $z = 1{,}96$, conforme se vê na Figura 7-3(a).

b. Em um teste unilateral esquerdo, o nível de significância $\alpha = 0{,}05$ é a área da região crítica à esquerda, de forma que o valor crítico corresponde a uma área de 0,4500 (0,5 − 0,05) à esquerda da média. Recorrendo à Tabela A-2, obtemos o valor crítico $z = -1{,}645$, conforme mostra a Figura 7-3(b).

c. Em um teste unilateral à direita, o nível de significância $\alpha = 0{,}05$ é a área da região crítica à direita, de forma que o valor crítico corresponde a uma área de 0,4500 (0,5 − 0,05) à direita da média. Com auxílio da Tabela A-2, obtemos o valor crítico $z = 1{,}645$, conforme Figura 7-3(c).



7-3 Determinação de valores críticos.

7-2 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-8, suponha que se faça um teste de hipótese da afirmação dada. Utilize μ para uma afirmação sobre a média, p para uma afirmação sobre uma proporção e σ para uma afirmação sobre uma variação.

a. *Expresse a afirmação em forma simbólica.*
b. *Identifique a hipótese nula H_0.*
c. *Identifique a hipótese alternativa H_1.*
d. *Identifique o teste como bilateral, unilateral esquerdo ou unilateral direito.*
e. *Identifique o erro tipo I para o teste.*
f. *Identifique o erro tipo II para o teste.*
g. *Supondo que a conclusão seja rejeitar a hipótese nula, enuncie-a em termos não-técnicos; certifique-se de que está se referindo à afirmação original. (Veja Figura 7-2.)*
h. *Supondo que a conclusão seja não rejeitar a hipótese nula, enuncie-a em termos não-técnicos; certifique-se de que está se referindo à afirmação original. (Veja Figura 7-2.)*

1. O advogado da Food and Drug Administration (FDA) afirma que a Medassist Pharmaceutical Company está fabricando comprimidos para gripe com quantidades de acetaminofena com média diferente dos 650 mg indicados no rótulo.
2. O Diretor de Admissões a uma faculdade afirma que o tempo que os estudantes levam para conquistar um diploma de bacharel tem média inferior a cinco anos.
3. O candidato republicano à presidência (dos EUA) afirma ter a preferência de mais da metade dos eleitores.
4. A Home Electronics Supply Company afirma que seus interruptores para uso doméstico têm níveis de falha com menor variação do que os interruptores fabricados por sua maior concorrente, que acusa uma variação dada por $\sigma = 0{,}4$ amp.
5. A NBC (National Broadcasting Company — EUA) afirma que tem 15% dos espectadores de 22 às 23 horas nas segundas-feiras.
6. A fim de atrair novos negócios, a Orange County Chamber of Commerce afirma que a renda média anual de sua região é superior a $50.000,00.
7. Um engenheiro da Ford alega que seu novo injetor de combustível eleva a quilometragem média do Taurus acima do nível atual de 30 mi/gal.
8. O comprador do New England Hospital Supply Company recomenda não comprar os novos termômetros digitais, porque variam mais do que os velhos termômetros, que acusam um desvio-padrão de 0,06°F.

Nos Exercícios 9-16, determine os valores críticos z para as condições dadas. Em cada caso, admita que se aplica a distribuição normal, podendo-se usar a Tabela A-2. Esboce também um gráfico mostrando o valor crítico e a região crítica.

9. Teste unilateral direito; $\alpha = 0{,}05$
10. Teste unilateral esquerdo; $\alpha = 0{,}05$
11. Teste unilateral esquerdo; $\alpha = 0{,}01$
12. Teste bilateral; $\alpha = 0{,}01$
13. Teste bilateral; $\alpha = 0{,}10$
14. Teste unilateral direito; $\alpha = 0{,}025$
15. Teste unilateral esquerdo; $\alpha = 0{,}025$
16. Teste bilateral; $\alpha = 0{,}02$

7-2 Exercícios B: Além do Básico

17. Alguém sugere que, no teste de hipóteses, é possível eliminar um erro tipo I fazendo-se $\alpha = 0$. Em um teste bilateral, que valores críticos correspondem a $\alpha = 0$? Se $\alpha = 0$, a hipótese nula pode ser rejeitada?
18. Suponha o leitor que esteja trabalhando com um nível de significância $\alpha = 0{,}05$ para testar a hipótese $\mu < 2$, e que sua amostra consista em 50 valores selecionados aleatoriamente. Determine a probabilidade de cometer um erro tipo II, dado que a população tem efetivamente distribuição normal com $\mu = 1{,}5$ e $\sigma = 1$. (*Sugestão*: Com H_0: $\mu \geq 2$, comece determinando os valores das médias amostrais que não levem à rejeição de H_0; ache em seguida a probabilidade de obter uma média amostral com um daqueles valores.)

7-3 Teste de uma Afirmação Sobre uma Média: Grandes Amostras

Vamos formalizar, nesta seção, o processo informal utilizado na seção precedente. Apresentaremos três métodos de teste de hipóteses que, embora aparentemente diferentes, são na realidade equivalentes, no sentido de que sempre levam às mesmas conclusões. O primeiro processo é o tradicional, e será o mais utilizado em todo o restante deste livro. O segundo processo se baseia nos valores P, e a ele vamos referir-nos freqüentemente; o terceiro processo, baseado em intervalos de confiança, será usado apenas em casos especiais.

Inicialmente identificaremos as duas hipóteses que se aplicam aos métodos desta seção.

Hipóteses para o Teste de uma Afirmação sobre a Média de uma (única) População

1. A amostra é grande ($n > 30$); pode-se aplicar o teorema central do limite e utilizar a distribuição normal.
2. Ao aplicar o teorema central do limite, podemos utilizar o desvio-padrão amostral s em substituição ao desvio-padrão populacional σ quando este não for conhecido e o tamanho da amostra for grande ($n > 30$).

O Método Tradicional de Teste de Hipóteses

A Figura 7-4 resume as etapas do **método tradicional** (ou **clássico**) do teste de hipóteses. Este processo utiliza as componentes descritas na Seção 7-2 como parte de um esquema para identificar um resultado amostral que é *significativamente* diferente do valor alegado. Uma estatística amostral importante (como $\bar{x}$) se converte em uma estatística de teste, que é comparada com um valor crítico. Para o teste de hipóteses desta seção, a estatística de teste é um escore z usado para padronizar a média $\bar{x}$, assim como o escore z foi utilizado no Capítulo 2 para padronizar um escore individual x. Calcula-se como segue a estatística de teste requerida no Passo 6:

Estatística de Teste para Afirmações sobre μ Quando $n > 30$

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\dfrac{\sigma}{\sqrt{n}}}$$

Note que, no Passo 7 da Figura 7-4, utiliza-se este critério de decisão:

Rejeitar a hipótese nula se a estatística de teste está na região crítica.
Não rejeitar a hipótese nula se a estatística de teste não está na região crítica.

Na Seção 7-2 apresentamos um exemplo informal de um teste de hipóteses; o exemplo que segue formaliza esse teste.

EXEMPLO Com os dados amostrais do início do capítulo ($n = 106$, $\bar{x} = 98{,}20°$, $s = 0{,}62$) e um nível de significância de 0,05, teste a afirmação de que a temperatura média do corpo de adultos sadios é 98,6°F. Aplique o método tradicional, seguindo o processo delineado na Figura 7-4.

SOLUÇÃO Recorra à Figura 7-4 e siga os passos abaixo:

Passo 1: A afirmação de que a média é igual a 98,6 é expressa simbolicamente como $\mu = 98{,}6$.

Passo 2: A alternativa (em forma simbólica) da afirmação original é $\mu \neq 98{,}6$.

Passo 3: A afirmação $\mu = 98{,}6$ contém a condição de igualdade, tornando-se, assim, a hipótese nula. Temos

H_0: $\mu = 98{,}6$ (afirmação original) $\quad H_1$: $\mu \neq 98{,}6$

Passo 4: Conforme especificado no enunciado do problema, o nível de significância é $\alpha = 0{,}05$.

Passo 5: Como a afirmação se refere à média populacional μ, a estatística amostral mais importante para este teste é $\bar{x} = 98{,}20$. E como $n > 30$, o teorema central do limite indica que a distribuição das médias amostrais pode ser aproximada por uma distribuição *normal*.

Passo 6: Ao calcular a estatística de teste, podemos utilizar $s = 0{,}62$ como uma estimativa razoável de σ (porque $n > 30$); obtemos a estatística de teste $z = -6{,}64$ transformando a média amostral $\bar{x} = 98{,}20$ em $z = -6{,}64$ como segue:

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\dfrac{\sigma}{\sqrt{n}}} = \frac{98{,}20 - 98{,}6}{\dfrac{0{,}62}{\sqrt{106}}} = -6{,}64$$

Obtemos os valores z notando primeiro que o teste é bilateral, porque uma média amostral significativamente inferior ou superior a 98,6 constitui forte evidência contra a hipótese nula $\mu = 98{,}6$. Dividimos, pois, $\alpha = 0{,}05$ igualmente entre as duas caudas, obtendo 0,025 em cada uma. Recorremos então à Tabela A-2 (porque o teorema do limite central permite supor que a distribuição é normal), obtendo o valor z correspondente a $0{,}5 - 0{,}025 = 0{,}4750$. Após achar $z = 1{,}96$, concluímos, por simetria, que o valor crítico esquerdo é $z = -1{,}96$. A Figura 7-5 mostra a estatística de teste, a região crítica e os valores críticos.

Passo 7: A média amostral $\bar{x} = 98{,}20$ é transformada na estatística de teste $z = -6{,}64$, que está na região crítica, o que nos leva a rejeitar a hipótese nula.

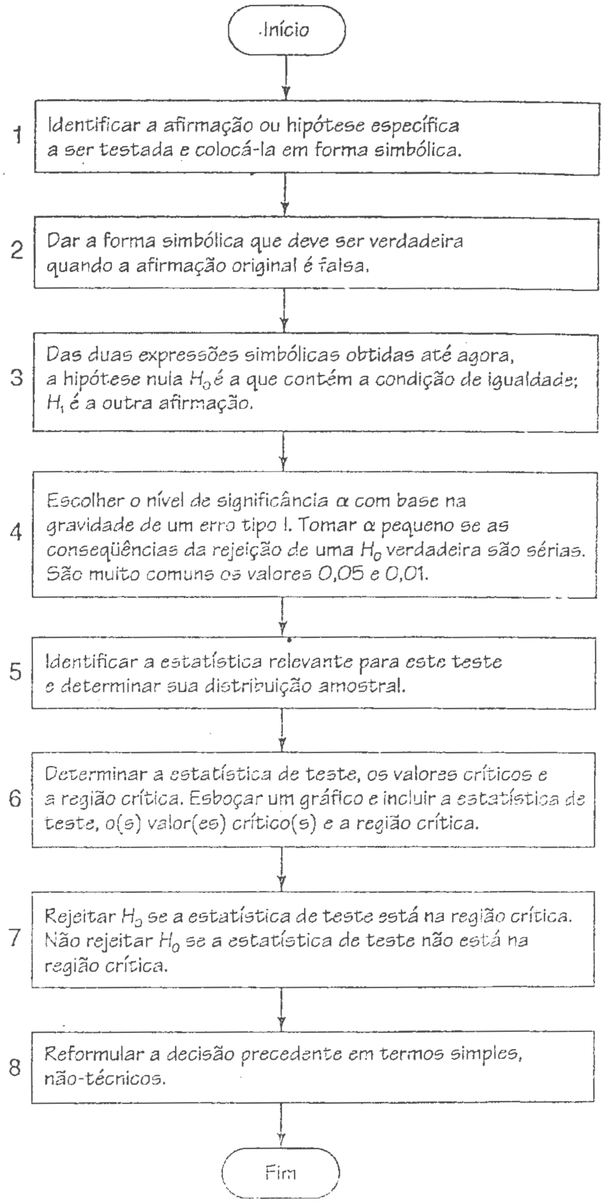


Fig. 7-4 Método tradicional de teste de hipóteses.

Passo 8: Para reformular a conclusão do Passo 7 em termos não-técnicos, podemos recorrer à Figura 7-2 da seção precedente. Estamos rejeitando a hipótese nula, que é a afirmação original. Concluímos que há evidência suficiente para justificar a rejeição da afirmação de que a temperatura média do corpo de um adulto sadio é 98,6°F.

No exemplo precedente, a média amostral de 98,20 conduziu a uma estatística de teste $z = -6,64$. Isto é, a média amostral dista 6,64 desvios-padrão da média alegada $\mu = 98,6$. No Capítulo 2 vimos que um valor $z = 6,64$ é bastante elevado e corresponde a um resultado incomum se, de fato, $\mu = 98,6$. Esta evidência é suficientemente forte para fazer-nos crer que a média seja efetivamente diferente de 98,6. Os valores críticos e a região crítica identificam claramente o intervalo de valores incomuns que fazem com que rejeitemos o valor alegado de μ.

O exemplo precedente ilustrou um teste bilateral; o exemplo que segue ilustra não só um teste unilateral direito como também a diferença entre *significância estatística* e *significância prática*.

EXEMPLO O Jack Wilson Health Club afirma, em um anúncio, que "você perderá peso após dois dias apenas com a dieta e o programa de exercícios Jack Wilson". O Bureau of Consumers Affairs do condado de Dade (EUA) faz um teste daquela afirmação selecionando aleatoriamente 33 pessoas que aderiram ao programa. Verificou-se que essas 33 pesso-

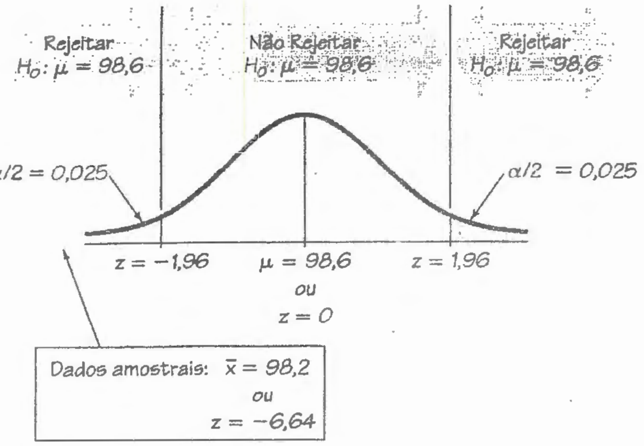


Fig. 7-5 Distribuição das temperaturas médias do corpo humano, supondo $\mu = 98,6$.

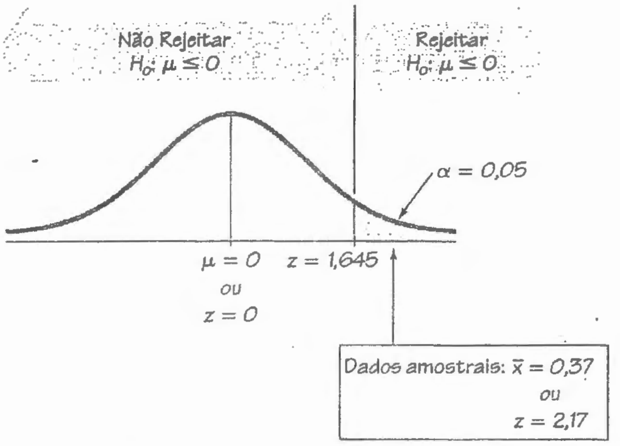


Fig. 7-6 Distribuição de perdas médias de peso.

as perderam, em média, 0,37 lb de peso, com um desvio-padrão de 0,98 lb. Teste a afirmação do anúncio ao nível de significância de 0,05.

SOLUÇÃO Sabemos que o peso das pessoas flutua naturalmente de um dia para outro; precisamos, pois, determinar se a perda média de 0,37 lb de peso representa um resultado significativo. Ou seja, a afirmação do clube será confirmada se a perda média de peso for significativamente maior do que 0. Observe a Figura 7-4 e siga os seguintes passos:

Passo 1: Expressamos como $\mu > 0$ uma perda média de peso significativamente maior do que 0.

Passo 2: O oposto da afirmação original (em forma simbólica) é $\mu \leq 0$.

Passo 3: A afirmação $\mu \leq 0$ contém a condição de igualdade, tornando-se, assim, a hipótese nula; e temos:

$$H_0: \mu \leq 0 \quad H_1: \mu > 0 \text{ (Afirmação original)}$$

Passo 4: Com o nível de significância de 0,05, temos $\alpha = 0,05$.

Passo 5: Deve-se utilizar a média amostral $\bar{x} = 0,37$ lb para testar uma afirmação a respeito da média populacional μ. Como $n > 30$, o teorema central do limite garante que a distribuição das médias amostrais pode ser aproximada por uma distribuição normal.

Passo 6: Ao calcular a estatística de teste, podemos utilizar $s = 0,98$ como estimativa razoável de σ (porque $n > 30$) e a estatística de teste $z = 2,17$ se obtém como segue:

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}} = \frac{0,37 - 0}{\frac{0,98}{\sqrt{33}}} = 2,17$$

O valor crítico $z = 1,645$ é encontrado na Tabela A-2 como o valor de z correspondente a uma área de 0,4500. (O teste é unilateral direito, porque a hipótese nula $\mu \leq 0$ é rejeitada se a média amostral $\bar{x}$ é significativamente *maior do que* 0.) A Figura 7-6 mostra a estatística de teste, a região crítica e o valor crítico.

Passo 7: Pela Figura 7-6, vemos que a média amostral 0,37 lb está efetivamente na região crítica, pelo que rejeitamos a hipótese nula H_0.

Passo 8: Há evidência suficiente para apoiar a afirmação de que a perda média de peso é superior a 0 lb; ou seja, o programa parece ser eficaz. Entretanto, mesmo que o resultado seja *estatisticamente significativo*, não parece ter *significado prático*, em virtude do diminuto valor (0,37 lb) da perda de peso.

Ao apresentar o resultado de um teste de hipóteses, nem sempre é necessário explicitar todos os passos detalhados na Figura 7-4. Entretanto, o resultado deve incluir a hipótese nula, a hipótese alternativa, o cálculo da estatística de teste, um gráfico como o da Figura 7-6, a conclusão inicial (rejeitar H_0 ou não rejeitar H_0) e a conclusão final em termos não-técnicos. O gráfico deve exibir a estatística de teste, o(s) valor(es) crítico(s), a região crítica e o nível de significância.

O Método do Valor *P* para o Teste de Hipóteses

Muitos artigos profissionais e programas de computador utilizam outra abordagem do teste de hipóteses, baseada no cálculo do valor de uma probabilidade, ou valor *P*. Dada uma hipótese nula e um conjunto de dados amostrais, o valor *P* reflete a plausibilidade de se obter tais resultados no caso de a hipótese nula ser, de fato, verdadeira. Um valor *P* muito pequeno (como 0,05 ou menos) sugere que os resultados amostrais são muito improváveis sob a hipótese nula; um valor tão pequeno de *P* constitui, pois, evidência contra a hipótese nula.

DEFINIÇÃO

Um ***valor P*** (ou **valor de probabilidade**) é a probabilidade de obter um valor da estatística amostral de teste *no mínimo tão extremo* como o que resulta dos dados amostrais, na suposição de a hipótese nula ser verdadeira.

Enquanto a abordagem tradicional resulta em uma conclusão do tipo "rejeitar/não rejeitar", os valores P dão o grau de confiança ao rejeitarmos uma hipótese nula. Por exemplo, um valor P de 0,0002 leva-nos a rejeitar a hipótese nula, mas pode também sugerir que os resultados amostrais sejam extremamente incomuns, se o valor alegado de μ é, de fato, correto. Em contraste, para um valor P de 0,40, não rejeitamos a hipótese nula, porque os resultados amostrais podem *facilmente* ocorrer se o valor alegado de μ é correto.

O processo de abordagem pelo valor P utiliza a maioria dos passos do método tradicional, mas os passos 6 e 7 são diferentes:

Passo 6: Achar o valor P.

Passo 7: Reportar o valor P. Alguns estatísticos preferem simplesmente reportar o valor P e deixar a conclusão para o leitor. Outros preferem usar o seguinte critério de decisão:

- *Rejeitar a hipótese nula* se o valor P é no máximo igual ao nível de significância α.
- *Não rejeitar a hipótese nula* se o valor P é maior do que o nível de significância α.

No Passo 7, se a conclusão se baseia somente no valor P, vale a seguinte orientação:

Valor P	Interpretação
Inferior a 0,01	Elevada significância estatística Evidência muito forte contra a hipótese nula
0,01 a 0,05	Estatisticamente significante Evidência adequada contra a hipótese nula
Superior a 0,05	Evidência insuficiente contra a hipótese nula

Cuidado com o Uso Indevido do Valor P

John P. Campbell, editor do *Journal of Applied Psychology*, escreveu o seguinte sobre os valores P: "Têm-se escritos livros para dissuadir as pessoas da noção de que valores P menores significam resultados mais importantes, ou que a significância estatística tenha algo a ver com a significância substantiva. É quase impossível dissociar os autores dos seus valores P, e quanto mais zeros após a vírgula decimal, mais as pessoas se apegam a eles." Embora possa ser necessário fazer a análise estatística dos resultados de um estudo, devemos dar forte ênfase à significância dos próprios resultados.

Muitos estatísticos consideram uma boa prática sempre escolher um nível de significância *antes* de proceder ao teste de uma hipótese, mormente quando se utilizam valores P, porque podemos nos sentir tentados a ajustar o nível de significância com base nos resultados. Por exemplo, com um nível de significância de 0,05 e um valor P de 0,06, deveríamos não rejeitar a hipótese nula, mas podemos ser levados a dizer que uma probabilidade de 0,06 é suficientemente pequena para justificar a rejeição da hipótese nula. Outros estatísticos acham que a seleção antecipada de um nível de significância reduz a utilidade dos valores P. Argumentam que não se deve especificar qualquer nível de significância, e que a conclusão deve ficar a cargo do leitor. Vamos utilizar o critério de decisão que envolve uma comparação entre um nível de significância e o valor P.

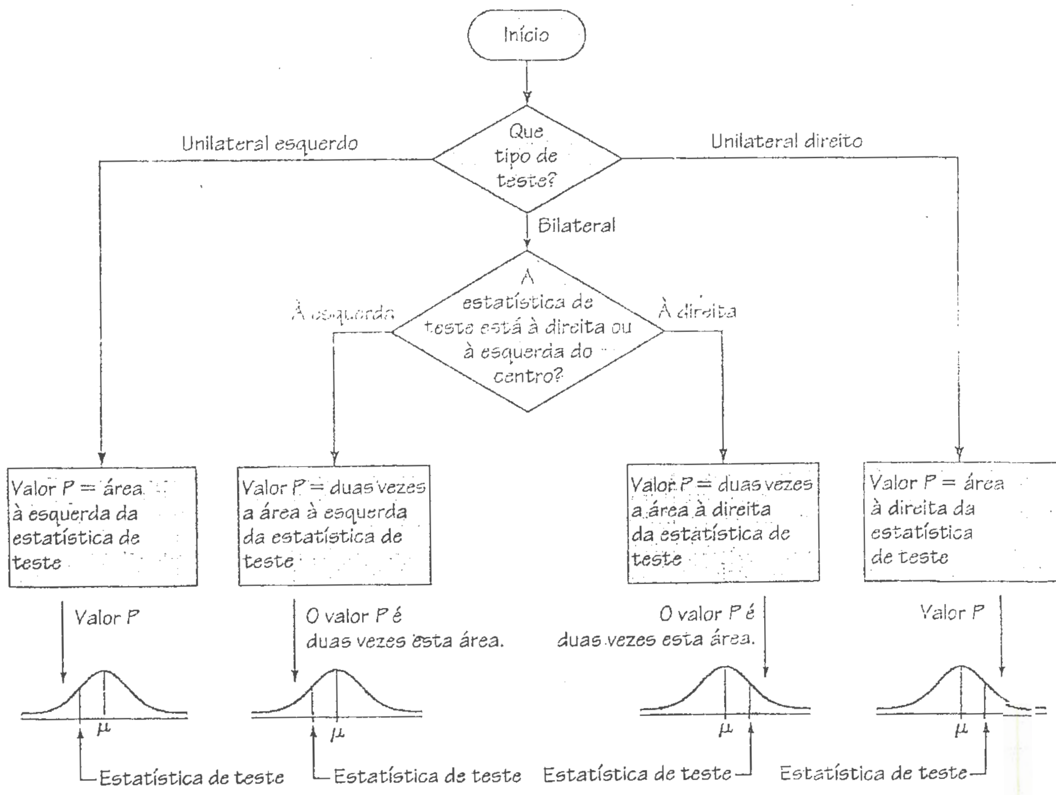


Fig. 7-7 Determinação dos valores P.

A Figura 7-7 detalha os passos e decisões que levam ao valor P. A figura mostra:

Teste unilateral direito:	O valor P é a área à direita da estatística de teste.
Teste unilateral esquerdo:	O valor P é a área à esquerda da estatística de teste.
Teste bilateral:	O valor P é *duas vezes* a área da região extrema delimitada pela estatística de teste.

A Figura 7-7 mostra como achar valores P e a Figura 7-8 resume o método do valor P para testar hipóteses. Uma comparação do método tradicional (resumido na Figura 7-4) com o método do valor P (Figura 7-8) mostra que eles são essencialmente o mesmo, diferindo apenas no critério de decisão. O método tradicional compara a estatística de teste com os valores críticos, enquanto o método do valor P compara o valor P com o nível de significância. Mas os métodos são equivalentes, no sentido de que sempre resultam na mesma conclusão.

EXEMPLO Com o auxílio do valor P, teste a afirmação de que a temperatura média dos adultos sadios é igual a 98,6°F. Como anteriormente, utilize um nível de significância de 0,05 e os dados amostrais apresentados na abertura do capítulo ($n = 106$, $\bar{x} = 98{,}20$, $s = 0{,}62$, distribuição em forma de sino).

SOLUÇÃO Afora os passos 6 e 7, a solução é a mesma que a dada anteriormente nesta seção utilizando o método tra-

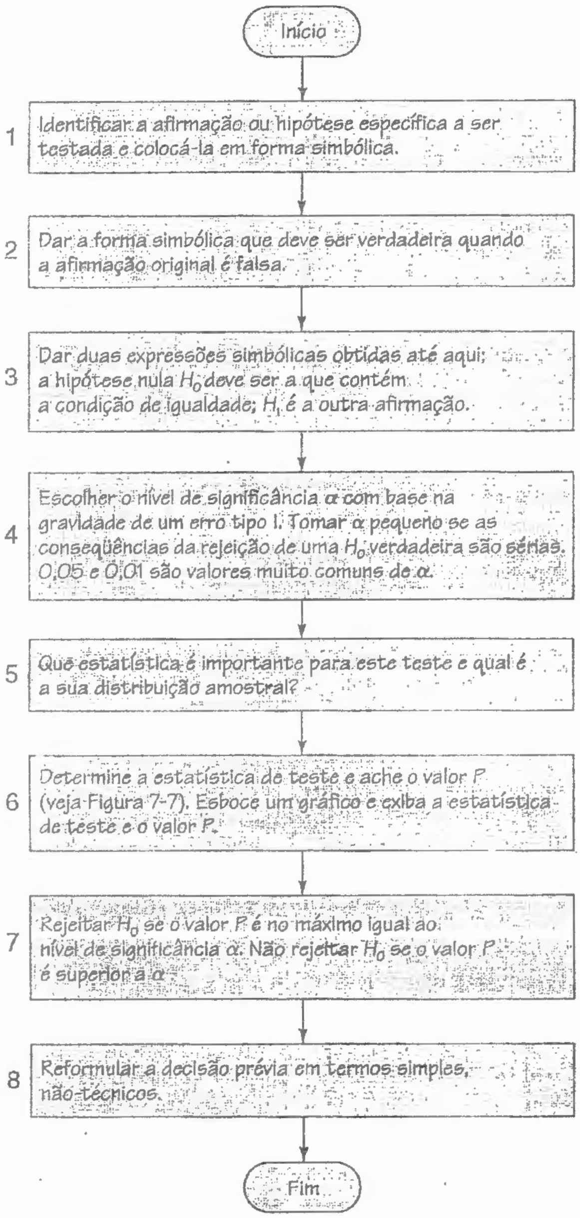


Fig. 7-8 O método dos valores P para o teste de hipóteses.

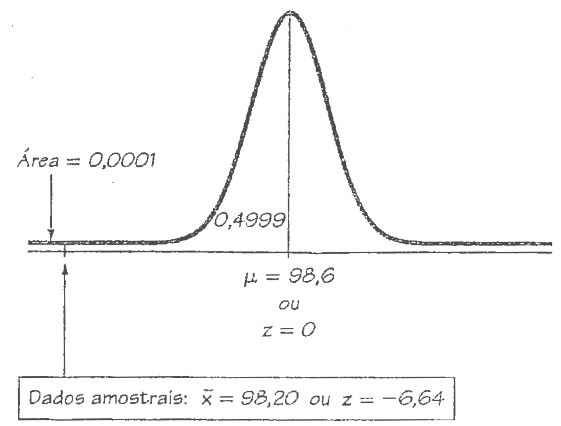


Fig. 7-9 O Método do valor P para testar H_0: $\mu = 98,6$
Como o teste é bilateral, o valor P é *duas vezes* a área sombreada.

dicional. Os passos 1, 2 e 3 conduzem às seguintes hipóteses:

H_0: $\mu = 98,6$ (afirmação original) H_1: $\mu \neq 98,6$

Nos passos 4 e 5 vimos que o nível de significância é $\alpha = 0,05$ e que o teorema central do limite indica o uso da distribuição normal. Passamos agora aos Passos 6 e 7.

Passo 6: A estatística de teste $z = -6,64$ já foi obtida em exemplo precedente. Podemos agora achar o valor P recorrendo à Figura 7-7. Como o teste é bilateral, o valor P é duas vezes a área à esquerda da estatística de teste $z = -6,64$. Na Tabela A-2, vemos que a área à esquerda de $z = -6,64$ é 0,0001, e assim o valor P é $2 \times 0,0001 = 0,0002$ (veja Figura 7-9).(Tabelas mais precisas do que a A-2 mostrariam que a área à esquerda de $z = -6,64$ é, na verdade, muito menor do que 0,0001.)

Passo 7: Como o valor P 0,0002 é menor do que o nível de significância $\alpha = 0,05$, rejeitamos a hipótese nula.

Tal como com o método tradicional, concluímos no Passo 8 que há evidência suficiente para justificar a rejeição da afirmação de que a temperatura média do corpo humano seja de 98,6°F.

O exemplo que segue aplica a técnica do valor P à dieta e ao programa de exercícios de Jack Wilson discutido em exemplo anterior desta seção.

EXEMPLO Com o método do valor P e o nível de 0,05 de significância, teste a afirmação de que o programa de dois dias de dieta e exercícios de Jack Wilson resulta em uma perda média de peso superior a 0 lb. Tenha em mente que os dados consistem em 33 indivíduos com perda média de 0,37 lb e desvio-padrão de 0,98 lb.

SOLUÇÃO Novamente aqui, recorra aos passos anteriores da resolução deste problema pelo método tradicional. Os passos 1, 2 e 3 tiveram como resultado as seguintes hipóteses:

H_0: $\mu \leq 0$ lb H_1: $\mu > 0$ lb (afirmação original)

Os passos 4 e 5 conduziram a um nível de significância $\alpha = 0,05$ e à decisão de que a distribuição normal é importante para este teste de uma afirmação sobre uma média amostral. Passemos agora aos Passos 6 e 7.

Passo 6: A estatística de teste $z = 2,17$ foi calculada na solução anterior deste mesmo problema. Para achar o valor P, recorremos à Figura 7-7, que indica que, para este teste unilateral direito, o valor P é a área à direita da estatística de teste. Na Tabela A-2, vemos que a área à direita de $z = 2,17$ é $0,5 - 0,4850 = 0,0150$. O valor P é, pois, 0,0150.

Passo 7: Como o valor P de 0,0150 é inferior ao nível de significância $\alpha = 0,05$, rejeitamos a hipótese nula $\mu \leq 0$.

Tal como na solução prévia deste mesmo problema, concluímos que há suficiente evidência para apoiar a afirmação de que a perda média de peso é superior a 0. O valor P 0,0150 mostra que é bastante improvável que obtivéssemos uma média amostral de 0,37 se a média fosse realmente 0.

O próximo processo de teste de hipóteses se baseia em intervalos de confiança, exigindo, assim, os conceitos abordados na Seção 6-2.

Teste de Afirmações com Intervalos de Confiança

Consideremos novamente o problema de teste de hipótese do início do capítulo. Queremos testar a afirmação de que a temperatura média dos adultos sadios é de 98,6°F. Os dados amostrais consistem em $n = 106$ temperaturas com média $\bar{x} = 98,20$ e desvio-padrão $s = 0,62$. No Capítulo 6 descrevemos métodos de construção de intervalos de confiança. Em particular, com os dados amostrais da temperatura do corpo humano, construímos o seguinte intervalo de 95% de confiança:

$$98,08 < \mu < 98,32$$

Temos 95% de confiança de que os limites 98,08 e 98,32 contêm a média populacional μ. (Isto significa que, se fôssemos repetir o experimento de coleta de 106 temperaturas, 95% das amostras teriam como resultado limites do intervalo de confiança que efetivamente conteriam a média populacional μ.) Este intervalo de confiança sugere que é muito pouco provável que a média populacional seja 98,6. Isto é, com base no intervalo de confiança obtido aqui, rejeitamos a suposição corrente de que a temperatura média dos adultos sadios seja de 98,6°F. Podemos generalizar este processo como segue: Primeiro construímos um intervalo de confiança com auxílio dos dados amostrais, e em seguida aplicamos o seguinte critério de decisão:

Uma estimativa intervalar de um parâmetro populacional contém os valores prováveis daquele parâmetro. Devemos, por conseguinte, rejeitar uma afirmação de que o parâmetro populacional tenha um valor que não está compreendido no intervalo de confiança.

Aplicando este critério, notamos que o intervalo de confiança dado aqui não contém o valor alegado de 98,6; conseqüentemente, rejeitamos a alegação de que a média populacional seja 98,6. (*Nota:* Só podemos estabelecer uma correspondência direta entre um intervalo de confiança e um teste de hipótese quando o teste é bilateral. Um teste de hipótese unilateral com nível de significância α corresponde a um intervalo de confiança com grau de confiança $1 - 2\alpha$. Por exemplo, um teste de hipóteses unilateral direito com 0,05 de significância corresponde a um intervalo de confiança de 90%.)

Esta utilização de intervalos de confiança nos fornece um método para identificar resultados altamente improváveis, de forma que podemos determinar quando há uma diferença *significativa* entre os resultados amostrais e um valor pretendido para o parâmetro.

Utilização de Computadores e Calculadoras no Teste de Hipóteses

O STATDISK, o Minitab e a calculadora TI-83 podem fazer testes de hipóteses para uma grande diversidade de circunstâncias. Mostram-se a seguir os resultados dados pelo STATDISK e pelo Minitab para o teste da afirmação de que $\mu = 98{,}6$, utilizando as 106 temperaturas da Tabela 7-1. Para o STATDISK, escolha na barra de menu principal o item `Analysis`, em seguida selecione `Hypothesis Testing`, a seguir selecione `Mean-One Sample`, e passe a introduzir os dados conforme solicitado.

Para o Minitab, introduza primeiro os dados na coluna C1, a seguir use os itens do menu `Stat`, `Basic Statistics`, em seguida `1-Sample z`, e introduza os dados solicitados. A caixa identificada como `alternative` é usada para selecionar a forma da hipótese alternativa, e pode incluir, `not equal`, `less than` ou `greater than`. O *default* de `not equal` é conveniente para o teste que estamos considerando; deixe, pois, a opção como está e clique em OK. Os resultados serão os exibidos.

Se estiver usando uma calculadora TI-83, acione STAT, selecione `TESTS` e escolha a primeira opção de `Z-Test`. Pode-se usar os dados originais ou as estatísticas resumo (`Stats`) provendo os elementos indicados na tela. (Como a amostra é grande, introduza o valor de $s = 0{,}62$ na linha que solicita a introdução de σ.) Os resultados da TI-83 incluirão a hipótese alternativa $\mu \neq 98{,}6$, a estatística de teste $z = -6{,}64$, o valor P (p = 3.1039E $^{-}$11), a média amostral e o tamanho da amostra.

No início desta seção observamos que estamos testando uma afirmação feita sobre a média de uma única população, e que o tamanho da amostra é grande ($n > 30$). Entretanto, há muitos casos importantes que envolvem afirmações sobre uma média, nos quais o tamanho da amostra é pequeno ($n \leq 30$). Tais casos serão abordados na seção seguinte. Nas seções e capítulos que seguem, aplicaremos métodos de teste de hipóteses a outras circunstâncias, como as que envolvem afirmações sobre proporções ou sobre desvios-padrão, ou as que abrangem mais de uma população. É fácil embaraçarmo-nos em uma teia complexa de passos sem sequer compreendermos o *rationale* subjacente do teste de hipóteses. A chave para entendermos isto é o conceito seguinte:

Ao testarmos uma afirmação, fazemos uma suposição (a hipótese nula) que contém a igualdade. Comparamos então a suposição com os resultados amostrais, e formulamos uma das conclusões seguintes.

STATDISK DISPLAY

File Edit Analysis Data Help

Claim	$\mu = \mu_{hyp}$
Null Hypothesis	$\mu = \mu_{hyp}$
Sample Size, n	106
Sample Mean, $\bar{x}$	98.20
Sample St Dev, s	0.6229
Large Sample	
Test Statistic, z	-6.6115
Critical z	±1.9600
P-Value	0.0000

95% Confidence Interval:
98.08 < μ < 98.32

Reject the Null Hypothesis

Sample provides evidence to reject the claim

MINITAB DISPLAY

```
TEST OF MU = 98.6000 VS MU N.E. 98.6000
          N      MEAN      STDEV     SE MEAN     T        P VALUE
C1      106    98.2000    0.6229    0.0605     -6.61    0.0000
```

- **Se os resultados amostrais podem ocorrer facilmente quando a suposição é verdadeira, atribuímos ao acaso a discrepância relativamente pequena entre a suposição e os resultados amostrais.**
- **Se os resultados amostrais não são suscetíveis de ocorrer com facilidade quando a suposição é verdadeira, explicamos a discrepância relativamente grande entre a suposição e os resultados amostrais concluindo que a suposição não é verdadeira.**

Ao interpretarmos os resultados de testes de hipóteses, devemos também ter em mente a distinção entre significância estatística e significância prática. Com uma amostra suficientemente grande, praticamente qualquer diferença entre uma média amostral e uma média populacional alegada se torna significativa. Como sempre, devemos usar o bom senso para interpretar resultados.

7-3 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-24, teste a afirmação utilizando o método tradicional do teste de hipóteses. Identifique também o valor P. *Suponha que todas as amostras tenham sido extraídas aleatoriamente.*

1. Teste a afirmação de que a média populacional é $\mu = 75$, dada uma amostra de $n = 100$ para a qual $\bar{x} = 78$ e $s = 15$. Faça o teste ao nível $\alpha = 0{,}05$ de significância.
2. Teste a afirmação $\mu > 750$ para uma amostra de tamanho n = 36 com $\bar{x} = 800$ e $s = 100$. Use o nível de significância $\alpha = 0{,}01$.
   2. TS: z = 3.00; CV: z = 2.33.
3. Teste a afirmação $\mu < 2{,}50$, para uma amostra de tamanho $n = 64$ para a qual $\bar{x} = 2{,}45$ e $s = 0{,}80$. Use o nível de significância $\alpha = 0{,}02$.
4. Teste a afirmação de que uma média populacional é diferente de 32,0, dada uma amostra de tamanho $n = 75$ para a qual $\bar{x} = 31{,}8$ e $s = 0{,}85$. Use o nível de significância $\alpha = 0{,}01$.
5. Quando os *home runs* são muitos em um jogo de beisebol, há suspeitas de que bolas novas estejam "viciadas" para que alcancem uma distância maior. Testes feitos em bolas velhas mostraram que, quando liberadas de uma altura de 24 ft sobre uma superfície de concreto, elas repicam uma média de 92,84 in. Em um teste com uma amostra de 40 bolas novas, as alturas dos repiques acusaram média de 92,67 in. e desvio-padrão de 1,79 in. (com base em dados do Brookhaven National Laboratory e de *USA Today*). Com um nível de 0,05 de significância, teste a afirmação de que as novas bolas apresentam repiques de alturas diferentes de 92,84 in. Essas bolas são "viciadas"?
6. A Hudson Valley Bottling Company distribui um tipo de cerveja sem álcool em garrafas que indicam o conteúdo de 32 oz. O Bureau of Weights and Measures seleciona aleatoriamente 50 dessas garrafas, mede seu conteúdo e obtém uma média amostral de 31,8 oz, com desvio-padrão de 0,75 oz. Ao nível de 0,01 de significância, teste a afirmação do Instituto de que a companhia está ludibriando os consumidores. Deve-se formalizar uma queixa?
7. Em um estudo de hábitos de consumidores, pesquisadores elaboraram um questionário para identificar os compradores compulsivos. Para uma amostra de consumidores que se declararam compradores compulsivos, os resultados dos questionários acusaram média de 0,83 e desvio-padrão de 0,24 (com base em dados de "A Clinical Screener for Compulsive Buying,", por Faber e Guinn, *Journal of Consumer Research*, Vol. 19). Suponha uma amostra de 32 indivíduos selecionados aleatoriamente. No nível de 0,01 de significância, teste a afirmação de que a população dos que se identificam como compradores compulsivos tem média superior a 0,21, a média geral da população. O questionário parece ser eficiente para identificar os compradores compulsivos?
8. Um artigo publicado no *New York Times* salientou que a duração média de vida de 35 regentes de orquestra do sexo masculino era de 73,4 anos, em contraste com a média de 69,5 anos para a população em geral. Supondo que os 35 homens tenham uma duração de vida com desvio-padrão de 8,7 anos, teste, ao nível de 0,05, a afirmação de que os regentes de orquestra têm duração média de vida diferente de 69,5 anos. (Veja também o Exercício 13 na Seção 1-3.)
9. Um estudo incluiu 123 crianças que estavam usando cinto de segurança ao se ferirem em colisões de veículos. O tempo gasto em uma unidade de tratamento intensivo acusa média de 0,83 dia e desvio-padrão de 0,16 dia [com base em dados de "Morbidity Among Pediatric Motor Vehicle Crash Victims: The Effectiveness of Seat Belts", por Osberg e Di Scala. *American Journal of Public Health*, Vol. 82, N.° 3]. Ao nível de 0.01 de significância, teste a afirmação de que a amostra com cinto de segurança provém de uma população com média inferior a 1,39 dias, que é a média para a população que não usava cinto quando se feriu em colisão de veículos. Os cintos de segurança parecem ser eficazes?
10. A renda média disponível *per capita* no Colorado é de $13.901 [com base em dados do U.S. Bureau of Economic Analysis]. Tom Phelps planeja abrir uma loja de revenda de carros Cadillac e quer verificar aquela cifra para determinada região do Colorado. Os resultados de uma pesquisa recentemente feita junto a 200 pessoas acusam média de $13.447 e desvio-padrão de $4883. Ao nível de 0,05 de significância, teste a afirmação de que a amostra foi extraída de uma população com renda média de $13.901. Há razão para Tom se preocupar com o nível de renda nessa região?
11. Estudou-se a eficiência de um curso de preparação para testes junto a uma amostra aleatória de 75 indivíduos, que fizeram o teste SAT antes e depois do treinamento. As diferenças entre as notas resultaram em um aumento médio de 0,6 e um desvio-padrão de 3,8. (Veja "An Analysis of the Impact of Commercial Test Preparation Courses on SAT Scores", por Sesnowitz, Bernhardt e Kwain. *American Education Research Journal*, Vol. 19, N.° 3.) Ao nível de 0,05 de significância, teste a afirmação de que o aumento da média populacional é maior do que 0, o que indica que o curso é eficiente para o aumento das notas. As pessoas devem fazer o curso?
12. Em um estudo das distâncias percorridas por ônibus antes da ocorrência do primeiro defeito grave no motor, uma amostra de 191 ônibus acusou média de 96.700 mi e desvio-padrão de 37.500 mi (com base em dados de *Technometrics*, Vol. 22, N.° 4). Ao nível de significância de 0,05, teste a afirmação do fabricante de que a distância média percorrida antes da ocorrência do primeiro defeito grave no motor é superior a 90.000 mi.
13. Uma pesquisa feita junto a 100 proprietários de automóvel selecionados aleatoriamente revelou que o tempo médio durante o qual eles pretendem conservar o carro é de 7,01 anos, com desvio-padrão de 3,74 anos (com base em dados de uma pesquisa Roper). O presidente da Newton Car Park está cogitando de lançar uma campanha de vendas tendo como alvo os proprietários de carro que pretendam comprar um carro diferente. Ao nível de significância de 0,05, teste a afirmação do gerente de vendas que, autoritariamente, afirma que o tempo médio durante o qual todos os proprietários de carros pretendem conservá-los é inferior a 7,5 anos.
14. Para 200 condenados por desfalque selecionados aleatoriamente, a duração média da pena de prisão é de 22,1 meses, com desvio-padrão de 8,6 meses (com base em dados do Ministério da Justiça

dos EUA). Kim Patterson é candidato a um cargo político e um dos pontos de sua plataforma eleitoral é um tratamento mais severo dos criminosos condenados. Teste sua afirmação de que os prazos de duração do confinamento desses criminosos têm média inferior a 2 anos. Adote o nível de 0,05 de significância.

15. O rótulo de remédio contra resfriados Dozenol indica a presença de 600 mg de acetaminofen em cada onça fluida. A Food and Drug Administration (FDA) selecionou aleatoriamente 65 amostras de uma onça e constatou que o conteúdo médio de acetaminofen é de 589 mg, com um desvio-padrão de 21 mg. Ao nível $\alpha = 0{,}01$, teste a afirmação da Medassist Pharmaceutical Company de que a média populacional é igual a 600 mg. O leitor compraria este remédio contra gripe?

16. A New England Insurance Company está revendo os hábitos de direção das mulheres na faixa etária entre 16 e 24 anos, para determinar se elas devem continuar pagando prêmios de seguro mais elevados do que as mulheres de outras faixas etárias. Em um estudo de 750 mulheres motoristas selecionadas aleatoriamente na faixa etária 16-24, a distância média percorrida em um ano é de 6047 mi, com desvio-padrão de 2944 mi (com base em dados da Federal Highway Administration). Ao nível de 0,01 de significância, teste a afirmação de que a média populacional para as mulheres com 16-24 anos é inferior a 7124 mi, que é a média conhecida para as mulheres de idade mais alta. Se as mulheres na faixa 16-24 dirigem menos, deveriam elas pagar prêmios de seguro menores?

17. O *Late Show with David Letterman* é visto por uma percentagem relativamente grande das residências, que gravam o show para assisti-lo posteriormente. O gerente de marketing do show afirma que a renda média das residências com videogravador é superior a $40.000. Teste essa afirmação, utilizando um nível de 0,005 de significância. Uma amostra de 1700 residências com vídeo acusa uma média amostral de $41.182 e um desvio-padrão de $19.990 (com base em dados da Nielsen Media Research).

18. O verdadeiro valor de um diploma de curso superior não pode ser medido quantitativamente, mas há maneiras de avaliá-lo em termos de renda. Os homens que têm apenas o curso secundário têm uma renda média anual de $21.652. Selecionados aleatoriamente 73 homens com diploma de faculdade, sua renda média anual é de $40.202 e o desvio-padrão é de $10.900 [com base em dados do Department of Labor — EUA]. Ao nível de 0,01, teste a afirmação de que os homens com diploma de faculdade têm renda média anual superior à dos que têm apenas o curso secundário. Como responderia a quem argüísse que o tamanho da amostra 73 é demasiadamente pequeno neste caso?

19. O tempo médio entre falhas de um rádio da Telektronic Company para aviões de pequeno porte é de 420 horas. Após terem sido modificados 35 aparelhos de rádio, em uma tentativa de melhorar sua confiabilidade, os testes acusaram um tempo médio entre falhas de 385 h para esta amostra, com um desvio-padrão de 24 h. Ao nível de 0,05, teste a afirmação de que as modificações melhoraram a confiabilidade. (Note que a confiabilidade melhorada deveria resultar em um tempo médio entre falhas *mais longo*.)

20. Como parte de uma campanha para atrair fazendeiros, o Iowa Farm Bureau afirma que a quantidade de precipitação pluvial no estado "é em média mais de dois pés e meio cada ano." Ao nível de significância de 0,01, teste a afirmação de que a precipitação média anual em Iowa é superior a 2,5 pés. Como dados amostrais, utilize os valores da precipitação do Conjunto de Dados 7 do Apêndice B. Há evidência suficiente para apoiar a informação divulgada? Há fundamento para uma acusação de propaganda enganosa?

21. Recorra ao Conjunto de Dados 1 do Apêndice B referente aos pesos *totais* de lixo descartado pelas residências em uma semana (com base em dados coletados como parte do Projeto do Lixo na Universidade do Arizona). Para esse conjunto de dados, a média é 27,44 lb e o desvio-padrão é 12,46 lb. Ao nível de significância de 0,01, teste a afirmação do supervisor da cidade de Providence de que o peso médio de todo o lixo descartado pelas residências em cada semana é inferior a 35 lb — o máximo que a cidade pode suportar. Com base no resultado, há motivo de preocupação com a possibilidade de haver excesso de lixo na localidade?

22. Verificando os pesos (em gramas) das moedas de 25 centavos relacionados no Conjunto de Dados 13 do Apêndice B, encontramos 50 pesos com média de 5,622 g e desvio-padrão de 0,068 g. O Ministério da Fazenda dos EUA alega que o processo usado para cunhar moedas de 25 centavos dá um peso médio de 5,670 g. Ao nível de 0,01 de significância, teste a afirmação de que o peso médio das moedas de 25 centavos em circulação é de 5,670 g. Se a alegação for rejeitada, qual é uma explicação possível para a discrepância?

23. A análise dos últimos algarismos de dados por vezes revela se os dados foram medidos e reportados com precisão. Se os últimos algarismos são distribuídos uniformemente de 0 a 9, a média deve ser 4,5. Os últimos algarismos nas extensões (em milhas) de 141 rios foram usados para testar a alegação de que tais dados provêm de uma população com média 4,5. Utilizando-se o Minitab para testar essa afirmação, o resultado é o apresentado no quadro a seguir [com base em dados de "Distribution of Final Digits in Data", por Preece, *The Statistician*, Vol. 30]. Interprete os resultados Minitab ao nível de significância de 0,05.

```
TEST OF MU = 4.500 VS MU N.E. 4.500
     N   MEAN    STDEV   SE MEAN  T       P VALUE
C1 141  2.319    2.899   0.244    -8.93   0.0000
```

24. Um pacote de confeitos M&M indica no rótulo um conteúdo de 1498 confeitos com o peso total de 1361 g, de modo que o peso médio de cada confeito é 1361/1498 = 0,9085 g. Em um teste para determinar se o consumidor está sendo prejudicado, seleciona-se uma amostra aleatória de 33 confeitos M&M marrons. (Veja Conjunto de Dados 11 do Apêndice B.) Usados os 33 pesos com Minitab, o resultado é o exibido a seguir. Interprete-o.

```
TEST OF MU = 0.90850 VS MU < 0.90850
    N   MEAN    STDEV   SE MEAN  T      P VALUE
C1  33  .91282  .03952  .00688   0.63   0.73
```

7-3 Exercícios B: Além do Básico

25. Um artigo reportou que uma hipótese nula de $\mu = 100$ fora rejeitada porque o valor P era inferior a 0,01. O tamanho da amostra era de 62, e a média amostral 103,6. Determine o maior desvio-padrão possível.

26. No Exercício 17, determine o menor valor da média amostral acima de $40.000 que apóie a afirmação de que a média é superior a $40.000. (Use o mesmo tamanho de amostra e o mesmo desvio-padrão.)

27. Em determinado teste de hipótese, a probabilidade α de um erro tipo I é fixa, enquanto a probabilidade β de um erro tipo II depende do valor particular de μ usado como alternativa da hipótese nula. Para testes de hipóteses do tipo encontrado nesta seção, podemos achar β como segue:

Passo 1: Determinar o(s) valor(es) de $\bar{x}$ que corresponde(m) ao(s) valor(es) crítico(s). Em

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

substituir z pelo(s) valor(es) crítico(s), introduzir os valores de $\mu_{\bar{x}}$ e $\sigma_{\bar{x}}$, e resolver em relação a $\bar{x}$.

Passo 2: Dado um valor particular de μ que seja uma alternativa da hipótese nula H_0, traçar a curva normal com este novo valor de μ no centro. Marcar também o(s) valor(es) de $\bar{x}$ achado(s) no Passo 1.

Passo 3: Recorrer ao gráfico do Passo 2 e achar a área da nova região crítica delimitada por $\bar{x}$. Esta é a probabilidade de rejeitar a hipótese nula quando o novo valor de μ é correto.

Passo 4: O valor de β é 1 menos a área do Passo 3. Esta é a probabilidade de não rejeitar a hipótese nula, quando o novo valor de μ é correto.

As etapas precedentes permitem-nos achar a probabilidade de não rejeitar H_0 quando ela é falsa. Estamos determinando a área sob a curva que exclui a região crítica em que rejeitamos H_0; esta área corresponde a não rejeitarmos uma H_0 falsa porque usamos um valor particular de μ que vai contra H_0. Para o exemplo das temperaturas do corpo humano discutido nesta seção, determine o valor de β correspondente a:

a. $\mu = 98{,}7$ b. $\mu = 98{,}4$.

28. O *poder* de um teste, $1 - \beta$, é a probabilidade de rejeitar uma hipótese nula falsa. Reveja o exemplo nesta seção referente à afirmação feita pelo Jack Wilson Health Club. Se o teste dessa afirmação tem poder 0,8, determine a média μ (veja Exercício 27).

7-4 Teste de uma Afirmação sobre uma Média: Pequenas Amostras

Os exemplos e exercícios de teste de hipóteses nas Seções 7-2 e 7-3 utilizam todos eles grandes amostras nos testes de afirmações sobre médias, e assim permitem o uso da distribuição normal. Para esses casos de grandes amostras, podemos aplicar o teorema central do limite para concluir que as médias amostrais se distribuem normalmente, independentemente da distribuição da população original. Todavia, não podemos utilizar o teorema central do limite quando as amostras são pequenas. Os casos de pequenas amostras foram considerados pela primeira vez na Seção 6-3, onde utilizamos a distribuição *t* de Student para estabelecer estimativas intervalares para uma média populacional μ. A Figura 7-10, baseada no mesmo raciocínio usado na Seção 6-3, mostra que algumas afirmações sobre médias são testadas com uma distribuição normal, e outras o são com uma distribuição *t* de Student. A Figura 7-10 resume as decisões a ser tomadas na escolha entre as distribuições normal e *t* de Student.

A Figura 7-10 resume as observações seguintes:

1. De acordo com o teorema central do limite, se obtemos amostras grandes ($n > 30$) (de qualquer população com qualquer distribuição), a distribuição das médias amostrais pode ser aproximada por uma distribuição normal.
2. Quando extraímos amostras (de qualquer tamanho) de uma população com distribuição normal, a distribuição das médias amostrais será aproximadamente normal com média μ e desvio-padrão $\sigma/\sqrt{n}$. Em um teste de hipóteses, o valor de

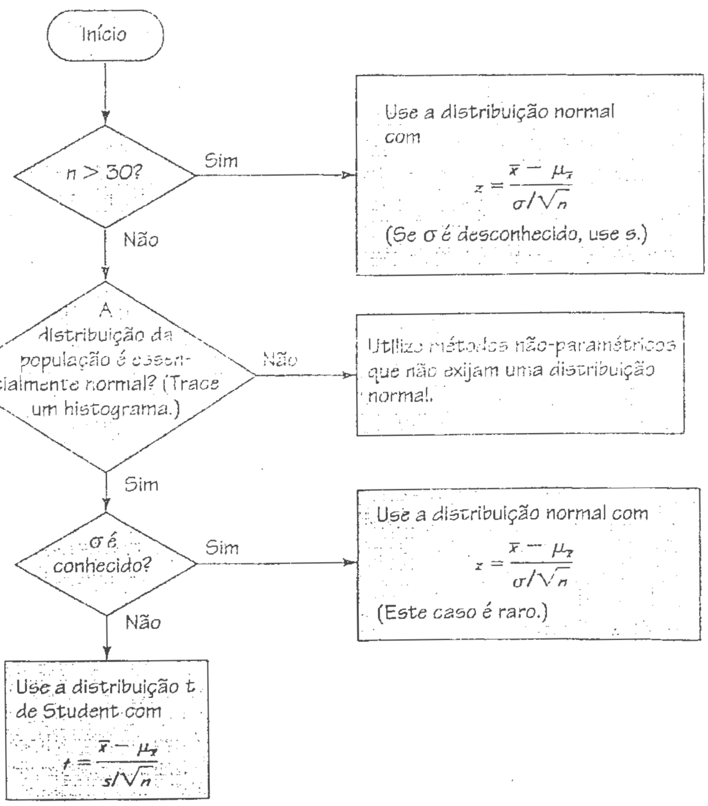


Fig. 7-10 Escolha entre a distribuição normal e a distribuição *t* ao testar uma afirmação sobre a média populacional μ.

μ corresponde à hipótese nula, e o valor do desvio-padrão populacional σ deve ser conhecido. Se σ é desconhecido e as amostras são grandes, podemos usar o desvio-padrão amostral s como substituto de σ, porque grandes amostras aleatórias tendem a representar a população.

3. As condições para utilizar a distribuição t de Student são as seguintes:
   a. A amostra é pequena ($n \leq 30$); e
   b. σ é desconhecido; e
   c. A população original tem distribuição essencialmente normal.
4. Se as amostras são pequenas, σ é desconhecido e a distribuição da população é sensivelmente não-normal, não podemos utilizar os métodos deste capítulo; devemos recorrer a métodos não-paramétricos, alguns dos quais são abordados no Capítulo 13.

Quando as condições relacionadas no item 3 são satisfeitas, utilizamos a distribuição t de Student, com a estatística de teste e os valores críticos dados a seguir:

Estatística de Teste para Afirmações sobre μ quando $n \leq 30$ e σ é Desconhecido

Se uma população é essencialmente normal, então a distribuição de

$$t = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}}$$

é essencialmente uma *distribuição t de Student* para todas as amostras de tamanho n. (A distribuição t de Student é chamada comumente *distribuição t*.)

Pesquisas e Psicólogos

Os resultados de uma pesquisa podem ser seriamente afetados pela maneira como são formuladas as questões. Uma expressão como "durante os últimos anos" é interpretada de modo diverso por diferentes pessoas. Durante os últimos anos (na realidade, desde 1980) pesquisadores e psicólogos têm trabalhado em conjunto para melhorar as pesquisas, reduzindo a tendenciosidade e aumentando a precisão. Em um caso, os psicólogos estudaram a afirmação de 10 a 15% dos pesquisados de terem votado na última eleição, quando, de fato, não o fizeram. Experimentaram teorias de falta de memória, um desejo de ser visto como responsável, e uma tendência daqueles que usualmente votam para dizer que votaram na eleição mais recente, mesmo que não o tenham feito. Constatou-se que somente a última teoria faz realmente parte do problema.

Valores Críticos

1. Na Tabela A-3 encontram-se os valores críticos.
2. Graus de liberdade $= n - 1$.
3. Após determinar o número de graus de liberdade, recorrer à Tabela A-3 e localizá-lo na coluna à esquerda. Com uma determinada linha de valores de t identificada, selecionar o valor crítico t que corresponde ao cabeçalho apropriado. Se um valor crítico está localizado na cauda esquerda, devemos considerá-lo como negativo.

Propriedades Importantes da Distribuição t de Student

1. A distribuição t de Student é diferente para cada tamanho de amostra (veja Figura 6-5 da Seção 6-3).
2. A distribuição t de Student tem a mesma forma geral de sino da distribuição normal padronizada; sua forma mais aberta reflete a maior variabilidade esperada em pequenas amostras.
3. A distribuição t de Student tem média $t = 0$ (tal como a distribuição normal padronizada, que tem média $z = 0$).
4. O desvio-padrão da distribuição t de Student varia com o tamanho da amostra, e é maior do que 1 (ao contrário da distribuição normal padronizada, em que $\sigma = 1$).
5. À medida que o tamanho n da amostra aumenta, a distribuição t de Student se aproxima da distribuição normal. Para valores de $n > 30$, as diferenças são tão pequenas que podemos usar os valores críticos z em lugar de elaborar uma tabela muito maior de valores críticos de t. (Os valores na base da Tabela A-3 são iguais aos valores críticos z correspondentes da tabela da distribuição normal padronizada.)

O exemplo que segue envolve uma amostra pequena extraída de uma população distribuída normalmente, cujo desvio-padrão σ não é conhecido. Essas condições exigem que utilizemos a distribuição t.

EXEMPLO Os sete valores relacionados a seguir são cargas axiais (em libras) da primeira amostra de sete latas de alumínio de 12 oz (veja Conjunto de Dados 15 no Apêndice B). A carga axial de uma lata é o peso máximo que seus lados podem suportar, e deve ser superior a 165 libras, porque esta é a pressão máxima aplicada quando se fixa a tampa no lugar. Ao nível de 0,01 de significância, teste a afirmação do engenheiro supervisor de que esta amostra provém de uma população com média superior a 165 libras.

270 273 258 204 254 228 282

SOLUÇÃO Utilizando os dados amostrais fornecidos, aplicamos os processos do Capítulo 2, obtendo $n = 7$, $\bar{x} = 252{,}7$ e $s = 27{,}6$. A média de 252,7 parece estar bem acima do valor desejado de 165, mas com apenas sete valores temos realmente evidência suficiente para apoiar a afirmação do supervisor? Façamos uma verificação, através de um teste formal de hipóteses, com os mesmos passos indicados na Figura 7-4.

Tal como na seção precedente, os Passos 1, 2 e 3 resultam nas seguintes hipóteses nula e alternativa:

H_0: $\mu \leq 165$
H_1: $\mu > 165$ (afirmação do supervisor)

Passo 4: O nível de significância é $\alpha = 0{,}01$.

Passo 5: Neste teste de uma afirmação sobre a média populacional, a estatística mais importante é a média amostral. Pela Figura 7-10, vemos que a amostra é pequena (porque $n = 7$, não excedendo 30); que é razoável concluir que a distribuição da população é normal (porque estamos lidando com medidas físicas de um produto fabricado sob condições padronizadas), e que σ é desconhecido. A Figura 7-10 mostra que devemos utilizar a distribuição t de Student (e não a distribuição normal).

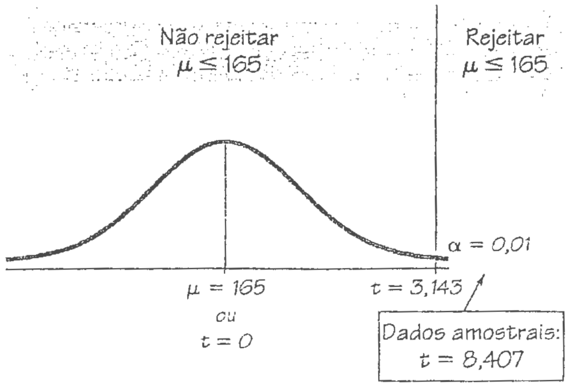


Fig. 7-11 Teste t da afirmação $\mu > 165$.

Passo 6: A estatística de teste é

$$t = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}} = \frac{252,7 - 165}{\frac{27,6}{\sqrt{7}}} = 8,407$$

O valor crítico $t = 3,143$ é encontrado na Tabela A-3. Localizamos primeiro $n - 1 = 6$ graus de liberdade na coluna à esquerda. Como este teste é unilateral direito com $\alpha = 0,01$, recorremos à coluna encabeçada por 0,01 (unilateral). O valor crítico $t = 3,143$ corresponde a 6 graus de liberdade e $\alpha = 0,01$ em uma cauda. A Figura 7-11 mostra a estatística de teste e o valor crítico.

Passo 7: Como a estatística de teste $t = 8,407$ está na região crítica, rejeitamos H_0.

Passo 8: (Recorra à Figura 7-2 para auxiliar na terminologia da conclusão final.) Há evidência suficiente para apoiar a afirmação do supervisor de que a amostra provém de uma população com média superior às 165 libras desejadas.

O valor crítico no exemplo precedente foi $t = 3,143$, mas se a amostra fosse grande ($n > 30$), o valor crítico teria sido $z = 2,33$. O maior valor crítico t de Student mostra que, com uma amostra pequena, a evidência amostral deve ser *mais extrema* para então considerarmos a diferença como significativa.

Valores P

O exemplo precedente seguiu a abordagem tradicional do teste de hipóteses, mas o STATDISK, o Minitab e a calculadora TI-83, assim como grande parte da literatura, apresentam valores P. Como a tabela de distribuição t (Tabela A-3) inclui apenas valores selecionados do nível de significância α, em geral não podemos achar o valor P específico na Tabela A-3. Em vez disso, podemos utilizar essa tabela para identificar limites que contenham o valor P. No último exemplo obtivemos a estatística de teste $t = 8,407$ e sabemos que o teste é unilateral com 6 graus de liberdade. Consultando a linha da Tabela A-3 correspondente a 6 graus de liberdade, vemos que a estatística de teste 8,407 excede o maior valor crítico naquela linha. Embora não possamos determinar um valor P *exato* na Tabela A-3, podemos concluir que deve ser inferior a 0,005. Ou seja, concluímos que valor $P < 0,005$. (Algumas calculadoras e programas permitem-nos achar valores P exatos. Para o exemplo precedente, o STATDISK apresenta um valor P de 0,00008, o Minitab dá um valor P 0,0001, e a calculadora TI-83 dá o valor P 0,000077613518.) Ao nível de significância de 0,01 e com um valor P inferior a 0,005, rejeitamos a hipótese nula (porque o valor P é inferior ao nível de significância), tal como fizemos com o método tradicional no exemplo precedente.

> **Melhores Resultados com Turmas Menores**
>
> Um experimento na Universidade de Nova York, em Stony Brook, revelou que os alunos apresentavam rendimento significativamente melhor em turmas limitadas a 35 alunos do que em grandes turmas com 150 a 200 alunos. Em um curso de cálculo, a percentagem de reprovação foi 19% para pequenas turmas, comparada com 50% para as turmas grandes. A percentagem de graus A foi de 24% para pequenas turmas e 3% para grandes turmas. Estes resultados sugerem que os estudantes se beneficiam com turmas pequenas, que permitem interação mais direta entre alunos e professores.

EXEMPLO Utilize a Tabela A-3 para achar o valor P correspondente a estes resultados: aplica-se a distribuição t de Student em um teste bilateral com uma amostra de $n = 10$ valores, obtendo-se a estatística de teste $t = 2,567$.

SOLUÇÃO Consulte a linha da Tabela A-3 referente a 9 graus de liberdade e observe que a estatística de teste 2,567 está entre os valores críticos 2,821 e 2,262. Como o teste é bilateral, consideramos os valores de α na parte superior que são identificados como "bilaterais". Os valores críticos 2,821 e 2,262 correspondem a 0,02 (bilateral) e 0,05 (bilateral), e assim expressamos como segue o valor P.

$$0,02 < \text{valor } P < 0,05$$

Com uma hipótese nula verdadeira, a chance de obter uma média amostral (de 10 valores amostrais) que se transforme em uma estatística de teste $t = 2,567$ está entre 0,02 e 0,05.

EXEMPLO Considere os 11 valores de depósito de nitrato de chuva ácida em Massachusetts do Conjunto de Dados 6 do Apêndice B (com base em dados do Ministério da Agricultura dos EUA). Ao nível de significância de 0,05, teste a afirmação de que a amostra provém de uma população com média igual a 5 kg/hectare.

SOLUÇÃO Nesta solução utilizamos a abordagem pelo valor P baseado em resultados de uma calculadora TI-83. Introduzimos primeiro os 11 valores amostrais na lista L1, acionamos então STAT e selecionamos `TESTS`. Escolhemos `T-Test` e passamos a introduzir os dados solicitados. Os resultados da TI-83 incluirão o valor `p = 0.0075851732`. (Pode-se utilizar também o STATDISK ou o Minitab para achar o valor P.) Como este valor P é inferior ao nível de significância de 0,05, rejeitamos a hipótese nula de que a média da população seja igual a 5. Há evidência suficiente para apoiar a rejeição da afirmação de que a média é igual a 5 kg/hectare.

Até aqui discutimos testes de hipóteses formuladas somente sobre médias populacionais. Na próxima seção testaremos hipóteses sobre proporções ou percentagens, e na última seção estudaremos afirmações sobre desvios-padrão ou variâncias.

7-4 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1 e 2, determine os valores críticos t para as hipóteses, tamanhos de amostra e níveis de significância dados.

1. a. H_0: $\mu = 98{,}6$ $n = 7$ $\alpha = 0{,}05$ b. H_0: $\mu \geq 100$ $n = 12$ $\alpha = 0{,}05$ c. H_1: $\mu > 32$ $n = 9$ $\alpha = 0{,}01$
2. a. H_0: $\mu \leq 1{,}07$ $n = 5$ $\alpha = 0{,}01$ b. H_1: $\mu < 75{,}2$ $n = 14$ $\alpha = 0{,}05$ c. H_1: $\mu \neq 64$ $n = 10$ $\alpha = 0{,}10$

Nos Exercícios 3 e 4, suponha a amostra extraída aleatoriamente de uma população com distribuição normal. Teste a afirmação feita, utilizando o método tradicional de teste de hipóteses.

3. Ao nível de 0,05 de significância, teste a afirmação de que $\mu \neq 64{,}8$. Os dados amostrais consistem em 12 valores com $\bar{x} = 59{,}8$ e $s = 8{,}7$.
4. Ao nível de 0,01 de significância, teste a afirmação de que $\mu < 927$. Os dados amostrais consistem em 10 valores com $\bar{x} = 874$ e $s = 57{,}3$.
5. Se utilizarmos o Minitab para os 11 valores de depósitos de *sulfato* de chuva ácida para a Pennsylvania constantes do Conjunto de Dados 6 do Apêndice B (com base em dados do Ministério da Agricultura dos EUA) e testarmos a afirmação de que a amostra provém de uma população com média superior a 12,00 kg/hectare, a tela apresentará a configuração a seguir. Tome um nível de 0,01 de significância e interprete os resultados Minitab.

```
Test of mu = 12.000 vs mu > 12.000
Variable N  Mean   StDev SE Mean T    P-value
C1       11 13.286 1.910 0.576   2.23 0.025
```

6. Considere os resultados dados por Minitab do Exercício 5. Se substituirmos a afirmação "maior do que 12,00" por "diferente de 12,00", como serão afetados os valores da linha inferior?

Para os testes de hipóteses dos Exercícios 7-24, use a abordagem tradicional resumida na Figura 7-4. Esboce um gráfico mostrando a estatística de teste e os valores críticos. Em cada caso, suponha que a população tem distribuição aproximadamente normal e que a amostra foi escolhida aleatoriamente. Cuidado: *Alguns dos exercícios exigem o uso da distribuição normal (estudada na seção precedente) em lugar da distribuição* t *de Student (estudada nesta seção); verifique cuidadosamente as condições para determinar a distribuição adequada.*

7. A Carolina Tobacco Company anunciou que seus cigarros mais vendidos contêm no máximo 40 mg de nicotina, mas a revista *Consumer Advocate* testou 10 cigarros escolhidos aleatoriamente e obteve $\bar{x} = 43{,}3$ mg e $s = 3{,}8$ mg. Outras evidências sugerem que a distribuição do conteúdo de nicotina é uma distribuição normal. A amostra é pequena porque o trabalho de laboratório exigido para extrair a nicotina é demorado e dispendioso. É temerário afirmar que o anúncio da companhia está errado e, assim, o editor da revista escolhe o nível de significância $\alpha = 0{,}01$ para testar a hipótese de que o conteúdo médio de nicotina é superior a 40 mg. Com um nível de 0,01 de significância, teste a suposição do editor de que a média seja superior a 40 mg.
8. A despesa com a transferência do pátio de armazenagem para a Consolidated Package Delivery Service (CPDS) só se justifica se ficar provado que o percurso médio diário é inferior a 214 mi. Em percursos de teste feitos com 12 caminhões de entrega, obtiveram-se a média de 198 mi e o desvio-padrão de 42 mi. Ao nível de 0,01 de significância, teste a afirmação de que a média é inferior a 214 mi. O pátio de armazenagem deve ser transferido?
9. Consulte o Conjunto de Dados 2 do Apêndice B. Utilizando apenas as 25 primeiras temperaturas relacionados para as 12 horas do dia 2, teste a afirmação de que a temperatura média do corpo de todos os adultos sadios é igual a 98,6°F. Use 0,05 como nível de significância. Para os primeiros 25 valores, $\bar{x} = 98{,}24$ e $s = 0{,}56$.
10. Consulte o Conjunto de Dados 2 do Apêndice B. Utilizando apenas as 35 primeiras temperaturas relacionadas para as 12 horas do dia 2, teste a afirmação de que a temperatura média de todos os adultos sadios é igual a 98,6°F. Use o nível de significância $\alpha = 0{,}05$. Para os 35 primeiros valores, $\bar{x} = 98{,}27$ e $s = 0{,}65$.
11. Determinou-se o custo de operação por cliente para cada uma de 12 organizações. Os 12 valores têm média de \$2133 e desvio-padrão de \$345 (com base em dados de "Organizational Communication and Performance", por Snyder e Morris, *Journal of Applied Psychology*, Vol. 69, No. 3). Ao nível de 0,01 de significância, teste a afirmação de um acionista, que se queixa de que a média para todas as organizações desse tipo excede \$1800 por cliente.
12. Testaram-se as propriedades de deslizamento de um pneu para neve, estabelecendo-se uma distância média de deslizamento de 154 ft para condições-padrão. Fabrica-se um pneu novo, mais dispendioso, mas os testes feitos com 17 desses pneus acusaram distância média de deslizamento de 148 ft e desvio-padrão de 12 ft. Em vista do custo envolvido, os novos pneus só serão adquiridos se ficar provado que, ao nível de 0,005 de significância, eles deslizam menos do que os pneus atuais. Com base na amostra, acha aconselhável comprar os novos pneus?
13. O Bank of New England está preocupado com o acúmulo de débito de clientes que utilizam cartões de crédito. A diretoria decidiu instituir um sistema dispendioso de monitoramento se a média para todos os clientes do banco for superior a \$2000. O banco selecionou aleatoriamente 50 portadores de cartão de crédito e determinou o valor de seus débitos. Para esse grupo, a média foi \$2177 e o desvio-padrão \$1257. Ao nível de significância de 0,025, teste a afirmação de que o débito médio é superior a \$2000. Com base no resultado, deve-se implementar o sistema de monitoramento?
14. Fez-se um estudo para determinar se um teste de escrita para funcionários necessitaria de revisão para uso em terminais de vídeo (VDTs). Os escores VDT de 22 indivíduos acusam média de 170,2 e desvio-padrão de 35,3 (com base em dados de "Modification of the Minnesota Clerical Test to Predict Performance on Video Display Terminals", por Silver e Bennett, *Journal of Applied Psychology*, Vol. 72, N.º 1). Com o nível de 0,05 de significância, teste a afirmação de que a média para todos os que fazem o teste VDT difere da média 243,5 para a versão padrão impressa do teste. Com base nesse resultado, acha que o teste deve ser revisto?
15. Em um estudo dos fatores que afetam o hipnotismo, obtiveram-se, para 16 indivíduos, as classificações sensoriais pela escala visual analógica (EVA). A média e o desvio-padrão para essas classificações amostrais foram 8,33 e 1,96, respectivamente (com base em dados de "An Analysis of Factors That Contribute to the Efficacy of Hypnotic Analgesia", por Price and Barber, *Journal of Abnormal Psychology*, Vol. 96, N.º 1). No nível de 0,01 de significância, teste a afirmação de que esta amostra provém de uma população com classificação média inferior a 10,00.

16. Um teste de leitura feito com 15 estudantes de terceiro grau acusou $\bar{x} = 31{,}0$ e $s = 10{,}5$. ("Os dados se baseiam em "A Longitudinal Study of the Effects of Retention/Promotion on Academic Achievement", por Peterson *et al.*, *American Educational Research Journal*, Vol. 24, N.° 1.) Esta média amostral de terceiro grau difere significativamente de uma média de 41,9 de uma população de primeiro grau? Tome $\alpha = 0{,}01$.

17. Kim Greco é estudante universitária e está preocupada com a faculdade, porque sabe que muitos estudantes levam mais de 4 anos para obter o grau de bacharel. Ao nível de 0,10 de significância, teste a afirmação do seu orientador de que o tempo médio é superior a 5 anos. Os dados amostrais consistem em uma média de 5,15 anos e um desvio-padrão de 1,68 anos para 80 bacharéis selecionados aleatoriamente (com base em dados do National Center for Education Statistics).

18. Com os pesos dos comprimidos de Bufferin do Conjunto de Dados 14 do Apêndice B, teste a afirmação de que o peso médio é igual a 650 mg. Tome um nível de 0,05 de significância. Para os dados amostrais, $\bar{x} = 665{,}41$ e $s = 7{,}26$.

19. Com os pesos apenas dos confeitos M&Ms *azuis* relacionados no Conjunto de Dados 11 do Apêndice B, teste a afirmação de que a média é no mínimo 0,9085 g, valor médio necessário para que os 1498 M&Ms tenham o peso total de 1361 g indicado no rótulo do pacote. Use o nível de 0,05 de significância. Para os M&Ms azuis, $\bar{x} = 0{,}9014$ g e $s = 0{,}0573$ g. Com base no resultado, pode-se concluir que o conteúdo do pacote está em desacordo com o peso indicado no rótulo?

20. Com os pesos somente dos M&Ms *marrons* relacionados no Conjunto de Dados 11 do Apêndice B, teste a afirmação de que o peso médio é superior a 0,9085 g, valor médio necessário para que os 1498 M&Ms tenham o peso total de 1361 g indicado no pacote. Utilize um nível de 0,05 de significância. Para os M&Ms marrons, $\bar{x} = 0{,}9128$ g e $s = 0{,}0395$ g. Com base no resultado, pode-se concluir que os pacotes têm peso superior ao indicado no rótulo?

21. Relacionam-se a seguir os valores do consumo de energia elétrica (em kWh) na casa do autor durante sete anos diferentes:

11.943 11.463 10.789 9907 9012 9942 11.153

A companhia fornecedora de energia afirma que o consumo anual médio é de 11.000 kWh e oferece um plano especial de pagamento baseado nessa quantidade. Ao nível de 0,05 de significância, teste a afirmação da companhia fornecedora de que a média é 11.000 kWh.

22. Relacionam-se a seguir os pesos, ao nascer (em quilogramas), de meninos nascidos de mães que ingeriram um suplemento especial de vitaminas (com base em dados da Secretaria de Saúde do Estado de Nova York). Ao nível de 0,05 de significância, teste a afirmação de que o peso médio, ao nascer, de todos os meninos de mães que tomaram o suplemento de vitaminas é igual a 3,39 kg, que é a média da população de todos os recém-nascidos. Com base no resultado, o suplemento de vitaminas parece ter algum efeito sobre o peso da criança ao nascer?

3,73	4,37	3,73	4,33	3,39	3,68	4,68	3,52
3,02	4,09	2,47	4,13	4,47	3,22	3,43	2,54

23. Rita Gibbons é uma comediante que faz gravações de suas apresentações e registra o total de vezes que deve interromper, até que os espectadores parem de rir, para então retomar a apresentação. Dão-se a seguir os tempos (em segundos) para 15 apresentações diferentes de um novo show. Teste a afirmação de que o tempo médio é superior a 63,2 s, o tempo médio de seu antigo show. Com base nos resultados, o novo show parece ser melhor do que o anterior?

86 45 44 78 52 79 86 66 61 57 98 44 61 99 87

24. Com referência ao conteúdo de nicotina dos cigarros (Conjunto de Dados 4 do Apêndice B), use os dados amostrais para testar a afirmação de que a média populacional é inferior a 1,0 mg. A conclusão se aplica ao nível médio de nicotina de todos os cigarros de 100 mm (não-mentol, nem *light*) fumados pelos consumidores? Por quê?

7-4 Exercícios B: Além do Básico

25. Para cada teste de hipóteses dado, que se pode concluir sobre o valor P utilizando apenas as Tabelas A-2 e A-3?

a. Exercício 7 b. Exercício 8 c. Exercício 9 d. Exercício 10

26. Em decorrência de certas condições, um teste de hipóteses exige a distribuição t de Student, descrita nesta seção. Suponha que, em lugar dela, tenha sido usada incorretamente a distribuição normal. Esse uso da distribuição normal torna o leitor mais, ou menos, propenso a rejeitar a hipótese nula? Ou não faz diferença? Explique.

27. Ao determinar valores críticos, às vezes necessitamos de níveis de significância que não constam da Tabela A-3. Alguns programas de computador aproximam os valores críticos de t por

$$t = \sqrt{\text{gl} \cdot (e^{A^2/\text{df}} - 1)}$$

onde

$$\text{df} = n - 1$$
$$e = 2{,}718$$
$$A = z\left(\frac{8\,\text{df} + 3}{8\,\text{df} + 1}\right)$$

e onde z é o escore z crítico. Com esta aproximação, determine o escore t crítico correspondente a $n = 10$ e um nível de 0,05 de significância em um teste unilateral direito. Compare os resultados com o valor crítico de t encontrado na Tabela A-3.

28. Com referência ao Exercício 7, suponha o leitor que esteja testando a alegação de que $\mu > 40$ mg. Determine β (a probabilidade de um erro tipo II), dado que o valor efetivo da média populacional é $\mu = 45{,}0518$ mg. (Veja Exercício 27 da Seção 7-3.)

Detectores de Mentiras

Por que não exigir que todos os suspeitos de crime passem por testes em um detector de mentiras, eliminando o julgamento por um júri? O Conselho de Assuntos Científicos da Associação Médica Americana afirma: Está confirmado que a classificação "culpado" pode ser feita com 75% a 97% de precisão, mas a taxa de falsos positivos é, não raro, suficientemente elevada para invalidar o uso desse teste (polígrafo) como único árbitro da decisão "culpado" ou "inocente." Um "falso positivo" é uma indicação de culpa quando o réu é, de fato, inocente. Mesmo com um grau de precisão tão alto como 97%, a percentagem de resultados falso-positivos pode chegar a 50%, de forma que metade dos inocentes é considerada culpada.

7-5 Teste de uma Afirmação sobre uma Proporção

Nas seções precedentes deste capítulo introduzimos os métodos básicos de teste de hipóteses; entretanto, eles se referiam apenas a afirmações sobre *médias* populacionais. Com os métodos desta seção, podemos testar uma afirmação sobre uma proporção, uma percentagem ou uma probabilidade, conforme ilustrado nos exemplos que seguem:

- Com base em uma pesquisa amostral, menos de 1/4 de todos os bacharéis fuma.

- A percentagem de espectadores de horários noturnos de TV que assistem a *The Late Show with David Letterman* é 18%.
- Se ocorre um acidente fatal com um automóvel, há uma probabilidade de 0,44 de envolver um motorista que bebeu.

Hipóteses Usadas ao Testar uma Afirmação sobre uma Proporção, uma Probabilidade ou uma Percentagem Populacional

1. São verificadas as condições para um *experimento binomial.* Isto é, temos um número fixo de provas independentes com probabilidade constante, e cada prova comporta dois resultados, que designamos "sucesso" e "falha".
2. As condições $np \geq 5$ e $nq \geq 5$ são ambas verificadas, de modo que **a distribuição binomial das proporções amostrais pode ser aproximada por uma distribuição normal com $\mu = np$** e $\sigma = \sqrt{npq}$ (conforme descrito na Seção 5-6).

Se essas hipóteses não forem todas satisfeitas, eventualmente poderemos utilizar outros métodos. Nesta seção, entretanto, vamos considerar apenas os casos em que as hipóteses são satisfeitas, de modo que a distribuição amostral das proporções amostrais pode ser aproximada pela distribuição normal. Usaremos a notação e a estatística de teste seguintes:

Notação

n = número de provas

$\hat{p} = \frac{x}{n}$ (proporção *amostral*)

π_0 p = proporção *populacional* (usada na hipótese nula)

$q = 1 - p$

Estatística para Teste de uma Afirmação sobre uma Proporção

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

π_0 π_0 $-(1-\pi_0)$

Obtém-se o valor crítico na Tabela A-2 (distribuição normal padronizada) pelos mesmos processos descritos na Seção 7-2. Por exemplo, em um teste bilateral com nível de significância $\alpha = 0{,}05$, dividimos α igualmente entre as duas caudas, e recorremos à Tabela A-2 para o escore z correspondente a uma área de $0{,}5 - 0{,}025 = 0{,}475$; o resultado é $z = 1{,}96$, de modo que os valores críticos são $z = -1{,}96$ e $z = 1{,}96$.

Para justificar a estatística de teste, notamos que, ao usar a distribuição normal para aproximar uma distribuição binomial, fazemos $\mu = np$ e $\sigma = \sqrt{npq}$, vindo

$$z = \frac{x - \mu}{\sigma} = \frac{x - np}{\sqrt{npq}}$$

Nesta expressão, x é o número de sucessos em n provas. Dividindo por n o numerador e o denominador desta última expressão, e substituindo x/n pelo símbolo $\hat{p}$, temos a estatística de teste que acabamos de dar. Em outras palavras, a estatística de teste é simplesmente $z = (x - \mu)/\sigma$ adaptada à notação binomial incorporando o fato que a distribuição de proporções amostrais $\hat{p}$ é uma distribuição normal com média p e desvio-padrão $\sqrt{pq/n}$.

Ao testar uma afirmação sobre uma proporção populacional p, é preciso ter cuidado em identificar corretamente a proporção amostral $\hat{p}$. A proporção amostral $\hat{p}$ às vezes é dada diretamente, como na afirmação "10% dos carros esporte observados são vermelhos", o que se expressa como $\hat{p} = 0{,}10$. Em outros casos, pode ser necessário calcular a proporção amostral, utilizando $\hat{p} = x/n$. Por exemplo, pela afirmação de que "96 das residências pesquisadas têm TV a cabo, e 54 não têm", podemos inicialmente determinar o tamanho da amostra como $96 + 54 = 150$; em seguida, calculamos o valor da proporção amostral das residências que têm TV a cabo, como segue:

$$\hat{p} = \frac{x}{n} = \frac{96}{150} = 0{,}64$$

Cuidado: O valor 0,64 é exato, mas quando o valor de $\hat{p}$ apresentado em uma calculadora ou em um computador resulta em muitas casas decimais, devemos utilizá-las todas ao calcular a estatística de teste z. O arredondamento demasiado de $\hat{p}$ pode resultar em grandes erros.

> **O Conquistador**
>
> *O Conquistador* é um filme de 1954 produzido próximo a St. George, Utah, um ano após ter sido detonada uma bomba atômica na vizinhança. Dentre as 220 pessoas que trabalharam no filme, John Wayne e outras 97 morreram de câncer. Com auxílio de métodos estatísticos, pode-se mostrar que a taxa de 45% de morte por câncer é significativamente mais elevada do que o que se poderia esperar em circunstâncias normais. Mas não é possível utilizar métodos estatísticos para mostrar que a causa das mortes por câncer foi a explosão da bomba atômica.

Consideremos agora um exemplo que ilustra uma aplicação prática dos conceitos precedentes. Este exemplo segue os mesmos processos básicos do teste de hipóteses dados na Seção 7-3. A afirmação no exemplo que segue envolve uma percentagem, mas devemos utilizar a forma decimal equivalente. Embora o método apresentado nesta seção possa ser usado para testar afirmações sobre proporções, probabilidades ou percentagens, devemos sempre realizar o teste com um valor de p entre 0 e 1; a soma de p e q deve ser exatamente 1.

EXEMPLO Em um estudo da eficácia do *air-bag* em automóveis, constatou-se que, em 821 colisões de carros de tamanho médio equipados com *air-bag*, 46 colisões resultaram em hospitalização do motorista (com base em dados do Highway Loss Data Institute). Ao nível de significância de 0,01, teste a afirmação de que a taxa de hospitalização nos casos de *air-bag* é inferior à taxa de 7,8% para colisões de carros de tamanho médio equipados com cintos automáticos de segurança.

SOLUÇÃO Aplicaremos o método tradicional de teste de hipóteses delineado na Figura 7-4. Em lugar de percentagens, esta solução utilizará a proporção amostral $\hat{p} = 46/821 = 0{,}0560292$ e a proporção populacional admitida $p = 0{,}078$.

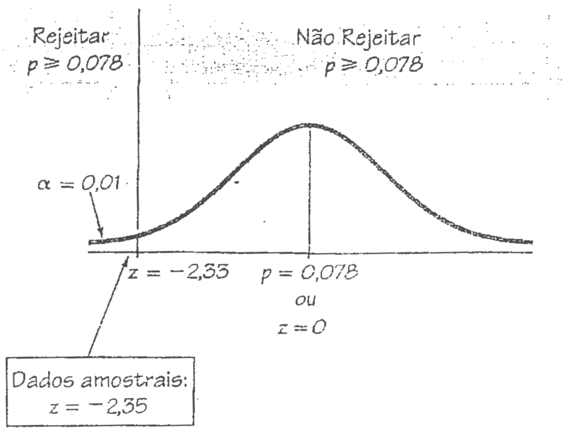


Fig. 7-12 Teste de hipóteses para a afirmação $p < 0,078$.

Passo 1: A afirmação original é de que a taxa de hospitalização no caso de *air-bag* é inferior a 7,8%. Em forma simbólica, temos $p < 0,078$.

Passo 2: O oposto da afirmação original é $p \geq 0,078$.

Passo 3: Como $p \geq 0,078$ contém a igualdade, temos

H_0: $p \geq 0,078$ (hipótese nula)
H_1: $P < 0,078$ (hipótese alternativa e afirmação original)

Passo 4: O nível de significância é $\alpha = 0,01$.

Passo 5: A estatística relevante para este teste é $\hat{p} = 46/821 = 0,0560292$. A distribuição das proporções amostrais é aproximada pela distribuição normal. (As exigências $np \geq 5$ e $nq \geq 5$ são ambas satisfeitas, com $n = 821$, $p = 0,078$ e $q = 0,922$.)

Passo 6: Obtém-se, como segue, a estatística de teste $z = -2,35$:

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0,0560292 - 0,078}{\sqrt{\frac{(0,078)(0,922)}{821}}} = -2,35$$

O valor crítico $z = -2,33$ é encontrado na Tabela A-2. Com $\alpha = 0,01$ na cauda esquerda, procuramos uma área de 0,4900 no corpo da Tabela A-2; o valor mais próximo é 0,4901, correspondente a $z = 2,33$; entretanto, deve ser negativo porque está à esquerda da média. A Figura 7-12 mostra a estatística de teste e o valor crítico.

Passo 7: Como a estatística de teste está dentro da região crítica, rejeitamos a hipótese nula.

Passo 8: Há evidência suficiente para apoiar a afirmação de que, para colisões de carros de tamanho médio, a taxa de hospitalização, no caso de haver o *air-bag*, é inferior à taxa de 7,8% verificada no caso de cintos de segurança automáticos.

EXEMPLO Em um teste de gosto de consumidores, 100 bebedores regulares de Pepsi recebem amostras cegas de Coca e Pepsi; 48 deles preferiram a Coca. Ao nível de significância de 0,05, teste a afirmação de que a Coca é preferida por 50% dos bebedores de Pepsi que participam de tais testes.

SOLUÇÃO Eis um resumo das componentes-chave do teste de hipóteses.

H_0: $p = 0,5$ (da afirmação de que a Coca é preferida por 50%)
H_1: $p \neq 0,5$

Estatística de teste:

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0,48 - 0,5}{\sqrt{\frac{(0,5)(0,5)}{100}}} = -0,40$$

A Figura 7-13 mostra a estatística de teste, os valores críticos e a região crítica. Como a estatística de teste não está na região crítica, não rejeitamos a hipótese nula. Não há evidência suficiente para rejeitar a afirmação de que 50% dos bebedores de Pepsi preferem Coca. (Os críticos desses testes de sabor alegam que os que se submetem ao teste freqüentemente não conseguem detectar diferenças, e apenas "palpitam" quando indicam suas preferências.)

Um Caso de Discriminação Utiliza a Estatística

A estatística desempenha freqüentemente papel-chave em casos de discriminação. Um desses casos envolveu Matt Perez e mais de 300 outros agentes do FBI, que ganharam uma ação, sob a alegação de que os hispânicos no FBI eram discriminados nas áreas de promoção, atribuições e ações disciplinares. O queixoso contratou o estatístico Gary Lafree, que mostrou que os postos de administração superior do FBI tinham proporções significativamente menores de empregados hispânicos. A estatística foi fundamental para a vitória do queixoso neste caso. (Veja o Programa 20 da série *Against All Odds: Inside Statistics* para uma discussão deste caso.)

O Método do Valor P

Os exemplos desta seção seguiram a abordagem tradicional do teste de hipóteses, mas teria sido mais fácil utilizar o método do valor *P*. Para achar o valor *P* correspondente a uma estatística de teste *z*, utilizaremos o processo seguinte, descrito na Seção 7-3:

Teste unilateral direito: Valor *P* = área à direita da estatística de teste *z*

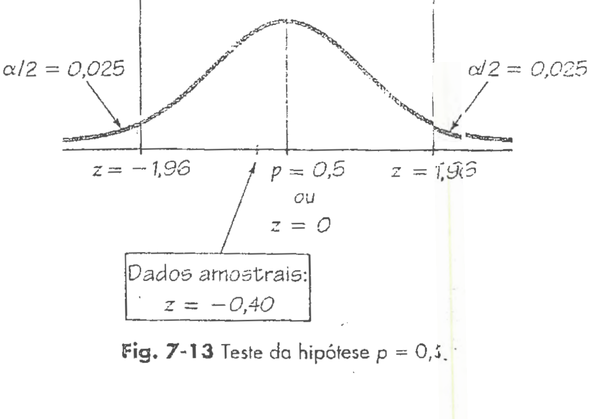


Fig. 7-13 Teste da hipótese $p = 0,5$.

Teste unilateral esquerdo:	Valor P = área à esquerda da estatística de teste z
Teste bilateral:	Valor P = *duas vezes* a área da região extrema delimitada pela estatística de teste z

Tal como na Seção 7-3, utilizamos o seguinte critério de decisão:

Rejeitar a hipótese nula se o valor P é inferior ou, no máximo, igual ao nível de significância α.

Como o último exemplo foi bilateral, o valor P é duas vezes a área à esquerda da estatística de teste $z = -0{,}40$. Pela Tabela A-2, vemos que a área entre $z = 0$ e $z = -0{,}40$ é 0,1554, de modo que a área à esquerda da estatística de teste $z = -0{,}40$ é $0{,}5 - 0{,}1554 = 0{,}3446$. O valor P é $2 \times 0{,}3446 = 0{,}6892$. Como o valor P supera o nível de significância de 0,05, não rejeitamos a hipótese nula e novamente concluímos que não há evidência suficiente para rejeitar a afirmação de que 50% dos bebedores de Pepsi preferem Coca. Mais uma vez, vemos que o método do valor P nada mais é do que outra maneira de chegarmos à mesma conclusão obtida com a abordagem tradicional.

Utilização de Computadores e Calculadoras

Podemos utilizar o STATDISK e a calculadora TI-83 para testar afirmações sobre uma proporção populacional. Em geral, pode-se usar também o Minitab.

7-5 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-20, aplique o método tradicional para testar a hipótese indicada. Inclua os passos relacionados na Figura 7-4 e esboce o gráfico apropriado. Identifique também o valor P. *Suponha que todas as amostras tenham sido extraídas aleatoriamente.*

1. Em um estudo sobre os *scanners* para a leitura de preços em lojas, verificaram-se 1234 itens, constatando-se que 20 deles acusavam excesso de preço (com base em dados de "UPC Scanner Pricing Systems: Are They Accurate?", por Goodstein, *Journal of Marketing* Vol 58). Ao nível de 0,05 de significância, teste a afirmação de que, com *scanners*, 1% das vendas acusam excesso de preço. (Antes da utilização de *scanners*, a taxa de sobrepreços era estimada em cerca de 1%.) Com base nesses resultados, acha que os *scanners* ajudam os consumidores?
2. A preocupação com o ambiente entra freqüentemente em conflito com a tecnologia moderna, como no caso dos pássaros que representam perigo para a aviação durante a decolagem. Um grupo ambiental afirma que tais acidentes com pássaros são tão raros que não se justifica matá-los. Um grupo de pilotos alega que entre as decolagens interrompidas que levam um avião a ultrapassar o final da pista, 10% são devidas a colisão com pássaros. Teste esta afirmação ao nível de significância de 0,05. Os dados amostrais consistem em 74 decolagens interrompidas, em que o avião ultrapassou o final da pista. Dentre esses 74 casos, 5 foram devidos à colisão com pássaros (com base em dados da Air Line Pilots Association and Boeing, relatado em *USA Today*).
3. O gerente de controle de qualidade da Telektronic Company considera a fabricação de secretárias eletrônicas como "fora de controle" quando a taxa geral de defeitos excede 4%. O teste de uma amostra aleatória de 150 secretárias eletrônicas acusou 9 defeituosas, o que corresponde a uma percentagem de 6% de defeitos. O gerente de produção alega tratar-se apenas de uma diferença casual, e que a produção realmente está sob controle, não sendo necessária qualquer medida corretiva. Teste a afirmação do gerente de produção, ao nível de 0,05 de significância. Afigura-se necessária uma ação corretiva?
4. Em 1990, 5,8% dos candidatos a emprego submetidos a teste de uso de drogas foram reprovados no teste. Ao nível de 0,01, teste a afirmação de que a taxa de reprovação agora é mais baixa, considerando que uma amostra aleatória de 1520 candidatos acusou 58 reprovações (com base em dados da American Management Association). O resultado sugere que, hoje, menos candidatos a emprego usam drogas?
5. Ralph Carter é um professor de história de curso secundário que diz que se os alunos não têm conhecimento do Holocausto, o currículo deve ser revisto para corrigir essa deficiência. Uma pesquisa Roper efetuada junto a 506 alunos do curso secundário revelou que 268 deles sequer sabiam que a palavra "Holocausto" se referia ao genocídio de cerca de 6 milhões de judeus, praticado pelos nazistas durante a II Guerra Mundial. Com os dados amostrais e um nível de 0,05 de significância, teste a hipótese de que mais de 50% dos estudantes não sabem a que se refere "Holocausto". Com base nesses resultados, Ralph Carter deve promover uma revisão do currículo?
6. Recorra ao Conjunto de Dados 8 do Apêndice B e considere apenas os estudantes de estatística com 21 anos de idade ou mais. Ache a percentagem de fumantes nesse grupo etário e teste a afirmação de que essa taxa é inferior a 32%, que é a taxa de fumantes para as pessoas com 21 anos ou mais de idade (com base em dados do U.S. National Institute on Drug Abuse — EUA). Há alguma razão para que os estudantes de estatísticas com idade de 21 anos ou mais fumem menos do que a população geral naquela faixa etária?
7. A Greybar Tax Company alega que seus clientes são selecionados para inspeção a uma taxa substancialmente superior a que se verifica na população geral. O Ministério da Fazenda reporta que fiscaliza 4,3% dos que ganham mais de $100.000, mas a verificação de 400 declarações da Greybar, selecionadas aleatoriamente, de clientes com renda superior a $100.000, acusou 56 inspecionados. Teste a afirmação da Greybar Tax Co., ao nível de 0,005 de significância. Se o leitor ganha mais de $100.000, utilizaria os serviços dessa companhia?
8. O presidente da Columbia Pictures, Mark Canton, afirma que 58% dos filmes produzidos têm classificação R. Teste essa alegação, ao nível de 0,10 de significância. Para dados amostrais, recorra ao Conjunto de Dados 10 do Apêndice B.
9. Um dirigente de televisão afirma que "menos da metade de todos os adultos são contra a violência exibida na televisão." Teste esta afirmação, utilizando os dados amostrais de uma pesquisa Roper em que 48% dos 1.998 adultos pesquisados manifestaram desagrado pela violência na televisão. Utilize o nível de 0,05 de significação.
10. "Os espectadores leigos fazem ressuscitação cardiopulmonar corretamente menos da metade das vezes," de acordo com um artigo na *USA Today*, que observou que, em 662 casos, 46% foram feitos corretamente. Teste a afirmação do artigo, ao nível de 0,025 de significância.
11. As companhias de seguro de automóvel estão cogitando de elevar os prêmios para aqueles que falam ao telefone enquanto dirigem. O National Consumers Group afirma que o problema realmente não é tão sério, porque apenas 10% dos motoristas usam telefone. A indústria do seguro faz um estudo e constata que, entre 500 motoristas selecionados aleatoriamente, 90 usam telefone (com base em

dados da revista *Prevention*). Teste a afirmação do grupo de consumidores ao nível de 0,02 de significância.

12. Bay Photo é uma firma de fotografia de São Francisco que deseja fazer uma campanha publicitária por telefone. Um dos sócios argumenta que clientes potenciais que não têm seus números no catálogo telefônico não serão atingidos, mas o gerente de marketing afirma que menos da metade dos residentes de São Francisco não têm os números no catálogo telefônico. Em uma amostra aleatória de 400 telefones residenciais de S. Francisco, 39% não estavam listados. Utilize um nível de significância de 0,005 para testar a afirmação do gerente de marketing.
13. Em estudos clínicos sobre o remédio Seldane contra alergia, 70 dos 781 pacientes experimentaram sonolência (com base em dados da Merrell Dow Pharmaceuticals, Inc.). Um concorrente alega que 8% dos usuários de Seldane experimentam sonolência. Teste a afirmação, ao nível de 0,10 de significância.
14. A Federal Aviation Administration está disposta a subvencionar uma pesquisa sobre desorientação espacial dos pilotos, se houver evidência amostral suficiente (ao nível de 0,01 de significância) para concluir que, entre os acidentes com aeronaves que envolvem essa desorientação, mais de três quartos são casos fatais. Um estudo de 500 acidentes de avião envolvendo desorientação espacial do piloto revelou que 91% desses acidentes resultaram em fatalidades (dados do Department of Transportation — EUA). Com base nesses resultados amostrais, a subvenção será aprovada?
15. A Dra. Kelly Roberts é diretora de uma escola de medicina e tem a seu cargo o planejamento de cursos para calouros. O presidente do conselho da escola está encorajando-a a aumentar a ênfase em pediatria, mas a Dra. Roberts alega que menos de 10% dos estudantes de medicina dos EUA preferem a pediatria, baseando essa alegação em dados amostrais que indicam que 64 dentre 1068 estudantes de medicina, selecionados aleatoriamente, escolheram a pediatria (com base em dados da Association of American Medical Colleges). Ao nível de 0,01 de significância, há evidência suficiente para apoiar o argumento da Dra. Roberts?
16. Um sistema de reservas da Air America acusa uma taxa de 7% de não-comparecimento. Adotou-se então um novo processo, pelo qual as reservas são confirmadas no dia anterior ao do vôo, fazendo-se um estudo de 5218 reservas pelo novo sistema, selecionadas aleatoriamente. Se se registraram 333 não-comparecimentos, teste a afirmação de que a taxa de não-comparecimento é menor no novo sistema. O novo sistema se afigura eficiente na redução do não-comparecimento?
17. Em um estudo de 71 fumantes que estavam procurando deixar de fumar utilizando uma terapia especial, 32 não estavam fumando um ano após o tratamento (com base em dados de "High-Dose Nicotine Patch Therapy," por Dale et al., *Journal of the American Medical Association*, Vol. 274, N.º 17). Ao nível de 0,10 de significância, teste a afirmação de que, dos fumantes que procuram deixar de fumar com aquela terapia, a maioria está fumando um ano após o tratamento. Esses resultados sugerem que a terapia não é eficaz?
18. A campanha presidencial Kennedy-Nixon foi extremamente acirrada. Kennedy venceu com 34.227.000 votos, contra 34.108.000 votos dados a Nixon. A proximidade dos resultados levou alguém a especular que o evento se assemelhou à jogada de uma moeda, e que Nixon poderia ter ganho em outro dia. Admita que os votos sejam dados amostrais selecionados aleatoriamente de uma grande população (embora, na realidade, não o sejam). Ao nível de 0,01 de significância, teste a afirmação de que a verdadeira proporção populacional pró-Kennedy excedeu 0,5. A diferença entre as votações se afigura casual ou significativa?
19. Uma pesquisa da Bruskin-Goldring Research junto a 1012 adultos mostrou que, entre os que compram tortas de frutas, 28% as comem, e 72% as destinam a outras finalidades, como calços de porta, alimento de pássaros e aterro. Isto surpreendeu os fabricantes de tortas de frutas, que supunham que essas tortas fossem um alimento apreciado. O presidente da Kansas Food Products Company alega que os resultados da pesquisa são falsos e que, na realidade, metade de todos os adultos consome suas tortas de frutas. Teste essa afirmação, ao nível de 0,01 de significância. Com base no resultado, acha que o fabricante das tortas de frutas deve cogitar de modificações para tornar seu produto mais atraente como alimento, ou mais adequado a outros usos?
20. Um artigo da revista *Prevention* abordou uma pesquisa feita junto a 1257 adultos. O artigo observou que, entre os pesquisados, 27% fumavam e 82% dos fumantes tinham tentado parar de fumar ao menos uma vez. Quantos, dentre os pesquisados, são os fumantes que tentaram deixar de fumar ao menos uma vez? Ao nível de 0,05 de significância, teste a afirmação de que menos de 25% dos adultos são fumantes que tentaram deixar o fumo ao menos uma vez.

7-5 Exercícios B: Além do Básico

21. Um repórter do *Providence Journal* afirma que 10% dos moradores da cidade acham que a prefeita está fazendo um bom trabalho. Teste essa afirmação se, em uma amostra aleatória de 15 moradores, nenhum aprova o trabalho da prefeita. Utilize o nível de significância de 0,05. Como $np = 1,5$ é menor do que 5, a distribuição normal não constitui uma aproximação adequada da distribuição de proporções amostrais e, assim, não se deve usar a estatística de teste dada nesta seção.
22. A Chemco, uma fornecedora de contêineres para resíduos químicos, constata que 3% de uma amostra de 500 unidades apresentam defeito. Sendo fundamentalmente desonesto, o gerente de produção da Chemco quer fazer uma declaração de que a taxa de unidades defeituosas não supera uma percentagem especificada, e não deseja ver sua afirmação rejeitada ao nível de 0,05 de significância, se os dados amostrais são usados. Qual é *a menor* percentagem de contêineres defeituosos que ele pode alegar nessas condições?
23. Com base no Conjunto de Dados 11 do Apêndice B, determine a proporção amostral dos confeitos M&M vermelhos. Use o resultado para testar a alegação da Mars Candy Company que 20% de seus confeitos M&M são vermelhos. Use o nível de 0,05 de confiança.
    a. Use o método tradicional.
    b. Use o método do valor P.
    c. Teste a afirmação construindo e interpretando um intervalo de 95% de confiança.
24. Veja o Exercício 9. Se o verdadeiro valor de p é 0,45, determine β, a probabilidade de um erro tipo II (veja Exercício 27 da Seção 7-3). (*Sug.:* No Passo 3, use os valores $p = 0.45$ e $pq/n = (0,45)(0,55)/1998$.)

7-6 Teste de uma Afirmação sobre um Desvio-Padrão ou uma Variância

Nas seções precedentes deste capítulo introduzir am-se métodos de teste de afirmações sobre médias e proporções populacionais. Nesta seção utilizam-se os mesmos procedimentos básicos para testar afirmações sobre o desvio-padrão σ ou a variância σ^2 de uma população. Para tais testes exige-se o seguinte:

Hipóteses para Testar Afirmações sobre σ ou σ^2

Ao testar uma hipótese sobre o desvio-padrão σ ou a variância σ^2 de uma população, admitimos que os valores da população sejam distribuídos normalmente.

Outros métodos de teste de hipóteses também exigiam uma população distribuída normalmente, mas os testes sobre desvios-padrão ou sobre variâncias não são tão ***robustos***, o que significa que as inferências podem ser bastante enganosas se a população não tem distribuição normal. Nesta seção, a condição de uma população distribuída normalmente é uma exigência muito mais estrita. Com a suposição de uma distribuição normal, a estatística de teste seguinte tem distribuição qui-quadrado com $n - 1$ graus de liberdade e valores críticos dados na Tabela A-4.

Estatística de Teste para Testar Hipóteses sobre σ ou σ^2

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

onde n = tamanho da amostra
s^2 = variância amostral
σ^2 = variância populacional (dada na hipótese nula)

A distribuição qui-quadrado foi introduzida na Seção 6-4, onde salientamos as seguintes propriedades importantes.

Propriedades da Distribuição Qui-Quadrado

1. Todos os valores de χ^2 são não-negativos, e a distribuição não é simétrica (veja Figura 7-14).
2. Há uma distribuição diferente para cada número de graus de liberdade (veja Figura 7-15).
3. Os valores críticos encontram-se na Tabela A-4, onde

$$\text{graus de liberdade} = n - 1$$

Obtêm-se os valores críticos na Tabela A-4 localizando primeiro a linha correspondente ao número adequado de graus de liberdade (onde gl $= n - 1$). Em seguida, utiliza-se o nível de significância α para determinar a coluna, conforme descrito a seguir.

Teste unilateral direito:	Localizar a área na parte superior da Tabela A-4 igual ao nível α.
Teste unilateral esquerdo:	Calcular $1 - \alpha$ e, em seguida, localizar a área na parte superior da Tabela A-4 igual a $1 - \alpha$.
Teste bilateral:	Calcular $1 - \alpha/2$ e $\alpha/2$ e localizar a área na parte superior da Tabela A-4 igual a $1 - \alpha/2$ (para o valor crítico esquerdo) e a área na parte superior igual a $\alpha/2$ (para o valor crítico direito). (Veja Figura 6-9 e o exemplo a ela relacionado.)

Fig. 7-14 Propriedades da distribuição qui-quadrado.

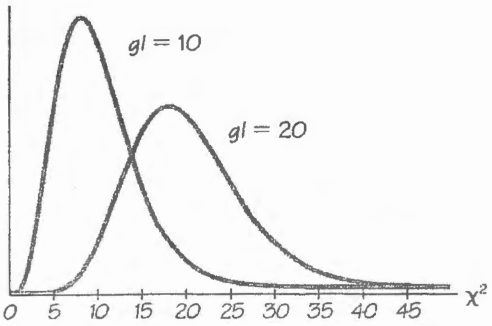


Fig. 7-15 Distribuição qui-quadrado para 10 e 20 graus de liberdade. Há uma distribuição diferente para cada número de graus de liberdade.

Ética em Experimentos

Os dados amostrais podem, em geral, ser obtidos simplesmente observando-se ou pesquisando-se elementos selecionados de uma população. Muitas outras situações exigem que, de alguma forma, manipulemos as circunstâncias para obter dados amostrais. Em ambos os casos podem surgir questões de ética. Pesquisadores em Tuskegee, Alabama, suspenderam o tratamento com penicilina de portadores de sífilis, a fim de que a doença pudesse ser estudada. Este experimento se prolongou por 27 anos!

Os técnicos de controle de qualidade querem não só garantir que um produto seja, em média, aceitável, como também fabricar produtos que tenham qualidade *consistente*, de modo que haja poucos itens defeituosos. Melhora-se a consistência reduzindo-se a variação. Por exemplo, a consistência dos altímetros para aviões é regida pela Federal Aviation Regulation (EUA) N.º 91.36, que exige que os altímetros sejam regulados e calibrados de modo a acusar uma leitura "a menos de 125 pés (em uma base de 95% de probabilidade)". Mesmo que a leitura da altitude média seja exatamente correta, um desvio-padrão excessivamente grande resultará em leituras individuais perigosamente altas ou baixas.

EXEMPLO A Stewart Aviation Products Company vem fabricando, com bons resultados, altímetros para aviões com erros distribuídos normalmente com média de 0 ft (obtida por calibragem) e desvio padrão de 43,7 ft. Após a instalação da

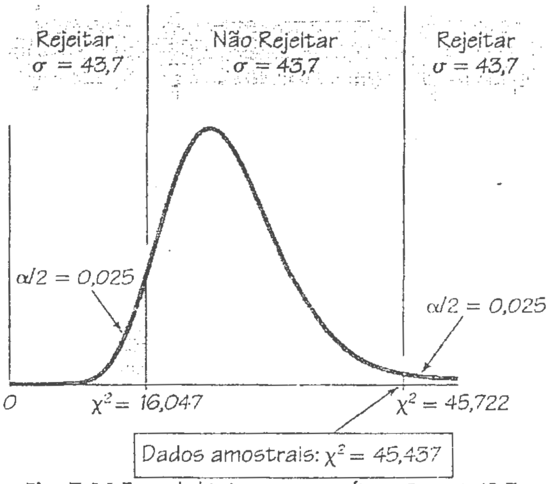


Fig. 7-16 Teste de hipótese para a afirmação $\sigma \neq 43,7$.

nova produção de equipamentos selecionaram-se aleatoriamente 30 altímetros da nova linha de produção. Esta amostra grupal acusou erros com desvio-padrão $s = 54,7$ ft. Ao nível de significância de 0,05, teste a afirmação de que os novos altímetros têm desvio-padrão diferente do valor anterior de 43,7 ft.

SOLUÇÃO Utilizaremos o método tradicional de teste de hipóteses, delineado na Figura 7-4.

Passo 1: A afirmação é expressa simbolicamente como $\sigma \neq 43,7$ ft.

Passo 2: Se a afirmação original é falsa, então $\sigma = 43,7$ ft.

Passo 3: Como a hipótese nula deve conter a igualdade, temos

$$H_0: \sigma = 43,7 \quad H_1: \sigma \neq 43,7 \text{ (afirmação original)}$$

Passo 4: O nível de significância é $\alpha = 0,05$.

Passo 5: Como se faz uma afirmação sobre σ, aplicamos a distribuição qui-quadrado.

Passo 6: A estatística de teste é

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(30-1)(54,7)^2}{43,7^2} = 45,437$$

Os valores críticos são 16,047 e 45,722; encontram-se na Tabela A-4, na 29.ª linha ($n - 1 = 29$ graus de liberdade) nas colunas correspondentes a 0,975 e 0,025. Veja a estatística de teste e os valores críticos ilustrados na Figura 7-16.

Passo 7: Como a estatística de teste não está na região crítica, não rejeitamos a hipótese nula.

Passo 8: Não há evidência suficiente para apoiar a afirmação de que o desvio-padrão seja diferente de 43,7 ft. (Para auxiliar no enunciado da conclusão final, veja a Figura 7-2.) Não obstante, seria aconselhável continuar monitorando e testando a nova linha de produção.

Há um aspecto no exemplo precedente que pode causar confusão: a referência a duas distribuições diferentes. A população de erros dos altímetros tem distribuição *normal*, mas no Passo 6 da solução precedente utilizamos a distribuição *qui-quadrado* porque a distribuição de $(n-1)s^2/\sigma^2$ é uma distribuição qui-quadrado. Os métodos desta seção exigem que a população original tenha distribuição normal, mas utilizamos a distribuição qui-quadrado para testar afirmações sobre desvios-padrão ou variâncias.

> **Ética em Relatórios**
>
> A American Association for Public Opinion Research elaborou um código voluntário de ética a ser usado no relato de resultados de pesquisas. Este código exige a inclusão do seguinte: (1) identificação do patrocinador da pesquisa, (2) data da realização da pesquisa, (3) tamanho da amostra, (4) natureza da população submetida a amostragem, (5) tipo de pesquisa utilizado, (6) fraseado exato do questionário. As pesquisas subvencionadas pelo governo dos EUA estão sujeitas a uma avaliação prévia do risco para os pesquisados, do mérito científico da pesquisa e da garantia do consentimento do pesquisado em participar.

O Método do Valor *P*

Em lugar da abordagem tradicional do teste de hipóteses, podemos também utilizar a abordagem pelo valor *P*, resumida nas Figuras 7-7 e 7-8. Em geral, não podemos achar valores *P exatos* porque a tabela da distribuição qui-quadrado (Tabela A-4) compreende apenas alguns valores de α. Em vez disto, utilizamos a tabela para identificar limites que contenham o valor *P*. A estatística de teste do último exemplo é $\chi^2 = 45,437$, e sabemos que o teste é bilateral com 29 graus de liberdade. Verificando a 29.ª linha da Tabela A-4, vemos que a estatística de teste 45,437 está entre os valores críticos 42,557 e 45,722, o que indica que a área à direita da estatística de teste está entre as áreas correspondentes de 0,05 e 0,025. A Figura 7-7 resume o processo geral para obter valores *P* e indica também que, em um teste bilateral com uma estatística de teste à direita do centro, o valor *P* é *duas vezes* a área à direita da estatística de teste. O valor *P* está, pois, entre 0,10 e 0,05, o que se expressa como

$$0,05 < \text{valor } P < 0,10$$

Como o valor *P* é superior ao nível de significância $\alpha = 0,05$, novamente não rejeitamos a hipótese nula. Vemos também, mais uma vez, que o método tradicional e o método do valor *P* são equivalentes, no sentido de que sempre levam à mesma conclusão.

A abordagem pelo valor *P* se torna particularmente importante quando usamos programas de computador que fornecem valores *P*, mas não valores críticos. Se o exemplo precedente for feito utilizando-se o STATDISK, o valor *P* 0,0533 está incluído nos resultados, mas o STATDISK fornece igualmente os valores críticos. O Minitab e a calculadora TI-83 não estão programados para testar afirmações que envolvem desvio-padrão ou variância, mas podem ser utilizados para achar as componentes necessárias para tal teste.

7-6 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1 e 2, use a Tabela A-4 para achar os valores críticos de χ^2, com base na informação dada.

1. a. H_0: $\sigma = 15$	b. H_1: $\sigma > 0{,}62$	c. H_1: $\sigma < 14{,}4$
$n = 10$	$n = 27$	$n = 21$
$\alpha = 0{,}01$	$\alpha = 0{,}01$	$\alpha = 0{,}05$
2. a. H_1: $\sigma < 1{,}22$	b. H_1: $\sigma > 92{,}5$	c. H_0: $\sigma = 0{,}237$
$n = 23$	$n = 12$	$n = 16$
$\alpha = 0{,}025$	$\alpha = 0{,}10$	$\alpha = 0{,}05$

Nos Exercícios 3-14, aplique o método tradicional para testar a hipótese levantada. Siga os passos indicados na Figura 7-4 e esboce o gráfico apropriado. Em todos os casos, admita que a população é distribuída normalmente e que a amostra foi selecionada aleatoriamente.

3. A Stewart Aviation Products Company utiliza um novo método de produção para fabricar altímetros para aviões. Uma amostra aleatória de 81 altímetros acusou erros com desvio-padrão $s = 52{,}3$ ft. Ao nível de 0,05 de significância, teste a afirmação de que a nova linha de produção tem erros com desvio-padrão diferente de 43,7 ft, que é o desvio-padrão do método de produção anterior. Se o desvio-padrão parece ter mudado, acha que o novo método é pior ou melhor do que o antigo?
4. Testes feitos em turmas passadas de estatística do autor acusaram notas com desvio-padrão de 14,1. Uma das turmas atuais tem agora 27 notas de teste com desvio-padrão de 9,3. Teste, ao nível de significância de 0,01, a afirmação de que a turma atual tem menor variação do que as turmas anteriores. Um desvio-padrão menor sugere um melhor desempenho da turma atual?
5. Com filas separadas em cada um de seus vários guichês, o Jefferson Valley Bank constatou que o desvio-padrão das filas de espera, normalmente distribuídas, nas tardes de sexta-feira, era de 6,2 min. O banco experimentou trabalhar com uma fila única e verificou que, para uma amostra de 25 clientes, os tempos de espera têm um desvio-padrão de 3,8 min. Com base em estudos prévios, podemos admitir que os tempos de espera tenham distribuição normal. Ao nível de 0,05 de significância, teste a afirmação de que uma fila única apresenta menor variação dos tempos de espera. (Os clientes tendem a ficar mais satisfeitos se os tempos de espera apresentam menor variação.) A adoção de uma fila única resulta realmente em menor espera?
6. As alturas dos homens variam mais do que as das mulheres? A enfermeira de um colégio feminino coletou dados de exames físicos. As alturas das mulheres acusaram um desvio-padrão de 2,5 in. O colégio agora tornou-se misto e os primeiros 20 estudantes do sexo masculino têm alturas com desvio-padrão de 2,8 in. Ao nível de significância de 0,025, teste a afirmação de que as alturas dos homens variam mais do que as das mulheres. Se a afirmação não for consistente, significa que as alturas dos homens não variam mais do que as das mulheres?
7. A Kansas Farm Products Company usa uma máquina que enche sacos de 50 lb de semente de milho. No passado, a máquina apresentava um desvio-padrão de 0,75 lb. Com o intuito de obter pesos mais compatíveis, os mecânicos substituíram algumas peças gastas da máquina. Uma amostra de 61 sacos, extraída da máquina já reparada, acusou média amostral de 50,13 lb e desvio-padrão amostral de 0,48 lb. Ao nível de 0,05 de significância, teste a suposição de que os pesos são "mais compatíveis" com a máquina reparada, em comparação com a situação anterior. Se os pesos são mais compatíveis, qual é o efeito sobre o desvio-padrão?
8. A Collins Investment Company constata que, se o desvio-padrão dos tempos semanais de interrupção de seu computador é de 2 h ou menos, então o computador é previsível, o que facilita o planejamento. Selecionaram-se aleatoriamente doze interrupções semanais do computador, calculando-se seu desvio-padrão, que foi calculado em 2,85 h. O gerente de operações alega que os tempos de acesso ao computador são imprevisíveis porque o desvio-padrão excede 2 h. Teste essa alegação ao nível de 0,025 de significância. A variação parece demasiadamente alta?
9. A Medassist Pharmaceutical Company utiliza uma máquina para encher frascos com um remédio, de tal maneira que o desvio-padrão dos pesos é de 0,15 oz. Testou-se uma nova máquina em 71 frascos e, para essa amostra, o desvio-padrão é 0,12 oz. A Dayton Machine Company, fabricante da nova máquina, afirma que ela enche os frascos com menor variação. Teste a afirmação da Dayton Machine Company, ao nível de 0,05 de significância. Se a máquina da Dayton está sendo usada como experiência, deve-se cogitar de sua aquisição?
10. Para adultos selecionados aleatoriamente, os QIs são distribuídos normalmente com média 100 e desvio-padrão 15. Uma amostra de 24 professores de faculdade, aleatoriamente escolhidos, acusou QIs com desvio-padrão de 10. Um psicólogo está convencido de que os professores de faculdade têm QIs com média superior a 100. Ele não entende muito bem o conceito de desvio-padrão e afirma que o desvio-padrão dos QIs dos professores de faculdade é igual a 15, o mesmo desvio-padrão da população em geral. Teste a afirmação ao nível de 0,05 de significância. Com base no resultado, que podemos concluir quanto ao desvio-padrão dos QIs dos professores de faculdade?
11. A pressão sistólica resulta da contração do coração. Comparando os níveis de pressão sistólica de homens e mulheres, a Dra. Jane Taylor obteve as mensurações para uma amostra aleatória de 50 mulheres, obtendo uma média amostral de 130,7 e um desvio-padrão amostral de 23,4. Se se sabe que os níveis de pressão sistólica dos homens acusam média e desvio-padrão de 133,4 e 19,7, respectivamente, teste a hipótese de que as mulheres apresentam maior variação. Use o nível de 0,05 de significância. (Todos os valores são em milímetros de mercúrio, e os dados se baseiam no National Health Survey.)
12. Se utilizamos as temperaturas do corpo humano ao meio-dia do dia 2, relacionadas no Conjunto de Dados 2 do Apêndice B, rejeitamos (ao nível de 0,05 de significância) a afirmação de que $\mu = 98{,}6°F$. Essas 106 temperaturas têm distribuição aproximadamente normal, com média $\bar{x} = 98{,}20°F$ e desvio-padrão $s = 0{,}62°F$. A estatística de teste indicará a rejeição de $\mu = 98{,}6°F$ desde que o desvio-padrão seja inferior a 2,11°F. Com a estatística amostral e um nível de 0,005 de significância, teste a afirmação de que $\sigma <$ 2,11°F.
13. Com base em dados do National Health Survey (EUA), os homens na faixa etária 25-34 têm alturas com desvio-padrão de 2,9 in. Teste a afirmação de que os homens na faixa etária 45-54 têm alturas com desvio-padrão diferente. Relacionam-se a seguir as alturas de 25 homens selecionados aleatoriamente na faixa 45-54:

66,80	71,22	65,80	66,24	69,62	70,49	70,00	71,46	65,72
68,10	72,14	71,58	66,85	69,88	68,69	72,77	67,34	68,40
68,96	68,70	72,69	68,67	67,79	63,97	67,19		

14. A seguir estão relacionados os pesos (em quilogramas), ao nascer, de meninos nascidos de mães que se submeteram a uma dieta especial suplementar de vitaminas (com base em dados do New York State Department of Health — EUA). Teste a afirmação de que esta amostra provém de uma população com desvio-padrão de 0,470 kg, que é o desvio-padrão dos pesos de meninos recém-nascidos em geral. Acha que a dieta suplementar de vitaminas afeta a variação dos pesos ao nascer?

3,73	4,37	3,73	4,33	3,39	3,68	4,68	3,52
3,02	4,09	2,47	4,13	4,47	3,22	3,43	2,54

7-6 Exercícios B: Além do Básico

15. Com auxílio da Tabela A-4, determine a amplitude dos valores P possíveis nos exercícios dados.

a. Exercício 3 **b.** Exercício 5 **c.** Exercício 9

16. Para grandes números de graus de liberdade, podemos aproximar valores críticos de χ^2 como segue:

$$\chi^2 = \frac{1}{2}(z + \sqrt{2k - 1})^2$$

Aqui, k é o número de graus de liberdade e z é o valor crítico, encontrado na Tabela A-2. Por exemplo, se queremos aproximar os dois valores críticos de χ^2 em um teste de hipótese bilateral, com $\alpha = 0{,}05$ e $n = 150$, tomamos $k = 149$ com $z = -1{,}96$, seguido de $k = 149$ e $z = 1{,}96$.

a. Utilize esta aproximação para estimar os valores críticos de χ^2 em um teste bilateral com $n = 101$ e $\alpha = 0{,}05$. Compare os resultados com os encontrados na Tabela A-4.

b. Com esta aproximação, estime os valores críticos de χ^2 em um teste bilateral com $n = 150$ e $\alpha = 0{,}05$.

17. Refaça o Exercício 16 usando esta aproximação (com k e z dados no Exercício 16):

$$\chi^2 = k\left(1 - \frac{2}{9k} + z\sqrt{\frac{2}{9k}}\right)^3$$

18. No Exercício 5, supondo σ efetivamente igual a 4,0, determine β (a probabilidade de um erro tipo II). Veja o Exercício 27 da Seção 7-3 e modifique o processo de modo que possa ser aplicado a um teste de hipóteses envolvendo σ em lugar de μ.

Vocabulário

hipótese	estatística de teste
teste de hipóteses	região crítica
teste de significância	valor crítico
hipótese nula	teste bilateral
hipótese alternativa	teste unilateral esquerdo
erro tipo I	teste unilateral direito
erro tipo II	método tradicional
nível de significância	método clássico
alfa (α)	valor P
beta (β)	valor de probabilidade

Revisão

Nos Capítulos 6 e 7 introduzimos dois conceitos importantes, utilizando dados amostrais para fazer inferências sobre dados populacionais: estimação de valores de parâmetros populacionais (Capítulo 6) e teste de afirmações feitas sobre parâmetros populacionais (Capítulo 7). Os parâmetros considerados nos Capítulos 6 e 7 são médias, proporções, desvios-padrão e variâncias.

Na Seção 7-2 apresentamos os conceitos fundamentais de um teste de hipóteses: hipótese nula, hipótese alternativa, erro tipo I, erro tipo II, estatística de teste, região crítica, valor crítico e nível de significância. Abordamos também testes bilaterais, testes unilaterais esquerdos, testes unilaterais direitos e a formulação de conclusões. Na Seção 7-3 utilizamos essas componentes para identificar três métodos diferentes para testar hipóteses:

1. O método tradicional (sintetizado na Figura 7-4)
2. O método do valor P (sintetizado na Figura 7-8)
3. Intervalos de confiança (abordados no Capítulo 6)

Nas Seções 7-3 a 7-6 discutimos métodos específicos para lidar com os diferentes parâmetros. Como é de suma importância fazer a escolha correta da distribuição e da estatística de teste, apresentamos a Figura 7-17, que resume as decisões-chave a serem tomadas.

Exercícios de Revisão

Nos Exercícios 1 e 2, ache os valores críticos apropriados. Em todos os casos, suponha que o desvio-padrão populacional σ é desconhecido.

1. a. H_1: $\mu < 27{,}3$, $n = 10$, $\alpha = 0{,}05$ b. H_1: $p \neq 0{,}5$, $n = 150$, $\alpha = 0{,}01$ c. H_1: $\sigma \neq 15$, $n = 20$, $\alpha = 0{,}05$
2. a. H_0: $\mu = 19{,}9$, $n = 40$, $\alpha = 0{,}10$ b. H_0: $\sigma \leq 0{,}93$, $n = 10$, $\alpha = 0{,}025$ c. H_0: $p \geq 0{,}25$, $n = 540$, $\alpha = 0{,}01$

Nos Exercícios 3 e 4, responda a cada item:

a. Dê a hipótese nula em forma simbólica.
b. Este teste é unilateral esquerdo, unilateral direito ou bilateral?
c. Descreva o erro tipo I em termos simples, destituídos de simbolismo e linguagem técnica.
d. Descreva o erro tipo II em termos simples, destituídos de simbolismo e linguagem técnica.
e. Identifique a probabilidade de cometer um erro tipo I.

3. A afirmação de que os consultores estatísticos têm uma renda média de \$90.000 deve ser testada ao nível de 0,05 de significância.
4. A afirmação de que a duração dos comerciais CBS tem desvio-padrão inferior a 15 s deve ser testada ao nível de 0,01 de significância.
5. Uma pesquisa *USA Today*/CNN/Gallup foi a base de uma afirmação de que "57% dos portadores de arma de fogo são a favor de leis mais severas para o porte de armas". Suponha que o grupo de pesquisa consista em 504 portadores de arma selecionados aleatoriamente. Teste a afirmação de que a maioria (mais de 50%) dos portadores de arma são a favor de leis mais severas sobre o porte de armas de fogo. Use o nível de 0,01 de significância.
6. Em uma máquina de raios X dental, um rótulo afirma que a máquina libera dosagens de radiação com menos de 5.00 miliroentgens. Os dados amostrais consistem em 36 observações selecionadas aleatoriamente, com média de 4,13 miliroentgens e desvio-padrão de 1,91 miliroentgens. Teste a afirmação feita no rótulo, utilizando o nível de 0,01 de significância.
7. Com os mesmos dados amostrais do Exercício 6, utilize o nível de significância de 0,01 para testar a afirmação de que as dosagens de radiação têm um desvio-padrão inferior a 2,50 miliroentgens. Qual é uma conseqüência prática de um desvio-padrão muito elevado neste caso?
8. Elvis ainda está vivo? O *USA Today* publicou uma reportagem sobre uma pesquisa da Universidade da Carolina do Norte feita junto a 1248 adultos do sul dos EUA. Constatou-se que 8% dos pesquisados crêem que Elvis Presley ainda está vivo. O artigo começa com a afirmação de que "quase 1 em cada 10" sulistas ainda acredita que Elvis esteja vivo. Ao nível de 0,01 de significância, teste a afirmação de que a verdadeira percentagem é inferior a 10%. Com base no resultado, determine se o resultado amostral de 8% justifica a expressão "quase 1 em cada 10."

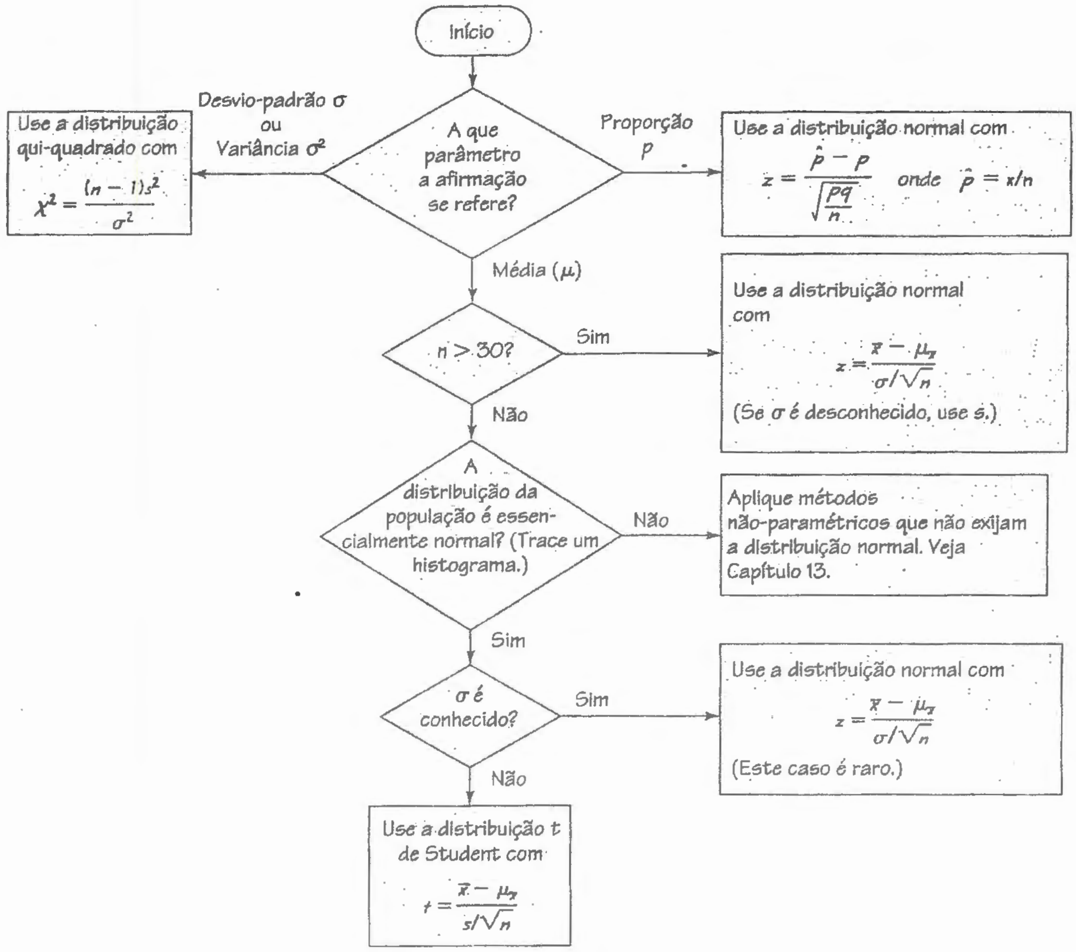


Fig. 7-17 Teste de uma afirmação sobre uma média, uma proporção, um desvio-padrão ou uma variância.

9. A Boston Bottling Company distribui cola em latas que anunciam o conteúdo de 12 oz. O Departamento de Pesos e Medidas (Bureau of Weights and Measures) selecionou aleatoriamente 24 latas, mediu seus conteúdos e obteve uma média de 11,82 oz, com desvio-padrão de 0,38 oz. Ao nível de 0,01 de significância, teste a afirmação de que a companhia está ludibriando os consumidores.

10. A Medassist Pharmaceutical Company fabrica um comprimido para crianças suscetíveis de convulsões. Admite-se que o comprimido contenha 20,0 mg de fenobarbital. Uma amostra aleatória de 20 comprimidos acusou os pesos (em mg) relacionados abaixo. Ao nível de significância $\alpha = 0,01$, esses comprimidos são aceitáveis?

27,5	26,0	22,9	23,4	23,0
23,9	32,6	20,9	22,9	24,3
24,8	16,1	24,3	17,3	18,9
20,7	33,0	15,6	24,3	23,3

Exercícios Cumulativos de Revisão

1. Para mulheres sadias na faixa etária 18-24, os valores da pressão sistólica (em mm Hg) têm distribuição normal com média 114,8 e desvio-padrão 13,1 (com base em dados do National Health Survey — EUA).
   a. Selecionada, aleatoriamente, uma mulher sadia dentre a população geral de todas as mulheres com idade entre 18 e 24, determine a probabilidade de obter uma com pressão sistólica superior a 124,23.
   b. Selecionadas aleatoriamente 16 mulheres, determine a probabilidade de a média de suas pressões sistólicas ser superior a 124,23.

2. Um pesquisador médico obtém os valores (em mm Hg) da pressão sistólica de mulheres na faixa etária 18-24, relacionados a seguir, que sofreram uma infecção causada por um vírus recentemente. (Como no Exercício 1, as mulheres naquela faixa etária têm pressão sistólica normalmente distribuídas, com média de 114,8 e desvio-padrão de 13,1.)
   a. Determine a média amostral $\bar{x}$ e o desvio-padrão s.
   b. Ao nível de 0,05 de significância, teste a afirmação de que a amostra provém de uma população com média igual a 114,8.
   c. Utilizando os dados amostrais, construa um intervalo de 95% de confiança para a média populacional μ. Os limites do intervalo abrangem o valor 114,8, que é a média para as mulheres sadias na faixa etária 18-24?
   d. Ao nível de 0,05 de significância, teste a afirmação de que a amostra provém de uma população com desvio-padrão igual a 13,1, que é o desvio-padrão para as mulheres sadias na faixa etária 18-24.
   e. Com base nos resultados precedentes, parece que o novo surto de infecção afeta a pressão sistólica?

134,9	78,7	108,9	133,0	123,7	96,1	126,9	89,8
132,0	134,7	132,1	121,7	112,3	150,2	158,3	154,4

3. Um estudante de psicologia planeja um experimento para testar a percepção extra-sensorial (PES). Nesse experimento, escolhe-se uma carta de um baralho bem misturado, e a pessoa, sem vê-la, deve adivinhar o seu naipe (paus, ouros, copas, espadas). Repete-se o experimento 25 vezes, repondo-se cada vez a carta no baralho.
   a. Para pessoas que adivinham, sem qualquer PES, determine o número médio de respostas corretas.
   b. Para pessoas que fazem suposições aleatórias, sem qualquer PES, ache o desvio-padrão do número de respostas corretas.
   c. Para pessoas que fazem suposições sem PES, determine a probabilidade de obter mais de 12 respostas corretas.
   d. Se uma pessoa deu mais de 12 respostas corretas, teste a afirmação de que ela fez suposições aleatórias. Utilize o nível de 0,05 de significância.
   e. Deseja-se fazer uma pesquisa para estimar a percentagem dos americanos adultos que acreditam que algumas pessoas tenham PES. Quantas pessoas devem ser entrevistadas para que se tenha 90% de confiança de que o erro da percentagem amostral não supere quatro pontos percentuais? 4. e. 423
4. Uma exigência importante nas Seções 7-4 e 7-6 é de que os dados amostrais devem provir de uma população distribuída normalmente.
   a. Construa um gráfico ramo-e-folhas para os comprimentos de cabeças de urso do Conjunto de Dados 3 do Apêndice B.
   b. Com base nos resultados, a distribuição parece ser normal? 4. b. Yes
   c. Que outro método poderia ser usado para determinar se a distribuição é normal? 4. c. a histogram
   d. Há algo particularmente digno de nota nesse conjunto de dados? 4. d. Yes

Projeto para Computador

Observe as temperaturas (em °F) resumidas no ramo-e-folhas a seguir.

```
97. | 01
97. | 589
98. | 00044
98. | 5566778
99. | 02
```

(A amostra é um subconjunto das temperaturas relacionadas na Tabela 7-1.) Use o STATDISK ou o Minitab (ou uma calculadora TI-83) e o nível de significância de 0,05 para os seguintes testes:

a. Teste a afirmação de que a amostra provém de uma população com média de 98,6°F.
b. Adicione 5 a cada temperatura original do ramo-e-folhas e teste a afirmação de que a amostra provém de uma população com média de 103,6. Compare os resultados com os obtidos na parte (a). De modo geral, o que varia quando uma mesma constante é adicionada a cada valor?
c. Após multiplicar por 10 cada temperatura original, teste a afirmação de que a amostra provém de uma população com média de 986. Compare os resultados com os achados na parte (a). De modo geral, o que varia quando cada valor é multiplicado pela mesma constante?
d. As temperaturas Fahrenheit podem ser convertidas para a escala Celsius mediante a fórmula $C = (5/9)(F - 32)$. Ou seja, primeiro subtraímos 32 de cada temperatura Fahrenheit, em seguida multiplicamos cada resultado por 5/9. Como a conversão da escala Fahrenheit para a escala Celsius envolve a adição da mesma constante (−32) e a multiplicação pela mesma constante (5/9), pode-se prever o que ocorre quando se convertem as temperaturas originais e se testa a hipótese $\mu = 37°C$? (Note que 98,6°F = 37°C.) Verifique sua predição (confirmando-a ou não) convertendo as temperaturas originais para a escala Celsius e testando a afirmação de que $\mu = 37°C$.
e. Ao testar a afirmação de que o peso médio dos homens é hoje superior ao peso médio conhecido dos homens que viveram há 100 anos, faz alguma diferença se os pesos são medidos em libras ou em quilogramas?

DOS DADOS À DECISÃO

Deve-se Abolir o Desconto no Seguro?

As companhias de seguro estão reestudando o desconto que costumam dar a carros equipados com um dispositivo antitrava de freios. Os carros avaliados em mais de $15.000 e não-equipados com esse dispositivo têm danos cujos valores acusam média de $8908 quando colidem. Uma amostra de carros de valor superior a $15.000 equipados com o dispositivo acusou os danos (em dólares) listados a seguir (com base em dados do Highway Loss Data Institute). Teste, ao nível de 0,05, a afirmação de que os valores dos danos não são diferentes quando os carros têm o dispositivo. Com base nesses resultados, as seguradoras devem abolir tais descontos? Em vista dos resultados, que decisão tomaria como gerente de uma companhia que faz seguro de carros?

9982	7509	9359	8934	6512	9761	11226	9332	10047	10200	8417	6948
11491	6902	9866	9227	6060	7260	9026	8651	8974	11184	8079	9371
9977	7731	8332	8750	6669	9545	7953	7483	7660	8891	7277	7467

ATIVIDADES EM GRUPO

1. *Atividade Extraclasse:* Uma atividade em grupo sugerida no Capítulo 6 envolvia a estimativa do erro médio (em segundos) dos relógios de pulso das pessoas. Use os mesmos dados coletados naquela ocasião para testar a afirmação de que o erro médio de todos os relógios de pulso é 0. Recorde que alguns erros são positivos (o relógio está adiantado) e outros são negativos (o relógio está atrasado). Descreva os detalhes do teste e inclua um gráfico mostrando a estatística de teste e os valores críticos. Coletivamente, andamos adiantados ou atrasados? Teste também a afirmação de que o desvio-padrão dos erros é inferior a 1 minuto. Quais são as implicações práticas de um desvio-padrão excessivamente grande?
2. *Atividade em Classe:* Em um grupo de três ou quatro pessoas, faça um experimento de PES, selecionando um dos membros do grupo. Trace um círculo em um pequeno pedaço de papel, e um quadrado em outro pedaço de papel do mesmo tamanho. Repita o experimento 20 vezes: Escolha aleatoriamente o círculo ou o quadrado e coloque-o na mão da pessoa de maneira que não possa ser vista (mão para trás), e peça à pessoa que identifique a forma do desenho (sem olhar para ele); registre se a resposta é correta. Teste a afirmação de que a pessoa tem PES porque a proporção de respostas corretas é superior a 0,5.
3. *Atividade em Classe:* Após dividir em grupos de 10 a 20 pessoas, cada membro do grupo deve registrar o número de batimentos cardíacos em um minuto. Após calcular $\bar{x}$ e s, cada grupo deve proceder ao teste da afirmação de que a média é superior a 59, que é o resultado do autor. (Quando a pessoa faz exercícios, tende a ter pulsações mais baixas, e o autor corre cinco milhas algumas vezes por semana.)

entrevista

Jay Dean

Vice-Presidente Sênior da San Francisco Office of Young & Rubicam Advertising

Jay Dean é diretor do Consumer Insights Department na sucursal de San Francisco da firma Young & Rubicam. Trabalhou para muitas empresas anunciadoras conhecidas, entre elas AT&T, Chevron, Clorox, Coors, Dr. Pepper, Ford Motor Co., General Foods, Gillette, Gulf Oil, H.J. Heinz, Kentucky Fried Chicken e Warner Lambert.

Até que ponto utiliza-se a estatística na Young & Rubicam?

Utilizamos a estatística diariamente. Fazemos um grande volume de pesquisas de consumo, bem como muitas pesquisas sobre atitudes e hábitos do consumidor. Para alguns clientes fazemos testes de produtos, testes de gasto, estudos de estratégia — qualquer coisa de que necessitem. Estou agora trabalhando em um planejamento típico, um teste de propaganda para os molhos de salada Take Heart, uma das marcas com que trabalhamos para a Clorox Company. Dois comerciais foram apresentados a amostras independentes de consumidores. Formulamos questões sobre o que cada comercial transmitia, percepções de marca, gastos e aversões etc. Faremos testes de significância para comparar os resultados e escolher o melhor comercial.

Em pesquisas mais amplas, utilizamos técnicas multivariadas como análise fatorial e regressão múltipla. O *marketing* está se tornando cada vez mais agressivo. Há maior número de marcas, maior concorrência de preços e consumidores mais sofisticados. Um ambiente mais agressivo exige mais pesquisa de mercado, e isto significa a utilização crescente da estatística.

Poderia citar um caso em que o uso da estatística foi fundamental para definir uma estratégia bem-sucedida?

Recentemente fizemos uma pesquisa para a Pine-Sol, outra marca do grupo Clorox. Produziram-se cerca de meia dúzia de esboços de comerciais, que foram apresentados a amostras independentes de consumidores. A estatística ajudou a identificar o melhor comercial para a nova campanha publicitária do Pine-Sol. Os primeiros resultados são muito encorajadores. Na Y&R acreditamos que a pesquisa ajude a planejar a propaganda mais eficiente. A estatística contribui para a tomada de decisões, com base nos resultados da pesquisa.

Como se costuma coletar dados para análise estatística?

Há muitas maneiras de coletar dados, mas, em geral, ou se faz uma amostragem telefônica aleatória para trabalhos de pesquisa, ou, se queremos mostrar às pessoas algo como um comercial, utilizamos o sistema de entrevista em um local central. Por exemplo, as pessoas são abordadas em um *shopping* por um entrevistador, sondadas quanto à aceitação, convidadas a ir a um local onde lhes é apresentado um comercial, etc.

Acha difícil obter amostras representativas e não-tendenciosas?

Sim. A indústria da pesquisa de mercado está agora muito preocupada com a crescente taxa de recusa do público em geral. Isto é devido, em parte, ao fato de que certos elementos disfarçam uma tentativa de venda sob a capa de uma pesquisa de mercado. As pesquisas pelo correio estão sujeitas a uma grande tendenciosidade devida à auto-seleção, e as taxas de resposta são em geral bem baixas. As entrevistas nos *shoppings* têm outros problemas. Nem todo mundo vai aos *shoppings*, e há certo grau de tendenciosidade por parte do entrevistador ao abordar as pessoas.

Qual é o seu tamanho típico de amostra?

Para uma pesquisa de âmbito nacional, a regra empírica indica cerca de mil pessoas, embora, em alguns casos, 600 sejam suficientes. Para um teste de propaganda, uma amostra da ordem de 200 seria um tamanho bastante satisfatório, mas às vezes utilizam-se apenas cem.

Em seu campo de trabalho, acha que os candidatos a emprego são favorecidos se tiverem estudado alguma estatística?

Sim, certamente. Todos têm de utilizar a estatística em algum nível. É muito importante para as pessoas admitidas conhecer bem a estatística, porque vão realizar grande parte de nosso trabalho de pesquisa. No início, cabe-lhes a tarefa de destrinchar e analisar dados.

Tem algum conselho para o estudante de hoje?

Se eu pudesse voltar à escola, certamente estudaria mais matemática, estatística e ciência da computação. Estudei muito essas matérias, mas gostaria de ter estudado ainda mais. Há uma grande explosão de dados nos negócios em dias atuais. Todos os ramos de negócio estão se tornando mais quantitativos do que jamais o foram no passado. Hoje há maior quantidade de informações do que podemos manejar, e devemos dispor dos instrumentos analíticos e do conhecimento para lidar com essas informações, se queremos ser bem-sucedidos.

Rayane Andrade

Triola

8 Inferências com Base em Duas Amostras

8-1 Aspectos Gerais

Nos Capítulos 6 e 7 apresentamos os dois tópicos principais da inferência estatística: a estimação de um parâmetro populacional e o teste de uma hipótese sobre uma população. Naqueles capítulos consideramos casos envolvendo uma única população; aqui, vamos lidar com duas populações. Descreveremos métodos de teste de hipóteses e a construção de intervalos de confiança para casos de duas populações.

8-2 Inferências sobre Duas Médias: Amostras Dependentes

Utilizam-se duas amostras dependentes (dados emparelhados) para testar hipóteses sobre as médias de duas populações, e apresentam-se métodos para a construção de intervalos de confiança como estimativas da diferença entre duas médias populacionais.

8-3 Inferências sobre Duas Médias: Amostras Grandes e Independentes

São utilizadas amostras independentes e grandes ($n > 30$) para testar hipóteses sobre as médias de duas populações. Apresentam-se métodos para a construção de intervalos de confiança usados para estimar a diferença entre duas médias populacionais.

8-4 Comparação de Duas Variâncias

Apresenta-se um método para testar hipóteses formuladas sobre duas variâncias ou desvios-padrão populacionais.

8-5 Inferências sobre Duas Médias: Amostras Independentes e Pequenas

Esta seção aborda casos em que duas amostras são independentes e pequenas ($n \leq 30$). Discutem-se processos para testar hipóteses sobre a diferença entre duas médias populacionais, e métodos para a construção de intervalos de confiança usados para estimar a diferença entre duas médias.

8-6 Inferências sobre Duas Proporções

Esta seção estuda métodos de inferência estatística aplicados a duas proporções populacionais, métodos para testar afirmações sobre a diferença entre duas proporções populacionais, e a construção de intervalos de confiança para estimar a diferença entre duas proporções.

Problema do Capítulo

Os filtros de cigarros realmente fazem diferença, ou são apenas truques de venda sem qualquer efeito real?

Na Tabela 8-1 relacionamos os conteúdos de alcatrão, nicotina e monóxido de carbono em uma amostra aleatória de cigarros tamanho padrão, com filtro e sem filtro. Todas as medidas são em miligramas, e os dados são da Federal Trade Commission. Na Figura 8-1 utilizamos diagramas em caixa para comparar os cigarros com filtro e os sem filtro. Esses diagramas sugerem diferenças essenciais, mas as duas amostras são relativamente pequenas ($n_1 = 21$ e $n_2 = 8$), de modo que podemos questionar se essas diferenças aparentes são realmente signifiticativas. Mais adiante nesta seção voltaremos aos dados da Tabela 8-1 ao testarmos a igualdade de duas médias populacionais.

Tabela 8-1 Alcatrão, Nicotina e Monóxido de Carbono em Cigarros

Com Filtro			Sem Filtro		
Alcatrão	Nicotina	CO	Alcatrão	Nicotina	CO
16	1,2	14	23	1,6	14
15	1,3	12	23	1,9	15
16	1,1	14	24	1,6	17
14	1,1	16	26	1,8	17
16	1,0	15	25	1,7	16
1	0,1	2	26	1,7	16
16	1,1	14	21	1,4	14
18	1,0	16	24	1,5	16
10	0,8	11			
14	1,0	13			
12	0,9	13			
11	0,8	12			
14	1,0	13			
13	1,0	12			
13	1,0	13			
13	0,9	14			
16	1,2	14			
16	1,1	14			
8	0,1	9			
16	1,2	17			
11	0,9	12			

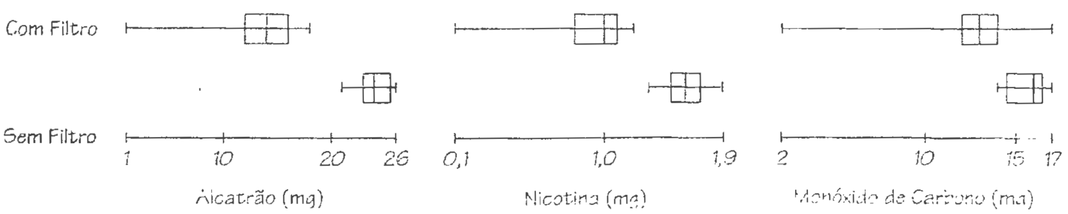


Fig. 8-1 Comparação entre cigarros com filtro e cigarros sem filtro.

8-1 Aspectos Gerais

No Capítulo 6 apresentamos um aspecto importante da inferência estatística: a utilização de amostras para construir intervalos de confiança, usados para estimar valores de parâmetros populacionais. No Capítulo 7 introduzimos um segundo aspecto igualmente importante da inferência estatística: a utilização de amostras para testar hipóteses sobre parâmetros populacionais. Nestes dois capítulos, os exemplos e exercícios envolviam o uso de *uma* amostra para fazer uma inferência sobre *uma* população. Na verdade, entretanto, há muitas situações importantes em que se faz necessário comparar *dois* conjuntos de dados amostrais. Seguem exemplos do tipo dos que serão encontrados neste capítulo, e que apresenta métodos de utilização de dados de duas amostras, permitindo fazer inferências sobre as populações de onde provêm.

- Determinar se há diferença entre o conteúdo médio de nicotina de cigarros com filtro e cigarros sem filtro.
- Determinar se as mulheres empregadas, expostas a éteres etil-glicóis, apresentam taxa de aborto (natural) diferente da taxa entre mulheres que não estão em contato com aqueles produtos químicos.

8-2 Inferências sobre Duas Médias: Amostras Dependentes

Apresentaremos nesta seção métodos para testar hipóteses e construir intervalos de confiança para duas amostras dependentes, o que significa que elas estão emparelhadas, ou ligadas. Começamos por definir formalmente *independente* e *dependente*.

DEFINIÇÃO

Duas amostras são **independentes** se a amostra extraída de uma das populações não tem qualquer relação com a amostra extraída da outra população. Se uma das amostras tem alguma relação com a outra, as amostras dizem-se **dependentes**. Tais amostras costumam ser chamadas **amostras emparelhadas**, ou **amostras ligadas** (porque obtemos dois valores para cada indivíduo, ou um valor de cada um de dois indivíduos que apresentam a mesma característica).

Consideremos os dados amostrais emparelhados apresentados a seguir. A amostra dos pesos antes do treinamento e a amostra dos pesos após o treinamento são *amostras dependentes*, porque cada par é formado de acordo com a pessoa envolvida. Os dados do tipo "antes/depois" são, em geral, emparelhados e dependentes.

Indivíduo	A	B	C	D	E	F
Peso antes do treinamento (kg)	99	62	74	59	70	73
Peso após o treinamento (kg)	94	62	66	58	70	76

(Com base em dados do *Journal of Applied Psychology*; Vol. 62, Nº 1.)

Para os dados a seguir, entretanto, as duas amostras são *independentes*, porque a amostra de mulheres não tem qualquer relação com a amostra de homens. Os dados não são emparelhados, como no caso anterior.

Pesos de mulheres (lb)	115	107	110	128	130		
Pesos de homens (lb)	128	150	160	140	163	155	175

Suposições Fazemos as seguintes suposições para os testes de hipóteses e intervalos de confiança dados nesta seção.

1. Devem-se escolher de maneira *aleatória* duas amostras dependentes de duas populações.
2. Ambas as populações devem ter *distribuição normal.* (Se elas se afastam radicalmente da distribuição normal, não devemos aplicar os métodos desta seção; poderemos utilizar os métodos não-paramétricos discutidos no Capítulo 13.)

Ao trabalharmos com duas amostras dependentes, baseamos nossos cálculos na diferença (d) entre os pares de dados, conforme ilustrado na tabela a seguir. (A simples comparação das duas médias $\bar{x}$ e $\bar{y}$ acarretaria a perda de informações importantes sobre os dados emparelhados.) A notação que segue se baseia nessas diferenças.

x	10	8	5	20
y	7	2	9	20
d	3	6	−4	0

Notação para Duas Amostras Dependentes

μ_d = média das diferenças d para a *população* de dados emparelhados

$\bar{d}$ = valor médio das diferenças d para os dados *amostrais* emparelhados (igual à média dos valores de $x - y$)

s_d = desvio-padrão das diferenças d para os dados *amostrais* emparelhados

n = número de *pares* de dados

Por exemplo, com os valores d iguais a 3, 6, −4, 0, extraídos da tabela apresentada, obtemos

$$\bar{d} = \frac{\Sigma d}{n} = \frac{3 + 6 + (-4) + 0}{4} = 1{,}25$$

$$s_d = \sqrt{\frac{n(\Sigma d^2) - (\Sigma d)^2}{n(n-1)}} = \sqrt{\frac{4(61) - (5)^2}{4(4-1)}} = 4{,}27$$

$$n = 4$$

Testes de Hipóteses

Vamos agora utilizar a notação precedente para descrever a estatística de teste a ser usada nos testes de hipóteses para afirmações sobre médias de duas populações, no caso de as amostras serem dependentes. Quando selecionamos aleatoriamente duas amostras dependentes de populações distribuídas normalmente, em que a média populacional das diferenças emparelhadas é μ_d, a estatística seguinte tem distribuição t de Student:

Estatística de Teste para Duas Amostras Dependentes

$$t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}}$$

onde graus de liberdade $= n - 1$

Se o número de pares de dados é grande ($n > 30$), o número de graus de liberdade será no mínimo 30, de forma que os valores críticos serão escores z (Tabela A-2) em lugar de valores t (Tabela A-3).

EXEMPLO Utilizando um cronometrador de reação, análogo ao descrito nas Atividades em Grupo do Capítulo 5, os indivíduos são submetidos a testes de reação com suas mãos esquerdas e suas mãos direitas. (Utilizaram-se somente indivíduos destros.) Os resultados (em milésimos de segundo) constam da tabela a seguir. No nível de 0,05 de significância, teste a afirmação de que há uma diferença entre a média dos tempos de reação da mão direita e da mão esquerda. Se um engenheiro está projetando a cabine de um jato de combate e deve colocar o ativador de ejeção do assento de modo a ser acessível tanto à mão direita como à mão esquerda, faz alguma diferença qual mão ele escolhe?

Pessoa	A	B	C	D	E	F	G	H	I	J	K	L	M	N
Direita	191	97	116	165	116	129	171	155	112	102	188	158	121	133
Esquerda	224	171	191	207	196	165	177	165	140	188	155	219	177	174

SOLUÇÃO Com o método tradicional, testaremos a afirmação de que há diferença entre os tempos de reação da mão direita e da mão esquerda. Como estamos lidando com dados emparelhados, começamos determinando as diferenças d = direita − esquerda. Seguimos então os passos indicados na Figura 7-4.

Passo 1: Se há diferença, é de esperar-se que a média dos valores d seja diferente de 0, o que, simbolicamente, se expressa como $\mu_d \neq 0$.

Passo 2: Se a afirmação original não é verdadeira, temos $\mu_d = 0$.

Passo 3: Como a hipótese nula deve conter a igualdade, temos

$$H_0\colon \mu_d = 0 \qquad H_1\colon \mu_d \neq 0 \text{ (afirmação original)}$$

Passo 4: O nível de significância é $\alpha = 0{,}05$.

Passo 5: Como estamos testando uma afirmação sobre as médias de dados dependentes emparelhados, aplicamos a distribuição t de Student.

Passo 6: Antes de achar o valor da estatística de teste, devemos achar os valores de $\bar{d}$ e s_d. Calculando a diferença d para cada indivíduo, obtemos as seguintes diferenças d = direita − esquerda:

−33, −74, −75, −42, −80, −36, −6, −10, −28, −86, 33, −61, −56, −41

Aplicamos a seguir as Fórmulas 2-1 e 2-4, como se segue:

$$\bar{d} = \frac{\Sigma d}{n} = \frac{-595}{14} = -42{,}5$$

$$s_d = \sqrt{\frac{n(\Sigma d^2) - (\Sigma d)^2}{n(n-1)}}$$

$$= \sqrt{\frac{14(39.593) - (-595)^2}{14(14-1)}} = 33{,}2$$

Com estas estatísticas e a suposição que $\mu_d = 0$, determinamos o valor da estatística de teste:

$$t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \frac{-42{,}5 - 0}{\frac{33{,}2}{\sqrt{14}}} = -4{,}790$$

Os valores críticos $t = -2{,}160$ e $t = 2{,}160$ encontram-se na Tabela A-3: localize a coluna para 0,05 (bilateral) e a linha de graus de liberdade para $n - 1 = 13$. A Figura 8-2 mostra a estatística de teste, os valores críticos e a região crítica.

Passo 7: Como a estatística de teste está na região crítica, rejeitamos a hipótese nula $\mu_d = 0$.

Passo 8: Há evidência suficiente para apoiar a afirmação de que há diferença entre os tempos de reação da mão direita e da mão esquerda. Como parece haver tal diferença, o engenheiro que planeja a cabine do jato de combate deve colocar o ativador do ejetor de assento de modo a ser prontamente acessível à mão mais rápida, que parece ser a direita, com tempos de reação aparentemente menores. (Poder-se-á exigir treinamento especial para pilotos canhotos se um teste análogo mostrar que sua mão principal é mais rápida.)

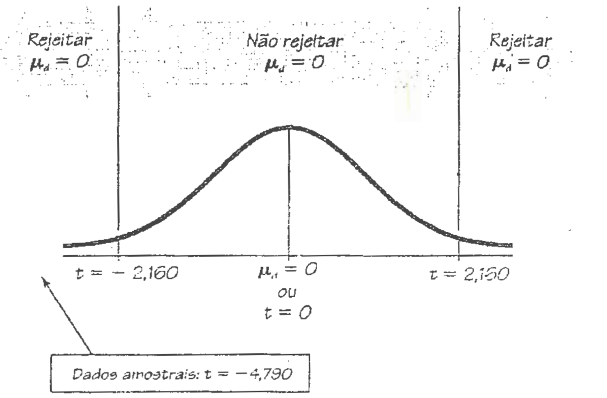


Fig. 8-2 Distribuição de diferenças entre tempos de reação da mão direita e da mão esquerda.

O Dentifrício Crest e Amostras Dependentes

No final da década de 1950, a Procter & Gamble introduziu o dentifrício Crest como o primeiro a conter flúor. Para testar a eficácia de Crest no combate à cárie, pesquisadores fizeram experiências com vários grupos de gêmeos. Um dos gêmeos em cada grupo usou o Crest com flúor, enquanto o outro gêmeo usou um dentifrício comum, sem flúor. Admitiu-se que cada par de gêmeos tivesse os mesmos hábitos alimentares, de higiene bucal, e mesmas características genéticas. Os resultados mostraram que os gêmeos que faziam uso do dentifrício Crest apresentavam um número de cáries significativamente menor do que os que não usavam. A utilização de gêmeos como amostras dependentes permitiu que os pesquisadores controlassem muitas das variáveis responsáveis pela cárie.

O exemplo anterior é um teste bilateral, mas os testes unilaterais direito e esquerdo seguem o mesmo processo básico. Por exemplo, se quisermos testar a afirmação de que os tempos de reação da mão direita são inferiores aos da mão esquerda, temos uma afirmação $\mu_d < 0$; isto é, os valores "direita − esquerda" devem ser negativos. Esta afirmação conduz a um teste unilateral esquerdo com valor crítico $t = -1{,}771$; a estatística de teste é ainda $t = -4{,}790$. Com a estatística de teste e o valor crítico dados, rejeitamos a hipótese nula $\mu_d \geq 0$. Este teste unilateral esquerdo conduz, pois, à conclusão de que há evidência suficiente para apoiar a afirmação de que as mãos direitas têm tempos de reação menores.

No exemplo bilateral precedente utilizamos o método tradicional; mas poderíamos igualmente ter usado o método do valor P, modificando os Passos 6 e 7. No Passo 6 utilizaríamos a estatística de teste $t = -4{,}790$ e, na 13ª linha da Tabela A-3, verificaríamos que 4,790 ultrapassa o maior valor da tabela, 3,012. O valor P no teste bilateral é, portanto, inferior a 2(0,005) ou 0,01; isto é, valor $P < 0{,}01$. No Passo 7, novamente rejeitaríamos a hipótese nula, porque o valor P é inferior ao nível de significância $\alpha = 0{,}05$.

Podemos estabelecer uma estimativa por intervalo de confiança, da diferença média populacional μ_d, utilizando a média amostral $\bar{d}$, o desvio-padrão das médias amostrais d (que é $s_d / \sqrt{n}$), e o valor crítico $t_{\alpha/2}$.

Intervalos de Confiança

A estimativa intervalar de confiança da diferença média μ_d é:

$$\bar{d} - E < \mu_d < \bar{d} + E$$

onde $E = t_{\alpha/2} \dfrac{s_d}{\sqrt{n}}$ e o número de graus de liberdade é $n - 1$.

EXEMPLO Com os dados amostrais do exemplo precedente, construa um intervalo de confiança de 95% para a estimativa de μ_d.

SOLUÇÃO Com os valores $\bar{d} = -42{,}5$, $s_d = 33{,}2$, $n = 14$ e $t_{\alpha/2} = 2{,}160$, determinamos primeiro a margem de erro E:

$$E = t_{\alpha/2} \frac{s_d}{\sqrt{n}} = 2{,}160 \frac{33{,}2}{\sqrt{14}} = 19{,}2$$

Podemos determinar agora o intervalo de confiança:

$$\bar{d} - E < \mu_d < \bar{d} + E$$
$$-42{,}5 - 19{,}2 < \mu_d < -42{,}5 + 19{,}2$$
$$-61{,}7 < \mu_d < -23{,}3$$

Este resultado costuma expressar-se também como $\mu_d = -42{,}5 \pm 19{,}2$ ou como $(-61{,}7; -23{,}3)$. A longo prazo, 95% de tais amostras levarão a limites do intervalo de confiança que contêm efetivamente a verdadeira média populacional das diferenças. Note que os limites do intervalo de confiança não contêm 0, o que indica que o verdadeiro valor de μ_d é significativamente diferente de zero; isto é, o valor médio das diferenças "direita − esquerda" é diferente de zero. Com base no intervalo de confiança, concluímos que há evidência suficiente para apoiar a afirmação de que há uma diferença entre os tempos de reação da mão direita e da mão esquerda. Esta conclusão está de acordo com a do exemplo precedente.

Utilização de um Computador e de uma Calculadora para Duas Amostras Dependentes

Os programas estatísticos como o STATDISK e o Minitab podem ser utilizados para testar hipóteses e construir intervalos de confiança para situações que envolvem duas amostras dependentes. Vêem-se a seguir as telas de resultados dados pelo STATDISK e pelo Minitab para o exemplo desta seção. Se estiver trabalhando com STATDISK, selecione `Analysis`, em seguida `Hypothesis Testing`, e `Mean-Two Dependent Samples`. O resultado será a tela mostrada aqui. (O STATDISK fornece automaticamente limites do intervalo de confiança para testes bilaterais.)

Se estiver utilizando o Minitab, a abordagem básica consiste em criar uma coluna das diferenças d, e fazer então um teste t da afirmação de que as diferenças provêm de uma população com média 0. Introduza os valores correspondentes à "direita" na coluna C1 e os valores correspondentes à "esquerda" na coluna C2. Selecione `Calc`, em seguida `Calculator` e introduza C3 para a coluna em que vai armazenar os resultados; introduza a expressão C1 − C2. (A coluna C3 contém agora as diferenças d.) Em seguida, selecione do menu principal o item `Stat`, depois `Basic Statistics` e `1-Sample t`. Introduza C3 para a variável, aponte no círculo `Test mean` e introduza o valor desejado para a média (0, neste caso). A caixa rotulada `Alternative` tem um *default* de "Not equal" (Não igual); deixe-o como está para este exemplo. Selecione OK e o resultado deve ser conforme mostrado. O Minitab também dará um intervalo de confian-

STATDISK DISPLAY

File Edit Analysis Data Help

Claim	$\mu_d = 0$
Null Hypothesis	$\mu_d = 0$
Sample Size, n	14
Difference Mean, $\bar{x}_d$	-42.50
Difference St Dev, s_d	33.17
Test Statistic, t	-4.7937
Critical t	±2.1604
P-Value	0.0004

95% Confidence Interval:
$-61.65 < \mu_d < -23.35$

Reject the Null Hypothesis

Sample provides evidence to reject the claim

MINITAB DISPLAY

```
TEST OF MU = 0.00 VS MU N.E. 0.00
      N   MEAN   STDEV  SE MEAN      T  P VALUE
C3   14 -42.50   33.17     8.87  -4.79   0.0004
```

ça, bastando selecionar `Stat`, `Basic Statistics`, `1-Sample t`, e a opção `Confidence interval`.

A calculadora TI-83 também pode ser usada para os métodos desta seção. Introduza os dados para a primeira variável em L1, os dados para a segunda variável em L2, limpe a tela e introduza L1 − L2 → L3. Acione então `STAT`, selecione `TESTS` e escolha a opção `T-Test`. Utilizando a opção `Data` de entrada (input), introduza os dados indicados e acione ENTER ao terminar. Se utilizar os dados emparelhados dos exemplos precedentes, a tela apresentará a estatística de teste $t = -4{,}793724854$, o valor P 0,0003507665, o valor de $\bar{d}$ (indicado como $\bar{x}$) e o valor de s_d (indicado como Sx). A TI-83 utiliza o método do valor P para testar hipóteses. Pode-se obter também um intervalo de confiança acionando `STAT`, selecionando `TESTS` e `TInterval`.

Utilizamos a Tabela A-3 para concluir que o valor P é inferior a 0,01, mas pelos resultados apresentados pelo STATDISK, Minitab e TI-83, vemos que o valor P é efetivamente 0,0004.

8-2 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1 e 2, suponha que você deseja testar a afirmação de que os dados amostrais emparelhados provêm de uma população para a qual a diferença média é $\mu_d = 0$. Adotando o nível de 0,05 de significância, ache (a) $\bar{d}$, (b) s_d, (c) a estatística de teste t e (d) os valores críticos.

1.

x	8	8	6	9	7
y	3	2	6	4	9

2.

x	20	25	27	27	23	29	30	26
y	20	24	25	29	20	29	32	29

3. Com os dados amostrais emparelhados do Exercício 1, construa um intervalo de confiança de 95% para a média populacional de todas as diferenças $x - y$.
4. Com os dados amostrais emparelhados do Exercício 2, construa um intervalo de confiança de 99% para a média populacional de todas as diferenças $x - y$.
5. Vale a pena fazer cursos preparatórios para testes padronizados como o SAT? Com os dados amostrais da tabela seguinte, teste a afirmação de que o Curso Preparatório Allan é eficiente para o SAT. Use o nível de 0,05 de significância.

Pessoa	A	B	C	D	E	F	G	H	I	J
Nota SAT antes do curso (x)	700	840	830	860	840	690	830	1180	930	1070
Nota SAT após o curso (y)	720	840	820	900	870	700	800	1200	950	1080

Com base em dados do College Board e em "An Analysis of the Impact of Commercial Test Preparation Courses on SAT Scores", de Sesnowitz, Bernhardt e Knaim, *American Educational Research Journal*, Vol. 19, Nº 3.

6. Com os dados amostrais do Exercício 5, construa um intervalo de 95% de confiança para a diferença média dos resultados "antes" menos "depois".
7. Captopril é um remédio destinado a baixar a pressão sistólica. Feito o teste com este remédio em pacientes, mediram-se suas pressões sistólicas (em mm de mercúrio) antes e depois de tomarem o remédio, com os resultados constantes da tabela a seguir. Com os dados amostrais, construa um intervalo de confiança de 99% para a diferença média entre os valores "antes" e "depois".

Pessoa	A	B	C	D	E	F	G	H	I	J	K	L
Antes	200	174	198	170	179	182	193	209	185	155	169	210
Depois	191	170	177	167	159	151	176	183	159	145	146	177

Com base em dados de "Essential Hypertension: Effect of an Oral Inhibitor of Angiotensin-Converting Enzyme", de MacGregor *et al.*, *British Medical Journal*, Vol. 2.

8. Considere os dados amostrais do Exercício 7. No nível de 0,05 de significância, teste a afirmação de que Captopril é eficaz para baixar a pressão sistólica.
9. Realizou-se um estudo para investigar a eficácia do hipnotismo na redução da dor. A tabela a seguir dá os resultados para pacientes selecionados aleatoriamente. No nível de 0,05 de significância, teste a afirmação de que as medidas sensoriais são inferiores após o hipnotismo. (Os valores se referem a "antes" e "depois" da hipnose; medidas em centímetros em uma escala de dor.) O hipnotismo parece eficaz na redução da dor?

Pessoa	A	B	C	D	E	F	G	H
Antes	6,6	6,5	9,0	10,3	11,3	8,1	6,3	11,6
Depois	6,8	2,4	7,4	8,5	8,1	6,1	3,4	2,0

Com base em "An Analysis of Factors That Contribute to the Efficacy of Hypnotic Analgesia", de Price e Barber, *Journal of Abnormal Psychology*, Vol. 96, Nº 1.

10. Com os dados do Exercício 9, construa um intervalo de confiança de 95% para a média das diferenças "antes − depois".
11. Costuma-se avaliar a inteligência das crianças dando-lhes blocos e pedindo-lhes que construam uma torre tão alta quanto possível. Repetiu-se um mês depois um experimento de construção com blocos, com os tempos (em segundos) dados na tabela a seguir. No nível de 0,01 de significância, teste a afirmação de que não há diferença entre os dois tempos.

Criança	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Primeira tentativa	30	19	19	23	29	178	42	20	12	39	14	81	17	31	52
Segunda tentativa	30	6	14	8	14	52	14	22	17	8	11	30	14	17	15

Com base em dados de "Tower Building", de Johnson e Courtney, *Child Development*, Vol. 3.

12. Com os dados amostrais do Exercício 11, construa um intervalo de confiança de 95% para a média das diferenças. Os limites do intervalo contêm 0, indicando que não há diferença significativa entre os tempos da primeira e da segunda prova?
13. Realizou-se um estudo para investigar alguns efeitos do treinamento físico. Os dados amostrais estão relacionados a seguir. (Veja "Effect of Endurance Training on Possible Determinants of VO_2, During Heavy Exercise," de Casaburi *et al.*, *Journal of Applied Physiology*, Vol. 62, Nº 1). no nível de 0,05 de significância, teste a afirmação de que o peso médio antes do treinamento é igual ao peso médio após o treinamento. Todos os

pesos são dados em quilogramas. Que se pode concluir quanto ao efeito do treinamento sobre o peso?

Antes do treinamento:	99	57	62	69	74	77	59	92	70	85
Após o treinamento:	94	57	62	69	66	76	58	88	70	84

14. Com os dados do Exercício 13, construa um intervalo de confiança de 95% para a média das diferenças entre os pesos antes e após o treinamento.
15. Com o Conjunto de Dados 2 do Apêndice B, use os dados emparelhados referentes às temperaturas de mulheres às 8 h e às 12 h do dia 2. Construa um intervalo de confiança de 95% para a diferença média entre as temperaturas às 8 horas e as temperaturas às 12 horas.
16. Com os dados amostrais do Exercício 15 e um nível de 0,05 de significância, teste a afirmativa de que, para aquelas temperaturas, a média das diferenças é 0. Com base no resultado, as temperaturas pela manhã e à noite parecem ser as mesmas?
17. A tela Minitab seguinte resultou de um experimento em que 10 indivíduos foram testados quanto ao enjôo por movimento, antes e depois de tomarem o remédio astemizole. A coluna Minitab C3 consiste em diferenças no número de movimentos da cabeça que os indivíduos pesquisados podem suportar sem enjoarem. (Obtiveram-se as diferenças subtraindo-se os valores "após" dos valores "antes".) No nível de significância de 0,05, teste a afirmação de que o astemizole produz efeito (para melhor ou para pior) sobre a vulnerabilidade ao enjôo por movimento. Com base no resultado, o leitor tomaria astemizole se temesse o enjôo durante um cruzeiro marítimo?

```
TEST OF MU = 0.0 VS MU N.E. 0.0
     N   MEAN   STDEV   SE MEAN      T   P VALUE
C3  10   -7.5    57.7      18.2  -0.41      0.69
```

18. Consulte a tela Minitab do Exercício 17. Que valor P resultaria de um teste da afirmação de que astemizole é eficaz para evitar o enjôo por movimento? Qual é a sua conclusão?
19. Use os resultados Minitab do Exercício 17 para construir um intervalo de confiança de 95% para a diferença média entre os números de movimentos de cabeça que o indivíduo pode suportar sem enjoar. Os limites do intervalo contêm 0, o que indicaria que o astemizole não tem efeito significativo sobre a vulnerabilidade ao enjôo por movimento?

8-2 Exercícios B: Além do Básico

20. Veja o exercício indicado e ache o valor P ou (pela Tabela A-3) os limites que contêm o valor P.

    a. Exercício 5   b. Exercício 9

21. O exemplo nesta seção utilizou tempos de reação dados em milésimos de segundo. Suponha que expressemos todos os tempos em segundos, como 0,191, 0,224 etc. O teste de hipóteses é afetado por essa mudança de unidade? E o intervalo de confiança? Como?
22. O intervalo de confiança de 95% para uma coleção de dados emparelhados é $0{,}0 < \mu_d < 1{,}2$. Com base neste intervalo de confiança, o método tradicional de teste de hipóteses conduz à conclusão de que a afirmação $\mu_d > 0$ se sustenta. Qual é o menor nível de significância possível para este teste de hipóteses?

8-3 Inferências sobre Duas Médias: Amostras Grandes e Independentes

Na Seção 8-2 apresentamos métodos de utilização da inferência estatística para dados dependentes (emparelhados); nesta seção, nossas duas amostras serão *independentes*. Naquela seção, definimos amostras como independentes se a amostra extraída de uma população não tem qualquer relação com a amostra extraída da outra população.

Esta seção se restringe a amostras não só independentes, como também grandes ($n > 30$); isto é, o tamanho n_1 da amostra 1 e o tamanho n_2 da amostra 2 serão ambos maiores do que 30. Na Seção 8-5 estudaremos o caso de pequenas amostras independentes. Para os objetivos desta seção, valem as seguintes suposições.

Suposições Ao testar hipóteses sobre duas médias populacionais e construir intervalos de confiança para a diferença entre duas médias, fazemos as seguintes suposições para os métodos utilizados neste capítulo.

1. As duas amostras são *independentes*.
2. Os tamanhos das duas amostras são grandes: $n_1 > 30$, $n_2 > 30$.

Diagramas em Caixas

Antes de passar ao teste formal de uma hipótese ou à construção de um intervalo de confiança, é conveniente construir um diagrama em caixas que permita a comparação das duas amostras. Por exemplo, a Figura 8-3 apresenta diagramas em caixa para as cargas axiais de uma amostra de latas de 0,0109 in. de espessura, e outra amostra de latas de 0,0111 in. de espessura. (A carga axial de uma lata é o peso máximo que seus lados podem suportar, e é medida utilizando-se uma placa para aplicar uma pressão crescente no topo da lata até que esta se deforme. Os dados originais constam do Conjunto de Dados 15 do Apêndice B.)

A observação dos valores dos dois conjuntos originais de dados revela que a carga axial de 504 lb (para as latas com 0,0111 de espessura) é um valor extremo (*outlier*), porque está muito distante de todos os outros valores. Se excluirmos esse *outlier*, a longa cauda direita na base do diagrama sofrerá uma redução para o valor identificado como 317. A comparação dos dois diagramas mostra que a dispersão dos dados é praticamente a mesma, mas os valores das latas de 0,0111 estão mais para a direita, o que sugere que aquelas latas comportam cargas axiais maiores,

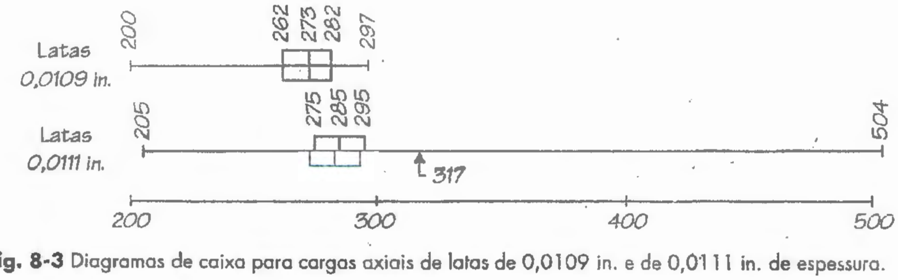


Fig. 8-3 Diagramas de caixa para cargas axiais de latas de 0,0109 in. e de 0,0111 in. de espessura.

sendo, por isso, mais resistentes. Podemos agora aplicar um teste de hipóteses para verificar se essa diferença aparente é significativa.

Teste de Hipóteses

Uma conclusão do teorema central do limite é que as médias amostrais tendem a distribuir-se normalmente. Mais adiante nesta seção justificaremos esta importante propriedade: as *diferenças* entre médias amostrais ($\bar{x}_1 - \bar{x}_2$) também tendem a distribuir-se normalmente. Com estas propriedades e as suposições feitas anteriormente, obtemos a seguinte estatística a ser utilizada em testes de hipóteses formuladas sobre as médias de duas populações. Esta estatística de teste, aliás, é análoga a várias outras estatísticas já encontradas. Observe que ela tem o mesmo formato básico

$$\frac{(\text{estatística amostral}) - (\text{parâmetro populacional afirmado})}{(\text{desvio-padrão da estatística amostral})}$$

Estatística de Teste para Duas Médias: Amostras Independentes e Grandes

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Tal como no Capítulo 7, se não conhecemos os valores de σ_1 e σ_2, podemos substituí-los por s_1 e s_2, desde que ambas as amostras sejam grandes. Se σ_1 e σ_2 são conhecidos utilizamos seus valores para o cálculo da estatística de teste, mas os casos reais em geral exigem o uso de s_1 e s_2. É raro conhecermos os valores de desvios-padrão populacionais se não conhecemos as médias populacionais.

EXEMPLO No nível de 0,01 de significância, teste a afirmação de que as latas de 0,0109 de espessura têm carga axial média inferior à das latas de 0,0111 de espessura. Os dados originais constam do Conjunto de Dados 15 do Apêndice B, e as estatísticas resumo encontram-se a seguir.

Cargas Axiais (lb)
de Latas de 0,0109 in.

$n_1 = 175$
$\bar{x}_1 = 267{,}1$
$s_1 = 22{,}1$

Cargas Axiais (lb)
de Latas de 0,0111 in.

$n_2 = 175$
$\bar{x}_2 = 281{,}8$
$s_2 = 27{,}8$

SOLUÇÃO

Passo 1: Podemos expressar a afirmação simbolicamente como $\mu_1 < \mu_2$.

Passo 2: Se a afirmação original é falsa, então $\mu_1 \geq \mu_2$.

Passo 3: A hipótese nula deve conter a igualdade, assim,

H_0: $\mu_1 \geq \mu_2$ $\quad$ H_1: $\mu_1 < \mu_2$ (afirmação original)

Passamos agora à hipótese $\mu_1 = \mu_2$ ou $\mu_1 - \mu_2 = 0$.

Passo 4: O nível de significância é $\alpha = 0{,}01$.

Passo 5: Como temos duas amostras independentes e grandes, e estamos testando uma afirmação sobre as duas médias populacionais, utilizamos a distribuição normal com a estatística de teste dada previamente nesta seção.

Passo 6: Os valores de σ_1 e σ_2 são desconhecidos, mas as amostras são ambas grandes; podemos, assim, utilizar os desvios-padrão amostrais como estimativas dos desvios-padrão populacionais, calculando-se a estatística de teste como se segue.

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} =$$

$$= \frac{(267{,}1 - 281{,}8) - 0}{\sqrt{\frac{22{,}1^2}{175} + \frac{27{,}8^2}{175}}} = -5{,}48$$

Como estamos utilizando uma distribuição normal, encontramos o valor crítico $z = -2{,}33$ na Tabela A-2. (Com $\alpha = 0{,}01$ na cauda esquerda, determinamos o valor z correspondente a uma área de $0{,}5 - 0{,}01$, ou 0,4900.) A Figura 8-4 exibe a estatística de teste, o valor crítico e a região crítica.

Passo 7: Como a estatística de teste está na região crítica, rejeitamos a hipótese nula $\mu_1 \geq \mu_2$.

Passo 8: Há evidência suficiente para apoiar a afirmação de que as latas de 0,0109 têm carga axial média inferior à das latas de 0,0111. Já era de se esperar que as latas menos espessas fossem mais fracas, mas agora sabemos que a diferença é estatisticamente significativa. A significação *prática* dessa diferença é um outro problema, não abordado por este teste de hipóteses.

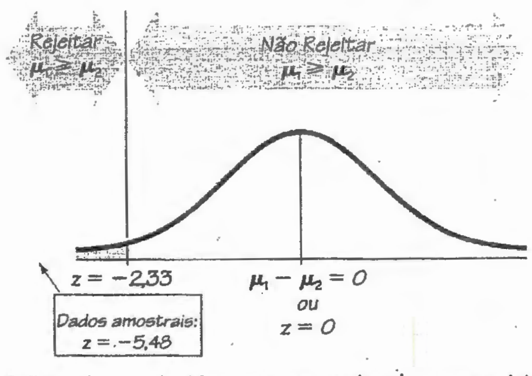


Fig. 8-4 Distribuição de diferenças entre médias de cargas axiais para latas de 0,0109 in. e 0,0111 in. de espessura.

Teste do Uso de Drogas: Falso-Positivo

Segundo a American Management Association, mais da metade das firmas fazem teste do uso de drogas. Para um candidato a emprego que é submetido ao teste, um *falso-positivo* é uma indicação de que a pessoa usa drogas, quando, de fato, não o faz. Um *falso-negativo* é uma indicação de que a pessoa não usa drogas, quando, de fato, as consome. A ***sensitividade do teste*** é a probabilidade de um resultado positivo para um usuário de drogas; a ***especificidade*** é a probabilidade de um resultado negativo para um não-usuário. Suponha que 3% dos candidatos a um emprego sejam viciados em drogas, e que um teste tenha uma sensitividade de 0,99 e uma especificidade de 0,98. Estas elevadas probabilidades fazem com que o teste se afigure confiável, mas 40% dos resultados positivos são falso-positivos.

Valores *P*

Aqui a determinação dos valores *P* é fácil porque as estatísticas de teste são escores *z* da distribuição normal padronizada. Simplesmente siga o método delineado nas Figuras 7-7 e 7-8. Para o teste de hipóteses esquerdo precedente, o valor *P* é a área à esquerda da estatística de teste $z = -5{,}48$. Consultando a Tabela A-2, vemos que a área à esquerda de $z = -5{,}48$ é 0,0001. (Como o valor *z* excede 3,09, usamos uma área de 0,4999, conforme se vê no rodapé da Tabela A-2.) Como o valor *P* de 0,0001 é inferior no nível de significância $\alpha = 0{,}01$, novamente rejeitamos a hipótese nula e concluímos que as latas de 0,0109 in. têm carga axial média inferior à das latas de 0,0111 in.

Intervalos de Confiança

O intervalo de confiança para a estimativa da diferença $\mu_1 - \mu_2$ é:

$$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$$

onde
$$E = z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

EXEMPLO Com os dados do exemplo precedente, construa um intervalo de confiança de 99% para a diferença entre as médias das cargas axiais das latas de 0,0109 in. e das latas de 0,0111 in.

SOLUÇÃO Inicialmente, determinamos a margem de erro *E*:

$$E = z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = 2{,}575\sqrt{\frac{22{,}1^2}{175} + \frac{27{,}8^2}{175}} = 6{,}9$$

Em seguida, estabelecemos como segue o intervalo de confiança desejado:

$$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$$
$$(267{,}1 - 281{,}8) - 6{,}9 < (\mu_1 - \mu_2)$$
$$< (267{,}1 - 281{,}8) + 6{,}9$$
$$-21{,}6 < (\mu_1 - \mu_2) < -7{,}8$$

Temos, pois, 99% de confiança de que os limites $-21{,}6$ e $-7{,}8$ contenham efetivamente a diferença entre as duas médias populacionais. Poderíamos apresentar este resultado de modo mais claro dizendo que μ_2 excede μ_1 por um valor entre 7,8 lb e 21,6 lb. Como estes limites não contêm 0, é muito pouco provável que as duas médias populacionais sejam iguais.

O Efeito *Placebo*

É crença comum que, no caso de pacientes que recebem um *placebo* (um tratamento sem qualquer valor medicinal), um terço deles acusa alguma melhora. Entretanto, um estudo mais recente de 6000 pacientes mostrou que, para as pessoas com problemas médicos brandos, os *placebos* aparentemente causaram melhora em cerca de dois terços dos casos. O efeito *placebo* parece ser mais forte quando o paciente está muito ansioso e gosta de seu médico. Como poderia ofuscar os estudos de novos tratamentos, o efeito *placebo* é minimizado utilizando-se o experimento do tipo *duplo cego*, em que nem o paciente, nem o médico sabem se o tratamento é um *placebo* ou um remédio real.

Por que as estatísticas de teste e os intervalos de confiança têm as formas particulares apresentadas? Ambas as formas se baseiam em uma distribuição normal dos valores de $\bar{x}_1 - \bar{x}_2$ com média $\mu_1 - \mu_2$ e desvio-padrão $\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}$. Isto é decorrência do teorema do limite central, introduzido na Seção 5-5, que afirma que as médias amostrais $\bar{x}$ se distribuem normalmente com média μ e desvio-padrão $\sigma/\sqrt{n}$. E quando o tamanho das amostras é 31 ou mais, a distribuição normal constitui aproximação razoável da distribuição das médias amostrais. Por um raciocínio análogo, os valores de $\bar{x}_1 - \bar{x}_2$ também tendem para uma distribuição normal com média $\mu_1 - \mu_2$. Quando ambas as amostras são grandes, a propriedade seguinte das variâncias levanos a concluir que os valores de $\bar{x}_1 - \bar{x}_2$ têm um desvio-padrão dado por

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

A variância das *diferenças* entre duas variáveis aleatórias independentes é igual à soma das variâncias dessas variáveis.

Isto é, a variância dos valores amostrais $\bar{x}_1 - \bar{x}_2$ tende a ser igual a

$$\sigma_{\bar{x}_1}^2 + \sigma_{\bar{x}_2}^2$$

desde que $\bar{x}_1$ e $\bar{x}_2$ sejam independentes. (Veja Exercício 21.) Para grandes amostras aleatórias, o desvio-padrão das médias amostrais é $\sigma/\sqrt{n}$, de forma que a variância das médias amostrais é σ^2/n. Combinando a propriedade aditiva das variâncias com a expressão do teorema central do limite para a variância de médias amostrais, temos:

$$\sigma_{\bar{x}_1 - \bar{x}_2}^2 = \sigma_{\bar{x}_1}^2 + \sigma_{\bar{x}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

Tomando a raiz quadrada, vemos que o desvio-padrão da diferença $\bar{x}_1 - \bar{x}_2$ pode expressar-se como

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Como *z* é um escore padronizado que corresponde em geral a

$$z = \frac{(\text{estatística amostral}) - (\text{média populacional})}{(\text{desvio-padrão da estatística amostral})}$$

vem

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

notando que os valores amostrais de $\bar{x}_1 - \bar{x}_2$ terão média $\mu_1 - \mu_2$ e o desvio-padrão obtido anteriormente.

Utilização de Computadores e Calculadoras para Testes de Hipóteses e Intervalos de Confiança

A construção de testes de hipóteses e intervalos de confiança desta seção pode efetuar-se com a utilização de programas estatísticos. Para usar o STATDISK, selecione `Analysis`, em seguida `Hypothesis Testing`, logo após `Mean-Two Independent Samples`. Os limites do intervalo de confiança estão incluídos nos resultados do teste de hipóteses. O resultado é mostrado a seguir. Vê-se que os resultados do STATDISK coincidem de forma bem satisfatória com os obtidos nesta seção.

O Minitab é planejado para dados brutos e não lida diretamente com estatísticas resumo, de modo que os dados originais devem ser introduzidos nas colunas C1 e C2. Após introduzir o primeiro conjunto de dados na Coluna C1 e o segundo conjunto de dados na coluna C2, selecione `Stat/Basic Statistics/2-Sample t`. Selecione a opção `Samples in different columns`, e introduza C1 para a primeira amostra e C2 para a segunda amostra. Na caixa rotulada como `Alternative`, selecione a opção `less than`. Introduza um nível de confiança de 99 (para um nível de significância de 0,01). A tela do Minitab deverá ser a que se vê em seguida.

Captura-Recaptura

Aos ecologistas interesso avaliar o tamanho de populações de espécies ameaçadas. Um método consiste em capturar uma amostra de algumas espécies, marcar todos os elementos dessa amostra e libertá-los em seguida. Mais tarde, captura-se outra amostra; a percentagem de elementos marcados, associada ao tamanho da primeira amostra, pode servir para estimar o tamanho da população. Este método de captura-recaptura foi usado, juntamente com outros métodos, para estimar a população de baleias azuis, e o resultado foi alarmante. A população estava reduzida a 1000. Isto levou a International Whaling Commission a proibir a matança de baleias azuis, a fim de impedir sua extinção.

Pode-se utilizar também a calculadora TI-83 para os métodos desta seção. Comece introduzindo os dados nas listas L1 e L2, acione STAT e selecione `TESTS`. Escolha a opção `2-SampZTest` e passe a introduzir os dados solicitados. (Se as estatísticas resumo são conhecidas, pode-se usar a opção `Stats`. Se os tamanhos de ambas as amostras são superiores a 30, introduza os desvios-padrão amostrais quando preparados para σ_1 e σ_2.) Com os dados do exemplo precedente obteremos a estatística de teste $z = -5.475651121$ e o valor P 2.1847990E⁻8, que é, na realidade, 0,0000000218. Para obter um intervalo de confiança, acione STAT, selecione `TESTS` e escolha a opção `2-SampZInt`.

STATDISK DISPLAY

File Edit Analysis Data Help

Claim	$\mu_1 < \mu_2$
Null Hypothesis	$\mu_1 >$ or $= \mu_2$
Large Samples	
Test Statistic, z	-5.4757
Critical z	-2.3263
P-Value	0.0000

99% Confidence Interval:
$\mu_1 - \mu_2 > -20.95$

Reject the Null Hypothesis

Sample provides evidence to support the claim

MINITAB DISPLAY

```
Two sample T for C1 vs C2
        N   Mean   StDev   SE Mean
C1    175  267.1    22.1       1.7
C2    175  281.8    27.8       2.1

99% CI for mu C1 - mu C2: (-21.6, -7.7)
T-Test mu C1 = mu C2 (vs <): T = -5.47 P = 0.0000
DF = 331
```

8-3 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1 e 2, utilize o nível de significância de 0,05 para testar a alegação de que as duas amostras provêm de populações com a mesma média. Em ambos os casos, as amostras são independentes e foram selecionadas aleatoriamente.

1.

Grupo de Controle	Grupo Experimental
$n_1 = 50$	$n_2 = 100$
$\bar{x}_1 = 75$	$\bar{x}_2 = 73$
$s_1 = 15$	$s_2 = 14$

2.

Elementos Tratados	Elementos Não-Tratados
$n_1 = 60$	$n_2 = 75$
$\bar{x}_1 = 8{,}75$	$\bar{x}_2 = 9{,}66$
$s_1 = 2{,}05$	$s_2 = 2{,}88$

3. Com os dados amostrais do Exercício 1, construa uma estimativa intervalar de 95% de confiança para a diferença entre as duas médias populacionais. Os limites do intervalo contêm 0? Que se pode concluir quanto à diferença entre o grupo de controle e o grupo experimental?
4. Com os dados amostrais do Exercício 2, construa uma estimativa intervalar de 95% de confiança para a diferença entre duas médias populacionais. O intervalo de confiança contém 0? Que se pode concluir quanto à diferença entre os elementos tratados e os não-tratados.
5. O estresse afeta a capacidade de memorização de testemunhas oculares? Este problema foi estudado em um experimento que testou a memória visual de uma testemunha uma semana após o interrogatório normal de um suspeito que cooperava, e um interrogatório exaustivo de um suspeito que não cooperava. Os números de detalhes lembrados uma semana após o incidente estão resumidos aqui (com base em "Eyewitness Memory of Police Trainees for Realistic Role Plays," por Yuille *et al.*, *Journal of Applied Psychology*, Vol.79, Nº 6). No nível de 0,01 de significância, teste a afirmação do artigo de que "o cansaço concorre para diminuir a quantidade de detalhes lembrados."

Sem Estresse	Com Estresse
$n_1 = 40$	$n_2 = 40$
$\bar{x}_1 = 53{,}3$	$\bar{x}_2 = 45{,}3$
$s_1 = 11{,}6$	$s_2 = 13{,}2$

6. Nielsen Media Research relatou que as mulheres adultas assistem à TV durante uma média de 5 h e 1 min por dia, comparado com uma média de 4 h e 17 min para homens adultos. Suponha que esses resultados tenham sido obtidos de uma amostra de 100 homens e 100 mulheres, e que os dois grupos tenham o mesmo desvio-padrão de 57 min. Construa um intervalo de 99% de confiança para $\mu_1 - \mu_2$, onde μ_1 é a média das mulheres adultas. Os limites do intervalo de confiança englobam 0? Isto sugere que há ou não diferença significativa entre as duas médias?
7. A Medassist Pharmaceutical Company deseja testar o Dozenol, um remédio novo contra resfriado para uso noturno. Os testes de tais produtos costumam incluir um "grupo de tratamento", que toma o remédio, e um "grupo de controle", que não o usa. Administra-se Dozenol a cinqüenta pessoas com gripe, e a 100 outras não-gripadas. Mede-se a pressão sistólica de cada pessoa obtendo-se as estatísticas amostrais mostradas. O chefe de pesquisa da Medassist afirma que o Dozenol não afeta a pressão sanguínea, ou seja, a média populacional do grupo tratado μ_1 e a média populacional do grupo de controle μ_2 são iguais. Teste a afirmação, no nível de significância de 0,01. Com base no resultado, recomendaria anunciar que o Dozenol não afeta a pressão sanguínea?

Grupo de Tratamento	Grupo de Controle
$n_1 = 50$	$n_2 = 100$
$\bar{x}_1 = 203{,}4$	$\bar{x}_2 = 189{,}4$
$s_1 = 39{,}4$	$s_2 = 39{,}0$

8. Considere os dados amostrais utilizados no Exercício 7 e construa um intervalo de 99% de confiança para $\mu_1 - \mu_2$, onde μ_1 e μ_2 representam a média do grupo de tratamento e do grupo de controle, respectivamente.
9. Alunos da faculdade do autor selecionaram aleatoriamente 217 carros de estudantes e constataram que a média de suas idades era de 7,89 anos, com desvio-padrão de 3,67 anos. Selecionaram também, aleatoriamente, 152 carros do corpo docente e do pessoal da administração, constatando uma média de 5,99 anos e um desvio-padrão de 3,65 anos. No nível de significância de 0,05, teste a afirmação de que os carros dos estudantes são mais velhos do que os dos professores e demais funcionários.
10. Recorra aos dados do Exercício 9 e construa um intervalo de confiança de 95% para a média $\mu_1 - \mu_2$, onde μ_2 é a idade média dos carros dos estudantes.
11. Como parte da National Health Survey, coletaram-se dados sobre os pesos de homens. Para 804 homens na faixa etária 25-34, a média foi 176 lb e o desvio-padrão 35,0 lb. Para 1657 homens na faixa 65-74, a média e o desvio-padrão foram 164 lb e 27,0 lb, respectivamente. No nível de significância de 0,01, teste a suposição de que os homens mais velhos provêm de uma população cuja média é inferior à média dos homens da faixa 25-34.
12. Com os dados do Exercício 11, construa um intervalo de confiança de 99% para a diferença entre as médias dos homens nas duas faixas etárias. Os limites do intervalo contêm 0? Isto indica que há, ou não, diferença significativa entre as duas médias?
13. Em um estudo sobre salários de comissários de bordo, selecionaram-se aleatoriamente salários pagos por duas companhias de aviação diferentes. Para 40 comissários de bordo da American Airlines, a média foi de \$23.870 e o desvio-padrão \$2960. Para 35 comissários da TWA, a média foi \$22.025 e o desvio-padrão \$3065 (com base em dados da Association of Flight Attendants). No nível de 0,10 de significância, teste a afirmação de que a American e a TWA pagam o mesmo salário médio. Com base no resultado, o salário deve ser um fator importante para um candidato a comissário de bordo escolher entre a American e a TWA?
14. Com os dados amostrais do Exercício 13, construa um intervalo de confiança de 90% para a diferença entre as médias populacionais ($\mu_1 - \mu_2$), onde μ_1 é o salário médio para todos os recepcionistas da American Airlines e μ_2 é o salário médio para os recepcionistas da TWA. Os limites do intervalo de confiança contêm 0? Isto sugere que não há diferença significativa entre as duas médias populacionais?
15. O Conjunto de Dados 11 do Apêndice B contém pesos de 100 confeitos M&M selecionados aleatoriamente. Esses pesos têm média de 0,9147 g e desvio-padrão de 0,0369 g. A edição anterior deste livro utilizou uma amostra diferente de 100 confeitos M&M, com média e desvio-padrão de 0,9160 g e 0,0433 g, respectivamente. A discrepância entre as duas médias amostrais é significativa?
16. Um estudo sobre o uso do cinto de segurança envolveu crianças hospitalizadas em conseqüência de acidentes com veículos motorizados. Para um grupo de 290 crianças que não usavam o cinto, o número de dias passados em uma UTI acusou média 1,39 e desvio-padrão 3,06. Para um grupo de 123 crianças que usavam o cinto de segurança, o número de dias na UTI acusou média de 0,83 e desvio-padrão de 1,77 (com base em dados de "Morbidity Among Pediatric Motor Vehicle Crash Victims: The Effectiveness of Seat Belts", de Osberg e Di Scala, *American Journal of Public Health*, Vol. 82, Nº 3). No nível de 0,01 de significância, teste a afirmação de que as crianças que não usam cinto de segurança acusam maior tempo de internação em UTI. Com base no resultado, há evidência significativa que favoreça o uso do cinto de segurança pelas crianças?

17. Considere o Conjunto de Dados 2 do Apêndice B e utilize somente as temperaturas às 12 h do Dia 1. Pretende-se testar, no nível de 0,05 de significância, a afirmação de que as pessoas da faixa etária 18-24 acusam a mesma média que as com 25 anos ou mais.
   a. Faça o teste utilizando o método tradicional.
   b. Faça o teste utilizando o método do valor *P*.
   c. Faça o teste construindo um intervalo de confiança de 95% para a diferença $\mu_1 - \mu_2$.
   d. Compare os resultados das partes (a), (b) e (c).
18. Considere as taxas de pulsação do Conjunto de Dados 8 do Apêndice B. Inicialmente, identifique e exclua os dois *outliers* que constituem taxas de pulso não-realísticas de estudantes vivos e razoavelmente sadios.
   a. No nível de significância de 0,05, teste, pelo método tradicional, a afirmação de que os estudantes de estatística de ambos os sexos têm a mesma taxa média de pulsação.
   b. Utilize o método do valor *P* para testar a alegação da parte (a).
   c. Construa um intervalo de confiança de 95% para a diferença $\mu_1 - \mu_2$, onde μ_1 e μ_2 são as taxas médias de pulsação dos estudantes homens e mulheres, respectivamente.
   d. Compare os resultados das partes (a), (b) e (c).

8-3 Exercícios B: Além do Básico

19. Os exemplos nesta seção utilizam as cargas axiais de uma amostra de latas de 0,0109 in. de espessura, e uma segunda amostra de latas de 0,0111 in. (Os dados originais constam do Conjunto de Dados 15 do Apêndice B.) Vê-se que a segunda amostra inclui um ponto extremo de 504 lb. Qual será o efeito sobre os diagramas da Figura 8-3, o teste de hipóteses e o intervalo de confiança se decidirmos que o ponto extremo deve ser excluído, por se tratar, efetivamente, de um erro?
20. Os exemplos desta seção utilizam as cargas axiais de latas dadas em libras. Qual o efeito sobre o diagrama em caixas, o teste de hipóteses e o intervalo de confiança, se convertermos todos os pesos para quilogramas?
21. a. Determine a variância desta *população* de valores x: 5, 10, 15. (Veja Seção 2-5 para a variância σ^2 de uma população.)
   b. Ache a variância desta *população* de valores y: 1, 2, 3.
   c. Relacione a *população* de todas as diferenças possíveis $x - y$ e ache sua variância.
   d. Com os resultados das partes (a), (b) e (c), verifique que $\sigma^2_{x-y} = \sigma^2_x + \sigma^2_y$. (Este princípio é utilizado para estabelecer a estatística de teste e o intervalo de confiança desta seção.)
   e. Qual é a relação da *amplitude* dos valores $x - y$ para a amplitude dos valores x e a amplitude dos valores y?
22. Veja o Exercício 11. Se a diferença entre as médias populacionais é, efetivamente, 6,0, determine β, a probabilidade de um erro tipo II. (Veja Exercício 27 da Seção 7-3.) (*Sug.*: No Passo 1, substitua $\bar{x}$ por $(\bar{x}_1 - \bar{x}_2)$, $\mu\bar{x}$ por 0 e $\sigma\bar{x}$ por

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

8-4 Comparação de Duas Variâncias

Como a variação entre dados é uma característica de extrema importância, apresentamos nesta seção um método que utiliza duas amostras para comparar as variâncias das populações das quais foram extraídas. O método exige as seguintes suposições:

Suposições Ao testarmos uma hipótese sobre as *variâncias* de duas populações, admitimos que

1. As duas populações são *independentes* uma da outra. (Recorde, da Seção 8-2, que duas amostras são independentes se a amostra extraída de uma população não tem qualquer relação com a amostra extraída da outra população.)
2. As duas populações são ambas *distribuídas normalmente*. (Esta suposição é importante, porque a estatística de teste é extremamente sensível a desvios da normalidade.)

Os cálculos desta seção se simplificarão grandemente se convencionarmos que s_1^2 representa *a maior* das duas variâncias amostrais. Na realidade, não faz diferença qual amostra é designada como amostra 1; assim, usaremos a notação seguinte:

Notação para Testes de Hipóteses com Duas Variâncias

s_1^2 = *a maior* das duas variâncias amostrais
n_1 = tamanho da amostra com a *maior* variância
σ_1^2 = variância da população da qual se extrai a amostra com *maior* variância.

Para a outra amostra e a outra população, empregam-se os símbolos s_2^2, n_2 e σ_2^2.

Para duas populações normalmente distribuídas com variâncias iguais (isto é, $\sigma_1^2 = \sigma_2^2$), a distribuição amostral da estatística de teste, a seguir, é a **distribuição F** mostrada na Figura 8-5, com valores críticos relacionados na Tabela A-5. Se repetirmos continuadamente o experimento que consiste em selecionar aleatoriamente amostras de duas populações normalmente distribuídas e com variâncias iguais, a distribuição da razão s_1^2/s_2^2 das variâncias amostrais é a distribuição *F*. Observe, na Figura 8-5, as seguintes propriedades da distribuição *F*:

- A distribuição *F* não é simétrica.
- Os valores da distribuição *F* não podem ser negativos.
- A forma exata da distribuição *F* depende de dois graus de liberdade diferentes.

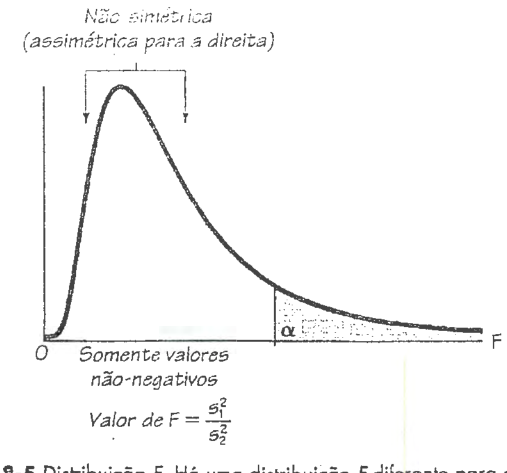


Fig. 8-5 Distribuição F. Há uma distribuição *F* diferente para cada par de graus de liberdade (numerador e denominador).

Estatística para Testes de Hipóteses com Duas Variâncias

$$F = \frac{s_1^2}{s_2^2}$$

Se as duas populações têm realmente variâncias iguais, então $F = s_1^2/s_2^2$ tende para o valor 1, porque s_1^2 e s_2^2 tendem para um mesmo valor. Mas se as duas populações têm variâncias radicalmente diferentes, s_1^2 e s_2^2 tendem a ser muito diferentes. Denotando por s_1^2 a maior das variâncias amostrais, vemos que a razão s_1^2/s_2^2 será um número grande se s_1^2 e s_2^2 tiverem valores muito diferentes. Conseqüentemente, um valor de F próximo de 1 favorece a conclusão de que $\sigma_1^2 = \sigma_2^2$, mas um valor grande de F constitui evidência contra a conclusão de igualdade das variâncias populacionais.

Esta estatística de teste se aplica a uma afirmação feita sobre duas variâncias, mas podemos aplicá-la igualmente a dois desvios-padrão populacionais. Qualquer afirmação sobre dois desvios-padrão populacionais pode ser reformulada em termos das variâncias correspondentes.

Comerciais

As redes de TV têm seus próprios departamentos de censura, para avaliar comerciais e verificar a veracidade de suas informações. A National Advertising Division, um departamento do Council of Better Business Bureau (EUA) analisa afirmações feitas em um anúncio. A Federal Trade Comission e advogados distritais locais também se vêem envolvidos. Em época passada, a Firestone teve de desmentir uma afirmação de que seus pneus permitiam uma parada 25% mais rápida, e a Warner Lambert precisou gastar $10 milhões para informar aos consumidores que Listerine não evita nem cura resfriados. Muitos anúncios enganosos são retirados voluntariamente, mas muitos escapam à fiscalização, simplesmente porque os mecanismos fiscalizadores não conseguem fazer face à enxurrada de comerciais.

Recorrendo à Tabela A-5, obtemos valores críticos de F determinados pelos três valores seguintes:

1. O nível de significância α. (A Tabela A-5 tem várias páginas com valores críticos para α = 0,01, 0,025 e 0,05.)
2. **Graus de liberdade do numerador** = $n_1 - 1$.
3. **Graus de liberdade do denominador** = $n_2 - 1$.

Para achar um valor crítico, localizamos a parte da Tabela A-5 correspondente a α (para um teste unilateral) ou $\alpha/2$ (para um teste bilateral); em seguida, procuramos a interseção da coluna que representa os graus de liberdade para s_1^2 com a linha que representa o número de graus de liberdade para s_2^2. Como estamos admitindo que a maior variância amostral é s_1^2, todos os testes unilaterais serão unilaterais à direita, e em todos os testes bilaterais deveremos achar apenas o valor crítico localizado à direita. Uma boa notícia: não precisamos achar um valor crítico que separe uma região crítica à esquerda. (Como a distribuição F não é simétrica e tem somente valores não-negativos, não se pode determinar um valor crítico à esquerda utilizando-se o negativo do correspondente valor crítico à direita; ao contrário, determina-se um valor crítico à esquerda tomando-se o inverso do valor à direita com os números de graus de liberdade permutados. Veja Exercício 15.)

Por vezes defrontamo-nos com números de graus de liberdade que não constam da Tabela A-5. Poderíamos, nesses casos, apelar para a interpolação linear a fim de obter uma aproximação do valor que falta, mas na maior parte dos casos isso não é necessário, porque a estatística de teste F ou é inferior ao menor valor crítico possível, ou é superior ao maior valor crítico possível. Por exemplo, na Tabela A-5 vemos que, para $\alpha = 0{,}025$ na cauda direita, 20 graus de liberdade no numerador e 34 graus de liberdade no denominador, o valor crítico de F está entre 2,0677 e 2,1952. Qualquer estatística de teste F inferior a 2,0677 resulta na não-rejeição da hipótese nula, qualquer estatística de teste F superior a 2,1952 conduz à rejeição da hipótese nula, e a interpolação é necessária somente se a estatística F está entre 2,0677 e 2,1952. (Veja Exercício 13.) A utilização de um programa estatístico como o STATDISK ou o Minitab elimina este problema, dando valores críticos ou valores P.

EXEMPLO Na seção precedente utilizamos uma amostra de 175 latas de alumínio de 0,0109 in. de espessura, e uma segunda amostra de 175 latas de alumínio de 0,0111 in. de espessura. O Conjunto de Dados 15 relaciona as cargas axiais das latas dessas amostras, e as estatísticas resumo são apresentadas a seguir.

Cargas Axiais (lb) das Latas de 0,0109 in.	Cargas Axiais (lb) das Latas de 0,0111 in.
$n_1 = 175$	$n_2 = 175$
$\bar{x}_1 = 267{,}1$	$\bar{x}_2 = 281{,}8$
$s_1 = 22{,}1$	$s_2 = 27{,}8$

No nível de 0,05 de significância, teste a afirmação de que as amostras provêm de populações com a mesma variância.

SOLUÇÃO Como estipulamos nesta seção que a maior variância seja denotada por s_1^2, invertemos a notação dos índices usados na seção precedente, e fazemos $s_1^2 = 27{,}8^2$, $n_1 = 175$, $s_2^2 = 22{,}1^2$ e $n_2 = 175$. Passamos agora à aplicação do método tradicional de teste de hipóteses delineado na Figura 7-4.

Passo 1: A afirmação de igualdade de variâncias populacionais expressa-se simbolicamente como $\sigma_1^2 = \sigma_2^2$.

Passo 2: Se a afirmação original é falsa, então $\sigma_1^2 \neq \sigma_2^2$.

Passo 3: Como a hipótese nula deve conter a igualdade, temos

$$H_0: \sigma_1^2 = \sigma_2^2 \text{ (original afirmação)} \qquad H_1: \sigma_1^2 \neq \sigma_2^2$$

Passo 4: O nível de significância é $\alpha = 0{,}05$.

Passo 5: Como este teste envolve duas variâncias populacionais, utilizamos a distribuição F.

Passo 6: A estatística de teste é

$$F = \frac{s_1^2}{s_2^2} = \frac{27{,}8^2}{22{,}1^2} = 1{,}5824$$

Para os valores críticos, notemos de início que se trata de um teste bilateral com 0,025 em cada extremidade. Como estamos impondo que a maior variância seja colocada no numerador da estatística de teste F, basta acharmos o valor crítico direito. Pela Tabela A-5, obtemos o valor crítico $F = 1{,}4327$, que corresponde a 0,025 na extremidade direita, com 174

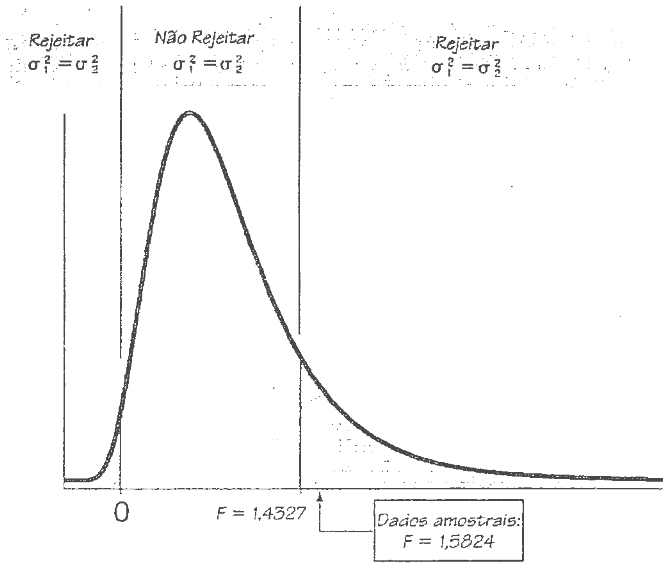


Fig. 8-6 Distribuição de s_1^2/s_2^2 para cargas axiais de latas de 0,0111 in. e 0,0109 in. de espessura.

graus de liberdade para o numerador e 174 graus de liberdade para o denominador. (Efetivamente, a tabela não inclui 174 graus de liberdade, de modo que escolhemos o valor mais próximo de 120 graus de liberdade. Se a diferença das variâncias amostrais é significativa com base em uma amostra de tamanho 121, a diferença certamente o seria com uma amostra de tamanho 175.)

Passo 7: A Figura 8-6 mostra que a estatística de teste $F = 1,5824$ está na região crítica; rejeitamos, pois, a hipótese nula.

Passo 8: Há evidência suficiente para apoiar a rejeição da afirmação de que as duas variâncias são iguais.

Menor Variação, Melhor Qualidade

A Ford e a Mazda vinham fabricando transmissões semelhantes, supostamente com as mesmas especificações. Mas as transmissões americanas necessitavam de mais reparos dentro da garantia do que as transmissões japonesas. Quando pesquisadores inspecionaram amostras de transmissões japonesas, a princípio julgaram que seus instrumentos de medição tivessem algum defeito, porque não estavam detectando variabilidade alguma entre as caixas de transmissão japonesas. Constataram que, embora as transmissões americanas estivessem dentro das especificações, as transmissões japonesas não só estavam dentro das especificações, como se apresentavam consistentemente próximas do valor desejado. Reduzindo a variabilidade entre suas caixas de transmissão, a Mazda reduziu os custos de produção, refugo e reparos dentro da garantia.

No exemplo precedente aplicamos um teste bilateral para a afirmação de igualdade de variâncias. Um teste unilateral direito da afirmação de que as latas de 0,0111 in. têm maior variância daria a mesma estatística de teste $F = 1,5824$, mas um valor crítico diferente $F = 1,3519$ (obtido na Tabela A-5 com $\alpha = 0,05$). Teríamos evidência suficiente para apoiar a afirmação de maior variância.

Descrevemos até aqui o método tradicional de teste de hipóteses sobre duas variâncias populacionais. O Exercício 14 estuda a abordagem pelo valor P e o Exercício 16 trata da construção de intervalos de confiança.

Utilização de Computadores e Calculadoras

O STATDISK e a calculadora TI-83 estão programados para fazer testes de hipóteses do tipo descrito nesta seção, mas o Minitab não faz estes testes. Quando utilizar o STATDISK, selecione `Analysis` do menu principal, selecione `Hypothesis Testing`, em seguida `Variance-Two Samples`. Se estiver trabalhando com uma calculadora TI-83, acione a tecla STAT, em seguida selecione `TESTS` e `2-SampFTEST`.

8-4 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1 e 2, teste a afirmação. Use o nível de 0,05 de significância e suponha que todas as populações tenham distribuição normal. Aplique o método tradicional de teste de hipóteses delineado na Figura 7-4 e trace os gráficos apropriados.

1. Afirmação: As populações A e B têm variâncias distintas ($\sigma_1^2 \neq \sigma_2^2$).
   Amostra A: $n = 25$, $\bar{x} = 175$, $s^2 = 900$
   Amostra B: $n = 31$, $\bar{x} = 200$, $s^2 = 2000$
2. Afirmação: A população A tem variância maior do que a população B.
   Amostra A: $n = 50$, $\bar{x} = 77,4$, $s = 18,1$
   Amostra B: $n = 15$, $\bar{x} = 75,7$, $s = 7,0$
3. Fez-se um experimento para investigar o efeito do cansaço sobre a capacidade de memorização de uma testemunha ocular. O experimento envolveu o interrogatório normal de um suspeito que cooperava e o interrogatório exaustivo de um suspeito que não cooperava. A seguir temos um resumo dos números de detalhes memorizados uma semana após o incidente (com base em dados de "Eyewitness Memory of Police Trainees for Realistic Role Plays",

de Yuille *et al.*, *Journal of Applied Psychology*, Vol. 79, Nº 6). No nível de 0,10 de significância, teste a afirmação de que as amostras provêm de populações com desvios-padrão diferentes.

Sem Estresse	Com Estresse
$n_1 = 40$	$n_2 = 40$
$\bar{x}_1 = 53{,}3$	$\bar{x}_2 = 45{,}3$
$s_1 = 11{,}6$	$s_2 = 13{,}2$

4. Estudam-se os tempos de espera dos clientes no Jefferson Valley Bank. Quando 25 clientes selecionados aleatoriamente entram em uma das várias filas, seus tempos de permanência na fila têm média de 6,896 min e desvio-padrão de 3,619 min. Quando 20 clientes selecionados aleatoriamente entram em uma única fila que é atendida pelos diversos guichês, seus tempos de espera têm média de 7,460 min e desvio-padrão de 1,841 min. No nível de significância de 0,01, teste a afirmação de que os tempos de espera da fila única têm menor desvio-padrão.
5. Como parte da National Health Survey (Pesquisa Nacional sobre a Saúde – EUA), coletaram-se dados sobre os pesos de homens. Para 804 homens com idades de 25 a 34, a média é de 176 lb e o desvio-padrão 35,0 lb. Para 1657 homens com idades de 65 a 74, a média e o desvio-padrão são 164 lb e 27,0 lb, respectivamente. No nível de significância de 0,01, teste a afirmação de que os homens mais velhos provêm de uma população cujo desvio-padrão é inferior ao do grupo de homens na faixa etária 25-34.
6. Em um estudo de mortes de pedestres no estado de Nova York, registraram-se os casos fatais mensais em dois períodos de tempo diferentes. Para o primeiro período, os dados estão resumidos nas seguintes estatísticas: $n = 12$, $\bar{x} = 46{,}42$, $s = 11{,}07$. Para o segundo período de tempo, os dados são: $n = 12$, $\bar{x} = 51{,}00$, $s = 10{,}39$ (com base em dados do New York State Department of Motor Vehicles, Nova York). No nível de 0,05 de significância, teste a afirmação de que ambos os períodos têm a mesma variância.
7. Em um estudo sobre salários de comissários de bordo, selecionaram-se aleatoriamente salários pagos por duas companhias diferentes. Para 40 comissários de bordo da American Airlines, a média é de \$23.870 e o desvio-padrão \$2960. Para 35 comissários de bordo da TWA, a média é \$22.025 e o desvio-padrão \$3065 (com base em dados da Association of Flight Attendants). No nível de 10% de significância, teste a afirmação de que os salários da American e da TWA têm o mesmo desvio-padrão.
8. Planeja-se um experimento para estudar a variabilidade de processos de atribuição de notas entre professores de química orgânica. Pede-se a dois professores diferentes que atribuam nota a um mesmo conjunto de 25 soluções de exames; suas notas acusam variâncias de 103,4 e 39,7, respectivamente. Com o nível de 0,05 de significância, teste a afirmação de que o critério do primeiro professor apresenta maior variância. Se um estudante se revela bastante fraco em química orgânica, e admitindo que as notas atribuídas por ambos os professores tenham a mesma média, há alguma vantagem para um estudante em escolher um dos professores em lugar do outro? Em caso afirmativo, qual deles? Por quê?
9. Testou-se a eficácia de um programa de treinamento mental em um programa de treinamento militar. Em um exame de artilharia antiaérea, registraram-se as notas para um grupo experimental e para um grupo de controle. Com os dados a seguir, teste a afirmação de que ambos os grupos provêm de populações com a mesma variância. Trabalhe com o nível de 0,05 de significância.

Grupo Experimental				Grupo de Controle			
60,83	117,80	44,71	75,38	122,80	70,02	119,89	138,27
73,46	34,26	82,25	59,77	118,43	54,22	118,58	74,61
69,95	21,37	59,78	92,72	121,70	70,70	99,08	120,76
72,14	57,29	64,05	44,09	104,06	94,23	111,26	121,67
80,03	76,59	74,27	66,87				

Dados de "Routinization of Mental Training in Organizations: Effects on Performance and Well-Being", de Larson, *Journal of Applied Psychology*, Vol. 72, nº 1.

10. Muitos estudantes passam pela experiência desagradável do pânico em um exame, porque a primeira questão se revela excepcionalmente difícil. Estudou-se a ordenação das questões de um teste em relação ao seu efeito sobre a ansiedade. Seguem-se as medidas da "ansiedade debilitante no teste" (vulgarmente chamada de pânico ou "branco"). Com o nível de 0,05 de significância, teste a afirmação de que as duas amostras dadas provêm de populações com a mesma variância.

Questões Dispostas em Ordem Crescente de Dificuldade					Questões Dispostas em Ordem Decrescente de Dificuldade			
24,64	39,29	16,32	32,83	28,02	33,62	34,02	26,63	30,26
33,31	20,60	21,13	26,69	28,90	35,91	26,68	29,49	35,32
26,43	24,23	7,10	32,86	21,06	27,24	32,34	29,34	33,53
28,89	28,71	31,73	30,02	21,96	27,62	42,91	30,20	32,54
25,49	38,81	27,85	30,29	30,72				

Com base em dados de "Item Arrangement, Cognitive Entry Characteristics, Sex and Test Anxiety as Predictors of Achievement in Examination Performance", de Klimko, *Journal of Experimental Education*, Vol. 52, Nº 4.

11. Coletaram-se dados amostrais para um estudo de suplementos de cálcio e seus efeitos sobre a pressão sanguínea. Iniciou-se o estudo com as medidas da pressão sanguínea de um grupo *placebo* e um grupo de cálcio (com base em dados de "Blood Pressure and Metabolic Effects of Calcium Supplementation in Normotensive White and Black Men", de Lyle *et al.*, *Journal of the American Medical Association*, Vol. 257, Nº 13). Com o nível de 0,10 de significância, teste a afirmação de que os dois grupos amostrais provêm de populações com o mesmo desvio-padrão. Se o experimento exige grupos com desvios-padrão iguais, os dois grupos do estudo são aceitáveis?

Placebo: 124,6 104,8 96,5 116,3 106,1 128,8 107,2 123,1 118,1 108,5 120,4 122,5 113,6

Cálcio: 129,1 123,4 102,7 118,1 114,7 120,9 104,4 116,3 109,6 127,7 108,0 124,3 106,6 121,4 113,2

12. Com os dados do Conjunto de Dados 11 do Apêndice B e o nível de 0,10 de significância, teste a afirmação de que os confeitos M&M vermelhos e amarelos têm pesos com desvios-padrão diferentes. Se o leitor fosse responsável pelo controle da produção, de modo que os confeitos vermelhos e amarelos tivessem pesos com a mesma variação, tomaria alguma providência de caráter corretivo?

8-4 Exercícios B: Além do Básico

13. Está sendo feito um teste da hipótese $\sigma_1^2 = \sigma_2^2$, com o nível de 0,05 de significância. Que se pode concluir de cada um dos casos abaixo?

    a. Estatística de teste $F = 2{,}0933$; $n_1 = 50$; $n_2 = 35$
    b. Estatística de teste $F = 1{,}8025$; $n_1 = 50$; $n_2 = 35$
    c. Estatística de teste $F = 2{,}3935$; $n_1 = 40$; $n_2 = 20$

14. Para testar uma afirmação sobre duas variâncias populacionais utilizando o método do valor P, primeiro determinamos a estatística de teste F, em seguida consultamos a Tabela A-5 para compará-la com os valores críticos para $\alpha = 0{,}01$, $\alpha = 0{,}025$ e $\alpha = 0{,}05$. Com referência ao Exercício 4, que se pode concluir do valor P?
15. Para os testes de hipóteses bilaterais desta seção, achamos apenas o valor crítico superior. Denotemos esse valor por F_R, onde o índice sugere o valor crítico à direita (*right*). Podemos achar como segue o valor crítico inferior F_L (left = esquerdo). Primeiro permutamos os graus de liberdade, em seguida tomamos o inverso do valor F resultante, obtido na Tabela A-5. (Costuma-se denotar F_R por $F_{\alpha/2}$, e F_L por $F_{1-\alpha/2}$.) Determine os valores críticos F_L e F_R para testes bilaterais, com base nos valores seguintes.

    a. $n_1 = 10$, $n_2 = 10$, $\alpha = 0{,}05$
    b. $n_1 = 10$, $n_2 = 7$, $\alpha = 0{,}05$
    c. $n_1 = 7$, $n_2 = 10$, $\alpha = 0{,}05$
    d. $n_1 = 25$, $n_2 = 10$, $\alpha = 0{,}02$
    e. $n_1 = 10$, $n_2 = 25$, $\alpha = 0{,}02$

16. Além de testar afirmações envolvendo σ_1^2 e σ_2^2, podemos também construir estimativas intervalares da razão σ_1^2/σ_2^2, utilizando a seguinte expressão.

$$\left(\frac{s_1^2}{s_2^2}\cdot\frac{1}{F_R}\right) < \frac{\sigma_1^2}{\sigma_2^2} < \left(\frac{s_1^2}{s_2^2}\cdot\frac{1}{F_L}\right)$$

F_L e F_R são os valores definidos no Exercício 15. Construa a estimativa intervalar de 95% de confiança para a razão da variância do grupo experimental para a variância do grupo de controle para os dados do Exercício 9.

17. Tem-se um conjunto de dados amostrais que consistem em temperaturas registradas para dois grupos diferentes de artigos produzidos por duas técnicas de produção diferentes. Uma especialista em controle de qualidade, que planeja analisar os resultados, começa por testar a igualdade dos dois desvios-padrão populacionais.
    a. Se ela adiciona a mesma constante a cada temperatura dos dois grupos, como fica afetada a estatística de teste F?
    b. Se ela multiplica pela mesma constante cada valor de ambos os grupos, como fica afetada a estatística de teste F?
    c. Se ela transforma todas as temperaturas da escala Fahrenheit para a escala Celsius, qual é o efeito sobre a estatística de teste F?

8-5 Inferências sobre Duas Médias: Amostras Independentes e Pequenas

Nesta seção vamos apresentar métodos de inferência estatística para situações que envolvem as médias de duas populações independentes, mas onde (ao contrário da Seção 8-3) pelo menos uma das amostras é pequena ($n \leq 30$). Estudaremos métodos para testar hipóteses e construir intervalos de confiança para situações em que prevalecem as seguintes suposições.

Suposições No teste de hipóteses sobre as médias de duas populações ou na construção de estimativas intervalares da diferença entre essas médias, os métodos desta seção aplicam-se a casos em que

1. As duas amostras são *independentes*.
2. As duas amostras são *extraídas aleatoriamente de* populações *distribuídas normalmente*.
3. Ao menos uma das duas amostras é pequena ($n \leq 30$).

Quando estas condições são satisfeitas, lançamos mão de um dos três processos diferentes correspondentes aos seguintes casos:

Caso 1: Os valores de ambas as variâncias populacionais são conhecidos. (Na realidade, este caso raramente ocorre.)

Caso 2: As duas populações parecem ter variâncias iguais. (Isto é, com base em um teste da hipótese $\sigma_1^2 = \sigma_2^2$, não rejeitamos a igualdade das duas variâncias populacionais.)

Caso 3: As duas populações parecem ter variâncias diferentes. (Isto é, com base em um teste da hipótese $\sigma_1^2 = \sigma_2^2$, rejeitamos a igualdade das duas variâncias populacionais.)

Esta seção aborda processos que são algo controvertidos, porque os estatísticos não são unânimes em concordar com um método único. Todavia, os métodos que apresentamos são comumente utilizados.

Caso 1: Ambas as Variâncias Populacionais São Conhecidas

Na realidade, o Caso 1 não ocorre com muita freqüência. Em geral, as variâncias populacionais são calculadas com base em dados populacionais conhecidos, e se pudéssemos determinar σ_1^2 e σ_2^2, poderíamos também achar os valores de μ_1 e μ_2, de forma que não haveria necessidade de testar afirmações ou construir intervalos de confiança. Se algum conjunto estranho de circunstâncias nos permite saber os valores de σ_1^2 e σ_2^2, mas não os de μ_1 e μ_2, o teste de afirmações sobre μ_1 e μ_2 pode fazer-se com a seguinte estatística de teste.

Estatística de Teste: Variâncias Populacionais Conhecidas

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$$

Como esta estatística de teste se refere à distribuição normal padronizada, é fácil acharmos valores P: Basta aplicarmos o processo resumido na Figura 7-7.

Intervalos de Confiança

Obtêm-se estimativas do intervalo de confiança para a diferença $\mu_1 - \mu_2$ utilizando-se

$$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$$

onde $$E = z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Esta estatística de teste e este intervalo de confiança refletem a propriedade de variâncias discutida na Seção 8-3: A variância das *diferenças* entre duas variáveis aleatórias independentes é igual à *soma* das variâncias da primeira e da segunda variável aleatória.

Como este caso de variâncias populacionais conhecidas é tão difícil de acontecer, não o ilustraremos com exemplos. Entretanto, os cálculos se aproximam estreitamente dos da Seção 8-3, que utilizam o mesmo formato da estatística de teste e do intervalo de confiança.

Se as três suposições feitas no começo desta seção são satisfeitas e se não conhecemos os valores dos desvios-padrão populacionais, adotamos o seguinte processo.

Escolha entre o Caso 2 e o Caso 3: Utilizar o teste F descrito na Seção 8-4, para testar a hipótese nula $\sigma_1^2 = \sigma_2^2$. Utilizar a conclusão do teste como se segue:

- ***Não rejeitar* $\sigma_1^2 = \sigma_2^2$: Tratar as populações como se tivessem variâncias iguais (Caso 2).**
- ***Rejeitar* $\sigma_1^2 = \sigma_2^2$: Tratar as populações como se tivessem variâncias diferentes (Caso 3).**

Os casos 2 e 3 exigem um teste F preliminar da afirmação $\sigma_1^2 = \sigma_2^2$, mas nem todos os estatísticos concordam com a realização deste teste F preliminar. Argumentam alguns que se aplicamos o teste F com certo nível de significância e então aplicamos um teste t em igual nível, o resultado global não terá o mesmo nível de significância. Outrossim, além de ser sensível a diferenças nas variâncias populacionais, a estatística F é muito sensível a afastamentos da distribuição normal, possibilitando assim a rejeição de uma hipótese nula pela razão errada. (Para um argumento contra o teste F preliminar, veja "Homogeneity of Variance in the Two-Sample Means Test", de Moser e Stevens, *The American Statistician*, Vol. 46, Nº 1.)

Caso 2: As Duas Populações Parecem Ter Variâncias Iguais (Porque Não Rejeitamos $\sigma_1^2 = \sigma_2^2$)

Se não temos evidência suficiente para rejeitar a igualdade das variâncias (isto é, se deixamos de rejeitar $\sigma_1^2 = \sigma_2^2$), calculamos uma **estimativa combinada de σ^2** comum a ambas as populações; denota-se por s_p^2 essa estimativa combinada, que é uma média ponderada de s_1^2 e s_2^2, conforme mostrado a seguir. (Minitab e a calculadora TI-83 são dois dentre os vários sistemas que utilizam a opção "combinada" deste caso.)

Estatística de Teste (Amostras Pequenas Independentes e Variâncias Iguais)

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_p^2}{n_1} + \dfrac{s_p^2}{n_2}}}$$

onde $$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)}$$

e o grau de liberdade é gl $= n_1 + n_2 - 2$.

Intervalo de Confiança (Amostras Pequenas Independentes e Variâncias Iguais)

$$(\bar{x}_1 - \bar{x}_2) - E < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + E$$

onde $$E = t_{\alpha/2}\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}$$

e s_p^2 é dado na estatística de teste.

EXEMPLO Considere os dados da Tabela 8-1 no começo do capítulo. No nível de 0,05 de significância, teste a afirmação de que a quantidade média de *nicotina* em cigarros tipo *king-size* com filtro é igual à quantidade média de *nicotina* presente em cigarros sem filtro. (Todas as medidas são em miligramas e os dados são da Federal Trade Commission.)

SOLUÇÃO Na Figura 8-1 utilizamos diagramas de caixas para comparar cigarros com filtro e sem filtro. Os dois diagramas que representam a quantidade de nicotina parecem indicar claramente que os cigarros sem filtro têm mais nicotina, mas vamos aplicar um teste formal da hipótese de que as duas quantidades médias são iguais. Utilizaremos as estatísticas amostrais a seguir.

Nicotina (mg)

Kings Com Filtro	Kings Sem Filtro
$n_1 = 21$	$n_2 = 8$
$\bar{x}_1 = 0{,}94$	$\bar{x}_2 = 1{,}65$
$s_1 = 0{,}31$	$s_2 = 0{,}16$

As duas amostras são independentes (porque são separadas e não-emparelhadas), os tamanhos das amostras são pequenos, e não conhecemos os valores de σ_1 nem de σ_2, de forma que estamos diante do Caso 2 ou do Caso 3. Aplicando um teste F preliminar (descrito na Seção 8-4) para escolher entre o Caso 2 e o Caso 3, testamos H_0: $\sigma_1^2 = \sigma_2^2$ com a estatística de teste

$$F = \frac{s_1^2}{s_2^2} = \frac{0{,}31^2}{0{,}16^2} = 3{,}7539$$

Com $\alpha = 0{,}05$ em um teste F bilateral e com 20 e 7 graus de liberdade para o numerador e o denominador, respectivamente, obtemos na Tabela A-5 o valor crítico de $F = 4{,}4667$. Como a estatística de teste calculada $F = 3{,}7539$ não está na região crítica, não rejeitamos a hipótese nula de igualdade de variâncias e utilizamos o processo delineado no Caso 2.

Por este processo, testamos a afirmação $\mu_1 = \mu_2$ com as seguintes hipóteses nula e alternativa:

$$H_0: \mu_1 = \mu_2 \text{ (ou } \mu_1 - \mu_2 = 0) \qquad H_1: \mu_1 \neq \mu_2$$

A estatística de teste para este caso de variâncias iguais exige o valor da variância combinada s_p^2; calculamos, pois, este valor em primeiro lugar:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)}$$

$$= \frac{(21 - 1) \cdot 0{,}31^2 + (8 - 1) \cdot 0{,}16^2}{(21 - 1) + (8 - 1)} = 0{,}078$$

Podemos agora obter o valor da estatística de teste:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_p^2}{n_1} + \dfrac{s_p^2}{n_2}}} = \frac{(0{,}94 - 1{,}65) - 0}{\sqrt{\dfrac{0{,}078}{21} + \dfrac{0{,}078}{8}}} =$$

$$-6{,}119$$

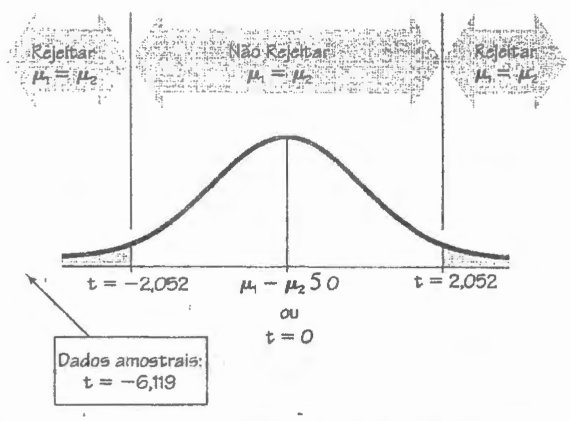


Fig. 8-7 Distribuição de diferenças entre médias de quantidades de nicotina em cigarros com filtro e cigarros sem filtro.

Os valores críticos $t = -2{,}052$ e $t = 2{,}052$ são obtidos da Tabela A-3 como interseção da coluna de $\alpha = 0{,}05$ (bilateral) e da linha correspondente a gl = 27 (21 + 8 − 2). A Figura 8-7 exibe a estatística de teste, os valores críticos e a região crítica. Como a estatística de teste está na região crítica, rejeitamos a hipótese nula $\mu_1 = \mu_2$. Há evidência suficiente para rejeitar a afirmação de que os cigarros com filtro e os sem filtro tenham a mesma quantidade média de nicotina. A conclusão subjetiva a que chegamos, analisando intuitivamente a Figura 8-1, é agora confirmada por um teste formal de hipótese.

EXEMPLO Com os dados do exemplo precedente, construa uma estimativa intervalar de 95% de confiança para $\mu_1 - \mu_2$.

SOLUÇÃO É fácil verificar que $\bar{x}_1 - \bar{x}_2 = 0{,}94 - 1{,}65 = -0{,}71$. Em seguida, achamos o valor da margem de erro E:

$$E = t_{\alpha/2}\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}} = 2{,}052\sqrt{\frac{0{,}078}{21} + \frac{0{,}078}{8}} = 0{,}238$$

Com $\bar{x}_1 - \bar{x}_2 = -0{,}71$ e $E = 0{,}238$, passamos a construir o intervalo de confiança:

$$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$$

$$-0{,}71 - 0{,}238 < (\mu_1 - \mu_2) < -0{,}71 + 0{,}238$$

$$-0{,}95 < (\mu_1 - \mu_2) < -0{,}47$$

Este formato final se mostra estranho com seus sinais negativos, mas podemos interpretá-lo dizendo que temos 95% de confiança de que o cigarros sem filtro têm um nível médio de conteúdo de nicotina que excede o dos cigarros com filtro em uma quantidade entre 0,47 mg e 0,95 mg.

Caso 3: As Duas Populações Parecem Ter Variâncias Desiguais (Porque Rejeitamos $\sigma_1^2 = \sigma_2^2$)

Se temos evidência suficiente para rejeitar a igualdade das variâncias (isto é, se rejeitamos $\sigma_1^2 = \sigma_2^2$), não há método exato para testar a igualdade de médias e construir intervalos de confiança. Um método *aproximado* consiste em utilizar a estatística de teste e o intervalo de confiança seguintes.

Estatística de Teste (Pequenas Amostras Independentes e Variâncias Desiguais)

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

onde gl = o menor dos dois $n_1 - 1$ e $n_2 - 1$.

Intervalo de Confiança (Pequenas Amostras Independentes e Variâncias Desiguais)

$$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$$

onde $E = t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

e gl = o menor dos dois $n_1 - 1$ e $n_2 - 1$.

Esta estatística de teste e este intervalo de confiança dão como número de graus de liberdade o menor dos dois, $n_1 - 1$ ou $n_2 - 1$, mas a aplicação da Fórmula 8-1 constitui uma alternativa mais conservativa e simplificada para o cálculo do número de graus de liberdade:

Fórmula 8-1 $$\text{gl} = \frac{(A + B)^2}{\frac{A^2}{n_1 - 1} + \frac{B^2}{n_2 - 1}}$$

onde $A = \frac{s_1^2}{n_1}$ e $B = \frac{s_2^2}{n_2}$

Obtêm-se resultados mais exatos com a aplicação da Fórmula 8-1, mas eles continuam a ser apenas aproximados. (*Nota:* O STATDISK utiliza a Fórmula 8-1 em lugar do "menor dos dois, $n_1 - 1$ ou $n_2 - 1$," de forma que os resultados do STATDISK são um pouco melhores do que os que constam deste texto.)

EXEMPLO Considere os dados listados na Tabela 8-1 no começo deste capítulo. No nível de 0,05 de significância, teste a afirmação de que a quantidade média de alcatrão em cigarros com filtro é menor do que a quantidade média de alcatrão em cigarros sem filtro. (Todas as medidas são em miligramas, e os dados são da Federal Trade Commission.)

SOLUÇÃO Na Figura 8-1, utilizamos diagramas em caixa para comparar os cigarros com filtro e os cigarros sem filtro. Os dois diagramas em caixa, representando as quantidades de alcatrão, parecem indicar claramente que os cigarros com filtro têm menos alcatrão, mas vamos justificar esta observação com um teste formal de hipótese baseado nas estatísticas amostrais a seguir.

Alcatrão (mg)	
Com Filtro	Sem Filtro
$n_1 = 21$	$n_2 = 8$
$\bar{x}_1 = 13{,}3$	$\bar{x}_2 = 24{,}0$
$s_1 = 3{,}7$	$s_2 = 1{,}7$

As duas amostras são independentes (porque são separadas e não-emparelhadas), os tamanhos das amostras são pequenos e não conhecemos os valores de σ_1 e σ_2; estamos, pois, diante do Caso 2 ou do Caso 3. Com um teste F preliminar, escolhemos entre o Caso 2 e o Caso 3, e testamos H_0: $\sigma_1^2 = \sigma_2^2$ com a estatística de teste

$$F = \frac{s_1^2}{s_2^2} = \frac{3{,}7^2}{1{,}7^2} = 4{,}7370$$

No nível de 0,05 em um teste F bilateral e com 20 e 7 graus de liberdade para o numerador e o denominador, respectivamente, achamos, na Tabela A-5, o valor crítico $F = 3{,}44$ Como a estatística de teste calculada $F = 4{,}7370$ está na região crítica, rejeitamos a hipótese nula de igualdade das variâncias e procedemos de acordo com o que vimos no Caso 3.

Ao aplicar o processo do Caso 3, testamos a afirmação $\mu_1 < \mu_2$ com as seguintes hipóteses nula e alternativa:

$$H_0: \mu_1 \geqslant \mu_2 \text{ (ou } \mu_1 - \mu_2 \geqslant 0)$$
$$H_1: \mu_1 < \mu_2 \text{ (afirmação original)}$$

A estatística de teste para este caso de variâncias desiguais é

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(13,3 - 24,0) - 0}{\sqrt{\frac{3,7^2}{21} + \frac{1,7^2}{8}}} = -10,630$$

Encontra-se o valor crítico $t = -1,895$ na Tabela A-3, na interseção da coluna para $\alpha = 0,05$ (unilateral) com a linha para gl $= 7$ (o menor dos dois, $21 - 1$ e $8 - 1$). Como este teste unilateral esquerdo tem uma estatística de teste $t = -10,630$ e um valor crítico $t = -1,895$, a estatística de teste está na região crítica; rejeitamos, pois, a hipótese nula $\mu_1 \geq \mu_2$. Há evidência suficiente para apoiar a afirmação de que os cigarros *king-size* com filtro têm nível médio de alcatrão inferior no nível médio dos cigarros sem filtro.

EXEMPLO Com os dados do exemplo precedente, construa uma estimativa intervalar de 95% de confiança para $\mu_1 - \mu_2$.

SOLUÇÃO Com $\bar{x}_1 - \bar{x}_2 = 13,3 - 24,0 = -10,7$, determinamos o valor da margem de erro E:

$$E = t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 2,365\sqrt{\frac{3,7^2}{21} + \frac{1,7^2}{8}} = 2,381$$

Com $\bar{x}_1 - \bar{x}_2 = -10,7$ e $E = 2,381$, construímos o intervalo de confiança como se segue:

$$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$$
$$-10,7 - 2,381 < (\mu_1 - \mu_2) < -10,7 + 2,381$$
$$-13,1 < (\mu_1 - \mu_2) < -8,3$$

Novamente aqui, este formato final se mostra estranho com seus sinais negativos, mas podemos interpretá-lo dizendo que temos 95% de confiança de que o conteúdo médio de alcatrão dos cigarros sem filtro excede a média correspondente dos cigarros com filtro em uma quantidade que está entre 8,3 e 13,1 mg.

Esta seção, a Seção 8-2 e a Seção 8-3 todas elas abordam inferências sobre as médias de duas populações. A determinação do processo correto pode ser difícil, porque devemos considerar aspectos como a independência das amostras, os tamanhos das amostras, e se as duas populações parecem ter variâncias iguais. Para evitar confusão, podemos utilizar os processos organizados e sistemáticos das Figuras 8-8 e 8-9 a seguir.

Utilização de Programas de Computador e Calculadoras para Duas Populações

Os cálculos dos exemplos desta seção são complicados, mas podemos simplificá-los utilizando uma calculadora ou um computador. O STATDISK já está programado para seguir automaticamente os processos indicados nesta seção. Basta selecionar os itens do menu `Analysis`, `Hypothesis Testing`, e `Mean-Two Independent Samples`. Quando relevante, a conclusão do teste F preliminar é incluída nos resultados STATDISK.

O Minitab não tem um comando para fazer um teste F da igualdade de duas variâncias. Se as duas variâncias populacionais parecem ser iguais, o Minitab possibilita o uso de uma estimativa conjunta, ou combinada, da variância comum. Após introduzir os dados amostrais nas colunas C1 e C2, selecionam-se as opções `Stat`, `Basic Statistics`, `2-Sample t`, clica-se `Samples in different columns` e passa-se à introdução de C1 para a primeira amostra e C2 para a segunda amostra. Na caixa identificada como `alternative`, seleciona-se a expressão para a hipótese alternativa (*not equal* ou *less than* ou *greater than* — não igual, menor que, maior que), introduz-se o nível de confiança apropriado para o teste (como 0,95 para $\alpha = 0,05$). Haverá uma caixa próxima a `Assume equal variances`; clique nessa caixa somente se as duas populações parecem ter variâncias iguais (Caso 2 desta seção). A apresentação Minitab a seguir corresponde ao primeiro exemplo (Caso 2) desta seção.

A calculadora TI-83 faz um teste da igualdade de variâncias populacionais, e dá a opção de utilizar variâncias "combinadas" (*pooled*) (se admitimos que $\sigma_1^2 = \sigma_2^2$) ou de não combinar variâncias. Para fazer testes do tipo encontrado nesta seção, acione STAT, selecione `TESTS` e escolha `2-SampTTest` (para um teste de hipóteses) ou `2-SampTInt` (para um intervalo de confiança).

MINITAB DISPLAY

```
Two sample T for C1 vs C2
        N      Mean      StDev   SE Mean
C1     21     0.943      0.309     0.067
C2      8     1.650      0.160     0.057

95% CI for mu C1 - mu C2: (-0.944, -0.470)
T-Test mu C1 = mu C2 (vs not =): T= -6.12
P=0.0000 DF= 27
Both use Pooled StDev = 0.278
```

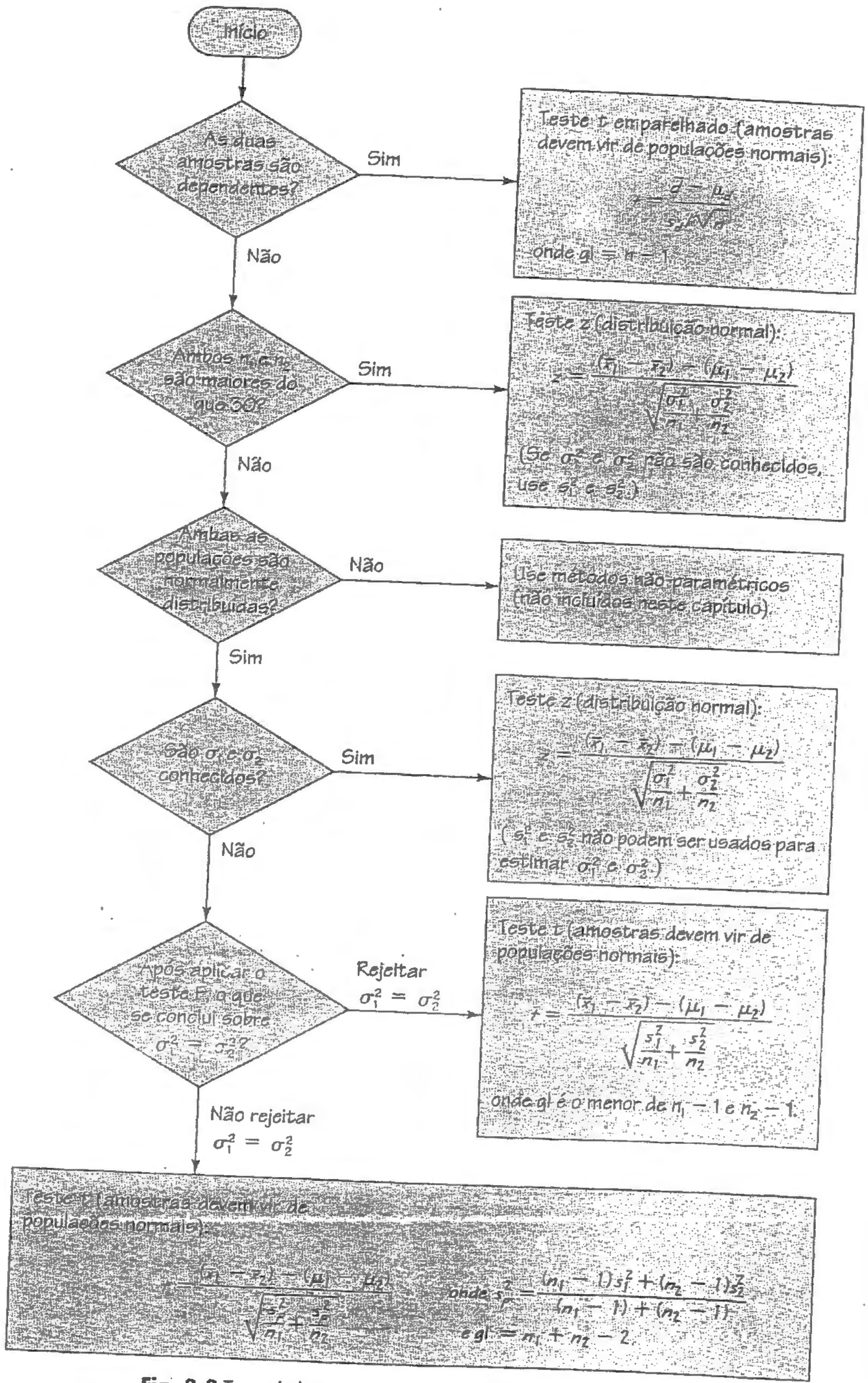


Fig. 8-8 Teste de hipóteses sobre médias de duas populações.

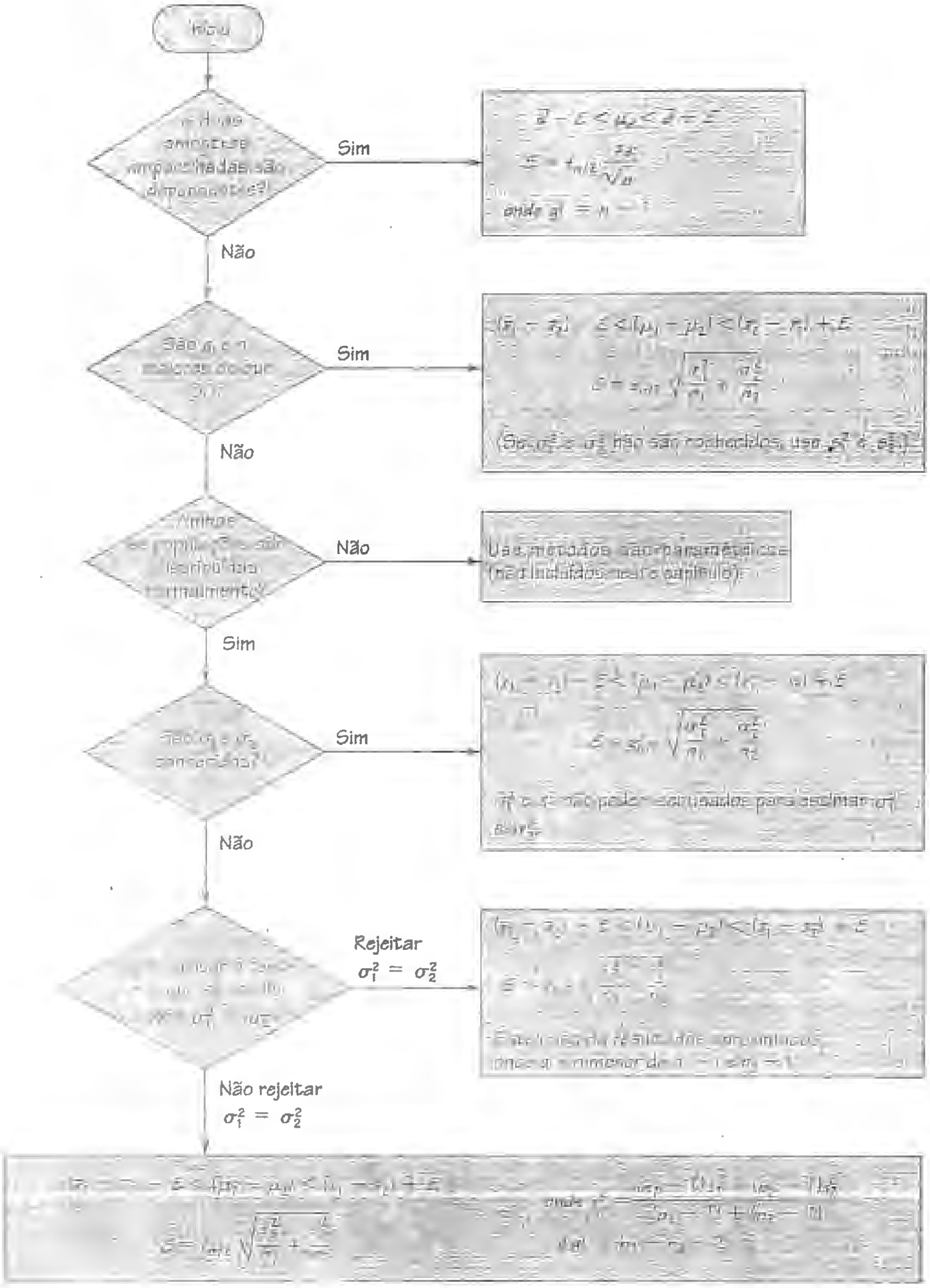


Fig. 8-9 Intervalos de confiança para a diferença entre duas médias populacionais.

8-5 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1 e 2, teste a afirmação, no nível de significância $\alpha = 0{,}05$, supondo que todas as populações tenham distribuição normal. Aplique o método tradicional de teste de hipóteses apresentado na Figura 7-4.

1. Afirmação: As populações A e B têm a mesma média ($\mu_1 = \mu_2$).
   Amostra A: $n_1 = 10$, $\bar{x}_1 = 75$, $s_1 = 15$
   Amostra B: $n_2 = 20$, $\bar{x}_2 = 80$, $s_2 = 12$
2. Afirmação: As populações A e B têm a mesma média ($\mu_1 = \mu_2$).
   Amostra A: $n_1 = 15$, $\bar{x}_1 = 150$, $s_1 = 40$
   Amostra B: $n_1 = 5$, $\bar{x}_2 = 140$, $s_2 = 10$
3. Com os dados amostrais do Exercício 1, construa uma estimativa intervalar de 95% de confiança para a diferença entre as médias das duas populações.
4. Com os dados amostrais do Exercício 2, construa uma estimativa intervalar de 95% de confiança para a diferença entre as médias das duas populações.
5. Testa-se a igualdade das concentrações médias de ozônio para os estados de Maine e Rhode Island (veja Conjunto de Dados 5 no Apêndice B), com os resultados Minitab apresentados aqui. (Um teste *F* preliminar mostra que não há diferença significativa entre as duas variâncias amostrais.) Construa um intervalo de confiança de 95% para a diferença $\mu_1 - \mu_2$, onde μ_1 denota a concentração média de ozônio para o estado do Maine. Os limites do intervalo de confiança contêm 0? Que se pode concluir quanto à igualdade das concentrações médias de ozônio nos estados do Maine e Rhode Island?

MINITAB DISPLAY

```
Two sample T for ME vs RI

        N       Mean        StDev      SE Mean
ME      6       2985         1352          552
RI      6       4789         2948         1203

95% C.I. for mu ME - mu RI: (-4754, 1148)
T-Test mu ME = mu RI (vs not =): T = -1.36
P = 0.20 DF = 10
Both use Pooled StDev = 2293
```

6. Com base nos resultados dados por Minitab do Exercício 5, teste, no nível de significância de 0,05, a afirmação de que os estados do Maine e Rhode Island têm a mesma concentração média de ozônio.
7. Os distúrbios psiquiátricos sérios estão relacionados com fatores biológicos que possam ser observados fisicamente? Em um estudo utilizou-se a tomografia computadorizada de raios X para coletar dados sobre os volumes do cérebro de um grupo de pacientes com distúrbios obsessivo-compulsivos e de outro grupo de controle constituído de pessoas sadias. Dão-se a seguir os resultados amostrais (em mL) para volumes do cordado direito (com base em dados de "Neuroanatomical Abnormalities in Obsessive-Compulsive Disorder Detected with Quantitative X-Ray Computed Tomography", de Luxenberg *et al.*, *American Journal of Psychiatry*, Vol. 145, Nº 9). Teste, no nível de 0,01, a afirmação de que os pacientes obsessivo-compulsivos e as pessoas sadias têm o mesmo volume cerebral médio. Com base nesses resultados, acha que os distúrbios obsessivo-compulsivos tenham uma base biológica?

   Pacientes obsessivo-compulsivos: $n = 10$, $\bar{x} = 0{,}34$, $s = 0{,}08$
   Grupo de controle: $n = 10$, $\bar{x} = 0{,}45$, $s = 0{,}08$

8. Com os dados amostrais do Exercício 7, construa uma estimativa intervalar de 99% de confiança para a diferença entre o volume cerebral médio do grupo de pacientes e o volume cerebral médio do grupo de controle.
9. O mesmo estudo citado no Exercício 7 resultou nos valores de volumes cerebrais (em mL) resumidos a seguir. Com essas estatísticas amostrais, construa um intervalo de confiança de 95% para a diferença entre o volume cerebral médio dos pacientes obsessivo-compulsivos e o volume cerebral médio das pessoas sadias.

   Pacientes obsessivo-compulsivos: $n = 10$, $\bar{x} = 1390{,}03$, $s = 156{,}84$
   Grupo de controle: $n = 10$, $\bar{x} = 1268{,}41$, $s = 137{,}97$

10. Com os dados amostrais do Exercício 9, teste, no nível de 0,05 de significância, a afirmação de que não há diferença entre a média dos pacientes obsessivo-compulsivos e a média das pessoas sadias. Com base neste resultado, parece que os distúrbios obsessivo-compulsivos têm base biológica?
11. O Conjunto de Dados 11 do Apêndice B inclui uma amostra de 21 M&Ms vermelhos, cujos pesos têm média de 0,9097 g e desvio-padrão de 0,0275 g. O conjunto apresenta também uma amostra de 8 M&Ms laranja com peso médio de 0,9251 g e desvio-padrão de 0,0472 g. No nível de 0,05 de significância, teste a igualdade das duas médias populacionais. Se o leitor fosse responsável pelo controle da produção, e desejasse que os M&Ms vermelhos e amarelos tivessem pesos com a mesma média, tomaria alguma providência corretiva?
12. Com base no Conjunto de Dados 11 do Apêndice B, teste a afirmação de que os M&Ms amarelos e os M&Ms verdes têm o mesmo peso médio. Se o leitor integrasse uma equipe encarregada de garantir que os M&Ms amarelos e os verdes tivessem o mesmo peso médio, tomaria alguma providência corretiva?
13. Com base no Conjunto de Dados 10 do Apêndice B, teste a afirmação de que não há diferença entre os comprimentos dos filmes classificados como R e os dos filmes classificados como G, PG ou PG-13.
14. Deu-se a um grupo de estudantes de estatística selecionados aleatoriamente o prazo de cinco minutos para estimar o valor de um produto de números, com os resultados constantes da tabela a seguir. No nível de significância de 0,05, teste a afirmação de que as duas amostras provêm de populações com a mesma média. A ordem dos números parece ter algum efeito sobre a estimativa? (Veja Atividades de Grupo, no final do Capítulo 2.)

    Estimativas de Estudantes para $1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8$

    1560 169 5635 25 842 40.320 5000 500 1110 10.000 200
    1252 4000 2040 175 856 42.200 49.654 560 800

    Estimativas de Estudantes para $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$

    100.000 2000 42.000 1500 52.836 2050 428 372 300 225
    64.582 23.410 500 1200 400 49.000 4000 1876 3600 354
    750 640

15. Com os dados referentes ao monóxido de carbono (CO) da Tabela 8-1, teste a afirmação de que os filtros nos cigarros são eficazes na redução do monóxido de carbono. Com base neste resultado e nos exemplos desta seção, acha que os filtros são eficazes, ou constituem apenas um truque de vendas sem qualquer efeito real?
16. Em um experimento destinado a testar os efeitos do álcool, registraram-se os erros em um teste de habilidade visual e motora para um grupo de tratamento que bebeu etanol, e outro grupo a quem foi dado um *placebo*. Os resultados constam da tabela a seguir. No nível de 0,05 de significância, teste a afirmação de que os dois grupos provêm de populações com a mesma média. Esses resultados apóiam a crença geral de que a bebida é prejudicial para motoristas, pilotos, capitães de navio e outros?

Grupo de Tratamento	Grupo *Placebo*
$n_1 = 22$	$n_2 = 22$
$\bar{x}_1 = 4{,}20$	$\bar{x}_2 = 1{,}71$
$s_1 = 2{,}20$	$s_2 = 0{,}72$

Com base em dados de "Effects of Alcohol Intoxication on Risk Taking, Strategy, and Error Rate in Visuomotor Performance", de Streufert *et al.* *Journal of Applied Psychology*, Vol. 77, Nº 4.

8-5 Exercícios B: Além do Básico

17. Em um experimento para testar os efeitos do álcool, mediram-se os níveis de álcool na respiração de um grupo de tratamento que bebeu etanol e outro grupo a quem foi dado um *placebo*. Os resultados constam da tabela a seguir. No nível de 0,05 de significância, teste a afirmação de que os dois grupos provêm de populações com a mesma média.

Grupo de Tratamento	Grupo *Placebo*
$n_1 = 22$	$n_2 = 22$
$\bar{x}_1 = 0{,}049$	$\bar{x}_2 = 0{,}000$
$s_1 = 0{,}015$	$s_2 = 0{,}000$

Com base em dados de "Effects of Alcohol Intoxication on Risk Taking, Strategy, and Error Rate in Visuomotor Performance", de Streufert *et al. Journal of Applied Psychology*, Vol. 77, Nº 4.

18. Suponha que duas amostras tenham o mesmo desvio-padrão e que ambas sejam independentes, pequenas e selecionadas aleatoriamente de populações com distribuição normal. Suponha também que queiramos testar a afirmação de que as amostras provêm de populações com a mesma média.
    a. É necessário fazer um teste F preliminar?
    b. Se ambas as amostras têm desvio-padrão s, qual é o valor de s_p^2 expresso em termos de s?
19. Considere os dois exemplos desta seção sobre a quantidade de alcatrão em cigarros. (Veja a subseção do Caso 3.) Como serão afetados o teste de hipóteses e o intervalo de confiança se o número de graus de liberdade for calculado pela Fórmula 8-1, e não pelo menor dos dois valores $n_1 - 1$ ou $n_2 - 1$?
20. Considere os dois exemplos desta seção sobre a quantidade de alcatrão em cigarros. (Veja a subseção do Caso 3.) Como serão afetados o teste de hipóteses e o intervalo de confiança se o ponto extremo é eliminado?

8-6 Inferências sobre Duas Proporções

Embora esta seção apareça em último lugar no capítulo, é provavelmente a mais importante, porque é aqui que abordamos métodos de utilização de duas proporções amostrais para fazer inferências (teste de hipóteses e construção de intervalos de confiança) sobre duas proporções populacionais.

Ao testar uma hipótese sobre duas proporções populacionais — como as proporções de pacientes curados em uma população que recebeu determinado tratamento e uma segunda população a quem foi dado um *placebo* — ou ao construir um intervalo de confiança para a diferença entre duas proporções populacionais, fazemos as seguintes suposições e adotamos a seguinte notação.

Suposições

1. Temos dois conjuntos *independentes* de dados amostrais selecionados aleatoriamente.
2. Em ambas as amostras verificam-se as condições $np \geq 5$ e $nq \geq 5$. (Em muitos casos, testaremos a afirmação de que duas populações têm proporções iguais e, assim, $p_1 - p_2 = 0$. Como supomos $p_1 - p_2 = 0$, não é preciso especificar o valor comum de p_1 e p_2. Em tais casos, podemos verificar as condições $np \geq 5$ e $nq \geq 5$ substituindo p pela proporção combinada estimada $\bar{p}$, que será definida mais tarde.)

Notação para Duas Proporções

Para a população 1, seja:

p_1 = proporção *populacional*
n_1 = tamanho da amostra
x_1 = número de sucessos na amostra

$$\hat{p}_1 = \frac{x_1}{n_1} \text{ (proporção amostral)}$$

$$\hat{q}_1 = 1 - \hat{p}_1$$

Atribuem-se significados análogos a p_2, n_2, x_2, $\hat{p}_2$, e $\hat{q}_2$, correspondentes à população 2.

Testes de Hipóteses

Na Seção 7-5 discutimos testes de hipóteses formuladas sobre uma única proporção populacional. Abordaremos agora testes de hipóteses feitas sobre duas proporções populacionais, *limitando-nos, porém, a afirmações do tipo* $p_1 = p_2$; utilizaremos a seguinte estimativa combinada do valor comum a p_1 e p_2. (Para afirmações de que a diferença entre p_1 e p_2 é igual a uma constante diferente de zero, veja o Exercício 17 desta seção.)

Estimativa Combinada de p_1 e p_2

A **estimativa combinada de p_1** e p_2, denotada por $\bar{p}$, é dada por

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

Denotamos por $\bar{q}$ o complemento de $\bar{p}$; assim, $\bar{q} = 1 - \bar{p}$.

Estatística de Teste para Duas Proporções

A seguinte estatística de teste se aplica a hipóteses nulas e alternativas que se enquadram em um dos três formatos:

$$H_0: p_1 = p_2 \qquad H_0: p_1 \geq p_2 \qquad H_0: p_1 \leq p_2$$
$$H_1: p_1 \neq p_2 \qquad H_1: p_1 < p_2 \qquad H_1: p_1 > p_2$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\dfrac{\bar{p}\,\bar{q}}{n_1} + \dfrac{\bar{p}\,\bar{q}}{n_2}}}$$

onde $p_1 - p_2 = 0$

$$\hat{p}_1 = \frac{x_1}{n_1} \quad \text{e} \quad \hat{p}_2 = \frac{x_2}{n_2}$$

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

$$\bar{q} = 1 - \bar{p}$$

O exemplo que segue ajudará a esclarecer os papéis de x_1, n_1, $\hat{p}_1$, $\bar{p}$ etc. Em particular, devemos reconhecer que, sob a hipótese de igualdade de proporções, se obtém a melhor estimativa da proporção comum, combinando ambas as amostras em uma única amostra grande, de modo que $\bar{p}$ passa a ser uma estimativa mais óbvia da proporção populacional comum.

EXEMPLO Pesquisadores da Johns Hopkins fizeram um estudo de empregadas da IBM que estavam grávidas. De 30 empregadas que lidavam com éter-glicol, 10 (ou 33,33%) tiveram aborto (espontâneo), mas, de 750 que não estavam expostas ao éter-glicol, apenas 120 (ou 16,0%) abortaram. No nível de 0,01 de significância, teste a afirmação de que as mulheres expostas ao éter-glicol apresentam maior taxa de aborto.

SOLUÇÃO Para fins de notação, a amostra 1 é o grupo que lida com éter-glicol, e a amostra 2 é o grupo não-exposto; as estatísticas amostrais podem, pois, resumir-se como se segue:

Expostas a Éter-Glicol	Não Expostas a Éter-Glicol
$n_1 = 30$	$n_2 = 750$
$x_1 = 10$	$x_2 = 120$
$\hat{p}_1 = \frac{10}{30} = 0{,}333$	$\hat{p}_2 = \frac{120}{750} = 0{,}160$

Aplicaremos agora o método tradicional de teste de hipóteses, resumido na Figura 7-4.

Passo 1: Podemos representar por $p_1 > p_2$ a afirmação de que as mulheres expostas ao éter-glicol têm maior taxa de aborto.

Passo 2: Se $p_1 > p_2$ é falsa, então $p_1 \leq p_2$.

Passo 3: Como nossa afirmação $p_1 > p_2$ não contém a igualdade, esta torna-se a hipótese alternativa; temos então:

$$H_0: p_1 \leq p_2 \qquad H_1: p_1 > p_2 \text{ (afirmação original)}$$

Passo 4: O nível de significância é $\alpha = 0{,}01$.

Passo 5: Usaremos a distribuição normal (com a estatística de teste dada previamente) como aproximação da distribuição binomial. Temos duas amostras independentes, e as condições $np \geq 5$ e $nq \geq 5$ são satisfeitas por ambas as amostras. Para confirmá-lo, notamos que, ao fazer este teste, supomos $p_1 = p_2$, onde seu valor comum é a estimativa combinada $\bar{p}$, calculada como

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{10 + 120}{30 + 750} = 0{,}1667$$

Com $\bar{p} = 0{,}1667$, temos $\bar{q} = 1 - 0{,}1667 = 0{,}8333$. Verificamos a seguir que $np \geq 5$ e $nq \geq 5$ para ambas as amostras:

Amostra 1

$$n_1 p = (30)(0{,}1667) = 5 \geq 5$$
$$n_1 q = (30)(0{,}8333) = 25 \geq 5$$

Amostra 2

$$n_2 p = (750)(0{,}1667) = 125 \geq 5$$
$$n_2 q = (750)(0{,}8333) = 625 \geq 5$$

Um Experimento com Pólio

Em 1954 fez-se um experimento para testar a eficácia da vacina Salk como proteção contra os efeitos devastadores da poliomielite. Aplicou-se a um grupo de cerca de 200.000 crianças uma solução salina inofensiva; o outro grupo de 200.000 crianças foi ministrada a vacina. O experimento foi do tipo "duplo cego", porque nem as crianças que estavam sendo vacinadas, nem os médicos, que aplicavam as vacinas e avaliavam os resultados, sabiam se estava sendo aplicada a vacina ou um *placebo*. Apenas 33 das 200.000 crianças vacinadas vieram a contrair a pólio paralisante, enquanto 115 das 200.000 que receberam o *placebo* contraíram a pólio paralisante posteriormente. A análise estatística desses e de outros resultados levou à conclusão de que a vacina Salk era de fato eficaz no combate à poliomielite paralisante.

Passo 6: Podemos agora achar o valor da estatística de teste:

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\,\bar{q}}{n_1} + \frac{\bar{p}\,\bar{q}}{n_2}}}$$

$$= \frac{(0{,}333 - 0{,}160) - 0}{\sqrt{\frac{(0{,}1667)(0{,}8333)}{30} + \frac{(0{,}1667)(0{,}8333)}{750}}} = 2{,}49$$

Obtemos o valor crítico $z = 2{,}33$ observando que se trata de um teste unilateral direito com $\alpha = 0{,}01$. O valor $z = 2{,}33$ encontra-se na Tabela A-2 como o escore z correspondente a uma área de $0{,}5 - 0{,}01 = 0{,}4900$.

Passo 7: Na Figura 8-10, vemos que a estatística de teste está na região crítica, o que nos leva a rejeitar a hipótese nula de que $p_1 \leq p_2$.

Passo 8: Concluímos que há evidência suficiente para apoiar a afirmação de que a taxa de aborto é maior entre as mulheres expostas ao etil-glicol.

Esses passos seguem a abordagem tradicional do teste de hipóteses, mas teria sido bastante fácil utilizar o método do valor P. No Passo 6, em lugar de achar o valor crítico de z, determinaríamos o valor P aplicando o processo resumido nas Figuras 7-7 e 7-8. Com a estatística de teste $z = 2{,}49$ e um teste unilateral direito, obtemos

$$\text{Valor } P = (\text{área à direita de } z = 2{,}49) \approx 0{,}0064$$

Novamente rejeitamos a hipótese nula porque o valor P 0,0064 é inferior ao nível de significância $\alpha = 0{,}01$.

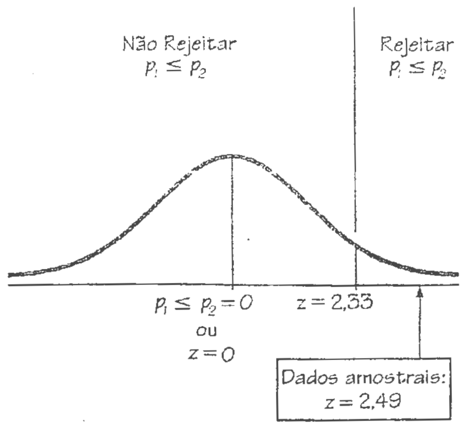


Fig. 8-10 Distribuição das diferenças entre proporções para grupo exposto a éter-glicol e grupo não exposto.

Com esta evidência, os pesquisadores da Johns Hopkins concluíram que as empregadas expostas ao éter-glicol "apresentam risco de aborto significativamente mais alto." Com base nesses resultados, a IBM alertou seus empregados para o perigo, notificou o Environmental Protection Agency e reduziu grandemente o uso de éter-glicol.

Como achamos os valores de x_1 e x_2? No exemplo precedente, com a informação de que dispúnhamos, pudemos ver facilmente que $x_1 = 10$ e $x_2 = 120$. Em outros casos, poderemos ter de calcular o valores de x_1 e x_2, que podem ser achados notando que $\hat{p}_1 = x_1/n_1$ implica

$$x_1 = n_1 \cdot \hat{p}_1$$

Se sabemos que a amostra 1 consiste em 500 pacientes tratados, dos quais 30% ficaram curados, o número efetivo de pacientes curados é $x_1 = 500.(0{,}30) = 150$. Neste caso, $n_1 = 500$, $\hat{p}_1 = 0{,}30$ e $x_1 = 150$.

Intervalos de Confiança

No exemplo precedente constatamos que parece haver maior taxa de aborto para mulheres expostas ao éter-glicol; passamos, pois, a estimar o valor da diferença. A diferença é suficientemente grande para justificar extensas e dispendiosas modificações nos processos de fabricação?

Com as mesmas suposições do princípio desta seção, podemos construir um intervalo de confiança para a diferença entre as proporções populacionais $p_1 - p_2$:

$$(\hat{p}_1 - \hat{p}_2) - E < (p_1 - p_2) < (\hat{p}_1 - \hat{p}_2) + E$$

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

EXEMPLO Com os dados amostrais do exemplo precedente, construa um intervalo de confiança de 99% para a diferença entre as duas proporções populacionais.

SOLUÇÃO Com um grau de confiança de 99%, $z_{\alpha/2} = 2{,}575$ (pela Tabela A-2). Calculamos primeiro a margem de erro E como se segue:

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} =$$

$$= 2{,}575\sqrt{\frac{(0{,}333)(0{,}667)}{30} + \frac{(0{,}160)(0{,}840)}{750}}$$

$$= 0{,}2242$$

Com $\hat{p}_1 = 0{,}333$, $\hat{p}_2 = 0{,}160$ e $E = 0{,}2242$, calculamos o intervalo de confiança:

$$(\hat{p}_1 - \hat{p}_2) - E < (p_1 - p_2) < (\hat{p}_1 - \hat{p}_2) + E$$

$$(0{,}333 - 0{,}160) - 0{,}2242 < (p_1 - p_2) < (0{,}333 - 0{,}160)$$

$$+ 0{,}2242 - 0{,}0512 < (p_1 - p_2) < 0{,}3972$$

O intervalo de confiança contém 0, o que sugere não haver diferença significativa entre as duas proporções. Isto parece contradizer a conclusão do exemplo anterior, que ilustra o processo de teste de hipóteses. Esta discrepância é atribuída a dois fatores: (1) O teste de hipóteses foi unilateral direito com nível de significância de 0,01, correspondendo, portanto, a um intervalo de confiança de 98% (mas o intervalo acima tem um grau de confiança de 99%); (2) A variância usada para o teste de hipóteses (baseada na suposição de proporções iguais) é diferente da variância usada para o intervalo de confiança (onde não se supõe que as variâncias sejam iguais).

Como sempre, devemos ser cuidadosos ao interpretar intervalos de confiança. Como p_1 e p_2 têm valores fixos, e não são variáveis, é errado afirmar que há 99% de chance de o valor de $p_1 - p_2$ estar entre $-0{,}0512$ e $0{,}3972$. O correto é afirmar que, se repetirmos o mesmo processo amostral e construirmos intervalos de confiança de 99%, a longo prazo, 99% dos intervalos conterão efetivamente o valor de $p_1 - p_2$.

Margem de Erro na Liderança

Os autores Stephen Ansolabehere e Thomas Belin escreveram, em seu artigo "Poll Faulting" (*Chance Magazine*) que "nossa maior crítica à maneira de reportar os resultados de uma pesquisa é que a margem de erro consiste em uma proporção única (em geral ±3%), quando a atenção da imprensa é claramente desviada para a *liderança* de um candidato". Os autores salientam que a liderança é, na realidade, a diferença entre duas proporções ($p_1 - p_2$) e explicam a seguir como estabeleceram a seguinte regra empírica: a liderança é aproximadamente $\sqrt{3}$ vezes maior do que a margem de erro para qualquer proporção. Para uma pesquisa pré-eleitoral típica, uma margem de erro reportada de ±3% se transforma em cerca de ±5% para a liderança de um candidato em relação ao outro. Afirmam os autores que o que se deve considerar é a margem de erro para a liderança.

Por que os processos desta seção funcionam? A estatística dada para os testes de hipóteses justifica-se como se segue:

1. Com $n_1p_1 \geq 5$ e $n_1q_1 \geq 5$, a distribuição de $\hat{p}_1$ pode ser aproximada por uma distribuição normal com média p_1, desvio-padrão $\sqrt{p_1q_1/n_1}$ e variância p_1q_1/n_1. Estas conclusões se baseiam nas Seções 5-6 e 6-4 e se aplicam também à segunda amostra.
2. Como $\hat{p}_1$ e $\hat{p}_2$ são cada um aproximados por uma distribuição normal, $\hat{p}_1 - \hat{p}_2$ também será aproximado por uma distribuição normal com média $p_1 - p_2$ e variância
$$\sigma^2_{(\hat{p}_1 - \hat{p}_2)} = \sigma^2_{\hat{p}_1} + \sigma^2_{\hat{p}_2} = \frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}$$
(Na Seção 8-2 vimos que a variância da diferença entre duas variáveis independentes é a soma de suas variâncias individuais.)
3. Como os valores de p_1, p_2, q_1 e q_2 são tipicamente desconhecidos, e como, pela hipótese nula, supomos $p_1 = p_2$, podemos combinar os dados amostrais. A estimativa combinada do valor comum de p_1 e p_2 é $\bar{p} = (x_1 + x_2)/(n_1 + n_2)$. Substituindo p_1 e p_2 por $\bar{p}$ e q_1 e q_2 por $\bar{q} = 1 - \bar{p}$, a variância, do Passo 2, conduz ao desvio-padrão abaixo:
$$\sigma_{(\hat{p}_1 - \hat{p}_2)} = \sqrt{\frac{\bar{p}\,\bar{q}}{n_1} + \frac{\bar{p}\,\bar{q}}{n_2}}$$
4. Sabemos agora que a distribuição de $p_1 - p_2$ é aproximadamente normal com média $p_1 - p_2$ e desvio-padrão conforme mostrado, de modo que a estatística de teste z tem a forma dada anteriormente.

A forma do intervalo de confiança exige uma expressão da variância diferente da que foi dada no Passo 3. Ali, estamos supondo $p_1 = p_2$, mas se não fizermos esta suposição (como na construção de um intervalo de confiança), estimamos a variância de $\hat{p}_1 - \hat{p}_2$ como

$$\sigma^2_{(\hat{p}_1 - \hat{p}_2)} = \sigma^2_{\hat{p}_1} + \sigma^2_{\hat{p}_2} = \frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}$$

e o desvio-padrão se escreve

$$\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

Na estatística de teste

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}}$$

fazemos z positivo e negativo (bilateral) e resolvemos em relação a $p_1 - p_2$. Os resultados são os limites do intervalo de confiança dados anteriormente.

Utilização de Computadores e Calculadoras para Inferências com Duas Proporções

Podemos utilizar o STATDISK em afirmações sobre duas proporções. Para fazer um teste de hipóteses, selecione `Analysis`, em seguida `Hypothesis Testing`, e `Proportion-Two Samples`. Os limites do intervalo de confiança estão incluídos nos resultados do teste de hipóteses.

A calculadora TI-83 também pode ser usada para testes de hipóteses e intervalos de confiança. Acione STAT e selecione `TESTS`. Escolha então a opção `2-PropZTest` (para um teste de hipóteses) ou `2-PropZInt` (para um intervalo de confiança). Ao testar uma hipótese, a calculadora TI-83 apresentará um valor P em lugar de valores críticos, o que nos leva a utilizar o método do valor P para fazer um teste de hipóteses.

8-6 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1 e 2, suponha um nível de significância $\alpha = 0{,}05$ para testar a afirmação de que $p_1 = p_2$. Com os tamanhos de amostra e os números de sucessos dados, determine (a) a estimativa combinada $\bar{p}$, (b) a estatística de teste z, (c) os valores críticos z, e (d) o valor P.

1.

Amostra 1	Amostra 2
$n_1 = 50$	$n_2 = 100$
$x_1 = 25$	$x_2 = 55$

2.

Amostra 1	Amostra 2
$n_1 = 500$	$n_2 = 200$
$x_1 = 300$	$x_2 = 150$

3. Um artigo do *USA Today* afirmou que "em um estudo de 200 pacientes de cor submetidos à cirurgia colorretal, 104 foram mantidos aquecidos com cobertores e fluidos intravenosos, e 96 foram mantidos sem aquecimento. Os resultados acusaram: apenas 6 dos que foram mantidos aquecidos apresentaram infecções em feridas, contra 18 dos que não tiveram aquecimento". No nível de 0,05 de significância, teste a afirmação contida no cabeçalho do artigo: "Os pacientes operados aquecidos se recuperam melhor." Se esses resultados forem confirmados, os pacientes cirúrgicos devem ser aquecidos rotineiramente?

4. Um relatório do Ministério da Justiça dos EUA (NCJ-156831) inclui a afirmação de que "em casos de crimes entre casais, as esposas acusadas têm menor probabilidade de ser condenadas do que os maridos acusados." Os dados amostrais consistiram em 277 condenações entre 318 maridos acusados, e 155 condenações entre 222 esposas acusadas. Teste a afirmação feita e identifique uma explicação possível para o resultado.

5. O chefe, recentemente nomeado, de um departamento de saúde mental afirma que os crimes cometidos por pessoas com menos de 21 anos de idade apresentam maior proporção de crimes violentos, quando comparados com os crimes cometidos por maiores de 21 anos. De 2750 prisões, selecionadas aleatoriamente, de criminosos menores de 21 anos de idade, 4,25% envolvem crimes violentos. E de 2200 prisões, selecionadas aleatoriamente, de criminosos com 21 anos ou mais, 4,55% envolvem crimes violentos (com base em dados do Uniform Crime Reports). Construa um intervalo de confiança de 95% para a diferença entre as duas proporções de crimes violentos. Os limites do intervalo de confiança contêm 0? Isto indica que não há diferença significativa entre as duas taxas de crimes violentos?

6. Nos testes iniciais da vacina Salk, 33 de 200.000 crianças vacinadas contraíram a poliomielite posteriormente. De 200.000 crianças vacinadas com um *placebo*, 115 vieram a ter a pólio posteriormente. Com o nível de 0,01 de significância, teste a afirmação de que a vacina Salk é eficaz na redução da taxa de incidência de poliomielite. A vacina parece realmente ser eficaz?

7. O *New York Times* publicou um artigo sobre um estudo em que a professora Denise Korniewicz e outros pesquisadores da Johns Hopkins submeteram luvas de laboratório a tensão. Entre 240 luvas de vinil, 63% deixaram passar vírus, e entre 240 luvas de látex, apenas 7% deixaram passar vírus. No nível de 0,005 de significância, teste a afirmação de que as luvas de vinil apresentam maior permeabilidade aos vírus do que as luvas de látex. (Pode parecer óbvio que 63% é muito maior do que 7%, mas devemos justificar nossa conclusão considerando os tamanhos das amostras, o nível de significância e a natureza da distribuição.)

8. Com os dados do Exercício 7, construa o intervalo de confiança de 99% para as diferenças entre as duas proporções de luvas permeáveis aos vírus. Esta diferença parece recomendar a decisão de utilizar as luvas de látex, mesmo que sejam mais caras?

9. Um técnico em relações públicas e consultor para a indústria aeronáutica está planejando uma estratégia para influenciar a percepção dos eleitores da regulamentação oficial das tarifas aéreas. Em uma pesquisa *New York Times*/CBS, verificou-se que 35% de 552 democratas acham que o governo deve regulamentar os preços de passagens aéreas, em comparação com 41% dos 417 republicanos pesquisados. Com o nível de 0,05 de significância, teste a afirmação de que não há diferença entre as proporções de democratas e republicanos que apoiam a regulamentação oficial das tarifas aéreas. Com base nos resultados, o analista deve cogitar de abordagens diferentes para democratas e republicanos?

10. Com os dados amostrais do Exercício 9, construa um intervalo de confiança de 95% para a diferença entre as proporções de democratas e republicanos que acham que o governo deve regulamentar os preços das passagens aéreas.

11. Karl Pearson, que elaborou muitos conceitos importantes em estatística, coletou dados sobre crimes em 1909. Dos condenados por incêndio criminoso, 50 bebiam e 43 eram abstêmios. Dos condenados por crime de fraude, 63 bebiam e 144 eram abstêmios. Com o nível de 0,01 de significância, teste a afirmação de que a proporção dos que bebem entre os incendiários é maior do que a proporção dos bebedores condenados por fraude. A bebida parece ter algum efeito sobre o tipo de crime? Por quê?

12. Como parte de uma campanha para a presidência, um candidato planeja incentivar o registro de eleitores. Ao considerar se deve visar a grupos etários específicos, este candidato constata que os resultados das pesquisas mostram que, entre 200 pessoas na faixa etária de 18-24 selecionadas aleatoriamente, 36,0% votaram. Na faixa etária de 25-44, em 250 pessoas, 54,0% votaram (com base

em dados do U.S. Bureau of Census). Construa um intervalo de confiança de 95% para a diferença entre as proporções de votantes nos dois grupos etários. A diferença entre as percentagens de votantes nos dois grupos é suficiente para que estes grupos sejam abordados diferentemente?

13. Com os dados amostrais do Exercício 12, teste a afirmação de que a proporção de eleitores é a mesma em ambas as faixas etárias. Utilize o nível de 0,05 de significância.

14. Os pesquisadores profissionais estão ficando preocupados com a crescente taxa de recusa entre potenciais pesquisados. Analisando o problema, há necessidade de saber se a taxa de recusa é universal ou se há diferença entre as taxas para residentes em cidades centrais e os que não residem em cidades centrais. Especificamente, verificou-se que, de 294 residentes em cidades centrais pesquisados, 28,9% recusaram-se a responder. Uma pesquisa de 1015 residentes de cidades não-centrais acusou uma taxa de recusa de 17,1% [com base em dados de "I Hear You Knocking But You Can't Come In" (estou ouvindo você bater, mas você não pode entrar), de Fitzgerald e Fuller, *Sociological Methods and Research*, Vol. 11, Nº 1]. Com o nível de 0,01 de significância, teste a afirmação de que a taxa de recusa em cidades centrais é a mesma de outras áreas.

15. Fez-se um estudo em 413 crianças hospitalizadas em conseqüência de acidentes com veículos. Entre 290 crianças que não estavam usando o cinto de segurança, 50 sofreram ferimentos graves. Entre 123 que estavam usando cintos, 16 sofreram ferimentos graves (com base em dados de "Morbidity Among Pediatric Motor Vehicle Crash Victims: The Effectiveness of Seat Belts" de Osberg e DiScala, *American Journal of Public Health*, Vol. 82, Nº 3). Há evidência amostral suficiente para concluir, no nível de 0,05 de significância, que a taxa de ferimentos graves é menor quando a criança está usando cinto de segurança? Com base no resultado, que providência deve ser tomada?

16. Em uma amostragem de jogos de uma temporada, verificou-se que o time da casa ganhou 127 de 198 jogos profissionais de *basquete*, e 57 de 99 jogos profissionais de *futebol* (americano) (com base em dados de "Predicting Professional Sports Game Outcomes from Intermediate Game Scores", de Cooper *et al.*, *Chance*, Vol. 5, Nºs 3-4). Construa um intervalo de confiança de 95% para a diferença entre as proporções de vitórias de casa. Os limites do intervalo de confiança contêm 0? Com base nos resultados, há diferença significativa entre as proporções de vitórias do time de casa? Que se pode concluir quanto à vantagem de jogar em casa?

8-6 Exercícios B: Além do Básico

17. Para testar a hipótese nula, de que a diferença entre duas proporções populacionais é igual a uma constante c, diferente de zero, vamos utilizar a estatística de teste

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - c}{\sqrt{\dfrac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \dfrac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

Desde que n_1 e n_2 sejam ambos grandes, a distribuição amostral da estatística de teste z será aproximadamente a distribuição normal padronizada. Com os dados amostrais do exemplo apresentado nesta seção, teste, no nível de 0,05 de significância, a afirmação de um médico de que, quando as mulheres são expostas aos efeitos do éter-glicol, sua percentagem de abortos é dez pontos percentuais maior do que no caso de mulheres não-expostas àqueles efeitos.

18. Extraem-se aleatoriamente dados amostrais de três populações independentes, cada uma de tamanho 100. As proporções amostrais são $\hat{p}_1 = 40/100$, $\hat{p}_2 = 30/100$ e $\hat{p}_3 = 20/100$.
Com o nível de 0,05 de significância, teste as hipóteses (a), (b) e (c):
    a. H_0: $p_1 = p_2$.
    b. H_0: $p_2 = p_3$.
    c. H_0: $p_1 = p_3$.
    d. De modo geral, se os testes levam à conclusão de que $p_1 = p_2$ e $p_2 = p_3$ são aceitáveis, decorre que $p_1 = p_3$ também é aceitável? Por quê ou porque não ?

19. Obtém-se, como se segue, o *tamanho da amostra* necessário para estimar a diferença entre duas proporções populacionais dentro de uma margem de erro E e um nível de confiança $1 - \alpha$. Na expressão

$$E = z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}$$

substituímos n_1 e n_2 por n (supondo que ambas as amostras tenham o mesmo tamanho) e substituímos por 0,5 cada um dos valores p_1, q_1, p_2 e q_2 (porque seus valores não são conhecidos). Finalmente, resolvemos em relação a n.

Com esta abordagem, determine o tamanho de cada amostra necessário para estimar a diferença entre as proporções de homens e mulheres que possuem carros. Suponha o leitor que deseja ter 95% de confiança de que seu erro não seja superior a 0,03.

20. Um artigo da Associated Press reportou que os pesquisadores "selecionaram aleatoriamente 100 motoristas de Nova York envolvidos em acidentes, e 100 que não tinham tido acidentes. Dos motoristas com acidentes, 13,7% possuíam telefone celular, enquanto apenas 10,6% dos motoristas sem acidente tinham celular. Analise esses resultados.

Vocabulário

amostras independentes
amostras dependentes
amostras emparelhadas
amostras combinadas
distribuição F
graus de liberdade do numerador
graus de liberdade do denominador
estimativa combinada de σ^2
estimativa combinada de p_1 e p_2

Revisão

Nos Capítulos 6 e 7 introduzimos dois conceitos relevantes da inferência estatística: a estimação de parâmetros populacionais e os métodos de teste de hipóteses sobre parâmetros populacionais. Enquanto naqueles capítulos abordamos apenas casos envolvendo uma única população, neste capítulo consideramos duas amostras extraídas de duas populações.

- Na Seção 8-2 consideramos inferências sobre duas populações dependentes (emparelhadas).
- Na Seção 8-3, tratamos de inferências sobre duas populações independentes, onde os tamanhos das amostras eram ambos grandes ($n > 30$).
- Na Seção 8-4 apresentamos métodos para testar afirmações sobre dois desvios-padrão ou duas variâncias populacionais.
- Na Seção 8-5 indicamos processos para fazer inferências sobre duas populações independentes, onde pelo menos uma das amostras era pequena ($n \leq 30$). Nas Seções 8-2, 8-3 e 8-5 abordamos uma diversidade de casos relativos a inferências sobre duas médias populacionais; esses casos foram resumidos na Figura 8-8 (para os testes de hipóteses) e na Figura 8-9 (para intervalos de confiança).
- A Seção 8-6 abordou processos de inferência sobre duas proporções populacionais.

Exercícios de Revisão

1. Em um estudo sobre as pessoas que param para ajudar motoristas com carros enguiçados, os pesquisadores admitiram que mais pessoas parariam para prestar auxílio se tivessem visto antes outro motorista com carro enguiçado obtendo auxílio. Em um experimento, 2000 motoris-

tas viram primeiro uma mulher sendo auxiliada no caso de um pneu furado e, em seguida, viram uma mulher sozinha, mais adiante na estrada, também com um pneu do carro furado; 2,90% daqueles 2000 motoristas pararam para auxiliar a segunda mulher. Entre outros 2000 motoristas que não viram a primeira mulher sendo socorrida, apenas 1,75% pararam para auxiliar a segunda mulher (com base em dados de "Help on the Highway", de McCarthy, *Psychology Today*). Com o nível de 0,05 de significância, teste a afirmação de que a percentagem dos que param após verem um motorista com carro enguiçado sendo auxiliado é maior do que a percentagem dos que param sem ter visto outra pessoa obtendo ajuda.

2. Testam-se doze tabletes de Dozenol quanto à solubilidade, antes e depois de serem armazenados por um ano. A tabela a seguir dá os índices de solubilidade:

Antes	472	487	506	512	489	503	511	501	495	504	494	462
Depois	562	512	523	528	554	513	516	510	524	510	524	508

   a. Com o nível de 0,05 de significância, teste a afirmação de que os tabletes de Dozenol são mais solúveis após o período de armazenagem.
   b. Construa uma estimativa intervalar de 95% de confiança para a diferença média "antes − depois".

3. Selecionam-se aleatoriamente doze comprimidos diferentes de cada um de dois remédios antigripais concorrentes, e faz-se um teste do conteúdo de acetaminofena em cada um. Os resultados (em miligramas) são os seguintes:

Dozenol	472	487	506	512	489	503	511	501	495	504	494	462
Niteze	562	512	523	528	554	513	516	510	524	510	524	508

   a. Com o nível de 0,05 de significância, teste a afirmação de que a quantidade média de acetaminofena é a mesma nas duas marcas.
   b. Construa uma estimativa intervalar de 95% de confiança da diferença entre a média para Dozenol e a média para Niteze.

4. Realizou-se um estudo sobre as relações entre diferentes tipos de escores de testes padronizados. No teste verbal da Graduate Record Examination, 68 mulheres obtiveram média de 538,82, com desvio-padrão de 114,16, e 86 homens obtiveram média de 525,23, com desvio-padrão de 97,23. (Veja "Equivalencing MAT and GRE Scores Using Simple Linear Transformation and Regression Methods", de Kagan and Stock, *Journal of Experimental Education*, Vol. 49, Nº 1.) Com o nível de 0,02 de significância, teste a afirmação de que os dois grupos provêm de populações com o mesmo desvio-padrão.

5. Com os dados amostrais do Exercício 4, teste a afirmação de que o escore médio das mulheres é igual ao dos homens. Utilize o nível de 0,02 de significância.

6. Uma questão de teste é considerada boa se permite discriminar entre estudantes preparados e estudantes não-preparados. A primeira questão de um teste foi respondida corretamente por 62 dentre 80 alunos preparados, e por 23 dentre 50 alunos não-preparados. Com o nível de 0,05 de significância, teste a afirmação de que esta questão foi respondida corretamente por uma proporção maior de estudantes preparados.

7. Em um estudo sobre técnicas usadas para medir a capacidade pulmonar, coletaram-se dados fisiológicos de 10 indivíduos. Os valores constantes da tabela a seguir são dados em litros, e representam a capacidade vital forçada dos 10 indivíduos, em posição sentada e em posição deitada. Com o nível de 0,05 de significância, teste a afirmação de que a posição não influi, de modo que a diferença média é zero.

Pessoa	A	B	C	D	E	F	G	H	I	J
Sentada	4,66	5,70	5,37	3,34	3,77	7,43	4,15	6,21	5,90	5,77
Deitada	4,63	6,34	5,72	3,23	3,60	6,96	3,66	5,81	5,61	5,33

Com base em dados de "Validation of Esophageal Balloon Technique at Different Lung Volumes and Postures", de Baydur *et al.*, *Journal of Applied Physiology*, Vol. 62, Nº 1.

8. A Air America está experimentando seu programa de treinamento para comissários de bordo. Com o programa tradicional de seis semanas, uma amostra aleatória de 60 comissários de bordo se submetem a um teste de competência, obtendo notas com média 83,5 e desvio-padrão 16,3. Com um novo programa de 10 dias, uma amostra aleatória de 35 comissários de bordo fazem o teste, obtendo nota média de 79,8 e desvio-padrão de 19,2. Com o nível de 0,01 de significância, teste a afirmação de que o programa de 10 dias resulta em notas com média inferior.

9. Pesquisadores estão testando sistemas comerciais de filtragem de ar, fabricados pela Winston Industrial Supply Company e pela Barrington Filter Company. Testam-se amostras aleatórias de cada companhia, registrando-se a eficiência da filtragem em uma escala padrão, com os seguintes resultados:

Winston	Barrington
$n = 18$	$n = 24$
$\bar{x} = 85{,}7$	$\bar{x} = 80{,}6$
$s = 2{,}8$	$s = 9{,}7$

(Escores mais altos correspondem a melhor filtragem.) Com o nível de 0,05 de significância, teste a afirmação de que ambos os sistemas têm a mesma média.

10. Em um estudo de mortes de pedestres no estado de Nova York, analisam-se os casos fatais mensais para dois períodos de tempo diferentes. Os dados do primeiro período estão resumidos nas estatísticas seguintes: $n = 12$, $\bar{x} = 46{,}42$, $s = 11{,}07$. Para o segundo período, os dados amostrais acusam: $n = 12$, $\bar{x} = 51{,}00$, $s = 10{,}39$ (com base em dados do New York State Department of Motor Vehicles).
   a. Com o nível de 0,05 de significância, teste a afirmação de que ambos os períodos de tempo têm a mesma média.
   b. Construa um intervalo de confiança de 95% para a diferença entre as duas médias populacionais.

Exercícios Cumulativos de Revisão

1. Os dados da tabela a seguir foram obtidos em uma pesquisa junto a indivíduos selecionados aleatoriamente.
   a. Se um dos indivíduos é selecionado aleatoriamente, ache a probabilidade de que alguém seja multado por excesso de velocidade.
   b. Se um dos indivíduos é selecionado aleatoriamente, ache a probabilidade de que um homem ou alguém seja multado por excesso de velocidade.
   c. Ache a probabilidade de que uma pessoa seja multada por excesso de velocidade, dado que a pessoa selecionada é homem.
   d. Ache a probabilidade de que uma pessoa seja multada por excesso de velocidade, dado que a pessoa selecionada é mulher.
   e. Com o nível de 0,05 de significância, teste a afirmação de que a percentagem de mulheres multadas por excesso de velocidade é inferior à dos homens. Pode-se concluir que os homens, de modo geral, correm mais do que as mulheres?

	Multadas por Excesso de Velocidade no Ano Passado	
	Sim	Não
Homens	26	224
Mulheres	27	473

Com base em dados de R. H. Bruskin Associates

2. A Newton Scientific Instrument Company fabrica balanças em seus turnos diurno e noturno. Testam-se rotineiramente amostras aleatórias, registrando-se os erros como positivos no caso de leituras com excesso, ou como negativos no caso de leituras com falta.

a. Testa-se uma amostra aleatória de 40 balanças fabricadas pelo turno diurno; os erros acusam média de 1,2 g e desvio-padrão de 3,9 g. Com o nível de 0,05 de significância, teste a afirmação de que a amostra provém de uma população com erro médio igual a 0.
b. Testa-se uma amostra aleatória de 33 balanças fabricadas pelo turno da noite; os erros acusam média de $-1{,}4$ g e desvio-padrão de 4,3 g. Com o nível de 0,05 de significância, teste a afirmação de que a amostra provém de uma população com erro médio igual a 0.
c. Com o nível de 0,05 de significância, teste a afirmação de que ambos os turnos fabricam balanças com o mesmo erro médio.
d. Construa um intervalo de confiança de 95% para o erro médio do turno do dia.
e. Construa um intervalo de confiança de 95% para o erro médio do turno da noite.
f. Construa um intervalo de confiança de 95% para a diferença entre os erros médios dos turnos do dia e da noite.
g. Se quisermos estimar o erro médio para o turno do dia, quantas balanças devem ser testadas para termos 95% de confiança de que a média não acusa desvio superior a 0,5 g? (*Sug.*: Use o desvio-padrão amostral da parte (a) como estimativa do desvio-padrão populacional σ.)

Projeto para Computador

Recorra ao Conjunto de Dados 1 do Apêndice B referentes aos pesos de metal e plástico descartados e use o STATDISK ou o Minitab para cada um dos casos seguintes:

a. Teste a afirmação de que os pesos do metal descartado e os pesos do plástico descartado têm a mesma média. Registre a estatística de teste e o valor *P*, e formule sua conclusão. Construa também um intervalo de confiança de 95% para a diferença $\mu_m - \mu_p$, onde μ_m é o peso médio do metal descartado e μ_p é o peso médio do plástico descartado.
b. Os pesos usados na parte (a) são dados em libras, e podem ser convertidos para quilogramas multiplicando-se cada valor por 0,4536. Converta todos os pesos de libras para quilogramas e refaça a parte (a). Compare os resultados com os resultados achados originalmente na parte (a). De modo geral, os resultados são afetados pela utilização de uma escala diferente?
c. O primeiro peso do metal descartado é 1,09 lb. Mude-o para 109 lb, omitindo a vírgula decimal. O valor 109 lb é obviamente um *outlier*, mas é apenas conseqüência de um erro no registro dos 62 pesos do metal. Após fazer aquela modificação, refaça a parte (a). Como são afetados os resultados por este *outlier*?

DOS DADOS À DECISÃO

A Aspirina Contribui para Evitar Ataques Cardíacos?

Em estudo recente de 22.000 médicos, metade tomou doses regulares de aspirina, e à outra metade foi administrado um *placebo*. O estudo se estendeu por seis anos, a um custo total de \$4,4 milhões. Entre os que tomaram aspirina, 104 tiveram ataques cardíacos, e dos que receberam um *placebo* 189 tiveram ataques. (As cifras se baseiam em dados do *Time* e do *New England Journal of Medicine*, Vol. 318, Nº 4.) Esses resultados mostram uma redução estatisticamente significativa dos ataques cardíacos no grupo dos que tomaram aspirina? O problema é de indiscutível importância, pois pode afetar muitas vidas.

Com os métodos deste capítulo, determine se o uso da aspirina contribui para evitar ataques cardíacos. Escreva um relatório resumindo suas conclusões. Mencione quaisquer fatores relevantes que possam, a seu ver, afetar a validade do estudo. Por exemplo, é fato importante que o estudo só tenha incluído médicos (homens)? Cabe destacar que a aspirina por vezes causa problemas de estômago?

ATIVIDADES EM GRUPO

1. *Atividades Extraclasse*: As Estimativas são Influenciadas por Números-âncora? Considere a atividade em grupo do Capítulo 2. Ali notamos que, de acordo com o autor John Rubin, quando alguém deve estimar um valor, sua estimativa não raro está "ancorada" a (ou é influenciada por) um número precedente. Naquele capítulo foi solicitado que as pessoas estimassem rapidamente o valor de $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$, enquanto a outro grupo de pessoas pediu-se que estimassem rapidamente o valor de $1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8$.

   No Capítulo 2 pudemos comparar os dois conjuntos de resultados utilizando estatísticas (como a média) e gráficos (como os diagramas em caixa). Agora, os métodos do Capítulo 8 permitem-nos comparar os resultados com um teste formal de hipóteses. Especificamente, vamos coletar dados amostrais e testar a afirmação de que, quando começamos com números mais elevados (como $8 \times 7 \times 6$), nossas estimativas tendem a ser maiores.
2. *Atividade em Classe*: No Exercício 18, da Seção 8-3, utilizamos as taxas de pulsação do Conjunto de Dados 8 do Apêndice B para testar a afirmação de que os estudantes de estatística de ambos os sexos têm a mesma taxa média de pulsação. Divida em grupos de acordo com o sexo, com 10 a 12 estudantes em cada grupo. Cada membro do grupo deve registrar sua taxa de pulsação, contando o número de batidas cardíacas por minuto, e calcular as estatísti-

ATIVIDADES EM GRUPO (Continuação)

cas dos grupos (n, $\bar{x}$, s). Os grupos devem permutar seus resultados e testar a mesma afirmação feita no Exercício 18 da Seção 8-3. Há alguma diferença? Por quê?

3. *Atividade em Classe*: Divida em grupos de cerca de 10 ou 12 estudantes e use o mesmo cronometrador de reação usado na Atividade em Grupo do Capítulo 5. Cada membro do grupo deve ser testado quanto ao tempo de reação da mão direita e da mão esquerda. Com os resultados do grupo, teste a afirmação de que não há diferença entre os tempos de reação de uma ou outra mão. Compare a conclusão com a conclusão formulada por outros grupos. Há alguma diferença?

Rayane Andrade

9

Triola

Correlação e Regressão

9-1 Aspectos Gerais

Este capítulo fornece os métodos para lidar com relações entre duas variáveis. Introduzem-se os conceitos fundamentais de correlação e regressão.

9-2 Correlação

Estuda-se a relação entre duas variáveis com auxílio de um gráfico (chamado diagrama de dispersão) e de uma medida (chamada coeficiente de correlação linear).

9-3 Regressão

Descrevem-se as relações lineares entre duas variáveis com o auxílio da equação e do gráfico de uma linha reta, chamada reta de regressão. Apresenta-se um método para determinar os valores preditos de uma variável.

9-4 Intervalos de Variação e de Predição

Introduz-se um método de análise das diferenças entre os valores preditos de uma variável e os valores efetivamente observados. Constroem-se intervalos de predição, que são estimativas intervalares de confiança para valores preditos.

9-5 Regressão Múltipla

Introduzem-se métodos para achar uma equação linear que relacione três ou mais variáveis. O coeficiente de determinação múltipla é apresentado como uma medida de quão bem os pontos amostrais se ajustam, ou aderem, à equação linear. Em vista dos cálculos envolvidos, enfatiza-se, nesta seção, a utilização de programas e a interpretação de resultados de computador.

Problema do Capítulo

Pode-se pesar um urso com uma fita métrica?

Pesquisadores têm estudado os ursos, anestesiando-os a fim de obterem medidas vitais como idade, sexo, comprimento e peso. (Não tente fazer isto em casa, porque a verificação do comprimento ou do sexo de um urso não-totalmente anestesiado pode ser uma experiência desagradável.) Como os ursos, em sua maioria, são bastante pesados e difíceis de serem levantados, os pesquisadores e os caçadores têm considerável dificuldade em pesar um urso na selva. Será possível determinarmos o peso de um urso a partir de outras medidas mais fáceis de obter?

Veja os dados da Tabela 9-1, que representam os oito primeiros ursos machos do Conjunto de Dados 3 do Apêndice B. (Por uma questão de clareza, utilizaremos este conjunto resumido de dados, mas é possível obter melhores resultados utilizando o conjunto mais completo de valores amostrais.) Com base nestes dados, parece haver alguma relação entre o comprimento de um urso e o seu peso? Em caso afirmativo, qual é esta relação? Se um pesquisador anestesia um urso e usa uma fita métrica para verificar que ele tem 71,0 in. de comprimento, como poderíamos usar este comprimento para predizer o peso do urso? Essas questões serão abordadas neste capítulo.

Tabela 9.1 Comprimentos e Pesos de Ursos Machos

	a	b	c	d	e	f	g	h
Comprimento x (in.) x	53,0	67,5	72,0	72,0	73,5	68,5	73,0	37,0
Peso y (lb) y	80	344	416	348	262	360	332	34
	4240	23220	29952	25056	19257	24660	24236	1258

= μ = 54,56
μ = [illegible]

Com base em dados de Minitab e Gary Alt.

9-1 Aspectos Gerais

No Capítulo 8 consideramos casos que envolviam uma variável e duas populações; neste capítulo vamos lidar com casos que envolvem duas variáveis em uma população — duas variáveis que correspondem a uma amostra de dados *emparelhados*, como os comprimentos e os pesos dos ursos da Tabela 9-1. Com esses dados emparelhados, vamos procurar determinar se há alguma relação entre as duas variáveis e, em caso afirmativo, identificar a relação. Por exemplo, com os dados da Tabela 9-1, procuraremos determinar se há relação entre o peso de um urso e seu comprimento. Se existe uma tal relação, vamos traduzi-la por uma equação, de modo que possamos predizer o peso de um urso medindo seu comprimento, em vez de pesá-lo efetivamente.

Iniciamos a Seção 9-2 considerando o conceito de correlação, utilizado para determinar se há alguma relação estatisticamente significativa entre duas variáveis. Pesquisamos essa relação com auxílio de um diagrama de dispersão (um gráfico) e de um coeficiente de correlação linear (uma medida da intensidade da associação linear entre duas variáveis). Na Seção 9-3 abordamos a análise de regressão, descrevendo a relação entre duas variáveis por meio de uma equação que as relaciona. Ali veremos como utilizar essa equação para predizer valores de uma variável.

Na Seção 9-4 analisamos a diferença entre valores preditos e valores efetivamente observados de uma variável. As Seções 9-2 a 9-4 tratam de relações entre duas variáveis, mas a Seção 9-5 lança mão de conceitos de regressão múltipla para descrever o relacionamento entre três ou mais variáveis. Em todo este livro trabalharemos apenas com relações lineares (linha reta) entre duas ou mais variáveis.

9-2 Correlação

Nesta seção procuraremos determinar se há algum relacionamento entre duas variáveis. Em estatística, tal relacionamento é chamado *correlação*.

DEFINIÇÃO
Existe uma **correlação** entre duas variáveis quando uma delas está, de alguma forma, relacionada com a outra.

A Tabela 9-1, por exemplo, consiste em dados emparelhados (às vezes chamados **dados bivariados**). Vamos determinar se há correlação entre a variável x (comprimento) e a variável y (peso). A importância de tal determinação decorre do fato de que a presença de uma correlação pode conduzir-nos a um método para estimar o peso de um urso medindo o seu comprimento.

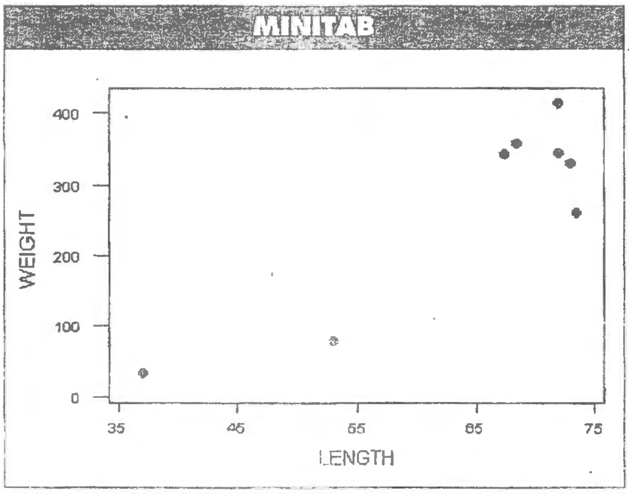
Quando trabalhamos com dados amostrais e estabelecemos métodos para formular inferências sobre populações, fazemos as seguintes suposições.

Suposições

1. A amostra de dados emparelhados (x, y) é *aleatória*.
2. Os pares de dados (x, y) têm uma **distribuição normal bivariada**. (As distribuições normais foram estudadas no Capítulo 5, mas a presente hipótese exige basicamente que, para qualquer valor fixo de x, os valores correspondentes de y tenham uma distribuição em forma de sino, e que, para qualquer valor fixo de y, os valores de x também tenham uma distribuição em forma de sino.)

Quanto à segunda suposição, em geral é difícil verificá-la; mas pode-se fazer uma verificação parcial determinando-se se os valores tanto de x como de y têm distribuições basicamente em forma de sino.

Não raro podemos formular conclusões intuitivas e qualitativas sobre dados emparelhados construindo um gráfico, usando o Minitab, semelhante ao apresentado a seguir, que representa os dados da Tabela 9-1. (Para obter este gráfico com o Minitab, introduzimos os valores x na coluna C1, os valores y correspondentes na coluna C2 e selecionamos no menu principal o item `Graph`; em seguida selecionamos `Plot` e passamos a introduzir C2 na caixa para a variável y e C1 na caixa para a variável x.)



A tela do Minitab é um exemplo de **diagrama de dispersão**, que é um gráfico de dados emparelhados (x, y) com o eixo x horizontal e o eixo y vertical. Como os pontos da figura aparentam um certo padrão, podemos concluir que há uma relação entre o comprimento de um urso e o seu peso. Esta conclusão é, em grande parte, subjetiva, porque se baseia na nossa percepção da existência de um padrão.

A Figura 9-1 mostra outros exemplos de diagramas de dispersão. Os gráficos das Figuras 9-1(a), (b) e (c) exibem um padrão de valores crescentes de y que correspondem a valores crescentes de x. No gráfico de (a) a (c), o padrão de pontos aproxima-se de uma linha reta, sugerindo uma relação mais forte entre x e y. Nos diagramas de dispersão (d), (e) e (f) os valores de y decrescem quando os valores de x crescem. Novamente aqui, ao procedermos de (d) a (f), a relação se torna mais forte. Em contraste com os seis primeiros gráficos, o diagrama (g) não apresenta qualquer padrão definido e sugere que não há correlação (ou relacionamento) entre x e y. Finalmente, o diagrama de dispersão (h) exibe um padrão que não é, entretanto, linear.

Como as conclusões tiradas de diagramas de dispersão tendem a ser subjetivas, necessitamos de métodos mais precisos e objetivos. Vamos utilizar o coeficiente de correlação linear para detectar padrões lineares, mas não padrões não-lineares, como o da Figura 9-1(h).

DEFINIÇÃO

O **coeficiente de correlação linear** r mede o grau de relacionamento linear entre os valores emparelhados x e y em uma *amostra*. Calcula-se seu valor com auxílio da Fórmula 9-1 a seguir. [O coeficiente de correlação linear é chamado às vezes **coeficiente de correlação momento-produto de Pearson,** em homenagem a Karl Pearson (1857-1936), que o estabeleceu.]

Fórmula 9-1
$$r = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{n(\Sigma x^2) - (\Sigma x)^2}\sqrt{n(\Sigma y^2) - (\Sigma y)^2}}$$

Como r é calculado com base em dados amostrais, é uma estatística amostral usada para medir o grau da correlação linear entre x e y. Se tivéssemos todos os pares de valores (x, y) para a população, a Fórmula 9-1 seria um parâmetro populacional, representado pela letra grega ρ (rô).

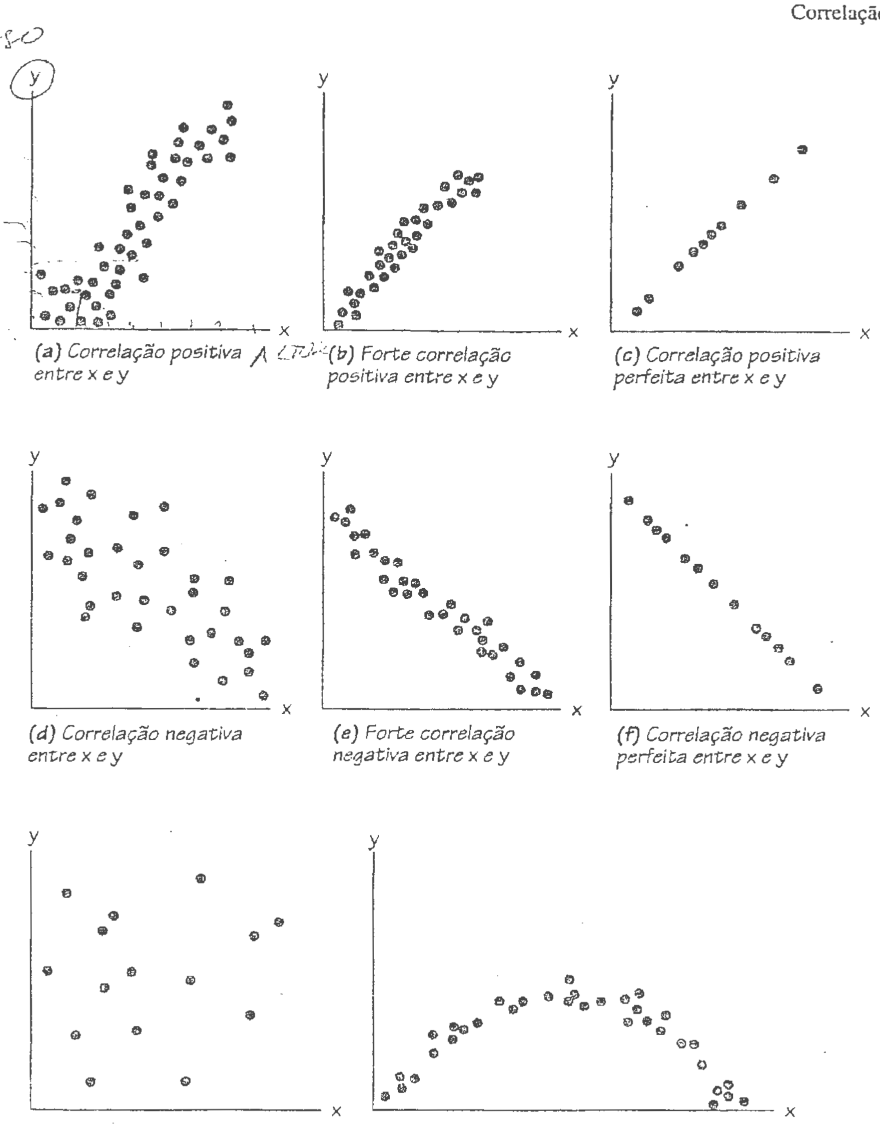
Mostraremos como calcular e interpretar o coeficiente de correlação linear r para uma lista de dados emparelhados, mas antes vamos estabelecer a notação adequada para a Fórmula 9-1. Mais adiante nesta seção apresentaremos a teoria fundamental que levou ao estabelecimento da Fórmula 9-1.

Notação para o Coeficiente de Correlação Linear

n	representa o número de pares de dados presentes.
Σ	denota a adição dos itens indicados.
Σx	denota a soma de todos os valores de x.
Σx^2	indica que devemos elevar ao quadrado cada valor de x e somar os resultados.
$(\Sigma x)^2$	indica que devemos somar os valores de x e elevar o total ao quadrado. É sumamente importante não confundir Σx^2 com $(\Sigma x)^2$.
Σxy	indica que devemos multiplicar cada valor de x pelo valor correspondente de y e somar então todos esses produtos.
r	representa o coeficiente de correlação linear para uma *amostra*.
ρ	representa o coeficiente de correlação linear para uma *população*.

Arredondamento do Coeficiente de Correlação Linear r

Devemos arredondar o coeficiente de correlação linear r para três decimais, a fim de que seu valor possa ser comparado com os



(a) Correlação positiva entre x e y

(b) Forte correlação positiva entre x e y

(c) Correlação positiva perfeita entre x e y

(d) Correlação negativa entre x e y

(e) Forte correlação negativa entre x e y

(f) Correlação negativa perfeita entre x e y

(g) Não há correlação entre x e y

(h) Correlação não-linear entre x e y

Fig. 9-1 Diagramas de dispersão.

valores críticos da Tabela A-6. Ao calcular r e outras estatísticas neste capítulo, o arredondamento no meio dos cálculos pode ocasionar erros sérios; recorra, pois, à memória de sua calculadora para armazenar os resultados intermediários, fazendo o arredondamento somente no final. Muitas calculadoras de preço reduzido têm a Fórmula 9-1 embutida, o que permite calcular r automaticamente após introduzir os dados amostrais.

EXEMPLO Com os dados da Tabela 9-1, calcule o coeficiente de correlação linear r. (Em um exemplo posterior utilizaremos este valor para determinar se há algum relacionamento entre os comprimentos dos ursos e seus pesos.)

SOLUÇÃO Para os dados amostrais emparelhados da Tabela 9-1, $n = 8$ porque há oito pares de dados. Os outros elementos necessários na Fórmula 9-1 se obtêm com os cálculos da Tabela 9-2. Note como esta disposição vertical facilita os cálculos.

Com os valores calculados e a Fórmula 9-1 podemos calcular r como se segue:

$$r = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{n(\Sigma x^2) - (\Sigma x)^2}\sqrt{n(\Sigma y^2) - (\Sigma y)^2}}$$

$$= \frac{8(151.879) - (516,5)(2176)}{\sqrt{8(34.525,75) - (516,5)^2}\sqrt{8(728.520) - (2176)^2}}$$

$$= \frac{91.128}{\sqrt{9433,75}\sqrt{1.093.184}} = 0,897$$

TABELA 9-2 Cálculo de r

Comprimento (in.) x	Peso (lb) y	$x \cdot y$	x^2	y^2
53,0	80	4.240	2809,00	6.400
67,5	344	23.220	4556,25	118.336
72,0	416	29.952	5184,00	173.056
72,0	348	25.056	5184,00	121.104
73,5	262	19.257	5402,25	68.644
68,5	360	24.660	4692,25	129.600
73,0	332	24.236	5329,00	110.224
37,0	34	1.258	1369,00	1.156
Total 516,5 ↑ Σx	2176 ↑ Σy	151.879 ↑ Σxy	34.525,75 ↑ Σx^2	728.520 ↑ Σy^2

Utilização de Computadores e Calculadoras para Correlação Linear

O STATDISK e o Minitab aceitam dados emparelhados como entrada (*input*) e dão o coeficiente de correlação linear como saída (*output*). Com o STATDISK, selecione `Analysis` na barra do menu principal e utilize a opção `Correlation and Regression`. Com o Minitab, introduza os dados emparelhados nas Colunas C1 e C2, selecione `Stat` na barra do menu principal, escolha `Basic Statistics`, em seguida escolha `Correlation`, e passe a introduzir C1 e C2 para as colunas a serem utilizadas. Com uma calculadora TI-83, introduza os dados emparelhados nas listas L1 e L2, acione `STAT` e selecione `TESTS`. A opção `LinRegTTest` resultará na apresentação de diversos valores, inclusive o coeficiente de correlação linear r.

Interpretação do Coeficiente de Correlação Linear

Devemos interpretar um valor calculado de r, como o valor $r = 0{,}897$ achado no exemplo precedente. Dada a maneira como foi estabelecida a Fórmula 9-1, o valor de r deve estar sempre entre -1 e $+1$, inclusive. Se o valor de r está próximo de 0, concluímos que não há correlação linear significativa entre x e y, mas se r está próximo de -1 ou $+1$, concluímos pela existência de correlação linear significativa entre x e y. Como a interpretação da expressão "próximo de" 0, ou 1, ou -1 é vaga, adotaremos o critério de decisão seguinte, bastante específico:

> **Se o módulo do valor calculado de r excede o valor na Tabela A-6, concluímos que há correlação linear significativa. Em caso contrário, não há evidência suficiente para apoiar a existência de uma correlação linear significativa.**

Quando não há realmente correlação linear entre x e y, a Tabela A-6 relaciona os valores dados como "críticos" no sentido seguinte: Eles separam valores *usuais* de r dos valores que são *raros*. Por exemplo, a Tabela A-6 mostra que, com $n = 8$ pares de dados amostrais, os valores críticos são 0,707 (para $\alpha = 0{,}05$) e 0,834 (para $\alpha = 0{,}01$). Nos Capítulos 6 e 7 descrevemos cuidadosamente os valores críticos e o papel de α. Eis como interpretamos esses números: Com oito pares de dados e sem qualquer correlação linear entre x e y, há uma chance de 5% de que o valor absoluto do coeficiente de correlação linear calculado r exceda 0,707. E com $n = 8$ e sem qualquer correlação linear, há uma chance de 1% de que $|r|$ exceda 0,834.

Correlação entre Linhas de Força e Câncer

Quando se constatam correlações entre variáveis, podem ocorrer resultados interessantes, surpreendentes e úteis. Vários estudos científicos sugerem a existência de uma correlação entre a exposição a campos eletromagnéticos e a incidência de câncer. Os epidemiologistas do Instituto Karolinska da Suécia pesquisaram 500.000 suecos que viviam a 300 metros de uma linha de alta tensão por um período de 25 anos, e constataram que as crianças apresentavam maior incidência de leucemia. Essas conclusões levaram o governo da Suécia a elaborar regulamentos que reduzissem o número de residências na proximidade de linhas de força de alta tensão. Em um artigo sobre esse estudo, a revista *Time* escreveu: "Embora a pesquisa não prove relação de causa e efeito, mostra uma correlação indiscutível entre o grau de exposição e o risco de leucemia infantil."

EXEMPLO Com os dados amostrais da Tabela 9-1, para os quais $r = 0{,}897$, recorra à Tabela A-6 para determinar se há correlação linear significativa entre comprimentos e pesos de ursos. Na Tabela A-6, tome o valor crítico para $\alpha = 0{,}05$. (Com $\alpha = 0{,}05$, concluiremos que há correlação linear significativa somente se a amostra é pouco provável, no sentido de que tal valor de r ocorre menos de 5% das vezes.)

SOLUÇÃO Na Tabela A-6, localize a linha para $n = 8$ (porque há oito pares de dados). Aquela linha contém os valores críticos 0,707 (para $\alpha = 0{,}05$) e 0,834 (para $\alpha = 0{,}01$). Com o valor crítico para $\alpha = 0{,}05$, vemos que há menos de 5% de chance de que, sem correlação linear alguma, o valor absoluto calculado de r exceda 0,707. Como $r = 0{,}897$, seu valor absoluto excede 0,707, e concluímos que há correlação linear significativa entre comprimentos e pesos de ursos.

Já observamos que, com o formato da Fórmula 9-1, o valor calculado de r está sempre entre -1 e $+1$, inclusive. Relacionamos esta propriedade juntamente com outras propriedades importantes.

Propriedades do Coeficiente de Correlação Linear r

1. O valor de r está sempre entre -1 e 1. Isto é,

$$-1 \leq r \leq 1$$

2. *O valor de r não varia se todos os valores de qualquer uma das variáveis são convertidos para uma escala diferente.* Por exemplo, se os pesos dos ursos da Tabela 9-1 são dados em quilogramas em vez de libras, o valor de r não se modificará.
3. *O valor de r não é afetado pela escolha de x ou y.* Permutando todos os valores de x e y, r permanecerá inalterado.
4. *r mede a intensidade, ou grau, de um relacionamento linear.* Não serve para medir a intensidade de um relacionamento não-linear.

Erros Comuns Que Envolvem a Correlação

Identificamos a seguir três dos erros mais comuns cometidos na interpretação de resultados que envolvem correlação.

1. *Devemos evitar a conclusão de que a correlação implica causalidade.* Um estudo mostrou uma correlação entre os salários de professores de estatística e o consumo individual de cerveja; mas essas duas variáveis são afetadas pelas condições econômicas, uma terceira variável oculta. (Define-se formalmente uma **variável oculta** como uma variável que afeta as variáveis em estudo mas não está incluída no estudo.)
2. *Surge outra fonte de erro potencial quando os dados se baseiam em taxas ou médias.* Quando utilizamos taxas ou médias para os dados, suprimimos a variação entre os indivíduos ou elementos, e isto pode levar a um coeficiente de correlação inflacionado. Um estudo acusou coeficiente de correlação linear de 0,4 para dados emparelhados relativos a renda e educação entre indivíduos, mas aquele coeficiente passou para 0,7 quando foram consideradas médias regionais.
3. *Um terceiro erro diz respeito à propriedade de linearidade.* A conclusão de que não há correlação linear significativa não quer dizer que x e y não estejam relacionados de alguma forma. Os dados ilustrados na Figura 9-2 conduzem a um valor $r = 0$, que é uma indicação da ausência total de correlação *linear* entre as duas variáveis; mas pode-se ver facilmente pela figura que o padrão dos dados reflete forte relacionamento *não-linear*. (A Figura 9-2 é um diagrama de dispersão do relacionamento entre a distância acima do solo e o tempo gasto por um objeto atirado para cima.)

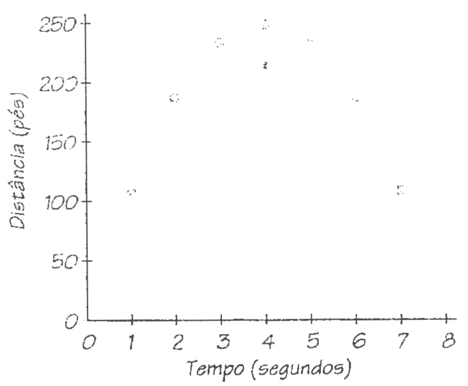


Fig. 9-2 Diagrama de dispersão da distância acima do solo e do tempo para um objeto lançado para cima.

Teste Formal de Hipóteses (Exige a Leitura do Capítulo 7)

Apresentamos dois métodos (resumidos na Figura 9-3) de utilização de um teste formal de hipóteses para determinar se existe correlação linear significativa entre duas variáveis. Alguns professores preferem o Método 1 porque reforça conceitos introduzidos em capítulos anteriores. Outros preferem o Método 2 porque exige cálculos mais fáceis.

A Figura 9-3 mostra que as hipóteses nula e alternativa se expressarão como segue:

H_0: $\rho = 0$ (Não há correlação linear significativa)
H_1: $\rho \neq 0$ (Correlação linear significativa)

Para a estatística de teste, utilizamos um dos métodos seguintes:

Método 1: A Estatística de Teste É t Este método segue o formato apresentado em capítulos anteriores. Utiliza a distribuição t de Student com uma estatística de teste da forma $t = (r - \mu_r)/s_r$, onde μ_r e s_r denotam, respectivamente, o valor alegado da média e o desvio-padrão amostral de valores de r. Como supomos que $\rho = 0$, decorre que $\mu_r = 0$. Mostra-se também que s_r, o desvio-padrão dos coeficientes de correlação linear, pode ser expresso como $[(1 - r^2)/(n - 2)]^{1/2}$. Podemos, pois, usar a seguinte estatística de teste.

Estatística de Teste t para Correlação Linear

$$t = \frac{r}{\sqrt{\dfrac{1 - r^2}{n - 2}}}$$

Valores críticos: Utilize a Tabela A-3 com graus de liberdade $= n - 2$.

Método 2: A Estatística de Teste É r Este método exige menos cálculos. Em lugar de calcular a estatística de teste, que acabamos de dar, usamos o valor calculado de r como estatística de teste. Os valores críticos encontram-se na Tabela A-6.

Estatística de Teste r para Correlação Linear

Estatística de teste: r

Valores críticos: Consulte a Tabela A-6.

A Figura 9-3 mostra que o critério de decisão consiste em rejeitar a hipótese nula $\rho = 0$ se o valor absoluto da estatística de teste excede os valores críticos; a rejeição de $\rho = 0$ significa que há evidência suficiente para apoiar a existência de uma correlação linear entre as duas variáveis. Se o valor absoluto da es-

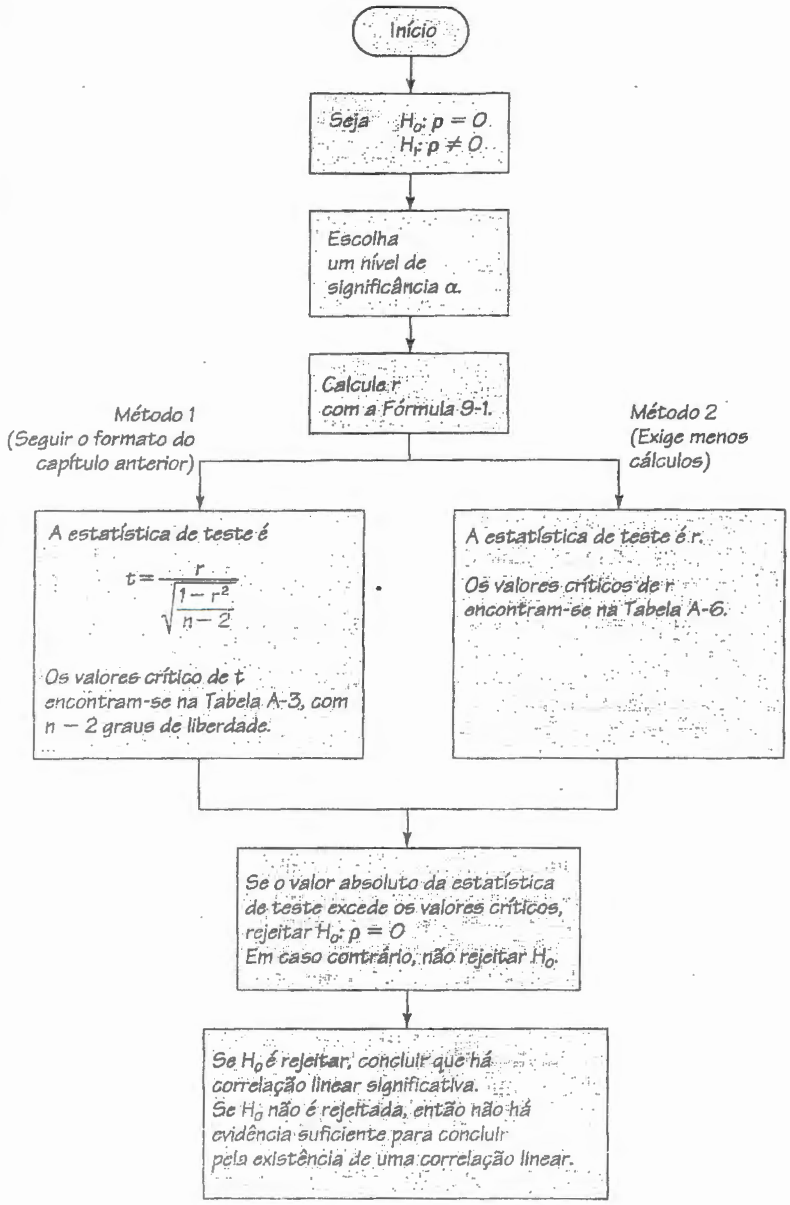


Fig. 9-3 Teste para correlação linear.

tatística de teste não excede os valores críticos, então não rejeitamos $\rho = 0$; isto é, não há evidência suficiente para concluir pela existência de uma correlação linear entre as duas variáveis.

EXEMPLO Com os dados amostrais da Tabela 9-1, teste a afirmação de que há uma correlação linear entre comprimentos e pesos de ursos. Para a estatística de teste, utilize ambos: (a) o Método 1 e (b) o Método 2.

SOLUÇÃO Veja a Figura 9-3. Afirmar que há correlação linear significativa equivale a afirmar que o coeficiente ρ de correlação linear populacional é diferente de 0. Temos, pois, as hipóteses seguintes:

H_0: $\rho = 0$ (Não há correlação linear significativa)
H_1: $\rho \neq 0$ (Correlação linear significativa)

Como não se especificou nível de significância, tomamos $\alpha = 0{,}05$.

Em exemplo precedente já obtivemos $r = 0{,}897$. Com este valor, determinamos agora a estatística de teste e o valor crítico, utilizando ambos os métodos que acabamos de descrever.

a. *Método 1:* A estatística de teste é

$$t = \frac{r}{\sqrt{\dfrac{1-r^2}{n-2}}} = \frac{0{,}897}{\sqrt{\dfrac{1-0{,}897^2}{8-2}}} = 4{,}971$$

Os valores críticos $t = -2{,}447$ e $t = 2{,}447$ são obtidos na Tabela A-3, onde 2,447 corresponde a 0,05 dividido entre duas caudas (com 0,025 em cada uma), e o número de graus de liberdade é $n - 2 = 6$.

Veja Figura 9-4 para o gráfico que inclui a estatística de teste e o valor crítico.

b. *Método 2:* A estatística de teste é $r = 0{,}897$. Os valores críticos $r = -0{,}707$ e $r = 0{,}707$ são obtidos na Tabela A-6 com $n = 8$ e $\alpha = 0{,}05$. Veja a Figura 9-5 para um gráfico que inclui esta estatística de teste e os valores críticos.

Com qualquer dos dois métodos, vemos que o valor absoluto da estatística de teste excede o valor crítico (Método 1: $4{,}971 > 2{,}447$; Método 2: $0{,}897 > 0{,}707$); ou seja, a estatística de teste está na região crítica. Rejeitamos, pois, H_0: $\rho = 0$. Há evidência suficiente para apoiar a existência de uma correlação linear entre comprimentos e pesos de ursos. O peso de um urso parece realmente corresponder ao seu comprimento.

O exemplo precedente e as Figuras 9-4 e 9-5 ilustram um teste de hipóteses bilateral. De modo geral, os exemplos e exercícios desta seção envolvem apenas testes bilaterais; mas podem ocorrer testes unilaterais no caso de afirmações sobre correlação linear positiva ou correlação linear negativa. Em tais casos, as hipóteses serão conforme se segue.

Afirmação de Correlação Negativa (Teste unilateral esquerdo)	Afirmação de Correlação Positiva (Teste unilateral direito)
H_0: $\rho \geq 0$	H_0: $\rho \leq 0$
H_1: $\rho < 0$	H_1: $\rho > 0$

Para esses testes unilaterais, o Método 1 pode ser aplicado como nos capítulos anteriores. Para o Método 2, ou calculamos o valor crítico, conforme indicado no Exercício 25, ou modificamos a Tabela A-6 substituindo os cabeçalhos das colunas $\alpha = 0{,}05$ e $\alpha = 0{,}01$ pelos valores críticos unilaterais $\alpha = 0{,}025$ e $\alpha = 0{,}005$, respectivamente.

Rationale: Apresentamos a Fórmula 9-1 para o cálculo de r e ilustramos sua aplicação; vamos agora justificá-la. A Fórmula 9-1 simplifica os cálculos nesta fórmula equivalente:

$$r = \frac{\Sigma(x - \bar{x})(y - \bar{y})}{(n - 1)s_x s_y}$$

Utilizaremos provisoriamente esta última versão da Fórmula 9-1, porque sua forma está relacionada mais diretamente com a teoria básica. Consideraremos os seguintes dados emparelhados, representados no diagrama de dispersão da Figura 9-6.

x	1	1	2	4	7
y	4	5	8	15	23

A Figura 9-6 exibe o ponto $(\bar{x}, \bar{y}) = (3, 11)$, chamado centróide dos pontos amostrais.

Leitura da Mão

Algumas pessoas acreditam que o comprimento da linha da vida de sua mão pode ser usado para predizer a longevidade. Em uma carta publicada no *Journal of the American Medical Association*, os autores M. E. Wilson e L. E. Mather refutam esta crença com o estudo de cadáveres. Registraram-se as idades na morte juntamente com os comprimentos das linhas de vida. Os autores concluíram que não há correlação significativa entre a idade na ocasião do falecimento e o comprimento da linha da vida.

DEFINIÇÃO

Dada uma coleção de dados emparelhados (x, y), o ponto $(\bar{x}, \bar{y})$ é chamado **centróide**.

Foi Karl Pearson quem primeiro elaborou a estatística r, às vezes chamada momento-produto de Pearson. Baseia-se no produto dos momentos $(x - \bar{x})$ e $(y - \bar{y})$; isto é, Pearson baseou sua medida de dispersão na estatística $\Sigma(x - \bar{x})(y - \bar{y})$. Em qualquer diagrama de dispersão, retas horizontal e vertical pelo centróide $(\bar{x}, \bar{y})$ dividem o diagrama em quatro quadrantes,

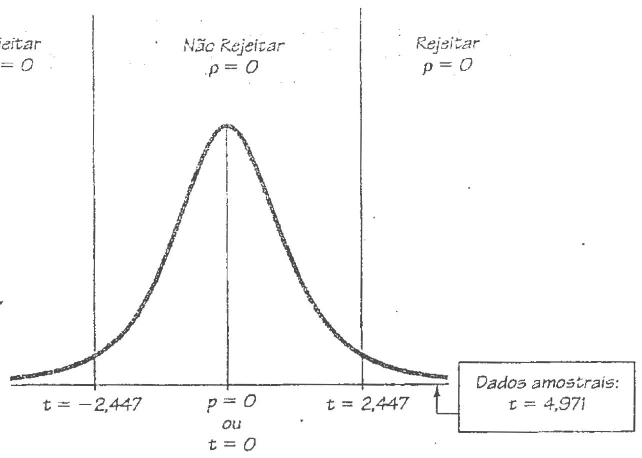


Fig. 9-4 Teste de H_0: $\rho = 0$ pelo método 1.

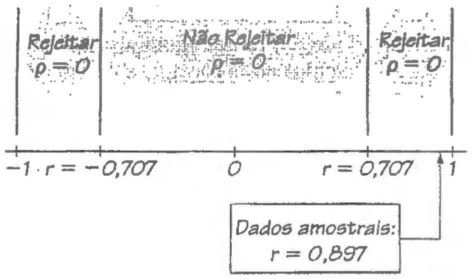


Fig. 9-5 Teste de H_0: $\rho = 0$ pelo método 2.

conforme a Figura 9-6. Se os pontos do diagrama de dispersão tendem para uma linha ascendente (como na figura), os valores individuais do produto $(x - \bar{x})(y - \bar{y})$ tendem a ser positivos, porque a maioria dos pontos está no primeiro e no terceiro quadrantes, onde os produtos de $(x - \bar{x})$ e $(y - \bar{y})$ são positivos. Se os pontos do diagrama de dispersão tendem para uma linha descendente, a maioria dos pontos está no segundo e no quarto quadrantes, onde $(x - \bar{x})$ e $(y - \bar{y})$ têm sinais opostos, de modo que $\Sigma(x - \bar{x})(y - \bar{y})$ é negativo. Os pontos que não seguem qualquer padrão linear tendem a dispersar-se pelos quatro quadrantes, e o valor de $\Sigma(x - \bar{x})(y - \bar{y})$ tende a ficar próximo de 0.

A soma $\Sigma(x - \bar{x})(y - \bar{y})$ depende da magnitude dos números usados. Por exemplo, transformando x de polegadas para pés, aquela soma se modificará. Para tornar r independente da escala particular usada, adotamos o desvio-padrão amostral seguinte:

$$r = \frac{\Sigma(x - \bar{x})(y - \bar{y})}{(n-1)s_x s_y}$$

Esta expressão pode ser transformada algebricamente para a forma da Fórmula 9-1.

Avaliação dos Professores pelos Alunos

Muitos estabelecimentos de ensino associam avaliação elevada dos alunos com ensino eficiente — uma equação freqüentemente incentivada pelo fato de que as avaliações dos estudantes são fáceis de administrar e medir.

Todavia, um estudo que comparou as avaliações dos professores, feitas por alunos, com a quantidade de matéria assimilada, revelou forte correlação *negativa* entre os dois fatores. Os professores mais bem avaliados pelos alunos pareciam induzir menor aprendizado.

Em um estudo relacionado, uma classe atribuiu elevada avaliação a um professor que dava muito pouca matéria, mas era interessante e divertido.

Nos capítulos precedentes estudamos métodos de inferência estatística apelando não só para métodos de teste de hipóteses como para métodos de construção de intervalos de confiança. Processo análogo pode ser aplicado para determinar intervalos de confiança para ρ. Todavia, como a construção de tais intervalos de confiança envolve transformações algo complicadas, este processo é apresentado no Exercício 28 (Além do Básico).

Podemos utilizar o coeficiente de correlação linear para determinar se há algum relacionamento linear entre duas variáveis. Com os dados da Tabela 9-1, concluímos que há uma correlação linear entre comprimentos e pesos de ursos. Tendo concluído pela existência de um relacionamento, interessa agora determinar qual é esse relacionamento, de modo que possamos calcular o peso de um urso, quando conhecemos seu comprimento. Na próxima seção abordamos este estágio da análise.

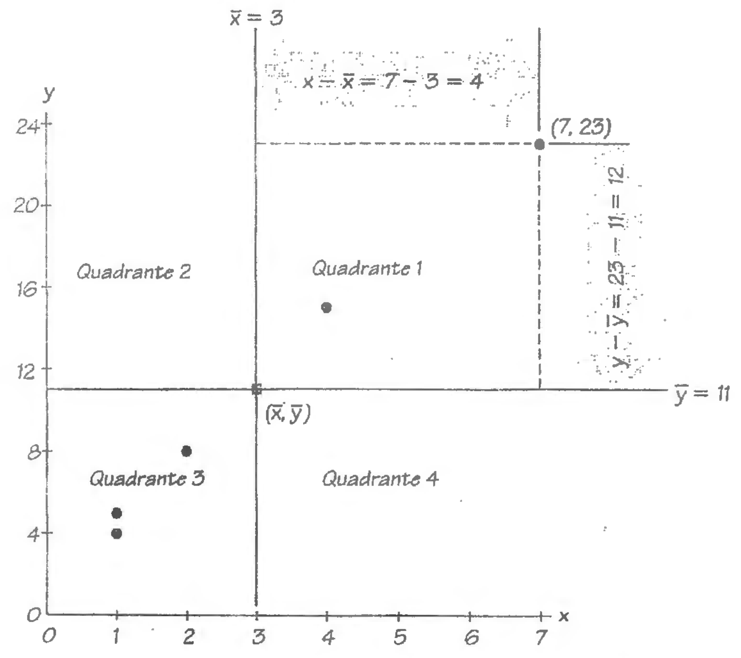


Fig. 9-6 Diagrama de dispersão particionado em quadrantes.

9-2 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1 e 2, suponha que uma amostra de n *pares de dados origine o valor dado de* r. *Consultando a Tabela A-6 (para* $\alpha = 0{,}05$*), determine se há correlação linear significativa entre* x *e* y. *(Veja o critério de decisão dado nesta seção.)*

1. a. $n = 32, r = 0{,}992$
   b. $n = 50, r = -0{,}333$
   c. $n = 17, r = 0{,}456$
2. a. $n = 22, r = -0{,}087$
   b. $n = 40, r = 0{,}299$
   c. $n = 25, r = -0{,}401$

Nos Exercícios 3 e 4, (a) com base no diagrama de dispersão, determine se há correlação linear significativa entre x *e* y, *e (b) ache os valores de* n, Σx, Σx^2, $(\Sigma x)^2$, Σxy *e o coeficiente de correlação linear* r.

3.

x	2	3	5	5	10
y	6	9	14	16	30

4.

x	2	3	5	5	10
y	6	0	15	5	2

Nos Exercícios 5-18,

a. Construa o diagrama de dispersão.
b. Ache o valor do coeficiente de correlação linear r.
c. Determine se há correlação linear significativa entre as duas variáveis. (Utilize somente $\alpha = 0{,}05$.*)*
d. Guarde seus resultados, pois na próxima seção utilizaremos estes mesmos dados.

5. Quando os ursos foram anestesiados, os pesquisadores mediram o perímetro (em polegadas) de seus tóraxes e obtiveram seus pesos (em libras). A seguir são apresentados os resultados para oito ursos machos. Com base nesses resultados, parece haver relação entre o peso e o perímetro do tórax dos ursos? Os resultados se modificarão se as medidas forem convertidas para pés (cada valor sendo dividido por 12)?

x Tórax (in.)	26	45	54	49	41	49	44	19
y Peso (lb)	90	344	416	348	262	360	332	34

Com base em dados de Minitab e Gary Alt.

6. A tabela a seguir relaciona os pesos (em centenas de libras) e as taxas de consumo de combustível em rodovia (em mi/gal) para uma amostra de carros de passeio novos. Com base nos resultados, espera um maior consumo de combustível se adquirir um carro mais pesado? Os resultados se modificam se os pesos forem dados como 2900, 3500, ..., 2400?

x Peso	29	35	28	44	25	34	30	33	28	24
y Combustível	31	27	29	25	31	29	28	28	28	33

Com base em dados da EPA.

7. A tabela a seguir dá os pesos (em libras) do plástico descartado por uma amostra de residências, juntamente com o tamanho destas. Há alguma correlação linear significativa? Este problema é importante para o Departamento do Censo, que financia projetos, porque a presença de uma correlação implica que podemos predizer o tamanho da população analisando o lixo descartado.

Plástico (lb)	0,27	1,41	2,19	2,83	2,19	1,81	0,85	3,05
Tamanho da Residência	2	3	3	6	4	2	1	5

Com base em dados fornecidos por Masakazu Tani e pelo Projeto do Lixo na Universidade do Arizona.

8. Os dados emparelhados a seguir consistem em pesos (em libras) de papel descartado e tamanhos de residências.

Papel	2,41	7,57	9,55	8,82	8,72	6,96	6,83	11,42
Tamanho da Residência	2	3	3	6	4	2	1	5

Com base em dados do Projeto do Lixo da Universidade do Arizona.

9. Os dados emparelhados a seguir consistem em pesos (em libras) de restos de comida e tamanhos de residências.

Comida	1,04	3,68	4,43	2,98	6,30	1,46	8,82	9,62
Tamanho da Residência	2	3	3	6	4	2	1	5

Com base em dados obtidos do Projeto do Lixo da Universidade do Arizona.

10. Os dados emparelhados a seguir consistem em pesos totais (libras) de lixo descartado e tamanhos de residências.

Peso Total	10,76	19,96	27,60	38,11	27,90	21,90	21,83	49,27	33,27	35,54
Tamanho da Residência	2	3	3	6	4	2	1	5	6	4

Com base em dados obtidos do Projeto do Lixo da Universidade do Arizona.

11. Fez-se um estudo para investigar a relação entre idade (em anos) e a CAS (Concentração de Álcool no Sangue) medida quando os motoristas intoxicados condenados foram presos pela primeira vez. Seguem dados amostrais para indivíduos selecionados aleatoriamente. Com base no resultado, parece haver relação entre o nível de CAS e a idade da pessoa testada?

Idade	17,2	43,5	30,7	53,1	37,2	21,0	27,6	46,3
CAS	0,19	0,20	0,26	0,16	0,24	0,20	0,18	0,23

Com base em dados do Programa STOP-DWI do Condado de Dutchess.

12. A tabela a seguir dá o número (em milhares) de armas automáticas registradas, juntamente com a taxa de criminalidade (em crimes por 100.000), para estados selecionados aleatoriamente. Consideram-se automáticas as armas que continuam disparando enquanto o gatilho está acionado. Os crimes com armas de fogo parecem estar relacionados com as armas automáticas? Uma correlação linear significativa implica que o aumento do número de armas automáticas resulta em maior número de crimes?

Armas automáticas	11,6	8,3	3,6	0,6	6,9	2,5	2,4	2,6
Taxa de criminalidade	13,1	10,6	10,1	4,4	11,5	6,6	3,6	5,3

Dados fornecidos pelo FBI e pelo Bureau of Alcohol, Tobacco and Firearms.

13. Use os dados emparelhados do Conjunto de Dados 16 do Apêndice B relativos às durações e intervalos das erupções do gêiser Old Faithful. Há alguma correlação linear significativa que sugira que o intervalo após uma erupção esteja relacionado com a duração da mesma?
14. Use os dados emparelhados do Conjunto de Dados 16 do Apêndice B relativos aos intervalos após erupções e alturas das erupções do gêiser Old Faithful. Há alguma correlação linear significativa que sugira que o intervalo após uma erupção esteja relacionado com a altura da mesma?
15. No Conjunto de Dados 4 do Apêndice B, considere os dados emparelhados que consistem em quantidades de alcatrão e nicotina. Com base no resultado, parece haver correlação linear significativa entre os conteúdos de alcatrão e de nicotina nos cigarros? Em caso afirmativo, os pesquisadores podem reduzir sua despesa de laboratório testando apenas uma dessas duas variáveis?
16. No Conjunto 4 do Apêndice B considere os dados emparelhados referentes às quantidades de monóxido de carbono e nicotina. Com base no resultado, parece haver correlação linear significativa entre os conteúdos de nicotina e de monóxido de carbono nos cigarros?
17. Considere os dados emparelhados do Conjunto de Dados 7 do Apêndice B que se referem à quantidade de precipitação pluviométrica total anual (PRECIP) e à quantidade de milho produzido (CORNPROD). Há correlação linear significativa (como seria de se esperar)?

18. Considere os dados emparelhados do Conjunto 7 do Apêndice B que se referem às temperaturas médias anuais em Iowa (AVTEMP) e às quantidades de milho produzido (CORNPROD). Há correlação linear significativa (como seria de se esperar)?

Nos Exercícios 19-22, indique o erro na conclusão. (Veja a lista de erros comuns incluída nesta seção.)

19. *Dado:* Os dados amostrais emparelhados das idades de indivíduos e seus resultados em um teste de raciocínio acusaram um coeficiente de correlação linear muito próximo de 0.
Conclusão: Os mais jovens tendem a obter resultados melhores.
20. *Dado:* Há uma correlação linear significativa entre a renda pessoal e o número de anos de escolaridade.
Conclusão: Mais instrução tem como resultado maior renda pessoal.
21. *Dado:* Indivíduos fazem um teste de habilidade verbal e um teste de destreza manual; os pares de observação acusam um coeficiente de correlação linear muito próximo de 0.
Conclusão: Não há qualquer relacionamento entre os escores dos dois testes.
22. *Dado:* Há correlação linear significativa entre a carga fiscal estadual e a renda média dos estados.
Conclusão: Há correlação linear significativa entre encargos individuais e rendas individuais.

9-2 Exercícios B: Além do Básico

23. Como é afetado o coeficiente de correlação linear r em cada um dos seguintes casos?
   a. Cada valor x é permutado com o valor y correspondente.
   b. Cada valor x é multiplicado pela mesma constante diferente de zero.
   c. Adiciona-se a mesma constante a cada valor x.

24. Além de testar a correlação linear entre x e y, podemos também utilizar *transformações* de dados para explorar outras relações. Por exemplo, podemos substituir cada valor x por x^2 e, com os métodos desta seção, determinar se há correlação linear entre y e x^2. Com os dados emparelhados da tabela a seguir, construa o diagrama de dispersão e teste se há correlação linear entre y e cada um dos valores seguintes. Qual caso dá maior valor de r?

a. x b. x^2 c. $\log x$ d. $\sqrt{x}$ e. $1/x$

x	1,3	2,4	2,6	2,8	2,4	3,0	4,1
y	0,11	0,38	0,41	0,45	0,39	0,48	0,61

25. Para achar os valores críticos de r na Tabela A-6, resolve-se

$$t = \frac{r}{\sqrt{\dfrac{1-r^2}{n-2}}}$$

em relação a r, obtendo-se

$$r = \frac{t}{\sqrt{t^2 + n - 2}}$$

onde o valor de t é dado pela Tabela A-3 supondo-se um caso bilateral com $n - 2$ graus de liberdade. A Tabela A-6 dá os resultados para valores selecionados de n e α. Aplique a fórmula de r dada aqui e a Tabela A-3 (com $n - 2$ graus de liberdade) a fim de achar os valores críticos de r para os casos dados.

   a. H_0: $\rho = 0$, $n = 50$, $\alpha = 0{,}05$
   b. H_1: $\rho \neq 0$, $n = 75$, $\alpha = 0{,}10$
   c. H_0: $\rho \geq 0$, $n = 20$, $\alpha = 0{,}05$
   d. H_0: $\rho \leq 0$, $n = 10$, $\alpha = 0{,}05$
   e. H_1: $\rho > 0$, $n = 12$, $\alpha = 0{,}01$

26. Utilize os dados amostrais do Exercício 5, mas teste a afirmação de que há correlação linear *positiva* entre os perímetros torácicos e os pesos de ursos. (Veja Exercício 25.)
27. O gráfico de $y = x^2$ é uma parábola, e não uma reta, e assim podemos esperar que o valor de r não reflita uma correlação linear entre x e y. Tomando $y = x^2$, faça uma tabela de valores de x e y para $x = 0, 1, 2, ..., 10$ e calcule o valor de r. Qual é sua conclusão? Como explica o resultado?
28. Dados n pares de dados para os quais se pode determinar o coeficiente de correlação r, aplique o processo seguinte para construir um intervalo de confiança para o parâmetro populacional ρ.

Passo a. Na Tabela A-2, localize o valor de $z_{\alpha/2}$ que corresponde ao grau de confiança desejado.

Passo b. Calcule os limites do intervalo w_E e w_D.

$$w_E = \frac{1}{2}\ln\left(\frac{1+r}{1-r}\right) - z_{\alpha/2}\cdot\frac{1}{\sqrt{n-3}}$$

$$w_D = \frac{1}{2}\ln\left(\frac{1+r}{1-r}\right) - z_{\alpha/2}\cdot\frac{1}{\sqrt{n-3}}$$

Passo c. Calcule agora os limites do intervalo de confiança na expressão a seguir:

$$\frac{e^{2w_E}-1}{e^{2w_E}-1} < \rho < \frac{e^{2w_D}-1}{e^{2w_D}-1}$$

Aplique este processo para construir um intervalo de confiança de 95% para ρ, dados 50 pares de valores para os quais $r = 0{,}600$.

9-3 Regressão

Na Seção 9-2 analisamos dados emparelhados com o objetivo de determinar se havia correlação linear significativa entre duas variáveis. Vamos agora descrever a relação traçando o gráfico e determinando a equação da reta que representa aquela relação. Essa reta é chamada reta de regressão, e sua equação é a equação de regressão. Sir Francis Galton (1822-1911) estudou o fenômeno da hereditariedade e mostrou que, quando casais altos ou baixos têm filhos, as alturas destes tendem a *regredir*, ou reverter para uma altura média mais típica. Continuaremos utilizando esta terminologia.

REGRESSAR

DEFINIÇÕES

Dada uma coleção de dados amostrais emparelhados, a **equação de regressão**

$$\hat{y} = b_0 + b_1 x$$

descreve a relação entre as duas variáveis. O gráfico da equação de regressão é chamado **reta de regressão** (ou *reta de melhor ajuste*, ou *reta de mínimos quadrados*).

Esta definição expressa uma relação entre x (chamada **variável independente** ou **variável preditora**) e $\hat{y}$ (chamada **variável dependente** ou **variável resposta**). Na definição precedente, a equação típica de uma reta ($y = mx + b$) é expressa na forma $\hat{y} = b_0 + b_1 x$, onde b_0 é o intercepto y e b_1 é o coeficiente angular. O quadro a seguir mostra que b_0 e b_1 são estatísticas amostrais usadas para estimar os parâmetros populacionais β_0 e β_1.

Notação para a Equação de Regressão

	Parâmetro Populacional	Estatística Amostral
Intercepto y da equação de regressão	β_0	b_0
Coeficiente angular da equação de regressão	β_1	b_1
Equação da reta de regressão	$y = \beta_0 + \beta_1 x$	$\hat{y} = b_0 + b_1 x$

Suposições: Para os métodos de regressão dados nesta seção, supomos que

1. Estamos investigando apenas relações lineares.
2. Para cada valor x, y é uma variável aleatória com distribuição normal (em forma de sino). Todas essas distribuições de y têm a mesma variância e, ainda, para um dado valor de x, a média da distribuição dos valores y está sobre a reta de regressão. (Os resultados não são afetados seriamente se os desvios da normalidade e da igualdade de variâncias não são grandes.)

Um importante objetivo desta seção é utilizar dados amostrais emparelhados para estimar a equação de regressão. Dispondo apenas dos dados amostrais, não podemos achar os valores exatos dos parâmetros populacionais β_0 e β_1, mas, com os dados amostrais, podemos estimá-los com b_0 e b_1, que se obtêm com as Fórmulas 9-2 e 9-3.

Fórmula 9-2 $$b_0 = \frac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2} \quad \text{intercepto-}y$$

Fórmula 9-3 $$b_1 = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2} \quad \text{(coeficiente angular)}$$

Estas fórmulas podem parecer assustadoras, mas estão programadas em muitas calculadoras, e assim podemos achar facilmente os valores de b_0 e b_1. Podemos utilizar também pacotes estatísticos como o STATDISK e o Minitab para achar b_0 e b_1. Nos casos em que devemos usar fórmulas em lugar de uma calculadora ou um computador, os cálculos necessários serão bastante facilitados se tivermos em mente as seguintes observações:

1. Se o coeficiente de correlação linear r já foi calculado pela Fórmula 9-1, então os valores de Σx, Σy, Σx^2 e Σxy também foram calculados e podem ser novamente utilizados na Fórmula 9-3. (Também, o numerador de r na Fórmula 9-1 é o mesmo numerador de b_1 na Fórmula 9-3; o denominador de r inclui o denominador de b_1. Se o cálculo de r foi esquematizado cuidadosamente, o cálculo de b_1 exige apenas a divisão de um número conhecido por outro.)
2. Se calcularmos primeiro o coeficiente angular b_1, poderemos utilizar a Fórmula 9-4 para achar o intercepto y b_0. [A reta de regressão passa sempre pelo centróide $(\bar{x}, \bar{y})$, de forma que a equação $\bar{y} = b_0 + b_1\bar{x}$ deve ser verdadeira, e esta equação pode expressar-se como na Fórmula 9-4. Em geral é mais fácil achar o intercepto y b_0 utilizando a Fórmula 9-4 do que com a Fórmula 9-2.]

Fórmula 9-4 $$b_0 = \bar{y} + b_1\bar{x}$$

Uma vez calculados b_0 e b_1, podemos identificar a equação estimada de regressão, que apresenta a seguinte propriedade especial: *A reta de regressão é a que melhor se ajusta aos pontos amostrais.* (O critério específico usado para determinar que reta se ajusta "melhor" é a propriedade dos mínimos quadrados, que será apresentada mais adiante.) Por ora discutiremos rapidamente o arredondamento e, a seguir, ilustraremos o processo para determinar e aplicar a equação de regressão.

Arredondamento do Intercepto y b_0 e do Coeficiente Angular b_1

É difícil dar uma regra simples, universal, para o arredondamento dos valores de b_0 e b_1, mas, de modo geral, sempre procuraremos arredondar cada um desses valores para *três algarismos significativos*, ou então utilizar os valores dados por um programa como STATDISK ou Minitab. Como esses valores são muito sensíveis a arredondamentos feitos em estágios intermediários dos cálculos, devemos procurar conservar ao menos seis algarismos significativos (ou usar valores exatos) nos estágios intermediários. Dependendo do vulto do arredondamento, as respostas dos exemplos e exercícios deste livro podem divergir ligeiramente dos resultados obtidos pelo leitor.

EXEMPLO Na Seção 9-2 utilizamos os dados da Tabela 9-1 (x = comprimentos de ursos; y = pesos de ursos) para achar o coeficiente de correlação linear $r = 0{,}897$, que indica uma correlação linear significativa. Vamos agora determinar a equação de regressão da reta que relaciona x e y.

SOLUÇÃO Encontraremos a equação de regressão usando as Fórmulas 9-3 e 9-4 e os valores já encontrados na Tabela 9-2 da Seção 9-2:

$$n = 8 \qquad \Sigma x = 516{,}5 \qquad \Sigma y = 2176$$
$$\Sigma x^2 = 34.525{,}75 \qquad \Sigma y^2 = 728.520 \qquad \Sigma xy = 151.879$$

Inicialmente, determinamos o coeficiente angular b_1 aplicando a Fórmula 9-3:

$$b_1 = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2}$$
$$= \frac{8(151.879) - (516{,}5)(2176)}{8(34.525{,}75) - (516{,}5)^2} = \frac{91.128}{9433{,}75} = 9{,}65979$$
$$= 9{,}66 \quad \text{(arredondado)}$$

Em seguida, determinamos o intercepto y b_0 pela Fórmula 9-4:

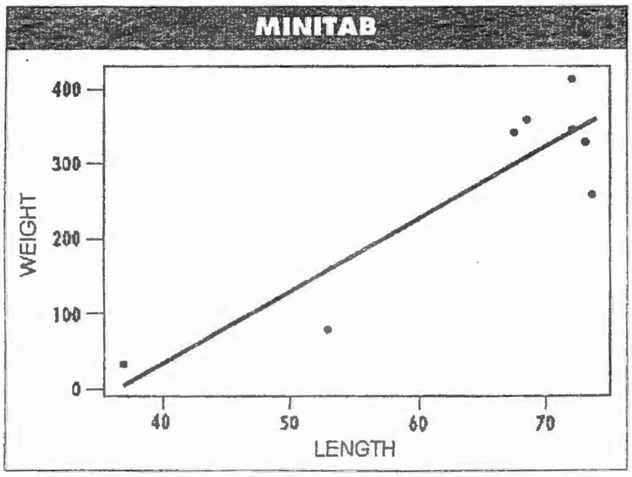
$$b_0 = \bar{y} - b_1\bar{x}$$
$$= \frac{2176}{8} - (9{,}65979)\frac{516{,}5}{8} = -352 \quad \text{(arredondado)}$$

Conhecidos o coeficiente angular b_1 e o intercepto y b_0, podemos expressar a equação estimada da reta de regressão como

$$\hat{y} = -352 + 9{,}66x$$

Devemos ter em mente que esta equação é uma *estimativa* da verdadeira equação de regressão $y = \beta_0 + \beta_1 x$. Esta estimativa se baseia em um conjunto particular de dados amostrais relacionados na Tabela 9-1; outra amostra extraída da mesma população provavelmente levaria a uma equação ligeiramente diferente.

A seguir mostra-se o gráfico da reta de regressão no diagrama de dispersão, obtido com o Minitab. Pode-se ver que essa reta se ajusta bem aos dados.



Variação Marginal

Com a equação de regressão, podemos ver o efeito sobre uma das variáveis, quando a outra sofre uma variação.

DEFINIÇÃO

Ao trabalharmos com duas variáveis relacionadas por uma equação de regressão, a **variação marginal** em uma delas é o quanto ela varia quando a outra variável sofre uma variação de exatamente uma unidade.

O coeficiente angular b_1 na equação de regressão representa a variação marginal resultante quando x varia de uma unidade. Para os dados dos ursos da Tabela 9-1, vemos que um aumento de uma unidade em x causa uma variação de 9,66 unidades em $\hat{y}$. Ou seja, se um urso cresce 1 polegada, seu peso predito terá um acréscimo de 9,66 libras.

Pontos Extremos (*Outliers*) e Pontos de Influência

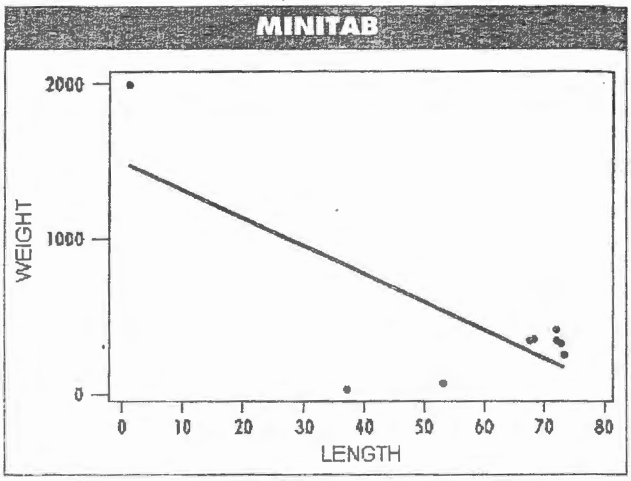
Uma análise de correlação/regressão de dados bivariados deve incluir uma investigação de *pontos extremos* (*outliers*) e *pontos de influência*, definidos como se segue.

DEFINIÇÕES

Em um diagrama de dispersão, um **ponto extremo** (*outlier*) é um ponto que está muito afastado dos demais pontos. Os dados amostrais emparelhados podem conter um ou mais **pontos de influência**, que são pontos que afetam fortemente o gráfico da reta de regressão.

Eis como determinar se um ponto é ponto de influência: Trace a reta de regressão correspondente aos dados com o ponto incluído e, em seguida, trace a reta de regressão excluindo aquele ponto. Se o gráfico varia consideravelmente, o ponto é um ponto de influência. Em geral, determinam-se os pontos de influência identificando os *outliers* que estão *horizontalmente* distantes dos outros pontos.

Veja, por exemplo, a reta obtida com o Minitab. Se incluirmos mais um urso com 35 in. de comprimento e 400 lb de peso, teremos um ponto extremo, porque o ponto correspondente ficaria no canto superior esquerdo do gráfico e estaria muito distante dos outros pontos. Este ponto não seria um ponto de influência porque não acarretaria grande modificação na reta de regressão. Já se o urso tivesse 1 in. de comprimento e pesasse 2000 lb (situação muito estranha), ter-se-ia um ponto de influência porque o gráfico da reta de regressão se modificaria consideravelmente, conforme ilustrado a seguir.



O Ozônio em Los Angeles

O South Coast Air Quality Management District monitora os níveis de ozônio para a região de Los Angeles. Os níveis de ozônio são afetados não só pelo tempo como pelo terrível fluxo de tráfego na área de Los Angeles. Um indicador adequado de um problema de ozônio é seu nível em partes por milhão para a pior hora do ano. A análise de regressão mostra uma tendência descendente naquele indicador, o que sugere que, a despeito do grande aumento de habitantes e de carros, os níveis de ozônio estão efetivamente em cerca da metade do que eram há 40 anos. Pode-se constatar a tendência descendente nos níveis da "pior hora" para seis anos consecutivos recentes: 0,32, 0,32, 0,25, 0,21, 0,22, 0,22. Essas análises estatísticas não só evidenciam o impacto de uma legislação antipoluição, como também suscitam problemas de custo-benefício da nova legislação.

Predições

As equações de regressão podem ser úteis quando usadas para *predizer* o valor de uma variável, dado um valor determinado da outra variável. Se a reta de regressão se ajusta bem aos dados, então tem sentido utilizar sua equação para fazer predições, desde que não ultrapassemos os limites dos valores disponíveis. Entretanto, *só devemos utilizar a equação da reta de regressão se r indica a existência de uma correlação linear significativa. Na ausência de uma tal correlação linear, não devemos utilizar a equação de regressão para projetar ou predizer; em vez disso, nossa melhor estimativa da segunda variável é simplesmente a sua média.*

Ao predizer um valor de y com base em determinado valor de x,

1. Se *não* há correlação linear significativa, o melhor valor predito de y é $\bar{y}$.
2. Se *há* correlação linear significativa, obtém-se o melhor valor predito de y substituindo-se o valor de x na equação de regressão.

A Figura 9-7 resume este processo, que é mais fácil de entender se encararmos r como uma medida de quão bem a reta de regressão se ajusta aos dados amostrais. Se r está próximo de -1 ou $+1$, então há um bom ajuste da reta de regressão aos dados, mas se r é vizinho de 0, o ajuste é fraco (e não deve ser usado para predições).

EXEMPLO Com os dados amostrais da Tabela 9-1, constatamos que há correlação linear significativa entre comprimentos e pesos de ursos, e obtivemos a equação de regressão $\hat{y} = -352 + 9{,}66x$. Se um urso tem comprimento de 71,0 in., vamos predizer seu peso.

SOLUÇÃO Sentimo-nos fortemente tentados a substituir diretamente x por 71,0 na equação de regressão, mas antes devemos considerar se há uma correlação linear significativa que justifique o emprego daquela equação. Neste exemplo temos realmente uma correlação linear significativa (com $r = 0{,}897$) e assim obtemos nosso valor predito como se segue:

$$\begin{aligned}\hat{y} &= -352 + 9{,}66x\\ &= -352 + 9{,}66(71{,}0) = 334\end{aligned}$$

O peso previsto para um urso de 71,0 in. de comprimento é 334 lb. (Se não houvesse uma correlação linear significativa, nossa melhor predição do peso seria $\bar{y} = 272$ lb.)

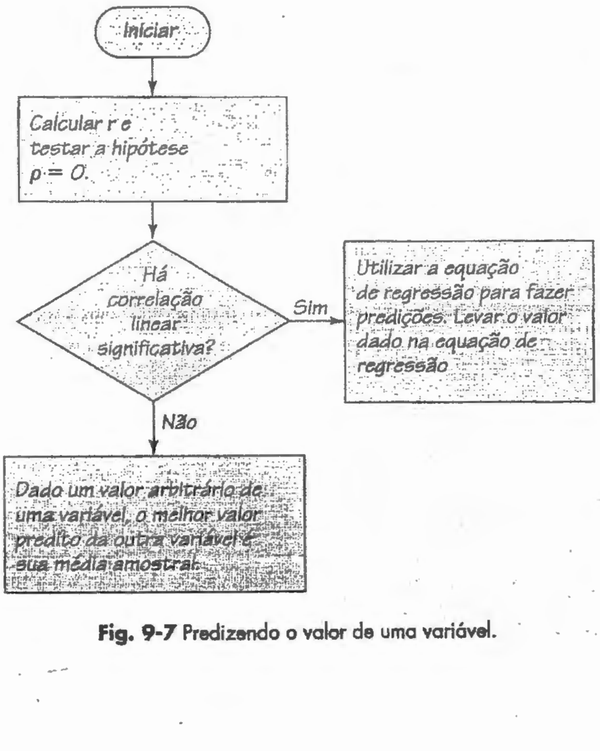


Fig. 9-7 Predizendo o valor de uma variável.

EXEMPLO Obviamente, não existe correlação linear significativa entre números de sapatos e QI de adultos. Dado que um adulto calça sapato número 9, determine a melhor predição de seu QI.

SOLUÇÃO Como não há correlação linear significativa, não utilizamos uma equação de regressão. A melhor predição do valor do QI é simplesmente o QI médio, que é 100.

Compare cuidadosamente as soluções dos dois exemplos precedentes e note que aplicamos a equação de regressão quando havia correlação linear significativa; na ausência de tal correlação, entretanto, a melhor predição do valor de y é simplesmente o valor da média amostral $\bar{y}$. Um erro comum consiste em aplicar a equação de regressão quando não há correlação linear significativa. Esse erro viola a primeira das diretrizes seguintes.

Diretrizes para Uso da Equação de Regressão

1. *Se não há correlação linear significativa, não use a equação de regressão para fazer predições.*
2. *Ao aplicar a equação de regressão para predições, mantenha-se dentro do âmbito dos dados amostrais.* Se acharmos uma equação de regressão relacionando as alturas das mulheres com os números de seus sapatos, é absurdo predizer o número do sapato de uma mulher que tenha 10 pés de altura.
3. *Uma equação de regressão baseada em dados passados não é necessariamente válida hoje.* A equação de regressão que relaciona preços de carros usados e idades de carros não é mais válida, se se baseia em dados da década de 1970.
4. *Não devemos fazer predições sobre uma população diferente daquela de onde provêm os dados amostrais.* Se coletamos dados amostrais sobre homens e estabelecemos uma equação de regressão relacionando idade e uso de controle remoto de TV, os resultados não se aplicam necessariamente às mulheres. Se utilizamos *médias* oficiais para estabelecer uma equação de regressão relacionando notas de matemática em SAT e escores verbais em SAT, os resultados não se aplicam necessariamente a *indivíduos*

Utilização de Computadores e Calculadoras na Correlação e Regressão

Em vista do volume de cálculos em jogo, em geral determinam-se o coeficiente angular e o intercepto y da reta de regressão com auxílio de programas de calculadora ou computador. Na Seção 9-2 descrevemos os processos para obter o valor do coeficiente de correlação linear r utilizando o STATDISK e o Minitab e uma calculadora TI-83. Na Seção 9-2 incluímos também um diagrama de dispersão para os dados da Tabela 9-1, utilizando o Minitab.

Para obter a equação de regressão com uma calculadora TI-83, introduzimos os dados emparelhados nas listas L1 e L2, acionamos STAT e selecionamos `TESTS`; escolhemos em seguida a opção `LinRegTTest` e os resultados apresentados incluirão o intercepto y e o coeficiente angular da equação de regressão. Em lugar de b_0 e b_1, a TI-83 apresenta esses valores como a e b.

Para obter a equação de regressão com o STATDISK, utilizamos as opções `Analysis` e `Correlation and Regression`. O STATDISK incluirá os resultados de correlação e regressão, conforme mostrado a seguir.

STATDISK DISPLAY

File Edit Analysis Data Help

Sample Size, n	8
Degrees Freedom	6
Correlation Results:	
Correlation Coeff, r	0.89735
Critical r	±0.70673
Reject the Null Hypothesis	
Sample provides evidence that the populations are correlated	
Regression Results:	
$Y = b_0 + b_1x$	
Y Intercept, b_0	-351.66
Slope, b_1	9.6598
Total Variation	136648
Explained Var	110035
Unexplained Var	26613
Standard Error	66.600
Coeff of Det, r^2	0.80524

Para obter a equação de regressão com o Minitab, introduzimos inicialmente os dados emparelhados nas colunas C1 e C2, em seguida selecionamos `Stat`, logo após `Regression`, e novamente `Regression`. Introduzimos agora C2 para a variável resposta e C1 para a variável preditora. Se utilizarmos os dados da Tabela 9-1, os resultados do Minitab começarão com:

```
The regression equation is
C2 = -352 + 9,66 C1
```

O Minitab inclui também outros resultados, alguns dos quais serão abordados na seção seguinte.

Resíduos e a Propriedade de Mínimos Quadrados

Já dissemos que a equação de regressão representa a reta que "melhor" se ajusta aos dados; daremos agora o critério para determinar a reta que é melhor do que todas as outras. Esse critério se baseia na distância vertical entre os pontos que representam os dados originais e a reta de regressão. Tais distâncias chamam-se *resíduos*.

DEFINIÇÃO

Dado um par de dados amostrais (x, y), um **resíduo** é a diferença ($y - \hat{y}$) entre um valor amostral observado y e o valor $\hat{y}$ predito com base na equação de regressão.

Esta definição pode parecer tão clara quanto as instruções de um formulário de imposto de renda, mas podemos entender facilmente os resíduos observando a Figura 9-8, que corresponde aos dados amostrais emparelhados a seguir.

x	1	2	4	5
y	4	24	8	32

Na Figura 9-8, os resíduos são representados por linhas tracejadas. Como exemplo específico, veja o resíduo indicado como 7, que está diretamente acima de $x = 5$. Levando o valor $x = 5$ na equação de regressão $\hat{y} = 5 + 4x$, obtemos o valor predito $\hat{y} = 25$. Quando $x = 5$, o valor *predito* de y é $\hat{y} = 25$, mas o valor amostral efetivamente *observado* é $y = 32$. A diferença $y - \hat{y} = 32 - 25 = 7$ é um resíduo.

A equação de regressão representa a reta que "melhor" se ajusta aos pontos de acordo com a *propriedade dos mínimos quadrados*.

DEFINIÇÃO

Uma reta verifica a **propriedade dos mínimos quadrados** se a soma dos quadrados dos resíduos é a menor possível.

Pela Figura 9-8, vemos que os resíduos são -5, 11, -13 e 7, e, assim, a soma de seus quadrados é

$$(-5)^2 + 11^2 + (-13)^2 + 7^2 = 364$$

Qualquer outra reta distinta de $\hat{y} = 5 + 4x$ dará resíduos cuja soma dos quadrados é maior do que 364.

Felizmente, não precisamos lidar diretamente com a propriedade de mínimos quadrados para achar a equação da reta de regressão. Utilizou-se o cálculo para condensar a propriedade de mínimos quadrados nas Fórmulas 9-2 e 9-3. Como a dedução dessas fórmulas exige o cálculo, não as incluiremos neste livro.

Transformações

Em muitos casos, há entre duas variáveis uma relação que não é linear. Por exemplo, a simples observação da tabela a seguir mostra que cada valor de y é o quadrado do valor de x correspondente, de forma que as duas variáveis estão relacionadas pela

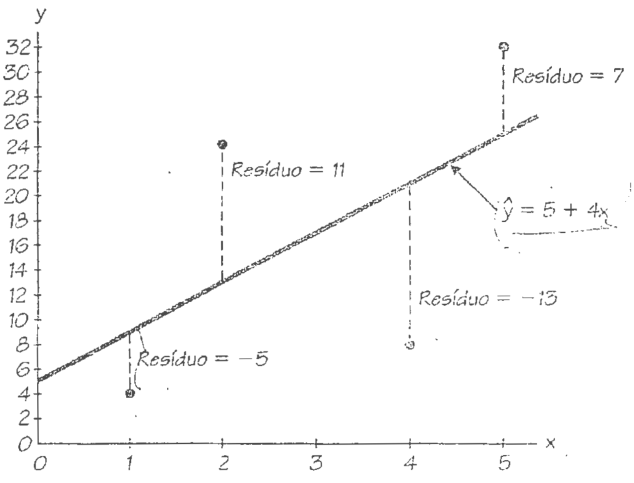


Fig. 9-8 Diagrama de dispersão com reta de regressão e resíduos.

equação $y = x^2$, e não por uma equação linear, que teria a forma $y = b_0 + b_1x$.

x	2	5	4	8	10
y	4	25	16	64	100

Hoje as calculadoras e os computadores facilitam a determinação de equações não-lineares. A calculadora TI-83 apresenta as seguintes opções:

`QuadReg:`	Acha a melhor equação quadrática ($y = ax^2 + bx + c$) que se ajusta aos dados amostrais.
`CubicReg:`	Acha o melhor polinômio de terceiro grau ($y = ax^3 + bx^2 + cx + d$) que se ajusta aos dados amostrais.
`QuartReg:`	Acha o melhor polinômio de quarto grau ($y = ax^4 + bx^3 + cx^2 + dx + e$) que se ajusta aos dados amostrais.
`LnReg:`	Acha a melhor equação logarítmica ($y = a + b \ln x$) que se ajusta aos dados amostrais.
`ExpReg:`	Acha a melhor equação exponencial ($y = ab^x$) que se ajusta aos dados amostrais.
`PwrReg:`	Acha a melhor equação potência ($y = ax^b$) que se ajusta aos dados amostrais.
`Logistic:`	Acha a melhor equação da forma $y = c/(1 + ae^{-bx})$ que se ajusta aos dados amostrais.

Pode-se achar a melhor equação de ajuste com auxílio de diagramas de dispersão (que podem revelar a natureza geral do relacionamento) e de uma calculadora ou de um computador cujas apresentações incluem valores de r ou outras estatísticas que permitem avaliar quão bem a equação se ajusta aos dados amostrais. Veja o Exercício 25.

9-3 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-4, determine a equação da reta de regressão para os dados apresentados.

1.

x	1	2	4	5
y	3	5	9	11

2.

x	5	3	2	1	0	2
y	−2	0	1	2	3	1

3.

x	2	3	5	5	10
y	6	9	14	16	30

4.

x	2	3	5	5	10
y	6	0	15	5	2

Nos Exercícios 5-18, utilize os mesmos dados dos exercícios da Seção 9-2. Em cada caso, determine a equação de regressão, considerando como independente a primeira variável (x). *Ache os valores preditos, quando pedidos.* Cuidado: *Ao achar valores preditos, siga o processo de predição delineado nesta seção.*

5.

x Tórax (in.)	26	45	54	49	41	49	44	19
y Peso (lb)	90	344	416	348	262	360	332	34

Ache o melhor peso predito para um urso com 52 in. de tórax.

6.

x Peso	29	35	28	44	25	34	30	33	28	24
y Combustível	31	27	29	25	31	29	28	28	28	33

Ache o melhor consumo de combustível predito para um carro que pesa 4200 lb. (Note que a tabela dá valores de x em centenas de libras.)

7.

Plástico (lb)	0,27	1,41	2,19	2,83	2,19	1,81	0,85	3,05
Tamanho da Residência	2	3	3	6	4	2	1	5

Qual é a melhor predição do tamanho de uma residência que descarta 0,50 lb de plástico?

8.

Papel	2,41	7,57	9,55	8,82	8,72	6,96	6,83	11,42
Tamanho da Residência	2	3	3	6	4	2	1	5

Qual é a melhor predição do tamanho de uma residência que descarta 10,00 lb de papel?

9.

Alimentação	1,04	3,68	4,43	2,98	6,30	1,46	8,82	9,62
Tamanho da Residência	2	3	3	6	4	2	1	5

Qual é a melhor predição do tamanho de uma residência que descarta 8,00 lb de alimentos?

10.

Peso Total	10,76	19,96	27,60	38,11	27,90	21,90	21,83	49,27	33,27	35,54
Tamanho da Residência	2	3	3	6	4	2	1	5	6	4

Qual é a melhor predição do tamanho de uma residência que descarta 50,00 lb de lixo?

11.

Idade	17,2	43,5	30,7	53,1	37,2	21,0	27,6	46,3
CAS	0,19	0,20	0,26	0,16	0,24	0,20	0,18	0,23

Qual é a melhor predição para o nível de álcool no sangue de uma pessoa de 21,0 anos de idade que foi condenada e presa por dirigir embriagada (DWI = Driving While Intoxicated)? (Mede-se o nível CAS quando a pessoa é presa pela primeira vez.)

12.

Armas automáticas	11,6	8,3	3,6	0,6	6,9	2,5	2,4	2,6
Taxa de criminalidade	13,1	10,6	10,1	4,4	11,5	6,6	3,6	5,3

Qual é a melhor predição para a taxa de criminalidade em um estado com 10.000 armas automáticas registradas? (Os números de armas automáticas desta tabela são dados em milhares.)

13. Com o Conjunto de Dados 16 do Apêndice B, utilize os dados emparelhados de duração e intervalo entre as erupções do gêiser Old Faithful. Qual é a melhor predição para o tempo que antecede a próxima erupção, dado que a erupção anterior durou 210 segundos?

14. Com o Conjunto de Dados 16 do Apêndice B, utilize os dados emparelhados referentes à altura e ao intervalo entre as erupções do Gêiser Old Faithful. Qual é a melhor predição do tempo antes da próxima erupção, sabendo-se que a erupção prévia teve uma altura de 275 pés?

15. Considere os dados emparelhados do Conjunto de Dados 4 do Apêndice B, referentes ao monóxido de carbono e à nicotina. Qual é a melhor predição para a quantidade de nicotina em um cigarro com 3 mg de alcatrão?

16. Considere os dados emparelhados do Conjunto 4 do Apêndice B, referente ao monóxido de carbono e à nicotina. Qual é a melhor predição da quantidade de nicotina para um cigarro com 18 mg de monóxido de carbono?

17. Com o Conjunto de Dados 7 do Apêndice B, considere os dados emparelhados referentes à precipitação pluviométrica anual (PRECIP) e à quantidade de milho produzida (CORNPROD). Qual é a melhor predição para a produção de milho em um ano cuja precipitação pluviométrica foi de 40,0 in.?

18. Com o Conjunto de Dados 7 do Apêndice B, use os dados emparelhados referentes às temperaturas médias anuais em Iowa (AVTEMP) e as quantidades de milho produzidas (CORNPROD). Qual é a melhor predição para a quantidade de milho produzida em um ano com temperatura média de 45,00°F?

19. Em cada um dos casos seguintes, ache a melhor predição para o valor de y, dado que $x = 3{,}00$. As estatísticas dadas constituem resumos para dados amostrais emparelhados.
 a. $r = 0{,}931$, $\bar{y} = 7{,}00$, $n = 10$, e a equação da reta de regressão é $\hat{y} = 4{,}00 + 2{,}00x$.
 b. $r = -0{,}033$, $\bar{y} = 2{,}50$, $n = 80$, e a equação da reta de regressão é $\hat{y} = 5{,}00 - 2{,}00x$.

20. Em cada um dos casos seguintes, ache a melhor predição para o valor de y, dado que $x = 2{,}00$. As estatísticas dadas constituem resumos para dados amostrais emparelhados.
 a. $r = -0{,}882$, $\bar{y} = 3{,}57$, $n = 15$, e a equação da reta de regressão é $\hat{y} = 23{,}00 - 8{,}00x$.
 b. $r = 0{,}187$, $\bar{y} = 9{,}33$, $n = 60$, e a equação da reta de regressão é $\hat{y} = 4{,}00 + 8{,}00x$.

21. Consulte os oito pares de dados relacionados na Tabela 9-1. Se incluirmos um nono urso com 120 in. de comprimento e 800 lb de peso, o novo ponto é um ponto extremo? É um ponto de influência?

22. Com os mesmos pares de dados relacionados na Tabela 9-1, se incluirmos um nono urso com 120 in. de comprimento e 50 lb de peso (afetivamente apelidado "magrela"), o novo ponto é um ponto extremo? É um ponto de influência?

9-3 Exercícios B: Além do Básico

23. Números grandes, como os da tabela a seguir, costumam causar problemas de cálculo. Com os valores dados, determine primeiro a equação da reta de regressão; em seguida ache a equação da reta de regressão após dividir por 1000 cada valor de x. Qual a influência, nos resultados, dessa variação em x? Como seriam afetados os resultados se cada valor de y fosse dividido por 1000?

x	924.736	832.985	825.664	793.427	857.366
y	142	111	109	95	119

24. De acordo com a propriedade dos mínimos quadrados, a reta de regressão minimiza a soma dos quadrados dos resíduos. Vimos que, com os dados emparelhados a seguir, a equação de regressão é $\hat{y} = 5 + 4x$, e que a soma dos quadrados dos resíduos é 364. Mostre que a equação $\hat{y} = 8 + 3x$ resulta em uma soma de quadrados maior do que 364.

x	1	2	4	5
y	4	24	8	32

25. Se o diagrama de dispersão revela um padrão não-linear que reconhecemos como outro tipo de curva, é possível aplicar os métodos desta seção. Para os dados apresentados, ache a equação linear ($y = b_0 + b_1x$) que melhor se ajuste aos dados amostrais; ache também a equação logarítmica ($y = a + b \ln x$) que melhor se ajuste àqueles dados. (*Sug.:* Comece substituindo cada valor de x por $\ln x$.) Qual dessas duas equações se ajusta melhor aos dados? Por quê?

x	2,0	2,5	4,2	10,0
y	12,0	18,7	53,0	225,0

26. Explique por que um teste da hipótese nula H_0: $\rho = 0$ é equivalente a um teste da hipótese nula H_0: $\beta_1 = 0$, onde ρ é o coeficiente de correlação linear para uma população de dados emparelhados, e β_1 é o coeficiente angular da reta de regressão para a mesma população.

9-4 Intervalos de Variação e de Predição

Na Seção 9-2 introduzimos o conceito de correlação linear e aplicamos o coeficiente de correlação linear r para determinar se havia correlação linear significativa entre duas variáveis. Além de servir como medida da correlação linear entre duas variáveis, o valor de r também pode dar-nos informações adicionais sobre a variação dos pontos amostrais em torno da reta de regressão. Começamos com o caso de uma amostra, que conduz a uma definição importante.

Suponhamos uma grande coleção de dados emparelhados, que apresente os seguintes resultados:

- Há uma correlação linear significativa.
- A equação da reta de regressão é $\hat{y} = 3 + 2x$.
- $\bar{y} = 9$
- Um dos pares de dados amostrais é $x = 5$ e $y = 19$.
- O ponto (5, 13) é um dos pontos da reta de regressão, pois fazendo $x = 5$ na equação de regressão obtemos

$$\hat{y} = 3 + 2x = 3 + 2(5) = 13$$

O ponto (5, 13) situa-se sobre a reta de regressão, mas o ponto (5, 19) pertence ao conjunto original de dados e não pertence à reta de regressão, porque não satisfaz a equação de regressão. Examine cuidadosamente a Figura 9-9 e note as diferenças definidas como segue.

DEFINIÇÕES

Suponha que tenhamos uma coleção de dados emparelhados que contenha o ponto (x, y), que $\hat{y}$ seja o valor predito de y (obtido pela equação de regressão), e que a média dos valores amostrais de y seja $\bar{y}$.

O **desvio total** (em relação à média) do ponto (x, y) é a distância vertical $y - \bar{y}$, que é a distância entre o ponto (x, y) e a reta horizontal que passa pela média amostral $\bar{y}$.

O **desvio explicado** é a distância vertical $\hat{y} - \bar{y}$, que é a distância entre o valor predito $\hat{y}$ e a reta horizontal que passa pela média amostral $\bar{y}$.

O **desvio não-explicado** é a distância vertical $y - \hat{y}$, ou seja, a distância vertical entre o ponto (x, y) e a reta de regressão. (A distância $y - \hat{y}$ também é chamada resíduo, definido na Seção 9-3.)

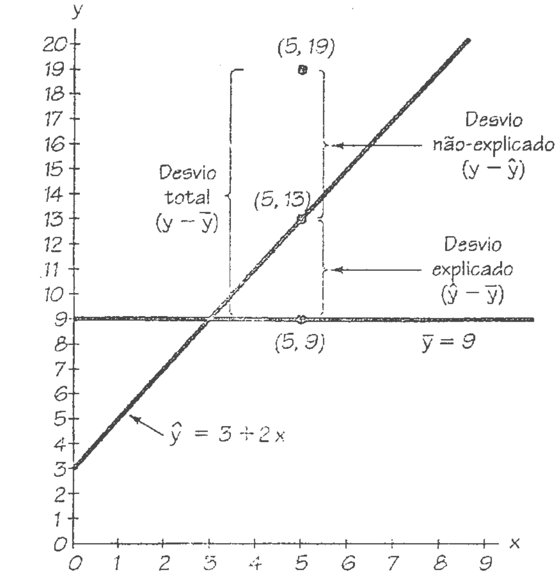


Fig. 9-9 Desvio não-explicado, desvio explicado, desvio total.

> **Preditores de Sucesso**
>
> Ao aceitar um(a) novo(a) aluno(a), uma faculdade gostaria de ter alguma indicação positiva de que o(a) mesmo(a) se sairá bem em seus estudos. São levados em conta testes SAT, testes padronizados de desempenho, classificação na turma, dificuldades no curso secundário, notas escolares elevadas e atividades extracurriculares. Em um estudo dos fatores que melhor indicam bom desempenho em uma faculdade, constatou-se que a classificação na turma e os testes padronizados de desempenho são melhores preditores do que os testes SAT. Uma equação de regressão múltipla com a nota média predita pela classificação na turma e pelo teste de desempenho não melhorou com a inclusão de outra variável referente ao escore SAT. Este estudo sugere que os escores SAT não devem ser incluídos nos critérios de admissão, mas os que os apóiam argumentam que os escores SAT servem para comparar estudantes de diferentes regiões geográficas e diferentes histórias escolares.

Para os dados específicos que estamos considerando, obtemos os seguintes resultados:

desvio total de (5, 19) $= y - \bar{y} = 19 - 9 = 10$

desvio explicado de (5, 19) $= \hat{y} - \bar{y} = 13 - 9 = 4$

desvio não-explicado de (5, 19) $= y - \hat{y} = 19 - 13 = 6$

Se ignorássemos totalmente os conceitos de correlação e regressão e ainda assim quiséssemos predizer um valor de y dados um valor de x e uma coleção de dados emparelhados (x, y), nossa melhor suposição seria $\bar{y}$. Mas não ignoramos totalmente aqueles dois conceitos: sabemos que, neste caso (com uma correlação linear significativa), a maneira de predizer o valor de y quando $x = 5$ é aplicar a equação de regressão, que dá $\hat{y} = 13$, conforme calculado anteriormente. Para explicar a discrepância entre $\bar{y} = 9$ e $\hat{y} = 13$, basta notar que existe uma correlação linear significativa cuja melhor descrição é a reta de regressão. Conseqüentemente, quando $x = 5$, *y deve ser* 13, e não 9. Mas, quando deveria ser 13, *y é realmente* 19. A discrepância entre 13 e 19 não pode ser explicada pela reta de regressão, e é chamada desvio não-explicado, ou resíduo. Pode-se generalizar como se segue o caso ilustrado na Figura 9-9:

$$(\text{desvio total}) = (\text{desvio explicado}) + (\text{desvio não-explicado})$$
$$\text{ou} \quad (y - \bar{y}) = (\hat{y} - \bar{y}) + (y - \hat{y})$$

Esta última expressão se aplica a um ponto particular (x, y), mas pode ser generalizada e modificada de modo a incluir todos os pares de dados amostrais, conforme mostra a Fórmula 9-5. Naquela fórmula, a **variação total** é expressa como a soma dos quadrados dos desvios totais, a **variação explicada** é a soma dos quadrados dos desvios explicados, e a **variação não-explicada** é a soma dos quadrados dos desvios não-explicados.

Fórmula 9-5

$$(\text{variação total}) = (\text{var. explicada}) + (\text{var. não-explicada})$$
$$\text{ou} \quad \Sigma(y - \bar{y})^2 = \Sigma(\hat{y} - \bar{y})^2 + \Sigma(y - \hat{y})^2$$

A definição que segue, de grande importância, utiliza as componentes desta última definição.

> **DEFINIÇÃO**
>
> O **coeficiente de determinação** é o valor da variação de y que é explicado pela reta de regressão. É dado por
>
> $$r^2 = \frac{\text{variação explicada}}{\text{variação total}}$$

Podemos calcular r^2 tanto pela definição que acabamos de dar com a Fórmula 9-5, como simplesmente elevando ao quadrado o coeficiente de correlação linear r, obtido pelos métodos descritos na Seção 9-2. Como exemplo, se $r = 0,8$, o coeficiente de determinação é $r^2 = 0,64$, o que significa que *64% da variação total de y podem ser explicados pela reta de regressão. Decorre que 36% da variação total de y permanecem não-explicados.*

EXEMPLO Com base nas medidas de ursos da Tabela 9-1, determine a percentagem da variação no peso y que pode ser explicada pela reta de regressão.

SOLUÇÃO Recorde que a Tabela 9-1 contém oito pares de dados amostrais, cada par consistindo nos comprimentos (em polegadas) e nos pesos (em libras) de oito ursos machos. Na Seção 9-2 vimos que o coeficiente de correlação linear é $r = 0,897$. O coeficiente de determinação é $r^2 = 0,897^2 = 0,805$, o que indica que a razão da variação explicada de y para a variação total de y é 0,805. Podemos, pois, afirmar que 80,5% da variação total de y podem ser explicados pela reta de regressão. Ou seja, 80,5% da variação total dos pesos dos ursos podem ser explicados pela variação em seus comprimentos; os restantes 19,5% são atribuídos a outros fatores.

Intervalos de Predição

Na Seção 9-3 recorremos aos dados amostrais da Tabela 9-1 para achar a equação de regressão $\hat{y} = -352 + 9,66x$, onde $\hat{y}$ representa o peso predito e x é o comprimento do urso. Com aquela equação, predissemos o valor de y, dado que $x = 71,0$ in. Obtivemos o valor de 334 lb como o melhor peso predito de um urso de 71,0 in. Como 334 é um valor único, é chamado *estimativa pontual*. No Capítulo 6 vimos que as estimativas pontuais têm a desvantagem de não dar qualquer idéia de sua precisão. Aqui, sabemos que o melhor valor predito é 334, mas não sabemos quão preciso ele é. No Capítulo 6 introduzimos as estimativas por intervalos de confiança, para superar essa desvantagem; nesta seção vamos adotar este último processo. Utilizaremos um **intervalo de predição**, que é uma estimativa intervalar de confiança de um valor predito de y.

O estabelecimento de um intervalo de predição exige uma medida da dispersão dos pontos amostrais em torno da reta de regressão. Recorde que o desvio não-explicado (ou resíduo) é a distância vertical entre um ponto amostral e a reta de regressão, conforme ilustrado na Figura 9-9. O erro-padrão da estimativa é uma medida coletiva da dispersão dos pontos amostrais em torno da reta de regressão; define-se formalmente como segue.

DEFINIÇÃO

O **erro-padrão da estimativa**, denotado por s_e, é uma medida das diferenças (ou distâncias) entre os valores amostrais y observados e os valores preditos $\hat{y}$ obtidos através da reta de regressão. É dado por

$$s_e = \sqrt{\frac{\Sigma(y - \hat{y})^2}{n - 2}}$$

onde $\hat{y}$ é o valor predito de y.

O desenvolvimento do erro-padrão da estimativa s_e acompanha o do desvio-padrão ordinário introduzido no Capítulo 2. Assim como o desvio-padrão mede o quanto os escores se desviam de sua média, o erro-padrão de estimativa s_e é uma medida de quanto os pontos amostrais se afastam da reta de regressão. O raciocínio que leva à divisão por $n - 2$ na fórmula apresentada equivale ao raciocínio que conduziu à divisão por $n - 1$ no caso do desvio-padrão ordinário; não nos aprofundaremos em detalhes. É importante notar que menores valores de s_e refletem pontos próximos da reta de regressão, enquanto maiores valores de s_e correspondem a pontos mais afastados daquela reta.

Também podemos aplicar a Fórmula 9-6 para calcular o erro-padrão da estimativa s_e. Algebricamente, equivale à expressão na definição, mas esta forma é mais fácil de manejar, porque não exige o cálculo de cada um dos valores preditos $\hat{y}$ mediante substituição na equação de regressão. A Fórmula 9-6, entretanto, exige a determinação do intercepto y b_0 e do coeficiente angular b_1 da reta de regressão estimada.

Fórmula 9-6

$$s_e = \sqrt{\frac{\Sigma y^2 - b_0\Sigma y - b_1\Sigma xy}{n - 2}} \quad \text{(erro padrão da estimativa)}$$

EXEMPLO Aplique a Fórmula 9-6 para achar o erro-padrão da estimativa das medidas dos ursos da Tabela 9-1.

SOLUÇÃO Na Seção 9-2 obtivemos, com a Tabela 9-1:

$$n = 8 \quad \Sigma y^2 = 728.520 \quad \Sigma y = 2176 \quad \Sigma xy = 151.879$$

Na Seção 9-3 recorremos aos dados da Tabela 9-1 para achar o intercepto y e o coeficiente angular da reta de regressão. Damos a seguir esses valores, com mais uma decimal por questão de precisão.

$$b_0 = -351{,}660 \qquad b_1 = 9{,}65979$$

Podemos agora, com esses valores, achar o erro-padrão da estimativa s_e:

$$s_e = \sqrt{\frac{\Sigma y^2 - b_0\Sigma y - b_1\Sigma xy}{n - 2}}$$

$$= \sqrt{\frac{728.520 - (-351{,}660)(2176) - (9{,}65979)(151.879)}{8 - 2}}$$

$$= 66{,}5994 = 66{,}6 \quad \text{(arredondado)}$$

Podemos avaliar a dispersão dos pontos amostrais em torno da reta de regressão com o erro-padrão da estimativa $s_e = 66{,}6$.

Com auxílio do erro-padrão da estimativa s_e, podemos construir estimativas intervalares que nos dirão quão confiáveis são realmente as nossas estimativas pontuais de y. Suponhamos que, para cada valor fixo de x, os valores amostrais correspondentes de y se distribuam normalmente em torno da reta de regressão, e que essas distribuições normais tenham a mesma variância. A estimativa intervalar seguinte se aplica a um y *individual*. (Para um intervalo de confiança usado para predizer a *média* de todos os valores de y para algum valor dado de x, veja Exercício 20.)

Intervalo de Predição para um y Individual

Dado um valor fixo x_0, o *intervalo de predição para um determinado* y é

$$\hat{y} - E < y < \hat{y} + E$$

onde a margem de erro E é

$$E = t_{\alpha/2}s_e\sqrt{1 + \frac{1}{n} + \frac{n(x_0 - \bar{x})^2}{n(\Sigma x^2) - (\Sigma x)^2}}$$

onde x_0 representa o valor dado de x, $t_{\alpha/2}$ tem $n - 2$ graus de liberdade, e s_e é dado pela Fórmula 9-6.

EXEMPLO Consulte os dados amostrais da Tabela 9-1 relativos aos comprimentos (x) e pesos (y) de ursos. Em seções anteriores, mostramos que

- Há uma correlação linear significativa (ao nível de 0,05 de significância).
- A equação de regressão é $\hat{y} = -352 + 9{,}66x$
- Para $x = 71{,}0$, o valor predito de y é 334 lb.

Construa um intervalo de predição de 95% para o peso de um urso que tem 71,0 in. de comprimento. Isto nos dará uma medida do grau de confiabilidade da estimativa de 334 lb.

SOLUÇÃO Já usamos os dados amostrais da Tabela 9-1 para achar os seguintes valores:

$$n = 8 \quad \bar{x} = 64{,}5625 \quad \Sigma x = 516{,}5 \quad \Sigma x^2 = 34.525{,}75$$

$$s_e = 66{,}5994$$

Pela Tabela A-3, temos que $t_{\alpha/2} = 2{,}447$. (Tomamos $8 - 2 = 6$ graus de liberdade com $\alpha = 0{,}05$ bilateral). Podemos agora calcular a margem de erro E fazendo $x_0 = 71{,}0$, porque queremos o intervalo de predição de y para $x = 71{,}0$.

$$E = t_{\alpha/2} s_e \sqrt{1 + \frac{1}{n} + \frac{n(x_0 - \bar{x})^2}{n(\Sigma x^2) - (\Sigma x)^2}}$$
$$= (2{,}447)(66{,}5994)\sqrt{1 + \frac{1}{8} + \frac{8(71{,}0 - 64{,}5625)^2}{8(34.525{,}75) - (516{,}5)^2}}$$
$$= (2{,}447)(66{,}5994)(1{,}07710) = 176$$

Com $\hat{y} = 334$ e $E = 176$, obtemos o intervalo de predição como segue:

$$\hat{y} - E < y < \hat{y} + E$$
$$334 - 176 < y < 334 + 176$$
$$158 < y < 510$$

Assim, para um urso de 71,0 in., temos 95% de confiança de que seu verdadeiro peso esteja entre 158 lb e 510 lb. Trata-se de um intervalo demasiadamente grande. (Um fator que contribui para isto é o pequeno tamanho (oito) da amostra.)

Além de sabermos que o peso predito é 334 lb, temos agora uma idéia da confiabilidade real daquela estimativa. O intervalo de predição de 95% mostra que a estimativa de 334 pode variar substancialmente.

Diferenças de Salários

Muitos artigos mencionam que, em média, uma trabalhadora em tempo integral ganha 70 centavos para cada dólar ganho por um trabalhador do sexo masculino. Pesquisadores do Institute for Social Research, da Universidade de Michigan, analisaram os efeitos de vários fatores e concluíram que cerca de um terço da discrepância pode ser explicado por diferenças de escolaridade, tempo de serviço, períodos de interrupção do trabalho e diferentes escolhas de emprego. Os outros dois terços permanecem inexplicados por esses fatores ligados ao trabalho.

O Minitab pode ser utilizado para gerar intervalos de predição de 95%. Para tanto, introduza os dados x na coluna C1, os dados y na coluna C2, selecione as opções `Stat`, `Regression` e `Regression`. Introduza C2 na caixa rotulada `Response` e introduza C1 na caixa rotulada `Predictors`. Clique na caixa `Options` e introduza 71,0 (ou qualquer outro valor desejado de x_0) na caixa rotulada `Prediction intervals for new observations`. Utilizando-se este processo com os dados da Tabela 9-1, o Minitab apresentará resultados que incluirão (158,6; 509,8) abaixo do cabeçalho `95,0% P.I.` Isto corresponde ao mesmo intervalo de predição achado no exemplo precedente, com pequena discrepância devida ao arredondamento.

9-4 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-4, utilizando o valor do coeficiente de correlação linear r, ache o coeficiente de determinação e a percentagem da variação total que pode ser explicada pela reta de regressão.

1. $r = 0{,}2$
2. $r = -0{,}6$
3. $r = -0{,}225$
4. $r = 0{,}837$

Nos Exercícios 5-8, ache (a) a variação explicada, (b) a variação não-explicada, (c) a variação total, (d) o coeficiente de determinação, e (e) o erro-padrão da estimativa s_e.

5. A tabela mostrada relaciona os números x de azulejos e os custos y (em dólares) de sua ajustagem e colocação. (A equação da reta de regressão é $\hat{y} = 2 + 3x$.)

x	1	2	3	5	6
y	5	8	11	17	20

6. Os dados emparelhados que se seguem consistem no perímetro torácico (em polegadas) e dos pesos (em libras) de uma amostra de ursos machos. (A equação de regressão é $\hat{y} = -187{,}462 + 11{,}2713x$.)

x Tórax (in.)	26	45	54	49	41	49	44	19
y Peso (lb)	90	344	416	348	262	360	332	34

7. Os dados emparelhados a seguir consistem nos pesos (em libras) de plástico descartado e tamanhos de residências. (A equação da reta de regressão é $\hat{y} = 0{,}549270 + 1{,}47985x$.)

Plástico (lb)	0.27	1,41	2,19	2:83	2,19	1,81	0,85	3.05
Tamanho da Residência	2	3	3	6	4	2	1	5

8. Os dados emparelhados mostrados consistem nos pesos totais (em libras) de lixo descartado e tamanhos de residências. (A equação da reta de regressão é $\hat{y} = 0{,}182817 + 0{,}119423x$.)

Peso Total	10.76	19.96	27.60	38,11	27,90	21,90	21,83	49.27	33.27	35.54
Tamanho da Residência	2	3	3	6	4	2	1	5	6	4

9. Considere os dados do Exercício 5 e suponha que estão satisfeitas as condições necessárias de normalidade e variância.
   a. Para $x = 4$, ache $\hat{y}$, o valor predito de y.
   b. Qual é o efeito do valor de s_e sobre a construção do intervalo de predição de 95% de y para $x = 4$?
10. Considere os dados do Exercício 6 e suponha como satisfeitas as condições necessárias de normalidade e variância.
    a. Para um urso com perímetro torácico de 52 in., ache $\hat{y}$, o peso predito.
    b. Ache o intervalo de predição de 99% de y para $x = 52$.
11. Considere os dados do Exercício 7 e suponha como satisfeitas as condições necessárias de normalidade e variância.
    a. Ache o tamanho predito de uma residência que descarta 2.50 lb de plástico.
    b. Ache o intervalo de predição de 95% para o tamanho de uma residência que descarta 2,50 lb de plástico.
12. Considere os dados do Exercício 8 e suponha como satisfeitas as condições necessárias de normalidade e variância.
    a. Ache o tamanho predito de uma residência que descarta 20,0 lb de lixo.
    b. Ache o intervalo de predição de 99% para o tamanho de uma residência que descarta 20,0 lb de lixo.

Nos Exercícios 13-16, utilize os dados amostrais da Tabela 9-1. Seja x o comprimento (em polegadas) de um urso, e y o seu peso. Construa uma estimativa intervalar da predição do peso de um urso que tem o comprimento dado. Trabalhe com o grau de confiança dado. (Veja o exemplo nesta seção.)

13. 50,0 in.; 95% de confiança
14. 65,0 in.; 90% de confiança
15. 49,7 in.; 90% de confiança
16. 68,0 in.; 99% de confiança

9-4 Exercícios B: Além do Básico

17. Podem-se obter intervalos de confiança para o intercepto y β_0 e o coeficiente angular β_1 da reta de regressão ($y = \beta_0 + \beta_1 x$) calculando-se os limites nos intervalos abaixo.

$$b_0 - E < \beta_0 < b_0 + E$$

onde $$E = t_{\alpha/2} s_e \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}}$$

$$b_1 - E < \beta_1 < b_1 + E$$

onde $$E = t_{\alpha/2} \cdot \frac{s_e}{\sqrt{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}}$$

Nestas expressões, o intercepto y b_0 e o coeficiente angular b_1 se obtêm a partir dos dados amostrais e $t_{\alpha/2}$ consta da Tabela A-3 com $n - 2$ graus de liberdade. Com os dados referentes aos ursos da Tabela 9-1, determine estimativas intervalares de 95% de confiança para β_0 e β_1.

18. **a.** Se uma coleção de dados emparelhados compreende ao menos três pares de valores, que se pode dizer sobre o coeficiente de correlação linear se $s_e = 0$?

b. Se uma coleção de dados emparelhados é tal que a variação explicada total é 0, que se pode dizer quanto ao coeficiente angular da reta de regressão?

19. **a.** Determine uma expressão para a variação não-explicada em termos do tamanho n da amostra e do erro-padrão da estimativa s_e.

b. Determine uma expressão para a variação explicada em termos do coeficiente de determinação r^2 e da variação não-explicada.

c. Seja uma coleção de dados emparelhados para a qual $r^2 = 0{,}900$ e a equação de regressão é $\hat{y} = 3 - 2x$. Determine o coeficiente de correlação linear.

20. Da expressão dada nesta seção para a margem de erro correspondente a um intervalo de predição para y, podemos obter

$$s_{\hat{y}} = s_e \sqrt{1 + \frac{1}{n} + \frac{n(x_0 - \bar{x})^2}{n(\Sigma x^2) - (\Sigma x)^2}}$$

que é o *erro padrão da predição* ao predizer um *valor isolado* de y, dado que $x = x_0$. Ao predizer a *média* de todos os valores de y para os quais $x = x_0$, a estimativa pontual $\hat{y}$ é a mesma, mas $s_{\hat{y}}$ é:

$$s_{\hat{y}} = s_e \sqrt{\frac{1}{n} + \frac{n(x_0 - \bar{x})^2}{n(\Sigma x^2) - (\Sigma x)^2}}$$

Com os dados da Tabela 9-1, prolongue o último exemplo desta seção para achar uma estimativa pontual e uma estimativa intervalar de 95% de confiança para o peso médio de todos os ursos com 71,0 in. de comprimento.

9-5 Regressão Múltipla

Até aqui, os exemplos e exercícios deste capítulo abrangeram o relacionamento entre duas variáveis, como comprimentos e pesos de ursos. Em muitos casos, entretanto, uma variável pode estar relacionada com outras duas ou mais variáveis. Por exemplo, em vez de predizer o peso de um urso com base somente em seu comprimento, poderíamos obter uma melhor predição utilizando o comprimento, o perímetro torácico e o tamanho do pescoço. Tal como nas seções anteriores deste capítulo, trabalharemos apenas com relações lineares; nosso objetivo é estabelecer uma equação de regressão múltipla, que expresse uma variável dependente y em termos de duas ou mais outras variáveis.

DEFINIÇÃO

Uma **equação de regressão múltipla** expressa um relacionamento linear entre uma variável dependente y e duas ou mais variáveis independentes ($x_1, x_2, \ldots, x_k$).

Utilizaremos a notação seguinte, que decorre de modo natural da notação usada na Seção 9-3.

Notação

$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \ldots + b_kx_k$ (Forma geral da equação de regressão múltipla estimada)

n = tamanho da amostra

k = número de variáveis *independentes*. (As variáveis independentes são chamadas também **variáveis preditoras** ou variáveis x.)

$\hat{y}$ = valor predito da variável dependente y (calculado a partir da equação de regressão múltipla)

$x_1, x_2, \ldots, x_k$ são as variáveis independentes

β_0 = intercepto y, ou valor de y quando todas as variáveis preditoras são 0.

b_0 = estimativa de β_0 baseada nos dados amostrais

$\beta_1, \beta_2, \ldots, \beta_k$ são os coeficientes das variáveis independentes $x_1, x_2, \ldots, x_k$

$b_1, b_2, \ldots, b_k$ são as estimativas amostrais dos coeficientes $\beta_1, \beta_2, \ldots, \beta_k$

Fazendo Música com Regressão Múltipla

A Sony fabrica milhões de discos compactos em Terre Haute, Indiana. Em um estágio do processo de fabricação, um laser expõe uma chapa fotográfica de modo que um sinal musical seja transferido para um sinal digital codificado com 0s e 1s. Este processo foi analisado estatisticamente, para identificar os efeitos de diferentes variáveis, como o tempo de exposição e a consistência da emulsão fotográfica. Os métodos de regressão múltipla mostraram que, entre todas as variáveis consideradas, quatro eram altamente significativas. O processo fotográfico foi ajustado para otimizar resultados com base nas quatro variáveis críticas. Como resultado, a percentagem de discos defeituosos diminuiu, mantendo-se a qualidade do tom. O emprego de métodos de regressão múltipla levou a custos mais baixos de produção e melhor controle do processo de fabricação.

O Uso do Software em Regressão Múltipla

Em seções prévias deste capítulo, desenvolvemos processos que podiam ser usados com uma calculadora, mas os cálculos necessários para a regressão múltipla são tão complexos que vamos admitir o uso de um programa estatístico. Detalharemos processos para introduzir dados e interpretar resultados obtidos por computador.

A calculadora TI-83 não dá equações de regressão múltipla, ao contrário do STATDISK e do Minitab. Com o STATDISK, utili-

TABELA 9-3 Dados Relativos a Ursos Machos Anestesiados

Variável	Coluna Minitab	Nome	Dados Amostrais							
y	C1	WEIGHT	80	344	416	348	262	360	332	34
x_2	C2	AGE	19	55	81	115	56	51	68	8
x_3	C3	HEADLEN	11,0	16,5	15,5	17,0	15,0	13,5	16,0	9,0
x_4	C4	HEADWDTH	5,5	9,0	8,0	10,0	7,5	8,0	9,0	4,5
x_5	C5	NECK	16,0	28,0	31,0	31,5	26,5	27,0	29,0	13,0
x_6	C6	LENGTH	53,0	67,5	72,0	72,0	73,5	68,5	73,0	37,0
x_7	C7	CHEST	26	45	54	49	41	49	44	19

zamos as opções de `Analysis e Multiple Regression`. O STATDISK dará a equação de regressão múltipla, o coeficiente de determinação múltipla R^2 e o valor de R^2 ajustado, expressões que serão logo descritas. O Minitab dá muito mais.

Com o Minitab, utilize o item `Statistics` do menu principal, selecione `Regression`, selecione `Regression` mais uma vez. Com os dados da Tabela 9-3, e com a especificação de uma variável resposta `WEIGHT`* e variáveis preditoras `HEADLEN` e `LENGTH`, a apresentação Minitab será a seguinte.

A apresentação Minitab tem três componentes-chave, identificadas por números dentro de círculos:

1. Equação de regressão múltipla:

```
WEIGHT = -374 + 18.8 HEADLEN + 5.87 LENGTH
```

2. R^2 ajustado = 0,759
3. Significância global da equação de regressão múltipla: p = 0,012.

*Weight = peso, age = idade, headlen, de head = cabeça e length = comprimento, headwidth, de head = cabeça e width = largura, neck = pescoço, length = comprimento, chest = tórax. (N. do T.)

Passamos a abordar rapidamente cada uma dessas três componentes. Como estamos focalizando a interpretação de resultados dados por computador e como a teoria subjacente é bastante complicada, incluiremos muito pouca teoria. Para uma abordagem mais ampla, o leitor deve consultar um livro mais aprofundado, como *Elementary Business Statistics*, de Triola e Franklin.

A Estatística no Tribunal

Os proprietários de um complexo de cinco edifícios de apartamentos na cidade de Nova York moveram uma ação em virtude de danos causados aos tijolos. O dano ocorreu quando a água foi absorvida pela parede de tijolos, seguindo-se ciclos de congelamento e descongelamento, o que fez com que a parede rachasse. Com cerca de 750.000 tijolos, não seria possível inspecionar um a um, o que levou à adoção de métodos de amostragem. Estatísticos aplicaram métodos de regressão para predizer o número total de tijolos danificados. As variáveis independentes incluíram qual dos cinco edifícios foi utilizado, a orientação da parede, a altura, e se a parede se voltava para um pátio interno ou era uma parede externa. A estimativa do dano total parece ter influenciado fortemente os acordos finais. (Veja "Bricks, Buildings, and the Bronx: Estimating Masonry Deterioration", de Fairley, Izenman, e Whitlock, *Chance*, Vol. 7, N.º 3.)

MINITAB DISPLAY

```
The regression equation is
WEIGHT = -374 + 18.8 HEADLEN + 5.87 LENGTH

Predictor      Coef    Stdev   t-ratio       p
Constant     -374.3    134.1     -2.79   0.038
HEADLEN       18.82    23.15      0.81   0.453
LENGTH        5.875    5.065      1.16   0.299

s = 68.56   R-sq = 82.8%   R-sq(adj) = 75.9%

Analysis of Variance

SOURCE        DF       SS      MS        F       p
Regression     2   113142   56571    12.03   0.012
Error          5    23506    4701
Total          7   136648

SOURCE        DF   SEQ SS
HEADLEN        1   106819
LENGTH         1     6323
```

1 Equação de regresso múltiplo

$R^2 = 0.828$

2 R^2Ajustado = 0,759

3 Significância global da equação de regrassão múltipla

Equação de Regressão Múltipla A equação de regressão múltipla

WEIGHT = −374 + 18.8 HEADLEN + 5.87 LENGTH

pode expressar-se em notação-padrão como

$$\hat{y} = -374 + 18{,}8x_3 + 5{,}87x_6$$

Esta equação é a que melhor se ajusta aos dados (de peso, comprimento da cabeça e comprimento total) de acordo com o critério de mínimos quadrados descrito na Seção 9-3. Isto é, quando utilizamos os dados da Tabela 9-3 para WEIGHT, HEADLEN e LENGTH, esta equação é a que melhor se ajusta aos dados; se tivéssemos utilizado outros dados, poderíamos achar outra equação que melhor se ajustasse a eles. (Mais adiante, mostraremos como estabelecer a equação de regressão que melhor se ajusta aos dados.) Se a equação se ajusta bem aos dados, pode ser usada para predições. Por exemplo, se constatamos que a equação é adequada para predições, e se temos um urso com 14,0 in. de comprimento de cabeça e 71,0 in. de comprimento total, podemos predizer seu peso levando esses valores na equação de regressão para obter um peso predito de 306 lb. Por outro lado, os coeficientes $b_3 = 18{,}8$ e $b_6 = 5{,}87$ podem servir para determinar uma *variação marginal*, descrita na Seção 9-3. O coeficiente $b_3 = 18{,}8$, por exemplo, mostra que, quando o comprimento total de um urso permanece constante, o peso predito aumenta de 18,8 lb para cada aumento de 1 in. no comprimento da cabeça.

R^2 Ajustado R^2 denota o **coeficiente de determinação múltipla,** que é uma medida do grau de ajustamento da equação de regressão múltipla aos dados amostrais. Um ajuste perfeito resultaria em $R^2 = 1$. Um ajuste muito bom acarreta um valor próximo de 1. E um ajuste fraco ocasiona um valor de R^2 próximo de 0. O valor $R^2 = 0{,}828$ no quadro Minitab indica que 82,8% da variação no peso dos ursos podem ser explicados pelo comprimento da cabeça x_3 e pelo comprimento total x_6. *O coeficiente múltiplo de determinação* R^2 é uma medida da aderência da equação de regressão aos dados amostrais, mas tem uma séria falha: na medida em que se incluem mais variáveis, R^2 aumenta. (Na verdade, R^2 poderia permanecer o mesmo, mas em geral aumenta.) Embora o maior valor de R^2 seja atingido simplesmente com a inclusão de todas as variáveis disponíveis, a melhor equação de regressão múltipla nem sempre inclui todas essas variáveis. Conseqüentemente, é melhor usar o coeficiente de determinação ajustado ao comparar diferentes equações de regressão múltipla, porque ele ajusta o valor de R^2 com base no número de variáveis e no tamanho da amostra.

DEFINIÇÃO

O **coeficiente de determinação ajustado** é o coeficiente múltiplo de determinação R^2 modificado de modo a levar em conta o número de variáveis e o tamanho da amostra. Calcula-se pela Fórmula 9-7.

Fórmula 9-7 $$R^2 \text{ ajustado} = 1 - \frac{(n-1)}{[n-(k+1)]}(1-R^2)$$

O Minitab apresenta o coeficiente de determinação ajustado como R-sq (adj) = 75.9%. Se aplicarmos a Fórmula 9-7 com o valor 0,828 para R^2, $n = 8$ e $k = 2$, veremos que o valor ajustado de R^2 é 0,759, confirmando o valor de 75,9% apresentado pelo Minitab. Para os dados da Tabela 9-3 referentes a peso, comprimento da cabeça e comprimento total, o valor de 82,8% de R^2 indica que 82,8% da variação no peso podem ser explicados pelo comprimento da cabeça x_3 e pelo comprimento total x_6, mas quando comparamos esta equação de regressão múltipla com outras é melhor usar o valor ajustado de R^2 de 75,9% (ou 0,759).

Modelo para Contribuições de Alunos

Em um estudo, elaborou-se uma equação de regressão múltipla que constitui um bom preditor de doações de alunos em uma faculdade de artes liberais. A faculdade esperava usar os resultados do estudo para melhorar sua estratégia de arrecadação de fundos. A variável dependente era a quantia total doada em um ano. As variáveis independentes incluíam renda, idade, estado civil, se o doador pertencia a alguma fraternidade ou irmandade, se o doador era participante ativo das atividades dos alunos, o grau de instrução do doador, a distância até a faculdade e a taxa de desemprego da nação (usada como medida da economia). Outras variáveis independentes (se o doador tinha filhos) foram excluídas por falta de significação estatística. (Veja "An Econometric Model of Alumni Giving: A Case Study for a Liberal Arts College", de Bruggink e Siddiqui, *The American Economist*, Vol. 39, Nº 2.)

Significância Global: O Valor P O Minitab apresenta o valor P 0,012. Este valor é uma medida da significância global da equação de regressão múltipla. Neste caso, o pequeno valor 0,012 sugere que a equação de regressão múltipla tem boa significância global, podendo ser usada para predições. Ou seja, tem sentido predizer pesos de ursos com base nos seus comprimentos de cabeça e nos seus comprimentos totais. Tal como o R^2 ajustado, este valor P é uma boa medida da aderência da equação aos dados amostrais. O valor 0,012 resulta de um teste da hipótese nula, $\beta_3 = \beta_6 = 0$. A rejeição de $\beta_3 = \beta_6 = 0$ implica que ao menos um dos valores β_3 ou β_6 é diferente de zero, o que sugere que esta equação de regressão é adequada para determinar os pesos de ursos.

Uma análise completa dos resultados Minitab poderia incluir outros elementos importantes como a significância dos coeficientes individuais, mas vamos limitar nossa discussão às três componentes-chave já identificadas.

Determinação da Melhor Equação de Regressão Múltipla

Até aqui consideramos apenas a equação de regressão múltipla relacionando pesos de ursos com as variáveis independentes comprimento da cabeça e comprimento total, mas há outras variáveis que poderiam ser incluídas. A Tabela 9-4 relaciona algumas combinações de variáveis.

Embora *a determinação da melhor equação de regressão múltipla seja em geral bastante difícil, ultrapassando o âmbito deste livro*, podemos obter algum auxílio com as seguintes diretrizes:

1. *Use o bom senso e considerações de ordem prática para incluir ou excluir variáveis.* Por exemplo, poderíamos excluir a variável idade porque um pesquisador inexperiente pode não saber como determinar a idade de um urso (e os ursos relutam em revelar suas idades).
2. Em vez de incluir quase todas as variáveis disponíveis, *inclua um número relativamente pequeno de variáveis independentes (x).* No processo de eliminação de variáveis independentes que não tenham influência na variável dependente, pode-

TABELA 9-4 Escolha da Melhor Equação de Regressão Múltipla

	COMPRIMENTO TOTAL	TÓRAX	COMPRIMENTO DA CABEÇA/ COMPRIMENTO TOTAL	IDADE/PESCOÇO/ COMPRIMENTO TOTAL/TÓRAX	IDADE COMPRIMENTO DA CABEÇA LARGURA DA CABEÇA/ PESCOÇO/ COMPRIMENTO TOTAL/ TÓRAX
R^2	0,805	0,983	0,828	0,999	0,999
R^2 ajustado	0,773	0,980	0,759	0,997	0,996
Significação global	0,002	0,000	0,012	0,000	0,046

ria ser aconselhável achar o coeficiente de correlação linear r para cada par de variáveis consideradas. Por exemplo, com os dados da Tabela 9-3, podemos ver que há uma correlação linear de 0,955 para o par de dados NECK/HEADLEN. Como há uma correlação tão elevada entre tamanho do pescoço e comprimento da cabeça, não há necessidade de incluir ambas essas variáveis. Ao escolher entre NECK e HEADLEN, devemos incluir NECK por esta razão: NECK é o melhor preditor de WEIGTH porque os dados emparelhados NECK/WEIGHT têm um coeficiente de correlação linear $r = 0,971$, que é maior do que $r = 0,884$ dos dados emparelhados HEADLEN/WEIGHT.

3. *Escolha uma equação que tenha um valor ajustado de R^2 com esta propriedade: Se se inclui uma variável independente adicional, o valor ajustado de R^2 não é aumentado substancialmente.* Por exemplo, a Tabela 9-4 mostra que, se utilizamos apenas a variável independente de CHEST, o R^2 ajustado é 0,980, mas quando incluímos todas as seis variáveis, o R^2 ajustado aumenta para 0,996. A utilização de seis variáveis em lugar de apenas uma é um preço muito alto a pagar por um aumento tão pequeno no valor ajustado de R^2. É mais conveniente trabalhar com a variável independente única CHEST do que utilizar todas as seis variáveis independentes.
4. *Para um dado número de variáveis independentes (x), escolha a equação com o maior valor ajustado de R^2.* Ou seja, escolha as variáveis com a propriedade de que nenhuma outra combinação do mesmo número de variáveis independentes dê um valor maior para R^2 ajustado. Por exemplo, se nos decidirmos por uma única variável independente, devemos escolher CHEST, porque seu R^2 ajustado de 0,980 é maior do que R^2 ajustado para qualquer outra variável isolada. Conseqüentemente, CHEST é o melhor preditor *isolado* de WEIGHT.
5. *Escolha uma equação que tenha uma significância global, tal como determinada pelo valor P na tela do computador.* Por exemplo, veja os valores da significância global na Tabela 9-4. A utilização de todas as seis variáveis independentes resulta em uma significância global de 0,046, que é apenas pouco significante no nível $\alpha = 0,05$; será melhor a variável única CHEST, que tem uma significância global de 0,000.

Com estas diretrizes em uma tentativa de achar a melhor equação de predição para os pesos dos ursos, vemos que, para os dados da Tabela 9-3, a melhor equação de regressão utiliza a variável independente única CHEST. A melhor equação de regressão parece ser

$$\text{WEIGHT} = -195 + 11.4\ \text{CHEST}$$

ou

$$\hat{y} = -195 + 11.4x_7$$

Para casos que envolvem um grande número de variáveis independentes, muitos pacotes estatísticos incluem um programa para **regressão passo a passo**, no qual são tentadas diferentes combinações, até atingir o melhor modelo. Por exemplo, com Minitab, se eliminamos a variável AGE (como no Passo 1), obtemos que a melhor equação de regressão é aquela em que CHEST é a única variável independente. (Se incluirmos todas as seis variáveis independentes, o Minitab selecionará uma equação de regressão com as variáveis independentes AGE, NECK, LENGTH e CHEST, com um valor ajustado de R^2 de 0,997 e significância global de 0,000.) Parece, pois, que podemos estimar o peso de um urso com base na medida de seu tórax, e a equação de regressão conduz à seguinte regra: A estimativa do peso de um urso é 11,4 vezes a medida do seu tórax (em polegadas) menos 195.

Quando abordamos a regressão na Seção 9-3, relacionamos quatro erros comuns que devem ser evitados ao fazermos predições com base em equações de regressão. Esses mesmos erros devem ser evitados ao utilizarmos equações de regressão múltipla. Devemos ter especial cuidado ao concluirmos pela existência de uma relação de causa e efeito.

9-5 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-4, consulte o quadro de resultados, dados pelo Minitab, a seguir, e responda às questões.

1. Identifique a equação de regressão múltipla que expressa o peso em termos da idade, da largura da cabeça e do tamanho do pescoço.
2. Identifique
   a. O valor P correspondente à significância global da equação de regressão múltipla.
   b. O valor do coeficiente R^2.
   c. O valor ajustado de R^2.

MINITAB

```
The regression equation is

WEIGHT = -285 -1.38 AGE -11.2 HEADWDTH +28.6 NECK

Predictor      Coef       Stdev      t-ratio          p
Constant    -285.21       78.45        -3.64      0.022
AGE         -1.3838      0.9022        -1.53      0.200
HEADWDTH     -11.24       20.88        -0.54      0.619
NECK         28.594       5.870         4.87      0.008

s = 32.49       R-sq = 96.9%    R-sq(adj) = 94.6%

Analysis of Variance

SOURCE       DF        SS        MS        F        p
Regression    3    132425     44142    41.81    0.002
Error         4      4223      1056
Total         7    136648

SOURCE       DF    SEQ SS
AGE           1     90527
HEADWDTH      1     16844
NECK          1     25054
```

3. Pode-se utilizar a equação de regressão múltipla para predizer o peso de um urso com base em sua idade, largura da cabeça e tamanho do pescoço? Por quê?
4. Um urso de 32 meses de idade tem largura de cabeça de 5,0 in. e 21,5 in. de tamanho de pescoço.
   a. Ache o peso predito do urso.
   b. O urso em questão pesava efetivamente 180 lb. Qual a precisão do peso predito na parte (a)?

Nos Exercícios 5-8, considere os dados do Conjunto 3 do Apêndice B. Seja WEIGHT *a variável dependente e sejam as variáveis independentes aquelas indicadas no exercício. (Os conjuntos de dados já estão armazenados no STATDISK.) Utilize um programa como STATDISK ou Minitab para responder às seguintes questões:*

a. Estabeleça a equação de regressão que expresse a variável dependente WEIGHT *em termos das variáveis independentes dadas.*
b. Identifique os valores do coeficiente múltiplo de determinação R^2, o valor ajustado de R^2 e (se estiver usando Minitab) o valor P correspondente à significância global.
c. A equação de regressão múltipla se afigura adequada para predizer o peso de um urso com base nas variáveis independentes dadas?

5. LENGTH e CHEST
6. NECK, LENGTH, e CHEST
7. AGE, NECK, LENGTH, e CHEST
8. HEADLEN, LENGTH, e NECK

Nos Exercícios 9-14, considere os pesos do lixo do Conjunto de Dados 1 do Apêndice B. Seja HHSIZE *(household size) a variável dependente, e sejam as variáveis independentes aquelas indicadas no exercício. (Os conjuntos de dados já estão armazenados no STATDISK.) Com um programa como STATDISK ou Minitab, responda às seguintes questões:*

a. Estabeleça a equação de regressão múltipla que expresse a variável dependente HHSIZE *em termos das variáveis independentes indicadas.*
b. Identifique os valores do coeficiente múltiplo de determinação R^2, o valor ajustado de R^2, e (se estiver usando Minitab) o valor P correspondente à significância global.
c. A equação de regressão múltipla parece adequada para fazer predições do tamanho de uma residência com base nas variáveis independentes dadas?

9. YARD
10. PLASTIC
11. PLASTIC e PAPER
12. METAL, PLASTIC, e FOOD
13. METAL e PLASTIC
14. PAPER e GLASS
15. Com base no Conjunto de Dados 7 do Apêndice B, determine a melhor equação de regressão múltipla que tenha a produção de milho como variável dependente. (Não inclua o ano como variável independente.) Esta equação é adequada para prever a produção de milho em Iowa? Por quê?
16. Com base no Conjunto de Dados 4 do Apêndice B, determine a melhor equação de regressão múltipla que tenha a nicotina como variável dependente. Esta equação é adequada para predizer o conteúdo de nicotina em um cigarro, com base no conteúdo de alcatrão e monóxido de carbono?

9-5 Exercícios B: Além do Básico

17. Em alguns casos, a equação de regressão múltipla de melhor ajuste tem a forma $\hat{y} = b_0 + b_1x + b_2x^2$. O gráfico de uma tal equação é uma parábola. Com os dados a seguir,

x	1	3	4	7	5
y	5	14	19	42	26

faça $x_1 = x$, $x_2 = x^2$ e determine a equação de regressão múltipla da parábola que melhor se ajuste aos dados. Com base no valor do coeficiente múltiplo de determinação, quão bem esta equação se ajusta aos dados?

Vocabulário

correlação
dados bivariados
distribuição normal bivariada
diagrama de dispersão
coeficiente de correlação linear
coeficiente de correlação momento-produto de Pearson
variável oculta
centróide
equação de regressão
reta de regressão
variável independente
variável preditora
variável dependente
variável resposta
variação marginal
outlier
ponto extremo
ponto de influência
resíduo
propriedade de mínimos quadrados
desvio total
desvio explicado
desvio não-explicado
variação total
variação explicada
variação não-explicada
coeficiente de determinação
intervalo de predição
erro-padrão da estimativa
equação de regressão múltipla
variáveis preditoras
coeficiente múltiplo de determinação
coeficiente ajustado de determinação
regressão passo a passo

Revisão

Neste capítulo apresentamos métodos básicos de utilização de dados emparelhados para pesquisar a relação entre duas variáveis correspondentes. Em todo o capítulo, limitamos nossa abordagem a relacionamentos lineares, porque o estudo de relações não-lineares exige recursos mais avançados de matemática.

- Na Seção 9-2 utilizamos diagramas de dispersão e o coeficiente de correlação linear para decidir se há uma correlação linear entre duas variáveis.
- Na Seção 9-3 apresentamos métodos para determinar a equação da reta de regressão que (pelo critério dos mínimos quadrados) melhor se ajuste aos dados emparelhados. Quando há correlação linear significativa, pode-se aplicar a reta de regressão para predizer o valor de uma das variáveis, dado um valor da outra variável.
- Na Seção 9-4 introduzimos o conceito de variação total, com as componentes de variação explicada e variação não-explicada. Definimos o coeficiente de determinação r^2 como o quociente obtido na divisão da variação explicada pela variação total. Estabelecemos também métodos para construir intervalos de predição, que permitem julgarmos a precisão de valores preditos.
- Na Seção 9-5 abordamos a regressão múltipla, que permite pesquisar relações entre diversas variáveis. Discutimos processos para obter uma equação de regressão múltipla, assim como os valores do coeficiente múltiplo de determinação R^2, o valor ajustado de R^2 e um valor P para a significância global da equação.

Exercícios de Revisão

Nos Exercícios 1-4, utilize os dados da tabela a seguir. Os dados provêm de um estudo do consumo de sorvete na primavera e no verão de três anos. O consumo de sorvete é dado em pints (medida de capacidade equivalente a 56,8 litros) per capita *por semana, o preço do sorvete é dado em dólares, a renda familiar é dada em dólares por semana, e a temperatura é dada em graus Fahrenheit.*

Consumo	0,386	0,374	0,393	0,425	0,406	0,344	0,327	0,288	0,269	0,256
Preço	1.35	1,41	1,39	1,40	1,36	1,31	1,38	1,34	1,33	1,39
Renda	351	356	365	360	342	351	369	356	342	356
Temperatura	41	56	63	68	69	65	61	47	32	24

Com base em dados de Kadiyala. *Econometrica*, Vol. 38.

1. a. No nível de 0,05 de significância, teste a correlação linear entre consumo e preço.
   b. Determine a equação da reta de regressão que expresse o consumo (y) em termos do preço (x).
   c. Qual é a melhor predição da quantidade consumida se o preço é \$1,38?
2. a. No nível de 0,05 de significância, teste a correlação linear entre consumo e renda.
   b. Determine a equação da reta de regressão que expresse o consumo (y) em função da renda (x).
   c. Qual é a melhor predição do consumo se a renda é \$365?
3. a. No nível de 0,05 de significância, teste a correlação linear entre consumo e temperatura.
   b. Determine a equação da reta de regressão que expresse o consumo (y) em termos da temperatura (x).
   c. Qual é a melhor predição do consumo para uma temperatura de 32°F?
4. Com o auxílio de um programa como o STATDISK ou o Minitab, determine a equação de regressão múltipla da forma $\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3$, onde a variável dependente y representa o consumo, x_1 representa o preço, x_2 representa a renda, e x_3 representa a temperatura. Identifique também o coeficiente de determinação múltipla R^2, o valor ajustado de R^2 e o valor P que representa a significância global da equação de regressão múltipla. Pode-se utilizar a equação de regressão para predizer o consumo de sorvete? Por acaso alguma das equações dos Exercícios 1-3 é melhor?

Nos Exercícios 5-8, utilize os dados amostrais da tabela a seguir. A tabela contém os "minutos antes da meia-noite do dia fatal", uma medida da ameaça de uma guerra nuclear estabelecida pelo Boletim dos Cientistas Atômicos. Os valores dos "periódicos" são índices da cobertura, pelos meios de comunicação, de tópicos relativos à guerra nuclear. Os valores das "poupanças" são expressos em termos de uma medida-padrão da taxa de poupança. Os dados são emparelhados para anos selecionados aleatoriamente.

Minutos	2.0	7.0	12,0	12.0	9,25	11,17	9.0	9,0	4,0
Periódicos	0,16	0,21	0,13	0,12	0,12	0,05	0.05	0,06	0,14
Poupança	7,2	10.2	13,5	13,5	12,0	13,1	11,6	11,6	8,5

Com base em dados de "Saving and the Fear of Nuclear War" (Poupança e o temor de uma guerra nuclear), por Slemrod, *Journal of Conflict Resolution*, Vol. 30, N.º 3.

5. a. Ao nível de 0,05 de significância, teste a correlação linear entre os minutos antes da meia-noite do dia fatal e o índice de poupança.
   b. Ache a equação da reta de regressão que expressa a poupança (y) em termos dos minutos antes da meia-noite (x).
   c. Qual é o melhor índice de predição da poupança para um ano em que o relógio do dia fatal marca 5,0 minutos antes da meia-noite?
6. a. No nível de 0,05 de significância, teste a correlação linear entre o índice dos periódicos e o índice de poupança.
   b. Determine a equação da reta de regressão que expressa o índice de poupança (y) em termos do índice dos periódicos (x).

c. Qual é a melhor predição do índice de poupança para um ano em que o índice dos periódicos é 0,20?

7. Utilize somente os dados emparelhados minutos/poupança. Para um ano em que o número de minutos antes da meia-noite no relógio do dia fatal é 5,0, determine uma estimativa intervalar de predição de 95% para o índice de poupança.

8. Sejam y = índice de poupança, x_1 = minutos antes da meia-noite e x_2 = índice dos periódicos. Utilize um programa como STATDISK ou Minitab para achar a equação de regressão múltipla da forma $\hat{y} = b_0 + b_1x_1 + b_2x_2$. Identifique também o valor do coeficiente R^2, o valor ajustado de R^2 e o valor P que representa a significância global da equação de regressão múltipla. Com base nos resultados, devemos usar a equação de regressão múltipla para fazer predições? Por quê ou porque não?

Exercícios Cumulativos de Revisão

1. Em 1970, o tempo médio entre erupções do gêiser Old Faithful era de 66 minutos. Considere os intervalos (em minutos) entre erupções para os dados recentes relacionados no Conjunto de Dados 16 do Apêndice B.
   a. Teste a afirmação do geólogo Rick Hutchinson, do Parque Nacional de Yellowstone, de que os intervalos entre as erupções são agora maiores do que em 1970.
   b. Construa um intervalo de confiança de 95% para o tempo médio entre erupções.

2. A Marc Michael Advertising Company elaborou dois comerciais de TV diferentes para *jeans* femininos marca Taylor. Um comercial é humorístico, o outro é sério. Um processo de teste envolve uma escala-padrão com valores mais altos indicando respostas mais favoráveis. Os resultados são apresentados a seguir. Com base nesses resultados, um dos comerciais parece ser melhor do que o outro? O aspecto da correlação é relevante para esta situação?

Consumidor	A	B	C	D	E	F	G	H
Comercial humorístico	26	33	19	18	29	27	23	24
Comercial sério	21	30	14	14	24	22	21	19

3. Ao estudar os efeitos da hereditariedade e do ambiente sobre a inteligência, decidiu-se analisar os QIs de gêmeos idênticos que foram separados pouco depois do nascimento. Gêmeos idênticos compartilham genes idênticos herdados do mesmo óvulo fertilizado. Estudando gêmeos idênticos criados separadamente, podemos eliminar a variável hereditariedade e isolar melhor os efeitos do ambiente. A tabela a seguir mostra os QIs de gêmeos idênticos (gêmeos mais velhos são x) criados separadamente. Com os dados amostrais determine se há algum relacionamento entre os QIs dos gêmeos. Escreva um resumo sobre o efeito da hereditariedade e do ambiente sobre a inteligência. Não esqueça que suas conclusões se basearão nesta amostra relativamente pequena de 12 pares de gêmeos idênticos.

x	107	96	103	90	96	113	86	99	109	105	96	89
y	111	97	116	107	99	111	85	108	102	105	100	93

Com base em dados de "IQs of Identical Twins Reared Apart", de Arthur Jensen, *Behavioral Genetics*.

Projeto de Computador

Com o conjunto completo de dados do Conjunto de Dados 3 do Apêndice B (armazenado em STATDISK), utilize o Minitab ou o STATDISK para determinar o coeficiente de correlação linear e a equação da reta de regressão quando a variável dependente peso é emparelhada com (a) comprimento da cabeça, (b) largura da cabeça, (c) tamanho do pescoço, (d) comprimento total e (e) tamanho do tórax. Qual das equações de regressão precedentes é melhor para predizer o peso de um urso, dados o comprimento da cabeça, a largura da cabeça, o tamanho do pescoço, o comprimento e o tamanho do tórax? Determine também a equação de regressão múltipla obtida quando a variável dependente peso é expressa em termos das variáveis independentes comprimento da cabeça, largura da cabeça, tamanho do pescoço e tamanho do tórax. Interprete os resultados do computador. A equação de regressão múltipla resultante é adequada para fazer predições do peso de um urso, com base no comprimento da cabeça, na largura da cabeça, no tamanho do pescoço e no tamanho do tórax? Por quê ou porque não?

Ao caminhar (com uma fita métrica mas sem balança) na trilha de Grizzly Lake, no Parque Nacional de Yellowstone, o autor certa vez encontrou um urso que ele conseguiu dominar e derrubar, deixando-o temporariamente inconsciente. (Nota do Editor: A verdade é que o autor viu um urso à distância de 250 jardas.) Dados todos os resultados precedentes, qual é o melhor peso predito para o urso que media 12,3 in. de comprimento de cabeça, 4,9 in. de largura da cabeça, 19,0 in. de comprimento do pescoço, 55,3 in. de comprimento e 31,4 in. de tórax?

DOS DADOS À DECISÃO

Que Computador o Leitor Deve Comprar?

A tabela a seguir relaciona preços (em dólares) emparelhados com taxas de desempenho e características de uma amostra aleatória de computadores de porte médio. Com os métodos deste capítulo, analise os dados amostrais. Determine se há correlação entre preço e classificação.

Computador	Classificação	Preço
Mem	81	2300
NMC	80	2000
Blue	86	2400
Swan	87	2600
HP-X	85	2800
GST	94	2500
Uni	90	2700
Sys	96	2300
EPS	96	2300
Mega	88	2100

Baseado em dados do *PC Magazine*.

Quanto deve custar um computador cuja taxa de desempenho/características é 105? Qual a classificação mais elevada possível, se o leitor só dispõe de $3000? Qual é a melhor aquisição: Um computador Dell Dimensão P133c (designado como Delc) ou um Robotech P133 (designado como Robo)? Por quê?

Computador	Classificação	Preço
IBM	109	3500
AMAX	104	4000
Gate	103	2425
Delx	103	2990
Robo	98	3300
Vekt	98	2300
HiQ	95	2750
Delc	100	2400
HP-V	94	3300
ACMA	96	3200

ATIVIDADES EM GRUPO

1. *Atividades Extra-classe:* Divida em grupos de 3 ou 4 pessoas. Investigue a relação entre duas variáveis coletando seus próprios dados amostrais emparelhados e aplicando os métodos deste capítulo para determinar se há correlação linear significativa. Identifique também a equação de regressão e descreva um processo para predizer valores de uma das variáveis, quando são dados valores da outra. Tópicos sugeridos:
   - Há alguma relação entre gosto e custo de diferentes marcas de tabletes de chocolate (ou colas)? (O gosto pode ser medido em uma escala, digamos, de 1 a 10.)
   - Há alguma relação entre os salários de jogadores profissionais de beisebol (ou basquetebol, ou rugby) e seus desempenhos em campo?
   - Taxas *versus* pesos: Há alguma relação entre as taxas de consumo de combustível e os pesos dos carros? Em caso afirmativo, qual é?
   - Há alguma relação entre os comprimentos dos pés dos homens (ou mulheres) e suas alturas?
   - Há alguma relação entre as notas médias dos estudantes e o tempo gasto vendo televisão? Em caso afirmativo, qual é?
   - Há alguma relação entre as alturas dos pais (ou mães) e as alturas de seus primeiros filhos (ou filhas)?
2. *Atividade em Classe:* Divida em grupos de 8 a 12 pessoas. Para cada grupo, *meça* sua altura e sua envergadura. Para a medição da envergadura, a pessoa deve ficar de pé com os braços distendidos como as asas de um avião. É fácil marcar a altura e a envergadura em um quadro-negro, e então medir as distâncias. Com os dados amostrais emparelhados, nota-se alguma correlação entre altura e envergadura? Em caso afirmativo, determine a equação de regressão, com a altura expressa em termos da envergadura. A envergadura pode ser usada como um preditor razoável da altura?
3. *Atividade em Classe:* Divida em grupos de 8 a 12 pessoas. Para cada membro de um grupo, registre a taxa de pulsação, anotando o número de batidas cardíacas em um minuto. Registre também a altura. Há alguma relação entre a taxa de pulsação e a altura? Em caso afirmativo, qual?
4. *Atividade em Classe:* Divida em grupos de 8 a 12 pessoas. Para cada membro de um grupo, use uma fita métrica e uma régua para medir a circunferência da cabeça e o comprimento do antebraço. Há alguma relação entre essas duas variáveis? Em caso afirmativo, qual?

entrevista

David Hall

Division Statistical Manager at the Boeing Commercial Airplane Group

David Hall é gerente da Divisão de Estatística, em Renton, do Boeing Commercial Airplane Group. Administra a Organização de Métodos Estatísticos, que focaliza a aplicação da estatística e de outras técnicas para melhorar continuamente a qualidade. Antes de entrar para a Boeing, Hall trabalhou no Battelle Pacific Northwest Laboratories, onde administrou um grupo de aplicações estatísticas.

Qual a extensão do uso da estatística na Boeing: Está aumentando, diminuindo ou permanece estável?

A utilização da estatística é ampla e está crescendo continuamente. Estou certa de que isto é verdade para a indústria aeronáutica em geral. Há uma conscientização da necessidade de melhorar, e para melhorar devemos dispor de dados. A estatística está alcançando cada vez mais destaque como instrumento na busca da melhoria de qualidade. A utilização de cartas de controle e, de modo mais geral, de controle de processos estatísticos, está crescendo. São comuns os experimentos planejados. No início, 99% dos instrumentos de estatística úteis são simples, muito simples. À medida que os processos se refinam e o conhecimento aumenta, tornam-se necessários instrumentos mais sofisticados. Análise de regressão e correlação, análise da variância, tabelas de contingência, teste de hipóteses, intervalos de confiança, análise de séries temporais — virtualmente todas as técnicas são empregadas em alguma ocasião. Mas à vista do estado atual da indústria americana, podem-se obter incríveis vantagens com os instrumentos mais simples, como os diagramas de Pareto.

Como se faz amostragem na Boeing?

Nosso Departamento de Controle de Qualidade utiliza muitos dos esquemas bem conhecidos de amostragem para inspeção. Verificamos os detalhes de tudo. Mas a amostragem está também incorporada às aplicações de controle de processos estatísticos. A amostragem pode ser bastante complexa e é tratada caso a caso.

Pode citar um exemplo específico de como o uso da estatística foi valioso para a Boeing?

Estivemos trabalhando com nossa Divisão de Fabricação para produzir peças mais consistentes pelo processo de hidroprensagem. Com a utilização de experimentos planejados, constatamos que o tipo de borracha colocada nos espaços vazios durante a formação poderia reduzir dramaticamente a variação de peça para peça. Isto é apenas um exemplo do que ocorre continuadamente. Exemplo de uma aplicação mais sofisticada foi o uso da reamostragem para estimar a variabilidade em um túnel de testes aerodinâmica.

Na manutenção da qualidade, o objetivo é zerar os defeitos ou é mais eficiente permitir a presença de defeitos que devam ser rejeitados ou retrabalhados?

Nunca é mais eficiente permitir defeitos que fiquem sujeitos à inspeção. É muito dispendioso. Frisamos constantemente que devemos produzir a fim de atingir o alvo. As coisas devem andar direito, e nosso maior empenho tem sido no sentido de reduzir a variação. Também em relação a fornecedores externos, temos um programa para fazer com que eles procurem continuamente reduzir a variação em relação ao valor alvo. Eventualmente, na indústria eletrônica, permite-se uma elevada taxa de defeitos, mas para nós isto é totalmente inaceitável.

Acha que os candidatos a empregos são favorecidos se tiverem estudado estatística?

Naturalmente temos um grande número de engenheiros na Boeing e quase todos eles estudaram alguma estatística. Esperamos que agora nossos administradores tenham maior familiaridade com a estatística do que no passado. Eles devem entender a variação e saber como usar os dados eficientemente.

Tem algum conselho para o estudante de hoje?

Quando admito estatísticos saídos diretamente da faculdade, quer se trate de B.S. (bacharéis), M.S. (mestres) ou Ph.D. (doutores), todos eles ainda têm um grande aprendizado pela frente, antes de se tornarem eficientes em nosso ambiente de trabalho. Eles necessitam sobretudo de habilidade ao lidar com as pessoas, espírito de equipe, planejamento e comunicação. Atualmente, a indústria americana está às voltas com a contratação de pessoas que compreendam a estatística e tenham capacidade de comunicar os resultados de suas aplicações.

Rayane Andrade

Triola

10 Experimentos Multinomiais e Tabelas de Contingência

10-1 Aspectos Gerais

Identificam-se os objetivos do capítulo. Este capítulo compreende dados amostrais que consistem em contagens de freqüências dispostas em uma linha com ao menos três categorias (tabela simples de freqüência), ou uma tabela com ao menos duas linhas e ao menos duas colunas (tabela de freqüência de dupla entrada).

10-2 Experimentos Multinomiais: Aderência

Apresenta-se o teste de hipóteses como processo de aderência para dados amostrais em uma tabela simples de freqüência. Este processo é usado para testar afirmações de que freqüências amostrais observadas se ajustam, ou aderem, a uma determinada distribuição.

10-3 Tabelas de Contingência: Independência e Homogeneidade

Definem-se e descrevem-se as tabelas de contingência (ou tabelas de freqüência de dupla entrada). Apresenta-se um método-padrão para testar afirmações de que, em uma tabela de contingência, a variável linha e a variável coluna são independentes. Apresenta-se também um método de teste de homogeneidade, pelo qual é testada uma afirmação de que diferentes populações têm as mesmas proporções de determinadas características.

Problema do Capítulo

As taxas de mortalidade não resultantes de combate são as mesmas para as tropas tanto em uma zona de combate como fora dela?

Fez-se uma comparação entre as taxas de mortalidade, não resultantes de combate, para tropas americanas em uma zona de combate e fora dela. A Tabela 10-1 resume em três categorias diferentes os dados amostrais relativos às mortes. Uma análise dos valores da tabela mostra que, entre os que morreram em conseqüência de ferimentos não-intencionais (acidentes de carro, de avião, explosão etc.), 19% estavam em zona de combate; dos que morreram de doença, 10% estavam em zona de combate; e dos que morreram em conseqüência de homicídio ou suicídio, 3% estavam em zona de combate. Essas diferenças são significativas?

Após as operações militares Desert Shield e Desert Storm, no Golfo Pérsico, alegou-se um aumento dramático da mortalidade entre as tropas americanas naquela região. Dados como os da Tabela 10-1 permitem-nos analisar a situação. Na Seção 10-3 estudaremos este conjunto particular de dados amostrais.

Tabela 10-1 Causas Extracombate da Morte de Tropas Americanas em Zona de Combate e Fora Dela

	Causa da Morte		
	Ferimento Não-intencional	Doença	Homicídio ou Suicídio
Em zona de combate	183	30	11
Fora da zona de combate	784	264	308

Os valores da tabela baseiam-se em dados do "Comparative Mortality Among U.S. Military Personnel in the Persian Gulf Region and Worldwide During Operations Desert Shield and Desert Storm", de Writer, DeFraites e Brundage, Journal of the American Medical Association, Vol. 275, N.º 2.

10-1 Aspectos Gerais

No Capítulo 1 definimos dados categóricos (ou qualitativos, ou atributos) como dados que podem ser separados em categorias diferentes (comumente chamadas **células**) que se distinguem por alguma característica não-numérica. Por exemplo, podemos separar uma amostra de M&Ms por categorias de cor: vermelho, laranja, amarelo, marrom, azul e verde. Obtida a freqüência de cada categoria, passamos a testar a afirmação de que as freqüências se ajustam (ou aderem) à distribuição de cores alegada pelo fabricante (Mars, Inc.). De modo geral, este capítulo focaliza a análise de dados categóricos que consistem em contagens de freqüências para as diferentes categorias. Na Seção 10-2 abordaremos os experimentos multinomiais, que consistem em valores dispostos em uma única linha ou coluna (chamadas tabelas simples de freqüência). Na Seção 10-3 consideraremos tabelas de contingência (ou tabelas de freqüência de dupla entrada), formadas por contagens de freqüências dispostas como na Tabela 10-1.

Veremos que os dois tópicos principais deste capítulo (experimentos multinomiais e tabelas de contingência) utilizam a mesma estatística de teste χ^2 (qui-quadrado). Recordemos as seguintes propriedades da distribuição qui-quadrado:

1. Ao contrário das distribuições normal e t de Student, a distribuição qui-quadrado não é simétrica. (Veja Figura 10-1.)
2. Os valores da distribuição qui-quadrado podem ser 0 ou positivos, mas nunca negativos. (Veja Figura 10-1.)
3. Há uma distribuição qui-quadrado diferente para cada número de graus de liberdade. (Veja Figura 10-2.)

Na Tabela A-4 encontram-se os valores críticos da distribuição qui-quadrado.

10-2 Experimentos Multinomiais: Aderência

Na Seção 4-3 definimos um experimento binomial como um experimento com um número fixo de provas independentes,

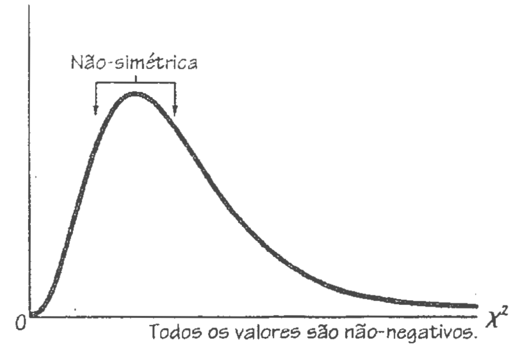


Fig. 10-1 A distribuição qui-quadrado.

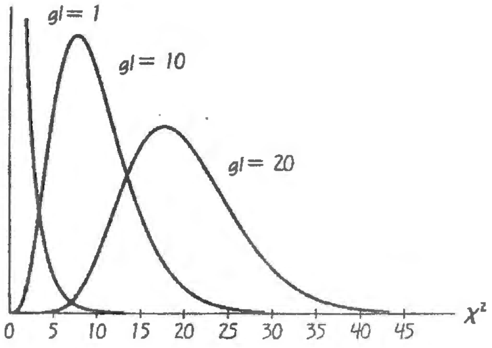


Fig. 10-2 Distribuições qui-quadrado para 1, 10 e 20 graus de liberdade.

probabilidades de resultados que permanecem constantes de uma prova para outra e pertencem a uma de duas categorias. A exigência de *duas* categorias se reflete no prefixo *bi* do termo *binomial*. Nesta seção vamos estudar experimentos multinomiais, em que cada prova dá resultados que pertencem a uma dentre várias (mais de duas) categorias. Afora esta diferença, os experimentos binomiais e multinomiais são essencialmente os mesmos.

DEFINIÇÃO

Um **experimento multinomial** é um experimento que verifica as seguintes condições:

1. O número de provas é fixo.
2. As provas são independentes.
3. Todos os resultados de cada prova devem ser classificados em exatamente uma dentre várias categorias.
4. As probabilidades para as diferentes categorias permanecem constantes (as mesmas) em cada prova.

Nesta seção apresentamos um método para testar a afirmação de que, em um experimento multinomial, as freqüências observadas nas diferentes categorias se ajustam a determinada distribuição. Como se trata de testar quão bem uma distribuição de freqüências observadas se adapta a uma distribuição teórica, este método costuma chamar-se um teste de aderência.

DEFINIÇÃO

Utiliza-se um **teste de aderência** para testar a hipótese de que uma distribuição de freqüências observadas se ajusta (ou adere) a determinada distribuição teórica.

Por exemplo, com uma calculadora TI-83, geramos aleatoriamente inteiros de 0 a 9 inclusive. Podemos testar uma amostra de resultados para determinar se eles se ajustam a uma distribuição uniforme discreta em que os algarismos 0, 1, 2, ..., 9 ocorrem com a mesma freqüência. Eis alguns outros exemplos da aplicação do teste de aderência:

- Teste a afirmação de que os números da loteria do Estado de Nova York (1, 2, ..., 54) ocorrem com a mesma freqüência.
- Teste a afirmação de Mars, Inc., de que seus confeitos M&M são feitos de maneira que 30% são marrons, 20% são amarelos, 20% vermelhos, 10% laranja, 10% verdes e 10% azuis.

Em nossos testes de aderência utilizamos a notação seguinte:

Notação

O representa a *freqüência observada* de um resultado.

E representa a *freqüência esperada* de um resultado.

k representa o *número de categorias, ou resultados, diferentes.*

n representa o *número total de provas.*

Em uma situação típica que exige um teste de aderência, temos freqüências *observadas* (denotadas por O) e devemos utilizar a distribuição teórica requerida para determinar as freqüências *esperadas* (denotadas por E). Em muitos casos podemos achar uma freqüência esperada multiplicando a probabilidade p de uma categoria pelo número n de provas diferentes:

$$E = np$$

Por exemplo, para testar se um dado é viciado ou não, jogando-o 60 vezes, temos $n = 60$ (porque há 60 provas) e $p = 1/6$ (porque um dado não é viciado se os seis resultados, ou faces, possíveis ocorrem com a mesma probabilidade 1/6). A freqüência esperada de cada categoria ou célula é, pois,

$$\begin{aligned} E &= np \\ &= (60)(1/6) = 10 \end{aligned}$$

Suposições Valem as seguintes suposições ao testarmos a proporção populacional alegada para cada uma de k categorias (em um experimento multinomial).

1. Os dados constituem uma amostra aleatória.
2. Os dados amostrais consistem em contagens de freqüências para as k categorias diferentes.
3. Para cada uma das k categorias, a freqüência *esperada* é, no mínimo, 5. (Não há qualquer exigência de que cada freqüência *observada* seja no mínimo igual a 5.)

Como as freqüências amostrais tipicamente acusam algum afastamento dos valores que teoricamente esperamos, apresentamos a questão-chave: As diferenças entre os valores efetivamente *observados O* e os valores teoricamente *esperados E* são estatisticamente significativas? Para responder a esta questão, aplicamos a seguinte estatística de teste, que mede a discrepância entre as freqüências observadas e esperadas.

Estatística de Teste de Aderência em Experimentos Binomiais

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

Valores Críticos

1. Na Tabela A-4 encontram-se os valores críticos, tomando-se $k - 1$ graus de liberdade, onde k é o número de categorias.
2. Os testes de hipótese de aderência são sempre unilaterais à direita.

A forma da estatística de teste χ^2 é tal que uma *estreita concordância* entre valores observados e esperados conduz a um valor *pequeno* de χ^2. Um grande valor de χ^2 indica forte discordância entre valores observados e esperados. Assim, um valor significativamente grande de χ^2 conduz à rejeição da hipótese nula, de que não há diferença entre as freqüências observadas e as esperadas. Nosso teste é, portanto, unilateral à direita, porque o valor crítico e a região crítica estão situados na extrema direita da distribuição. Ao contrário dos testes de hipótese anteriores, em que devíamos determinar se o teste era unilateral esquerdo, unilateral direito ou bilateral, os testes de aderência são sempre unilaterais à direita.

Desde que saibamos como achar valores da estatística de teste e valor crítico, podemos testar hipóteses aplicando o mesmo processo introduzido no Capítulo 7 e resumido na Figura 7-4.

Os Assentos Mais Seguros nos Aviões

Muitos de nós acreditamos que os assentos traseiros dos aviões são os mais seguros em caso de acidente. Os técnicos de segurança, entretanto, não concordam em que qualquer parte do avião seja mais segura do que outra. Alguns aviões batem primeiro com o nariz quando aterrissam, mas outros batem primeiro com a cauda na decolagem. Matt McCormick, técnico em sobrevivência do National Transportation Safety Board, afirmou à revista *Travel* que "não há um lugar seguro para sentar". Pode-se aplicar um teste de aderência, com a hipótese nula de que todas as partes do avião são igualmente seguras. Os aviões acidentados poderiam ser divididos nas partes dianteira, média e traseira. As freqüências observadas de acidentes fatais poderiam então ser comparadas com as freqüências esperadas em uma distribuição uniforme de acidentes. A estatística de teste χ^2 reflete o vulto das discrepâncias entre freqüências observadas e esperadas, e pode revelar se algumas partes do avião são mais seguras do que outras.

EXEMPLO Muitas pessoas acreditam que, quando um cavalo inicia uma corrida, tem mais chance de ganhar se sua posição na linha de partida está mais próxima do limite interno da pista. A posição de partida 1 está mais próxima do limite interno, seguida pela posição 2, e assim por diante. A Tabela 10-2 relaciona o número de vitórias de cavalos nas diferentes posições de partida. Teste a afirmação de que as probabilidades de vitória não são as mesmas para as diferentes posições de partida.

TABELA 10-2 Posições de Partida de Cavalos Vencedores

	Posição de Partida							
	1	2	3	4	5	6	7	8
Número de vitórias	29	19	18	25	17	10	15	11

Com base em dados do *New York Post*.

SOLUÇÃO A Tabela 10-2 relaciona os resultados para 144 vitórias. Se a chance de vitória em cada posição de partida é a mesma, a probabilidade de vitória para cada posição é $p = 1/8$ e o número esperado de vitórias para cada posição é

$$E = np = (144)(1/8) = 18$$

As freqüências observadas (denotadas por O) são 29, 19, 18, ..., 11, mas as freqüências esperadas (denotadas por E) são 18, 18, 18, ..., 18. Tendo identificado as freqüências observadas e esperadas, passemos ao nosso processo-padrão para teste de hipóteses.

Passo 1: A afirmação original é que as probabilidades de vitória são diferentes, conforme as posições de partida. Isto é, ao menos uma das probabilidades $p_1, p_2, ..., p_8$ é diferente das outras.

Passo 2: Se a afirmação original é falsa, então todas as probabilidades são iguais; isto é, $p_1 = p_2 = ... = p_8$.

Passo 3: A hipótese nula deve conter a condição de igualdade; portanto,

$$H_0: p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = p_7 = p_8$$

H_1: Ao menos uma das probabilidades é diferente das outras.

Passo 4: Como não se especificou nível de significância, tomamos $\alpha = 0{,}05$, uma escolha bastante comum.

Passo 5: Como estamos testando uma afirmação sobre a distribuição das vitórias para as diferentes posições de largada, aplicamos o teste de aderência descrito nesta seção. Utiliza-se a distribuição χ^2 com a estatística de teste dada anteriormente.

Passo 6: Com as freqüências observadas O e as freqüências esperadas E, calculamos o valor da estatística de teste χ^2, conforme Tabela 10-3. A estatística de teste é $\chi^2 = 16{,}333$ (arredondada). O valor crítico é $\chi^2 = 14{,}067$ (encontrado na Tabela A-4, com $\alpha = 0{,}05$ na extremidade direita e $k - 1 = 7$ graus de liberdade). A Figura 10-3 exibe a estatística de teste e o valor crítico.

Passo 7: Como a estatística de teste está na região crítica, há evidência suficiente para rejeitar a hipótese nula.

Passo 8: Há evidência suficiente para apoiar a afirmação de que as probabilidades de vitória são diferentes, conforme a posição de largada. Parece que a posição de largada deveria ser levada em conta ao tentar escolher qual cavalo ganhará uma corrida.

Podem-se usar as técnicas desta seção para testar se uma distribuição de freqüências observadas se ajusta a alguma distribuição de freqüência teórica. Para a posição do vencedor no caso da corrida de cavalos, aplicamos um teste de aderência para decidir se as freqüências observadas se ajustam a uma distribuição uniforme, e constatamos que não havia tal ajuste. Como muitas análises estatísticas exigem uma população distribuída normalmente, podemos aplicar o teste qui-quadrado desta seção para

TABELA 10-3 Cálculo da Estatística de Testes χ^2 para os Dados da Corrida de Cavalos

Posição de Largada	Freqüência Observada O	Freqüência Esperada E	$O - E$	$(O - E)^2$	$\frac{(O-E)^2}{E}$
1	29	18	11	121	6,7222
2	19	18	1	1	0,0556
3	18	18	0	0	0
4	25	18	7	49	2,7222
5	17	18	−1	1	0,0556
6	10	18	−8	64	3,5556
7	15	18	−3	9	0,5000
8	11	18	−7	49	2,7222
	144	144			$\chi^2 = \Sigma\frac{(O-E)^2}{E} = 16{,}3334$

(Esses dois totais devem ser iguais.)

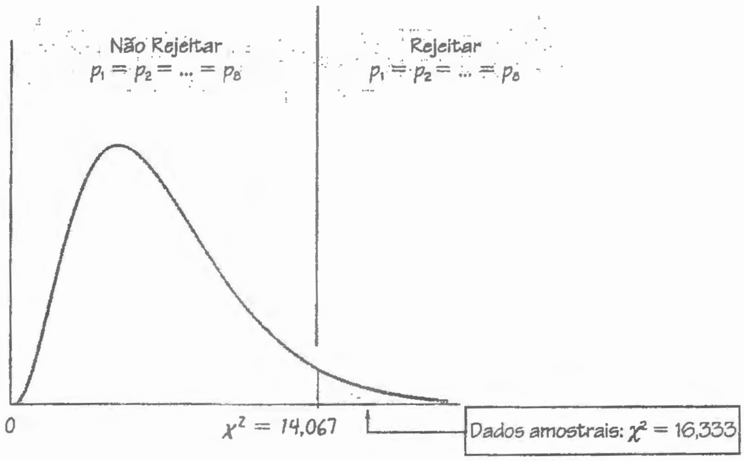


Fig. 10-3 Teste de aderência de $p_1 = p_2 = \ldots = p_8$.

determinar se as amostras dadas são extraídas de populações assim distribuídas (veja Exercício 22).

O exemplo precedente baseou-se na hipótese nula de que as probabilidades são todas iguais para as diferentes posições de largada. Podemos aplicar os métodos desta seção também no caso de as probabilidades (ou freqüências) teóricas supostas serem diferentes, como no exemplo a seguir.

Será que Mendel Adulterou Seus Dados?

R. A. Fisher analisou os resultados dos experimentos de Mendel com hibridização e constatou que os dados observados estavam demasiadamente próximos dos resultados esperados. Diz Fisher: "Os dados foram sem dúvida sofisticados sistematicamente, e após examinar várias possibilidades, não tenho dúvida de que Mendel foi ludibriado por um auxiliar de jardinagem, que sabia muito bem o que seu chefe esperava de cada experimento." Fisher aplicou o teste qui-quadrado e concluiu que havia uma probabilidade de apenas 0,00004 de uma concordância tão grande entre resultados esperados e resultados observados.

EXEMPLO Mars, Inc. afirma que seus confeitos M&M apresentam a seguinte distribuição de cores: 30% marrom, 20% amarelo, 20% vermelho, 10% laranja, 10% verde e 10% azul. As cores dos confeitos M&M relacionadas no Conjunto de Dados 11 do Apêndice B estão resumidas na Tabela 10-4. Com os dados amostrais e um nível de 0,05 de significância, teste a afirmação de que a distribuição de cores é a afirmada por Mars, Inc.

TABELA 10-4 Freqüências dos Confeitos M&M

	Marrom	Amarelo	Vermelho	Laranja	Verde	Azul
Freqüência Observada	33	26	21	8	7	5
Freqüência Esperada	30	20	20	10	10	10

SOLUÇÃO Ampliamos a Tabela 10-4 de modo a incluir as freqüências esperadas, que se calculam como se segue. Para n, usamos o número total de provas (100), que é o número total de M&Ms observados na amostra. Para as probabilidades, tomamos as frações decimais equivalentes às percentagens alegadas (30%, 20%, ..., 10%).

Marrom: $E = np = (100)(0,30) = 30$

Amarelo: $E = np = (100)(0,20) = 20$

.
.
.

Azul: $E = np = (100)(0,10) = 10$

Ao testar a afirmação, os passos 1, 2 e 3 resultam nas seguintes hipóteses:

H_0: $p_{mr} = 0,3$, $p_{am} = 0,2$, $p_{vrm} = 0,2$, $p_{lar} = 0,1$, $p_{vrd} = 0,1$ e $p_{az} = 0,1$
H_1: Ao menos uma das proporções acima é diferente do valor alegado.

Os passos 4, 5 e 6 levam-nos a aplicar o teste de aderência com o nível de 0,05 de significância e uma estatística de teste calculada pela Tabela 10-5.

A estatística de teste é $\chi^2 = 5,950$. O valor crítico de χ^2 é 11,071, que se obtém da Tabela A-4 (com $\alpha = 0,05$ à direita e $k - 1 = 5$ graus de liberdade). A Figura 10-4 mostra a estatística de teste e o valor crítico. Como a estatística de teste não está na região crítica, não há evidência suficiente para rejeitarmos a afirmação de que as cores se distribuem com as percentagens alegadas por Mars, Inc.

Na Figura 10-5 marcamos as proporções alegadas de 0,30, 0,20, ..., 0,10, juntamente com as proporções observadas de 33/100, 26/100, 21/100, 8/100, 7/100, 5/100, o que nos permite visualizar a discrepância entre a distribuição alegada e as freqüências efetivamente observadas. Os pontos ao longo da linha sólida representam as proporções alegadas, e os pontos ao longo da linha tracejada representam as proporções observadas. Os pares de pontos correspondentes são todos razoavelmente próximos, o que mostra que todas as freqüências esperadas estão bastante próximas das freqüências observadas correspondentes. De modo geral, gráficos como os da Figura 10-5 ajudam a comparar visualmente freqüências esperadas e freqüências observadas, sugerindo também quais categorias são responsáveis pelas principais discrepâncias.

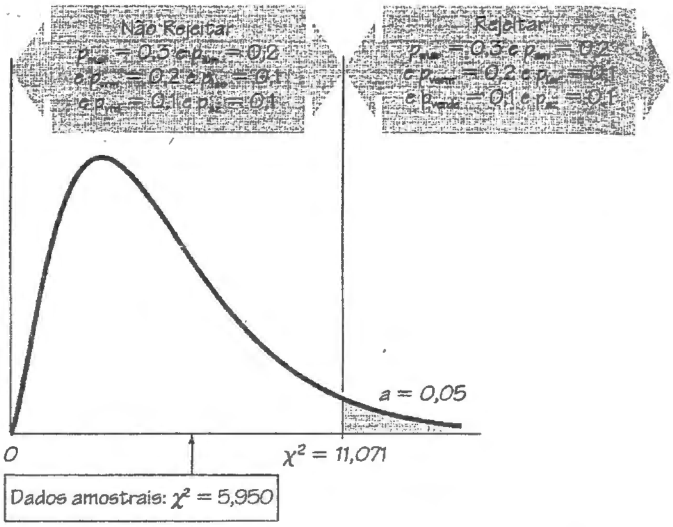


Fig. 10-4 Teste de aderência de $p_{mr} = 0,3$ e $p_{am} = 0,2$ e $p_{vrm} = 0,2$ e $p_{lar} = 0,1$ e $p_{vrd} = 0,1$ e $p_{az} = 0,1$.

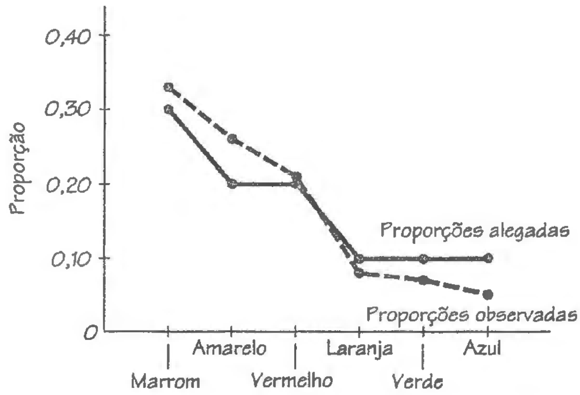


Fig. 10-5 Comparação entre proporções alegadas e observadas.

TABELA 10-5 Cálculo da Estatística de Teste χ^2 para os Dados M&M

Categoria Cor	Freqüência Observada O	Freqüência Esperada $E = np$	$O - E$	$(O - E)^2$	$\frac{(O-E)^2}{E}$
Marrom	33	30	3	9	0,3000
Amarelo	26	20	6	36	1,8000
Vermelho	21	20	1	1	0,0500
Laranja	8	10	−2	4	0,4000
Verde	7	10	−3	9	0,9000
Azul	5	10	−5	25	2,5000
	100	100			↑

$$\chi^2 = \Sigma\frac{(O - E)^2}{E} = 5,9500$$

Os exemplos precedentes contribuem para dar um sentido ao papel da estatística de teste χ^2. O que pretendemos medir é o grau de discrepância entre freqüências observadas e esperadas. A simples soma das diferenças entre valores observados e esperados não resulta em uma medida efetiva, porque essa soma é sempre 0, como se vê a seguir.

$$\Sigma(O - E) = \Sigma O - \Sigma E = n - n = 0$$

Com os quadrados de $O - E$ temos uma estatística melhor, que reflete as diferenças entre freqüências observadas e esperadas. (Elevamos ao quadrado os valores de $O - E$ essencialmente pelas mesmas razões por que tomamos os quadrados de $x - \bar{x}$ na fórmula do desvio-padrão.) O valor de $\Sigma(O - E)^2$ mede apenas a magnitude das diferenças, mas precisamos achar a magnitude das diferenças em relação ao que é esperado. Obtém-se essa magnitude relativa através da divisão pelas freqüências esperadas, como na estatística de teste.

A distribuição teórica de $\Sigma(O - E)^2/E$ é uma distribuição discreta, porque o número de valores possíveis é limitado. A distribuição pode ser aproximada por uma distribuição qui-quadrado, que é contínua. Esta aproximação é em geral considerada aceitável, desde que todos os valores de E sejam pelo menos 5. Esta exigência foi incluída nas hipóteses que se aplicam a esta seção. Na Seção 5-6 vimos que a distribuição de probabilidade normal, contínua, é uma aproximação razoável da distribuição de probabilidade binomial, discreta, desde que np e nq sejam ambos no mínimo 5. Vemos agora que a distribuição qui-quadrado, contínua, pode ser uma aproximação razoável da distribuição discreta de $\Sigma(O - E)^2/E$, desde que todos os valores de E sejam no mínimo 5. (Há maneiras de contornar o problema de uma freqüência esperada ser inferior a 5, como por exemplo combinar categorias de modo que a soma das freqüências esperadas não seja inferior a 5.)

O número de graus de liberdade reflete o fato de podermos atribuir livremente freqüências a $k - 1$ categorias antes de determinar a freqüência de cada categoria. Embora possamos atribuir "livremente" freqüências a $k - 1$ categorias, não podemos ter freqüências negativas nem tampouco freqüências tão grandes que sua soma exceda o total das freqüências observadas para todas as categorias combinadas.

Retardando a Morte

David Phillips, sociólogo da Universidade da Califórnia, estudou a capacidade de certas pessoas de retardarem sua morte até depois de um acontecimento importante. Analisando as taxas de mortalidade de homens judeus que morreram na proximidade da Páscoa, constatou que a taxa de mortalidade caía sensivelmente na semana antes da Páscoa, elevando-se a partir da semana seguinte. Phillips observou fenômeno análogo entre as mulheres sino-americanas; sua taxa de mortalidade caía na semana anterior ao Festival Lunar da Colheita, para se elevar na semana seguinte.

Valores *P*

Nos exemplos desta seção utilizou-se a abordagem tradicional do teste de hipóteses, mas é possível também abordar o problema através dos valores *P*. Esses valores podem ser obtidos pelos mesmos métodos descritos nas Seções 7-3 e 7-6. Assim é que o exemplo precedente resultou em uma estatística de teste $\chi^2 = 5{,}950$. Como havia $k = 6$ categorias, tínhamos $k - 1 = 5$ graus de liberdade. Consultando a Tabela A-4, vemos que, para a linha de 5 graus de liberdade, a estatística de teste 5,95 é inferior ao valor crítico direito de 9,236, de forma que o valor *P* é superior a 0,10. Se fizermos o exemplo precedente utilizando STATDISK, o resultado apresentará um valor *P* de 0,31111. Com uma calculadora TI-83, obtemos, para $\chi^2 = 5{,}950$ e 5 graus de liberdade, o valor *P* 0,311.

O elevado valor *P* sugere que a hipótese nula não deve ser rejeitada. Recorde que rejeitamos a hipótese nula somente quando o valor *P* não supera o nível de significância.

Programas de Computador para Experimentos Multinomiais

STATDISK é planejado para analisar experimentos multinomiais, mas estes não podem ser analisados com Minitab ou com a calculadora TI-83. Para utilizar STATDISK, selecione `Analysis` na barra de menu principal, e em seguida a opção `Multinomial Experiments`.

10-2 Exercícios A: Habilidades e Conceitos Básicos

1. Uma forma comum de testar a autenticidade de dados é analisar a freqüência de algarismos. Quando as pessoas se pesam e seus pesos são arredondados para a libra mais próxima, é de se esperar que os últimos algarismos 0,1,2,...,9 ocorram aproximadamente com a mesma freqüência. Em contraste, quando se pergunta a uma pessoa qual é o seu peso, os algarismos 0 e 5 tendem a ocorrer com maior freqüência. O autor selecionou aleatoriamente 80 estudantes, obteve seus pesos e registrou apenas o último algarismo de cada um, com os resultados apresentados na tabela a seguir. Com o nível de significância de 0,01, teste a afirmação de que os últimos algarismos ocorrem com a mesma freqüência. Os resultados sugerem que os estudantes tenham sido realmente pesados, ou se limitaram a reportar seus pesos?

Último algarismo	0	1	2	3	4	5	6	7	8	9
Freqüência	35	0	2	1	4	24	1	4	7	2

2. Usa-se uma calculadora TI-83 para gerar aleatoriamente números inteiros entre 0 e 9 inclusive, obtendo-se os resultados a seguir. Ao nível de 0,05 de significância, teste a afirmação de que a calculadora TI-83 gera algarismos distribuídos uniformemente.

Algarismo	0	1	2	3	4	5	6	7	8	9
Freqüência	32	29	37	21	41	24	33	32	25	26

3. Um exercício da Seção 3-4 envolveu quatro estudantes que perderam uma prova e procuraram se justificar alegando um pneu furado em seu carro. Para apurar a veracidade da alegação, o professor pediu aos estudantes que dissessem qual pneu tinha furado. Se eles realmente não tiveram qualquer pneu furado, seria possível que todos eles indicassem o mesmo pneu? O autor pediu a 40 estudantes que identificassem o pneu que escolheriam. Os resultados constam da tabela a seguir (à exceção de um estudante, que escolheu o sobressalente). No nível de 0,05 de significância, teste a hipótese do autor de que os dados se ajustam a uma distribuição uniforme. O que sugere o resultado quanto à possibilidade de os estudantes mencionarem o mesmo pneu, quando na realidade não tiveram nenhum pneu furado?

Pneu	Dianteiro esquerdo	Dianteiro direito	Traseiro esquerdo	Traseiro direito
Número de estudantes	11	15	8	6

4. É comum crença de que ocorre um maior número de acidentes fatais com automóveis em determinados dias da semana, como sexta-feira ou sábado. Seleciona-se aleatoriamente uma amostra de mortes em acidentes com automóveis no estado de Montana, em ano recente. Relacionam-se abaixo os números de casos fatais para os diferentes dias da semana. No nível de 0,05 de significância, teste a afirmação de que a freqüência da ocorrência de acidentes é a mesma em qualquer dia da semana.

Dia	Dom	Seg	Ter	Qua	Qui	Sex	Sab
Número de ocorrências fatais	31	20	20	22	22	29	36

Com base em dados do Insurance Institute por Highway Safety.

5. Muitas pessoas acreditam que os acidentes fatais com carros em conseqüência de DWI (*Driving While Intoxicated*) são devidos a bebedores ocasionais que abusam das bebidas nas noites de sexta-feira e sábado, enquanto outros acham que tais acidentes são devidos a alcoólatras, que ingerem bebida diariamente. Em um estudo de acidentes fatais com carros, selecionaram-se aleatoriamente 216 casos em que o motorista acusava percentual de álcool no sangue superior a 0,10. Esses casos foram separados pelos diversos dias da semana, com os resultados exibidos na tabela abaixo. No nível de 0,05 de significância, teste a afirmação de que esse tipo de acidente ocorre com a mesma freqüência em qualquer dia da semana. A evidência apóia a teoria de que os casos fatais de DWI são devidos a bebedores ocasionais ou a bebedores contumazes?

Dia	Dom	Seg	Ter	Qua	Qui	Sex	Sab
Número	40	24	25	28	29	32	38

Com base em dados do Programa STOP-DWI do Condado de Dutchess.

6. Fez-se um estudo de 147 acidentes industriais que exigiram tratamento médico. Desses acidentes, 31 ocorreram na segunda-feira, 42 na terça, 18 na quarta, 25 na quinta, e 31 na sexta (com base em dados de "Counted Data CUSUM's", por Lucas, *Technometrics*, Vol. 27, N.° 2). Teste a afirmação de que os acidentes ocorrem com a mesma proporção nos cinco dias da semana. Em caso contrário, que fatores podem explicar a diferença?

7. Com o nível de 0,05 de significância e os dados de acidentes industriais do Exercício 6, teste a afirmação de um técnico de segurança de que os acidentes se distribuem pelos dias úteis como se segue: 30% na segunda-feira, 15% na terça, 15% na quarta, 20% na quinta e 20% na sexta. A rejeição dessa afirmação contribui para uma correção do problema de acidentes industriais?

8. O gerente do Supermercado Gleason deve decidir a quantidade de cada sabor de sorvete que deve estocar a fim de atender à demanda dos consumidores, sem que haja perda de sabores menos procurados. O fornecedor de sorvete afirma que, entre os sabores mais populares, os clientes têm as seguintes preferências: 62% preferem baunilha, 18% preferem chocolate, 12% preferem napolitano e 8% preferem baunilha com calda. Uma amostra de 200 clientes acusou os resultados a seguir. Com o nível de 0,05 de significância, teste se o fornecedor identificou corretamente as preferências dos consumidores.

Sabor	Baunilha	Chocolate	Napolitano	Baunilha em Calda
Clientes	120	40	18	22

Dados baseados em resultados da International Association por Ice Cream Manufacturers.

9. O número π é um número irracional com a propriedade de que, quando procuramos expressá-lo em forma decimal, é necessário um número infinito de casas decimais, sem que haja qualquer padrão de repetição. Na representação decimal de π, os 100 primeiros algarismos ocorrem com as freqüências dadas na tabela a seguir. Com o nível de 0,05 de significância, teste se os algarismos se distribuem de maneira uniforme.

Algarismo	0	1	2	3	4	5	6	7	8	9
Freqüência	8	8	12	11	10	8	9	8	12	14

10. O número 22/7 é análogo a π no sentido de que ambos exigem um número infinito de casas decimais. Todavia, 22/7 é um número racional porque pode ser expresso como quociente de dois inteiros, o que não ocorre com π. Quando expressamos em forma decimal números como 22/7, há um padrão de repetição. Na representação decimal de 22/7, os 100 primeiros algarismos ocorrem com as freqüências dadas na tabela a seguir. Com o nível de 0,05 de significância, teste se os algarismos se distribuem uniformemente. Qual a diferença entre este resultado e o do Exercício 9?

Algarismo	0	1	2	3	4	5	6	7	8	9
Freqüência	0	17	17	1	17	16	0	16	16	0

11. Entre os motoristas que tiveram um acidente de carro no ano passado, selecionam-se aleatoriamente 88, que são classificados por idade, com os resultados constantes da tabela a seguir. Se todas as idades têm a mesma taxa de acidente, é de se esperar (em virtude da distribuição das idades dos motoristas licenciados) que as categorias dadas abranjam 16%, 44%, 27% e 13% dos indivíduos, respectivamente. Com o nível de 0,05 de significância, teste se a distribuição de acidentes está de acordo com a distribuição de idades. Há algum grupo etário que acuse taxa aparentemente desproporcional de acidentes?

Idade	Menos de 25	25-44	45-64	Mais de 64
Motoristas	36	21	12	19

Com base em dados do Insurance Information Institute.

12. Em uma análise de impactos de bombas V-1 na Segunda Guerra Mundial, o sul de Londres foi subdividido em regiões, cada uma com área de 0,25 km². Na Seção 4-5 apresentamos um exemplo e incluímos uma tabela das freqüências efetivas de impactos e das freqüências esperadas pela distribuição de Poisson. Com base nos valores relacionados aqui, teste se as freqüências observadas se ajustam a uma distribuição de Poisson. Tome o nível de 0,05 de significância.

N.° de ataques a bomba	0	1	2	3	4 ou mais
N.° real de regiões	229	211	93	35	8
N.° esperado de regiões (da distribuição de Poisson)	227,5	211,4	97,9	30,5	8,7

13. Consulte o Conjunto de Dados 16 do Apêndice B referente ao gêiser Old Faithful, e teste se os intervalos de tempo se distribuem uniformemente entre as cinco categorias 55-64, 65-74, 75-84, 85-94, 95-104.

14. Consultando o Conjunto de Dados 8 do Apêndice B, registre apenas os últimos algarismos das contagens de pulso. Verifique a autenticidade das contagens, testando se os últimos algarismos ocorrem com a mesma freqüência. As contagens de pulso parecem ser autênticas? (Veja o Exercício 1.)

15. Com base no Conjunto de Dados 10 do Apêndice B, categorize os filmes ali constantes como fracos (0,0 a 1,5 estrelas), razoáveis (2,0 ou 2,5 estrelas), bons (3,0 a 3,5 estrelas) ou excelentes (4,0 estrelas). O crítico de cinema James Harrington afirma que os filmes se distribuem igualmente pelas quatro categorias. Teste essa afirmação, com o nível de 0,05 de significância.

16. Maryland seleciona honestamente os números de sua loteria? Consulte o Conjunto de Dados 12 do Apêndice B e use os 150 algarismos relacionados para a Loteria Pick Three. Utilizando 10 categorias cor-

respondentes aos 10 algarismos possíveis, ache a freqüência de cada categoria. Ao nível de 0,05, teste a afirmação de que a loteria é honesta, no sentido de que os 10 algarismos ocorrem com a mesma freqüência.

10-2 Exercícios B: Além do Básico

17. Que sabe o leitor sobre o valor P para o teste de hipótese no Exercício 16?
18. Ao fazer um teste de aderência conforme descrito nesta seção, suponha que multipliquemos cada freqüência observada pelo mesmo inteiro positivo maior do que 1. Como é afetado o valor crítico? Como é afetada a estatística de teste?
19. Neste exercício mostraremos que um teste de hipótese envolvendo um experimento multinomial com apenas duas categorias é equivalente a um teste de hipótese para uma proporção (Seção 7-5). Suponha que um experimento multinomial particular tenha apenas dois resultados possíveis A e B com freqüências observadas f_1 e f_2, respectivamente.
    a. Determine uma expressão para a estatística de teste χ^2, e ache o valor crítico para um nível de 0,05 de significância. Suponha que estamos testando se ambas as categorias têm a mesma freqüência $(f_1 + f_2)/2$.
    b. Usa-se a estatística

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

para testar se uma proporção populacional é igual a um certo valor p. Com $p = 0{,}5$, $\alpha = 0{,}05$ e

$$\hat{p} = \frac{f_1}{f_1 + f_2}$$

mostre que z^2 é equivalente a χ^2 [da parte (a)]. Mostre também que o quadrado do escore crítico z é igual ao valor crítico χ^2 da parte (a)

20. A seguir tem-se uma distribuição de freqüências observadas:

N.º de sucessos	0	1	2	3
Freqüência	89	133	52	26

    a. Supondo uma distribuição binomial com $n = 3$ e $p = 1/3$, aplique a fórmula da probabilidade binomial para achar a probabilidade correspondente a cada categoria da tabela.
    b. Com as probabilidades obtidas na parte (a), determine a freqüência esperada de cada categoria.
    c. Com o nível de 0,05 de significância, teste se as freqüências observadas se ajustam a uma distribuição binomial com $n = 3$ e $p = 1/3$.

21. Em ano recente, houve 116 mortes por homicídio em Richmond, Virginia (com base em "A Classroom Note On the Poisson Distribution: A Model for Homicidal Deaths in Richmond, VA for 1991" *Mathematics and Computer Education*, por Winston A. Richards). Se as freqüências de mortes nos diferentes dias se ajustam a uma distribuição de Poisson, se apresentarão como na tabela a seguir. Com o nível de 0,05 de significância, teste se as freqüências observadas se ajustam a uma distribuição de Poisson. (*Cuidado*: Nem todas as freqüências esperadas são no mínimo 5.)

	Número de Homicídios				
	0	1	2	3	4
Número efetivo de dias	268	79	17	1	0
Número esperado de dias	265,6	84,4	13,4	1,4	0,1

22. A seguir temos uma distribuição de freqüências observadas para uma amostra de QI's:

Escore de QI	Menos de 80	80-95	96-110	111-120	Mais de 120
Freqüência	20	20	80	40	40

    a. Supondo uma distribuição normal com $\mu = 100$ e $\sigma = 15$, aplique os métodos do Capítulo 5 para achar a probabilidade de um indivíduo selecionado aleatoriamente pertencer a cada classe. (Utilize as fronteiras de classe de 79,5, 95,5, 110,5, 120,5.)
    b. Com as probabilidades obtidas na parte (a), ache a freqüência esperada de cada categoria.
    c. Com o nível de 0,01 de significância, teste se os escores de QI foram selecionados aleatoriamente de uma população distribuída normalmente com $\mu = 100$ e $\sigma = 15$.

10-3 Tabelas de Contingência: Independência e Homogeneidade

Os exemplos e exercícios da Seção 10-2 envolvem freqüências dispostas em uma única linha (ou coluna), de acordo com a categoria. Nesta seção também estão em jogo freqüências por categoria, mas aqui vamos considerar casos com pelo menos duas linhas e pelo menos duas colunas. Veja, por exemplo, a Tabela 10-1, que tem duas linhas (em combate e fora de combate) e três colunas (as três categorias de causa de morte). As tabelas como a Tabela 10-1 são geralmente chamadas *tabelas de contingência*, ou *tabelas de freqüência de dupla entrada*.

DEFINIÇÃO

Uma **tabela de contingência** (ou **tabela de freqüência de dupla entrada**) é uma tabela em que as freqüências correspondem a duas variáveis. (Uma variável categoriza as linhas, a outra categoriza as colunas.)

As tabelas de contingência têm importância especial porque são usadas freqüentemente para analisar resultados de pesquisas. Por isso, os métodos apresentados nesta seção estão entre os mais usados.

Nesta seção apresentamos dois tipos de teste de hipótese baseados em tabelas de contingência. Consideramos primeiro testes de independência, usados para determinar se uma variável linha de uma tabela de contingência é independente de sua variável coluna. Abordamos em seguida testes de homogeneidade para testar se diferentes populações têm a mesma proporção de determinada característica. Uma boa notícia: Os dois tipos de teste de hipótese utilizam *os mesmos* métodos básicos. Começamos com testes de independência.

Teste de Independência

DEFINIÇÃO

Um **teste de independência** testa a hipótese nula de que a variável linha e a variável coluna em uma tabela de contingência não estão relacionadas, isto é, são independentes.

É de suma importância reconhecer que, neste contexto, a palavra contingência se refere a dependência, mas trata-se apenas de uma dependência estatística, e não pode ser usada para estabelecer uma ligação direta de causa e efeito entre as duas variáveis em questão. Por exemplo, após analisar os dados da Tabela 10-1, poderíamos concluir pela existência de uma relação entre a causa da morte e o deslocamento do pessoal militar para uma zona de combate, mas isto não significa que o deslocamento afete diretamente a causa da morte.

Suposições Ao testarmos a hipótese nula de independência entre as variáveis linha e coluna em uma tabela de contingência, aplicam-se as seguintes suposições. (Note que estas suposições não exigem que a população original tenha distribuição normal nem qualquer outro tipo de distribuição.)

1. Os dados amostrais são selecionados aleatoriamentre.
2. A hipótese nula H_0 é a afirmação de que as variáveis linha e coluna são *independentes*; a hipótese alternativa H_1 afirma que as variáveis linha e coluna são dependentes.
3. Para cada célula na tabela de contingência, a freqüência *esperada* E é no mínimo 5. (Não há tal exigência para as freqüências observadas.)

Nosso teste de independência entre as variáveis linha e coluna utiliza a seguinte estatística de teste:

Estatística para um Teste de Independência

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

Valores Críticos

1. Os valores críticos encontram-se na Tabela A-4, com

   graus de liberdade $= (r - 1)(c - 1)$

   onde r é o número de linhas e c é o número de colunas.
2. Os testes de independência com tabelas de contingência envolvem apenas regiões críticas unilaterais à direita.

A estatística de teste permite-nos medir o grau de discordância entre as freqüências efetivamente observadas e as freqüências que deveríamos esperar teoricamente no caso de as variáveis serem independentes. Pequenos valores da estatística de teste χ^2 indicam acentuada concordância entre as freqüências observadas e as freqüências esperadas com variáveis linha e coluna independentes. Grandes valores de χ^2 encontram-se à direita da distribuição qui-quadrado e refletem diferenças significativas entre freqüências observadas e esperadas. Em grandes amostragens repetidas, a distribuição da estatística de teste χ^2 pode ser aproximada pela distribuição qui-quadrado, desde que todas as freqüências esperadas sejam no mínimo 5. O número de graus de liberdade $(r-1)(c-1)$ traduz o fato de que, em virtude de conhecermos o total das freqüências em uma tabela de contingência, podemos atribuir livremente freqüências a apenas $r - 1$ linhas e $c - 1$ colunas antes de determinar a freqüência de cada célula. [Entretanto, não podemos ter freqüências negativas, nem freqüências tão grandes que a soma de qualquer linha (ou coluna) exceda o total de freqüências observadas para aquela linha (ou coluna).]

Na seção anterior, conhecíamos as probabilidades correspondentes e podíamos facilmente determinar os valores esperados, mas a tabela de contingência típica não apresenta as probabilidades de interesse. Assim, temos de estabelecer um método para obter os valores esperados correspondentes. Primeiramente iremos descrever o processo para achar os valores das freqüências esperadas, para em seguida justificá-lo. A freqüência esperada E de cada célula da tabela de freqüências pode ser calculada com auxílio da equação abaixo:

Freqüência Esperada para uma Tabela de Contingência

$$\textit{Freqüência esperada} = \frac{(\textit{total de linhas})\,(\textit{total de colunas})}{(\textit{total geral})}$$

Aqui, o *total geral* se refere ao total de todas as freqüências observadas na tabela. Por exemplo, a freqüência esperada da célula superior esquerda da Tabela 10-6 (uma reprodução da Tabela 10-1 com as freqüências esperadas inseridas entre parênteses) é 137,09, que se obtém notando que o total de todas as freqüências para aquela linha é 224, o total das freqüências da coluna é 967 e o total de todas as freqüências da tabela é 1580; obtemos assim uma freqüência esperada de

$$E = \frac{(\textit{total de linhas})(\textit{total de colunas})}{(\textit{total geral})} = \frac{(224)(967)}{1580} = 137{,}09$$

EXEMPLO A freqüência esperada da célula superior esquerda da Tabela 10-6 é 137,09. Ache a freqüência esperada da célula inferior esquerda, admitindo independência entre a causa da morte e a condição da pessoa (em situação de combate ou não).

TABELA 10-6 Freqüências Observadas (e Freqüências Esperadas)

	Causa da Morte			
	Ferimento Acidental	Doença	Homicídio ou Suicídio	**Totais de linhas**
Em zona de combate	183 (137,09)	30 (41,68)	11 (45,23)	224
Fora da zona de combate	784 (829,91)	264 (252,32)	308 (273,77)	1356
Totais de Colunas:	**967**	**294**	**319**	**Total Geral: 1580**

SOLUÇÃO A célula inferior esquerda está na segunda linha (com o total de 1356) e na primeira coluna (com o total de 967). A freqüência esperada é

$$E = \frac{(total\ de\ linhas)(total\ de\ colunas)}{(total\ geral)} = \frac{(1356)(967)}{1580} = 829,91$$

Seria conveniente parar aqui e verificar que as outras freqüências esperadas são 41,68, 45,23, 252,32, e 273,77.

Para melhor entender a lógica da determinação de freqüências esperadas por este processo, suponha que conheçamos apenas os totais por linha e por coluna, e que devamos preencher as freqüências esperadas das células admitindo a independência (ou ausência de relacionamento) entre as duas variáveis em jogo — ou seja, suponhamos que são conhecidos apenas os totais de linhas e de colunas mostrados na Tabela 10-6. Comecemos com a célula no canto superior esquerdo. Como 224 dentre as 1580 foram mobilizadas, temos P(mobilizada) = 224/1580. Analogamente, 967 pessoas morreram em conseqüência de ferimento acidental; assim, P(morte por ferimento acidental) = 967/1580. Como estamos admitindo a independência entre a causa da morte e a condição da pessoa (mobilizada ou não), podemos aplicar a regra da multiplicação das probabilidades, obtendo

$$P(\text{mobilizado e morto por ferimento acidental}) = \frac{224}{1580} \cdot \frac{967}{1580}$$

Esta equação é uma aplicação da regra de multiplicação para eventos independentes, que se expressa em geral como se segue: $P(A \text{ e } B) = P(A) \cdot P(B)$. Encontrada uma expressão para a probabilidade de estar na célula superior esquerda, passamos agora à determinação do *valor esperado* para aquela célula, que se obtém multiplicando-se a probabilidade da mesma pelo número total de pessoas, conforme a equação a seguir:

$$E = p \cdot n = \left[\frac{224}{1580} \cdot \frac{967}{1580}\right] \cdot 1580 = 137,09$$

A forma deste produto sugere uma maneira geral de obter a freqüência esperada de uma célula:

$$frequência\ esperada\ E = \frac{(total\ de\ linhas)}{(total\ geral)} \cdot \frac{(total\ de\ colunas)}{(total\ geral)} \cdot (total\ geral)$$

Esta expressão se reduz a

$$E = \frac{(total\ de\ linhas) \cdot (total\ de\ colunas)}{(total\ geral)}$$

Passamos agora a utilizar dados de uma tabela de contingência para testar hipóteses, como no exemplo a seguir, que usa os dados do Problema do Capítulo.

EXEMPLO Com os dados da Tabela 10-1 e com o nível de significância de 0,05, teste a afirmação de que a causa de morte fora de combate é independente da condição de a pessoa estar mobilizada em uma zona de combate.

SOLUÇÃO As hipóteses nula e alternativa são:

H_0: A causa da morte é independente de a pessoa estar mobilizada em uma zona de combate.

H_1: Há dependência entre a causa da morte e a condição de a pessoa estar mobilizada.

O nível de significância é $\alpha = 0,05$.

Como os dados estão em forma de tabela de contingência, aplicamos a distribuição χ^2 com esta estatística de teste:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$
$$= \frac{(183 - 137,09)^2}{137,09} + \frac{(30 - 41,68)^2}{41,68} + \frac{(11 - 45,23)^2}{45,23} + \frac{(784 - 829,91)^2}{829,91} + \frac{(264 - 252,32)^2}{252,32} + \frac{(308 - 273,77)^2}{273,77}$$
$$= 15,3748 + 3,2731 + 25,9052 + 2,5397 + 0,5407 + 4,2798 = 51,913$$

(Obtém-se a estatística de teste mais precisa 51,905 mantendo-se um maior número de decimais nos cálculos intermediários. A calculadora TI-83, STATDISK e Minitab estão todos acordes em que 91,505 é um resultado melhor.) O valor crítico é $\chi^2 = 5,991$, que se obtém na Tabela A-4 notando que $\alpha = 0,05$ na cauda direita e que o número de graus de liberdade é dado por $(r-1)(c-1) = (2-1)(3-1) = 2$. A Figura 10-6 mostra a estatística de teste e o valor crítico. Como a estatística de teste está na região crítica, rejeitamos a hipótese nula de que a causa da morte seja independente do fato de a pessoa estar mobilizada em uma zona de combate. Parece haver uma relação entre esses dois fatos.

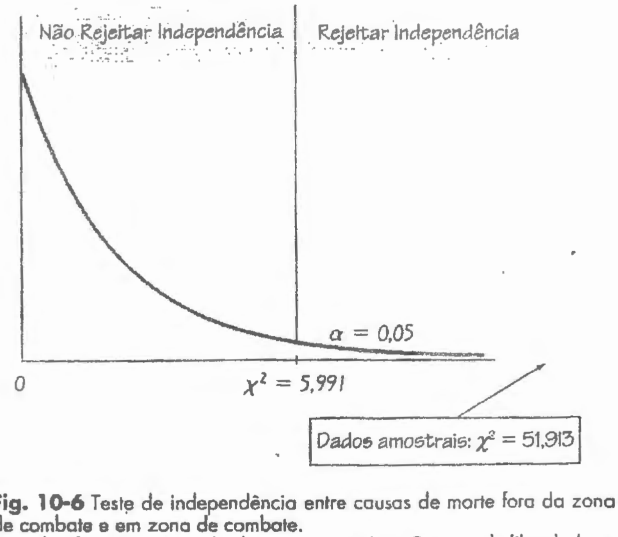


Fig. 10-6 Teste de independência entre causas de morte fora da zona de combate e em zona de combate.
Se a distribuição qui-quadrado tem apenas 1 ou 2 graus de liberdade, a forma da distribuição é a apresentada aqui.

Utilização de Computadores e Calculadoras com Tabelas de Contingência

Em vista do volume de cálculo exigido, é conveniente utilizar uma calculadora ou um programa estatístico para testes do tipo discutido nesta seção. A calculadora TI-83 permite introduzir os dados como uma matriz, calcula e exibe os seguintes valores: a estatística de teste χ^2,o valor P e o número de graus de liberdade. As apresentações STATDISK e Minitab precedentes mostram os resultados obtidos para os dados da Tabela 10-1. Se o leitor estiver usando STATDISK, deve selecionar `Analysis` na barra do menu principal, em seguida selecionar `Contingency Tables` e passar a introduzir os dados conforme solicitados. Se estiver usando Minitab, primeiro introduza as freqüências observadas em colunas, em seguida selecione `Stat` da barra do menu principal, a seguir a opção `Tables`, e passe a introduzir os nomes das colunas que contêm as freqüências observadas, tais como C1 C2. Minitab fornece a estatística de teste e o valor P, mas não o valor crítico nem a conclusão. Os resultados de STATDISK incluem a estatística de teste, o valor crítico, o valor P e a conclusão.

Valores P

No exemplo precedente utilizamos a abordagem tradicional do teste de hipóteses, mas podemos facilmente utilizar o processo do valor P. A TI-83, STATDISK e Minitab fornecem todos o mesmo valor P 0,000 (quando arredondado). Utilizando o critério dado na Seção 7-3, rejeitamos a hipótese nula porque o valor P é inferior ao nível de significância $\alpha = 0{,}05$.

Se não dispusermos de uma calculadora adequada ou de um programa estatístico, poderemos estimar os valores P utilizando os mesmos métodos introduzidos anteriormente. O exemplo precedente resultou em uma estatística de teste $\chi^2 = 51{,}913$, e o valor crítico se baseou em dois graus de liberdade. Consultando a Tabela A-4, nota-se que, para a linha de 2 graus de liberdade, a

STATDISK

File Edit Analysis Data Help

Degrees Freedom	2
Test Statistic, χ^2	51.905
Critical χ^2	5.9915
P-Value	0.000000

Reject the Null Hypothesis

Data provides evidence that the rows and columns are related

MINITAB

```
Expected counts are printed below observed counts

                 C1       C2       C3      Total
    1           183       30       11        224
             137.09    41.68    45.23

    2           784      264      308       1356
             829.91   252.32   273.77

Total           967      294      319       1580

ChiSq =      15.372 + 3.274 + 25.901 +
              2.539 + 0.541 +  4.279 =    51.905
df = 2.      p = 0.000
```

estatística de teste $\chi^2 = 51,913$ está à direita do maior valor crítico 10,597, superando-o, de forma que o valor P deve ser inferior a 0,005, o que expressamos como: valor $P < 0,005$. Com base neste valor P relativamente pequeno, novamente rejeitamos a hipótese nula e concluímos que há evidência amostral suficiente para rejeitar a afirmação de que a causa de morte e a mobilização são independentes. Se o valor P fosse maior do que o nível de significância de 0,05, não rejeitaríamos a hipótese nula de independência.

Teste de Homogeneidade

No exemplo precedente, ilustramos um teste de independência com uma amostra de 1580 indivíduos de uma única população de pessoal militar. Em alguns casos, as amostras são extraídas de populações diferentes; devemos então determinar se essas populações têm as mesmas proporções das características em estudo.

DEFINIÇÃO

Em um **teste de homogeneidade**, testa-se a afirmação de que diferentes populações apresentam as mesmas proporções de determinadas características.

Como um teste de homogeneidade utiliza dados amostrais extraídos de populações diferentes, temos totais predeterminados, ou para as linhas ou para as colunas, na tabela de contingência. Conseqüentemente, um teste de homogeneidade envolve escolhas aleatórias feitas de modo que ou os totais de linhas, ou os totais de colunas, são predeterminados. Para fazer distinção entre um teste de homogeneidade e um teste de independência, podemos colocar a seguinte questão:

Os tamanhos das amostras usadas para diferentes populações foram predeterminados (teste de homogeneidade), ou extraiu-se uma grande amostra de modo que tanto os totais de linhas como os de colunas foram determinados aleatoriamente (teste de independência)?

Como exemplo de teste de homogeneidade, suponha que queiramos testar a afirmação de que a proporção de eleitores Republicanos é a mesma nos estados de Nova York, Califórnia, Texas e Iowa. Se optarmos por determinar a filiação política de 200 habitantes de Nova York, 250 californianos, 100 texanos e 80 de Iowa, então na tabela de contingência que resume os resultados estarão predeterminados os totais por linhas ou os totais por coluna (na verdade, os que representam os diferentes estados); os valores serão 200, 250, 100, 80.

Ao fazermos um teste de homogeneidade, podemos aplicar os mesmos processos já apresentados nesta seção, conforme ilustrado no exemplo a seguir.

EXEMPLO O sexo de um pesquisador tem influência nas respostas dadas por homens a uma pesquisa? Em um artigo do *U.S.News & World Report* sobre pesquisas, afirmou-se que "Em assuntos delicados, as pessoas preferem dar respostas 'aceitáveis' a dar respostas honestas; suas respostas podem depender do sexo ou da raça do entrevistador." Para fundamentar esta afirmação, compilaram-se dados para uma pesquisa do Instituto Eagleton, na qual os pesquisados deveriam responder se estavam de acordo com a seguinte afirmação: "O aborto é um problema de natureza privada, cuja decisão deve ficar a critério da mulher, sem interferência do governo." Analisaremos o efeito do sexo somente sobre pesquisados do sexo masculino. A Tabela 10-7 se baseia nas respostas de homens pesquisados. Suponha que a pesquisa tenha sido planejada de modo que os entrevistadores homens devam obter 800 respostas de homens, e as entrevistadoras mulheres devam obter 400 respostas também de homens. Com o nível de 0,05 de significância, teste a afirmação de que as proporções de resposta concordo/discordo são as mesmas tanto para os entrevistados por homens como para os entrevistados por mulheres.

TABELA 10-7 Sexo e Respostas à Pesquisa

	Sexo do Entrevistador	
	Homem	Mulher
Homens que concordam	560	308
Homens que discordam	240	92

MINITAB DISPLAY

```
Expected counts are printed below observed counts

            C1        C2     Total
  1        560       308      868
        578.67    289.33

  2        240        92      332
        221.33    110.67

Total      800       400     1200

ChiSq =   0.602 + 1.204 +
          1.574 + 3.149 = 6.529
df = 1, p = 0.011
```

SOLUÇÃO Como temos totais de coluna predeterminados, de 800 pesquisados entrevistados por homens e 400 entrevistados por mulheres, testamos a homogeneidade com as seguintes hipóteses:

H_0: As proporções de respostas concordo/discordo são as mesmas tanto para os entrevistados por homens como para os entrevistados por mulheres.

H_1: As proporções são diferentes.

O nível de significância é $\alpha = 0{,}05$. Aplicamos a mesma estatística de teste χ^2 já mencionada, e calculada pelo mesmo processo. Em lugar de relacionar os detalhes do cálculo, damos a apresentação Minitab que resulta dos dados da Tabela 10-7.
A apresentação Minitab (veja página anterior) mostra as freqüências esperadas de 578,67, 289,33, 221,33 e 110,67. Mostra, também, a estatística de teste $\chi^2 = 6{,}529$ e o valor P 0,011. Com a abordagem do teste de hipóteses pelo valor P, rejeitamos a hipótese nula de proporções iguais (homogêneas). Há evidência suficiente para rejeitar a afirmação de que as proporções são as mesmas. Parece haver dependência entre a resposta do pesquisado e o sexo do entrevistador. Aparentemente os homens são influenciados pelo sexo do entrevistador, embora esta conclusão seja uma afirmação de causalidade que não é justificada pela análise estatística.

10-3 Exercícios A: Habilidades e Conceitos Básicos

1. A Tabela 10-7 resume os dados referentes a pesquisados do sexo masculino, mas a tabela abaixo resume os dados relativos a uma amostra de mulheres. Com o nível de 0,01 de significância, e supondo que os tamanhos amostrais de 800 homens e 400 mulheres sejam predeterminados, teste a afirmação de que as proporções de respostas concordo/discordo são as mesmas tanto para os entrevistados por homens como para os entrevistados por mulheres.

	Sexo do Entrevistador	
	Homem	Mulher
Mulheres que concordam	512	336
Mulheres que discordam	288	64

Com base em dados do Eagleton Institute.

2. No caso judicial *Estados Unidos vs. Cidade de Chicago*, foram postas em dúvida as práticas honestas de emprego. Um grupo minoritário (A) e um grupo majoritário (B) fizeram o exame para capitão do Corpo de Bombeiros, com os seguintes resultados:

	Aprovados	Reprovados
Grupo A	10	14
Grupo B	417	145

Com os resultados acima e com um nível de significância de 0,05, teste a afirmação de que o sucesso no teste é independente do grupo.

3. Nicorette é um chiclete que ajuda fumantes a deixarem de fumar cigarros. A tabela a seguir mostra os resultados de testes feitos para detectar reações negativas. Com o nível de 0,05 de significância, teste a afirmação de que o tratamento (remédio ou placebo) é independente da reação (se o paciente experimenta, ou não, irritação na boca ou na garganta). Se o leitor está pensando em recorrer a Nicorette para deixar de fumar, deve se preocupar com aqueles efeitos colaterais?

	Remédio	Placebo
Irritação na boca ou garganta	43	35
Nenhuma irritação na boca ou garganta	109	118

Com base em dados de Merrell Dow Pharmaceuticals, Inc.

4. Fez-se um estudo de 531 pessoas feridas em acidentes de bicicleta; os resultados de uma amostra aleatória constam da tabela abaixo. Com o nível de 0,05 de significância, teste a afirmação de que o uso do capacete não reduz a possibilidade de ferimentos no rosto. Com base nos resultados, acha que o uso do capacete ajuda a reduzir o risco de ferimentos no rosto?

	Com Capacete	Sem Capacete
Com ferimentos faciais	30	182
Todos os ferimentos não-faciais	83	236

Com base em dados de "A Case-Control Study of the Effectiveness of Bicycle Safety Helmets in Preventing Facial Injury", por Thompson, Thompson, Rivara e Wolf, *American Journal of Public Health*, Vol. 80, Nº 12.

5. Fez-se uma pesquisa para determinar se há restrições, quanto ao sexo, na confiança que o povo deposita na polícia. Os resultados amostrais constam da tabela a seguir. Com o nível de 0,05 de significância, teste a afirmação de que não há tal restrição.

	Confiança na Polícia		
	Muita	Alguma	Muito pouca ou nenhuma
Homens	115	56	29
Mulheres	175	94	31

Com base em dados do Ministério da Justiça dos EUA e da Gallup Organization.

6. Em um estudo de monitoramento em lojas de departamentos, utilizaram-se amostras de compras para comparar os preços monitorados com os preços marcados. A tabela a seguir resume os resultados de 819 artigos. Quando as lojas utilizam monitores para conferir os produtos, as taxas de erro são as mesmas para os artigos com preços normais e para os artigos em promoção? Como reagiriam os consumidores em face de um excesso desproporcional nos preços dos artigos em promoção?

	Artigos com Preço Regular	Artigos com Preço Especial
Preço reduzido	20	7
Preço aumentado	15	29
Preço correto	384	364

Com base em dados de "UPC Scanner Pricing Systems: Are They Accurate?", por Ronald Goodstein, *Journal of Marketing*, Vol. 58.

7. A tabela a seguir relaciona resultados de uma pesquisa obtidos de uma amostra aleatória de vítimas de diferentes crimes. Com o nível de 0,05 de significância, teste a afirmação de que o tipo de crime é independente do fato de o criminoso ser um estranho.

	Homicídio	Roubo	Assalto
O criminoso era um estranho	12	379	727
O criminoso era conhecido ou parente	39	106	642

Com base em dados do Ministério da Justiça dos EUA.

8. Um estudo de usuários e não-usuários do cinto de segurança resultou nos dados amostrais aleatórios resumidos na tabela a seguir. Teste a afirmação de que a quantidade de fumo é independente do

uso do cinto de segurança. Uma teoria plausível é que as pessoas que fumam mais estão menos preocupadas com a sua saúde e segurança, sendo, assim, menos propensas a usar os cintos. Os dados amostrais apóiam esta teoria?

	Número de Cigarros Fumados por Dia			
	0	1-14	15-34	35 ou mais
Usam cinto de segurança	175	20	42	6
Não usam cinto de segurança	149	17	41	9

Com base em dados de "What Kinds of People Do Not Use Seat Belts?", por Helsing and Comstock, *American Journal of Public Health*, Vol. 67, N.º 11.

9. Muitas pessoas acreditam que os criminosos que se confessam culpados tendem a sofrer penas mais leves do que os que são condenados em julgamento. A tabela a seguir resume dados amostrais aleatórios de acusados de roubo em San Francisco. Todos eles já tinham sofrido penas de prisão anteriormente. Com o nível de 0,05 de significância, teste a afirmação de que a sentença (prisão ou liberdade) é independente da alegação (do acusado). Se o leitor fosse advogado de defesa de um acusado culpado, aconselharia uma confissão de culpa com base nesses resultados?

	Declaração de Culpa	Declaração de Inocência
Enviado à prisão	392	58
Mantido em liberdade	564	14

Com base em dados de "Does it Pay To Plead Guilty? Differential Sentencing and the Functioning of Criminal Courts", por Brereton and Casper, *Law and Society Review*, Vol. 16, Nº 1.

10. Um estudo de acidentes de automóvel e motoristas que usam telefone celular selecionados aleatoriamente acusou os seguintes dados amostrais. Com o nível de 0,05 de significância, teste a afirmação de que a ocorrência de acidentes é independente do uso de telefone celular. Com base nesses resultados, parece que a utilização de celulares afeta a segurança da direção?

	Com Acidente no Ano Passado	Sem Acidente no Ano Passado
Usa telefone celular	23	282
Não usa telefone celular	46	407

Com base em resultados de AT&T e da Automobile Association of America.

11. Um estudo feito para determinar a taxa de fumantes entre pessoas de diferentes grupos etários originou os dados amostrais aleatórios resumidos na tabela a seguir. Com o nível de 0,05 de significância, teste a afirmação de que o fumo é independente do grupo etário Com base nesses dados, tem sentido dirigir a propaganda de cigarros a grupos etários específicos?

	Idade (anos)			
	20-24	25-34	35-44	45-64
Fuma	18	15	17	15
Não fuma	32	35	33	35

Com base em dados do National Center for Health Statistics.

12. A tabela abaixo apresenta dados relativos ao time vencedor em diferentes esportes. Com o nível de 0,10 de singnificância, teste a afirmação de que as vitórias casa/visitante são independentes do esporte.

	Basquete	Beisebol	Hockey	Futebol
O time de casa ganha	127	53	50	57
O time visitante ganha	71	47	43	42

Com base em "Predicting Professional Sports Game Outcomes from Intermediate Game Scores", por Copper, DeNeve e Mosteller, *Chance*, Vol. 5, n.º 3-4.

13. A tabela a seguir apresenta dados amostrais usados pelo estatístico Karl Pearson em 1909. O tipo de crime aparenta ter alguma relação com o fato de o criminoso beber ou ser abstêmio? Há crimes que parecem estar associados à bebida?

	Incêndio criminoso	Estupro	Violência	Furto	Falsificação	Fraude
Bebedor	50	88	155	379	18	63
Abstêmio	43	62	110	300	14	144

14. Um estudo de pessoas que se recusaram a responder a questões de uma pesquisa originou os dados amostrais aleatórios da tabela a seguir. Com o nível de 0,01 de significância, teste a afirmação de que a cooperação do indivíduo (resposta, recusa) é independente da faixa etária. Algum grupo etário em particular parece ser menos cooperante?

	Idade					
	18-21	22-29	30-39	40-49	50-59	60 ou mais
Responde	73	255	245	136	138	202
Recusa	11	20	33	16	27	49

Com base em dados de "I Hear You Knocking But You Can't Come In", por Fitzgerald e Fuller, *Sociological Methods and Research*, Vol. 11, N.º 1.

15. Com base no Conjunto de Dados 8 do Apêndice B, teste se o sexo do estudante de estatística é independente do fato de ele ser fumante.

16. Com base no Conjunto de Dados 8 do Apêndice B, teste se o exercício físico dos estudantes de estatística é independente do sexo.

10-3 Exercícios B: Além do Básico

17. A distribuição qui-quadrado é contínua, enquanto a estatística de teste usada nesta seção é discreta. Alguns estatísticos usam a *correção de continuidade de Yates* em células com freqüência esperada inferior a 10 ou em todas as células de uma tabela de contingência com duas linhas e duas colunas. Com a correção de Yates, substituímos

$$\sum \frac{(O-E)^2}{E} \quad \text{por} \quad \sum \frac{(|O-E|-0{,}5)^2}{E}$$

Dada a tabela de contingência do Exercício 10, determine o valor da estatística de teste χ^2 com a correção de Yates e sem ela. De modo geral, qual é o efeito da correção de Yates sobre o valor da estatística de teste?

18. **a.** Para uma tabela de contingência com duas linhas e duas colunas, freqüências a e b na primeira linha e freqüênias c e d na segunda linha, verifique que a estatística de teste se escreve

$$\chi^2 = \frac{(a+b+c+d)(ad-bc)^2}{(a+b)(c+d)(b+d)(a+c)}$$

b. Seja $\hat{p}_1 = a/(a+c)$ e $\hat{p}_2 = b/(b+d)$. Mostre que a estatística de teste

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\frac{\bar{p}\,\bar{q}}{n_1} + \frac{\bar{p}\,\bar{q}}{n_2}}}$$

onde

$$\bar{p} = \frac{a+b}{a+b+c+d}$$

e

$$\bar{q} = 1 - \bar{p}$$

é tal que $z^2 = \chi^2$ [mesmo resultado da parte (a)]. (Este resultado mostra que o teste qui-quadrado envolvendo uma tabela 2×2 é equivalente ao teste da diferença entre duas proporções, descrito na Seção 8-6.)

Vocabulário

células
experimento multinomial
teste de aderência
tabela de contingência
tabela de freqüência de dupla entrada
teste de independência
teste de homogeneidade

Revisão

Na Seção 10-2 descrevemos o teste de aderência para determinar se uma linha (ou coluna) de freqüências tem determinada distribuição. Na Seção 10-3 descrevemos as tabelas de contingência e o teste de independência entre as variáveis linha e coluna, assim como o teste de homogeneidade, para testar se diferentes populações têm as mesmas proporções de certas características. Seguem alguns elementos-chave dos métodos discutidos neste capítulo.

- *Seção 10-2 (Teste de aderência):*

  A estatística de teste é $\chi^2 = \sum \frac{(O-E)^2}{E}$

  O teste é unilateral à direita com $k-1$ graus de liberdade.
  Todas as freqüências esperadas devem ser no mínimo 5.
- *Seção 10-3 (Teste de independência ou homogeneidade em uma tabela de contingência):*

  A estatística de teste é $\chi^2 = \sum \frac{(O-E)^2}{E}$

  O teste é unilateral à direita com $(r-1)(c-1)$ graus de liberdade.
  Todas as freqüências esperadas devem ser no mínimo 5.

Exercícios de Revisão

1. Um estudo do uso de cinto de segurança em táxis abrangeu 77 táxis em Nova York, 129 em Chicago e 72 em Pittsburgh, com os resultados apresentados na tabela a seguir. Um porta-voz do sindicato de motoristas de táxi argumenta que, embora pareça que poucos táxis têm cintos em condições de uso, o tamanho da amostra extraída das três cidades foi de apenas 278, de modo que os resultados não são significativos. Com o nível de 0,05 de significância, teste a afirmação de que as três cidades têm a mesma proporção de táxis com cinto de segurança em condições de uso; isto é, teste a homogeneidade das proporções de cintos de segurança em condições de uso nas três cidades.

		Nova York	Chicago	Pittsburgh
O táxi tem cinto de segurança	sim	3	42	2
em condições de uso	não	74	87	70

Com base em "The Phantom Taxi Seat Belt", por Welkon and Reisinger, *American Journal of Public Health*, Vol. 67, N.º 11.

2. Quando a revista *Time* pesquisou as mortes nos EUA em conseqüência de armas de fogo no período de uma semana, obteve os resultados da tabela a seguir. Com o nível de 0,05 de significância, teste a afirmação de que as taxas de morte por arma de fogo são as mesmas para os diferentes dias da semana. Há alguma justificativa para a teoria de que ocorrem mais mortes por arma de fogo nos fins de semana, quando mais pessoas estão em casa?

Dia da semana	Seg	Ter	Qua	Qui	Sex	Sab	Dom
Número de mortes por arma de fogo	74	60	66	71	51	66	76

3. Em um experimento de percepção extra-sensorial, pediu-se às pessoas que identificassem o mês mostrado em um calendário na sala vizinha. Com os resultados acusados abaixo, teste se os meses foram selecionados com a mesma freqüência. Admita o nível de 0,05 de significância. Se os meses parecem não ter sido selecionados com a mesma freqüência, há alguma base para a afirmação de que as pessoas têm percepção extra-sensorial?

Mês	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
Número de Pessoas	8	12	9	15	6	12	4	7	11	11	5	20

4. A tabela a seguir resume os resultados para 100 vôos selecionados aleatoriamente de cada uma das três companhias de aviação. Com o nível de 0,05 de significância, teste se a USAir, a American e a Delta têm as mesmas proporções de vôos no horário.

	USAir	American	Delta
Chegada no horário	80	77	76
Atrasado	20	23	24

Com base em dados do Ministério dos Transportes (EUA).

5. A tabela a seguir resume os resultados de testes clínicos feitos com o antialérgico Seldane. Com o nível de 0,05 de significância, teste a afirmação de que a ocorrência de dores de cabeça é independente do grupo (Seldane, placebo, controle). Com base nesses resultados, os usuários de Seldane devem preocupar-se com a dor de cabeça?

	Usuários de Seldane	Usuários de Placebo	Controle
Dor de cabeça	49	49	24
Sem dor de cabeça	732	616	602

Com base em dados de Merrell Dow Pharmaceuticals, Inc.

Exercícios Cumulativos de Revisão

1. Suponha que a Tabela 10-8 relacione escores de quatro pessoas, onde o escore x se refere a um teste de memória e o escore y se refere a um teste de raciocínio. Teste a afirmação de que há relação entre os escores x e y.
2. Suponha que a Tabela 10-8 relacione escores de teste de quatro pessoas onde o escore x se refere a um teste feito antes de uma sessão de treinamento para melhora de memória e o escore y se refere a um teste feito após o treinamento. Teste a afirmação de que a sessão de treinamento contribui para elevar os escores.
3. Suponha que na Tabela 10-8 as letras A, B, C, D representam as escolhas na primeira questão de um teste de múltipla escolha. Admita também que x representa homens e y representa mulheres, e que os valores da tabela são contagens de freqüências, de modo que 66 homens escolhem A, 77 mulheres escolhem A, 80 homens escolhem B e assim por diante. Teste a afirmação de que homens e mulheres escolhem as diferentes respostas nas mesmas proporções.
4. Suponha que, na Tabela 10-8, as letras A, B, C, D representam diferentes versões do mesmo teste de raciocínio. Os escores x fo-

TABELA 10-8

	A	B	C	D
x	66	80	82	75
y	77	89	94	84

ram obtidos por quatro homens selecionados aleatoriamente, e os escores y foram obtidos por quatro mulheres selecionadas aleatoriamente. Teste se os homens e as mulheres têm o mesmo escore médio.

Projeto para Computador

Use sua calculadora para gerar aleatoriamente 100 algarismos entre 0 e 9 inclusive; use então um programa estatístico (como STATDISK ou Minitab) para gerar aleatoriamnte outros 100 algarismos entre 0 e 9 inclusive. Registre os resultados na tabela a seguir, e teste se o algarismo selecionado é independente do fato de ter sido usada a calculadora ou o programa estatístico. Teste a afirmação utilizando STATDISK ou Minitab, e obtenha uma cópia impressa da apresentação. Interprete os resultados do computador e escreva suas conclusões.

	0	1	2	3	4	5	6	7	8	9
Calculadora										
STATDISK ou Minitab										

DOS DADOS À DECISÃO

A tabela resume mortes selecionadas aleatoriamente de homens na faixa etária 45-64. Teste a afirmação de que o fumo é independente de causa da morte. Outros dados mostram que apenas 45% dos homens naquela faixa etária são fumantes. Com os dados amostrais, teste a afirmação de que a proporção de mortes entre fumantes é $p = 0{,}45$. (*Sug.:* Consulte a Seção 7-5.) Entre os 1000 homens na faixa etária 45-64 que morreram, os fumantes estão representados de maneira desproporcional?

	Causa da Morte		
	Câncer	Doença Cardíaca	Outros
Fumante	135	310	205
Não-fumante	55	155	140

Com base em dados de "Chartbook on Smoking, Tobacco and Health", USDHEW publicação CDC75-7511.

ATIVIDADES EM GRUPO

1. *Atividade Extraclasse:* Divida a turma em grupos de quatro ou cinco estudantes. Cada membro do grupo deve pesquisar ao menos 15 estudantes do sexo masculino e 15 do sexo feminino na mesma faculdade, formulando duas perguntas: (a) Que partido político merece sua simpatia? (2) Se se tratasse de justificar uma ausência como conseqüência de um pneu furado, que pneu escolheria se o professor perguntasse? (Veja Exercício 3 da Seção 10-2.) Peça ao indivíduo que escreva as duas respostas em um cartão, registre o sexo do indivíduo e se ele escreveu com a mão direita ou com a mão esquerda. Aplique os métodos deste capítulo para analisar os dados coletados.
2. *Atividade Extraclasse:* Em uma amostragem aleatória, cada elemento da população tem a mesma chance de ser escolhido. Quando o autor pediu a 60 alunos que escolhessem "aleatoriamente" três algarismos cada um, obtiveram-se os resultados a seguir. Utilize esta amostra de 180 algarismos para testar a afirmação de que os estudantes selecionaram os algarismos aleatoriamente. Os estudantes tiveram êxito em escolher seus próprios números aleatórios? Escolha sua própria amostra de estudantes e peça a cada um que selecione três algarismos aleatoriamente. Teste a aleatoriedade de seus resultados. Teste também seus resultados para ver se concordam com os relacionados a seguir. Escreva um relatório descrevendo seu experimento e enuncie com clareza suas conclusões.
3. *Atividades em Classe:* Divida a turma em grupos de três ou quatro estudantes. Cada grupo deve receber um dado, com a recomendação de testá-lo. O dado é viciado ou não? Descreva a análise e os resultados.
4. *Atividade em Classe:* Divida em grupos de seis ou mais estudantes. Cada membro do grupo deve fornecer os algarismos de seu número no seguro social. (Os algarismos devem ser reordenados de forma a manter a privacidade.) Construa uma relação combinada desses algarismos e teste a afirmação de que os algarismos 0, 1, 2, ..., 9 ocorrem com a mesma freqüência.

213	169	812	125	749	137	202	344	496	348	714	765
831	491	169	312	263	192	584	968	377	403	372	123
493	894	016	682	390	123	325	734	316	357	945	208
115	776	143	628	479	316	229	781	628	356	195	199
223	114	264	308	105	357	333	421	107	311	458	007

Rayane Andrade

Triola

11 Análise da Variância

11-1 Aspectos Gerais

Identificam-se os objetivos do capítulo, que introduz métodos para testar a hipótese de que três ou mais populações tenham a mesma média. No Capítulo 8 testamos a igualdade entre as médias de duas populações e utilizamos estatísticas de teste que tinham distribuição normal ou distribuição *t* de Student. Os métodos deste capítulo são, entretanto, muito diferentes. Em virtude da natureza dos cálculos envolvidos, recomenda-se o uso de programas de computador e da calculadora TI-83.

11-2 ANOVA* de Um Critério

Nesta seção introduzimos o método básico de ANOVA, utilizado para testar a afirmação de que três ou mais populações têm a mesma média. ANOVA de um critério compreende diferentes amostras, categorizadas de acordo com uma única característica.

11-3 ANOVA de Dois Critérios

Na Seção 11-2 trabalhamos com um único fator (ou característica usada para estabelecer diferenças entre populações), mas nesta seção vamos considerar dois fatores. Como os cálculos são muito difíceis, daremos ênfase à interpretação de resultados obtidos por computador.

*Daqui por diante iremos nos referir à Análise da Variância abreviadamente como ANOVA, de *Analysis of Variance*. (N. do T.)

Problema do Capítulo

Será que o Old Faithful está se tornando menos confiável?

Uma das atrações naturais mais populares dos Estados Unidos é o gêiser Old Faithful, no Parque Nacional de Yellowstone. O nome do gêiser provém da regularidade dos intervalos de tempo entre duas erupções; mas estará o Old Faithful se modificando? A Tabela 11-1 relaciona os intervalos de tempo (em minutos) entre erupções para quatro anos diferentes. As quatro amostras têm médias de 63,3 min, 74,3 min, 81,7 min e 83,8 min, respectivamente, o que sugere que o tempo médio entre erupções não seja o mesmo nos quatro anos. A Figura 11-1 inclui os diagramas em caixa para as quatro amostras, e esses diagramas também sugerem que as quatro amostras se originem de populações com médias diferentes. Considerando, entretanto, que cada amostra consiste apenas em 12 valores, serão essas diferenças estatisticamente significativas?

O Capítulo 8 apresenta métodos para testar a igualdade entre duas médias populacionais. Queremos agora testar a igualdade entre quatro médias populacionais, ou seja, queremos testar a afirmação de que $\mu_{1951} = \mu_{1985} = \mu_{1995} = \mu_{1996}$. Neste capítulo introduziremos métodos para testar tais afirmações e analisaremos os dados da Tabela 11-1.

TABELA 11-1 Intervalo de Tempo (em min) entre Erupções do Gêiser Old Faithful

1951	1985	1995	1996
74	89	86	88
60	90	86	86
74	60	62	85
42	65	104	89
74	82	62	83
52	84	95	85
65	54	79	91
68	85	62	68
62	58	94	91
66	79	79	56
62	57	86	89
60	88	85	94
$n_1 = 12$	$n_2 = 12$	$n_3 = 12$	$n_4 = 12$
$\bar{x}_1 = 63{,}3$	$\bar{x}_2 = 74{,}3$	$\bar{x}_3 = 81{,}7$	$\bar{x}_4 = 83{,}8$
$s_1 = 9{,}4$	$s_2 = 14{,}2$	$s_3 = 13{,}7$	$s_4 = 10{,}9$

Com base nos dados do geólogo Rick Hutchinson e do National Park Service.

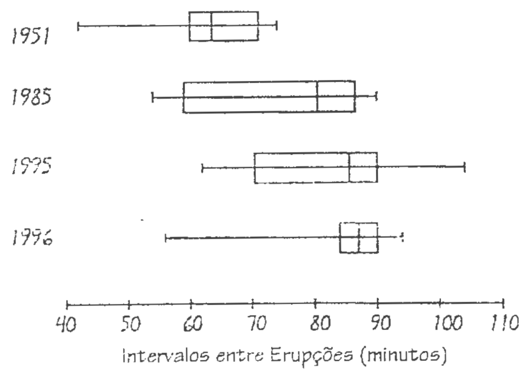


Fig. 11-1 Intervalos entre erupções (minutos).

11-1 Aspectos Gerais

Nas Seções 8-2, 8-3 e 8-5 introduzimos processos para testar a hipótese de igualdade das médias de *duas* populações (H_0: $\mu_1 = \mu_2$). Na Seção 11-2 apresentamos um procedimento para testar a igualdade das médias de *três* ou *mais* populações. (É claro que os métodos utilizados neste capítulo também podem ser usados para a igualdade entre duas médias, mas os métodos do Capítulo 8 são muito mais eficazes.) Uma hipótese nula típica na Seção 11-2 será H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$; a hipótese alternativa é H_1: Ao menos uma das médias é diferente. O método que vamos aplicar se baseia em uma análise de variâncias amostrais.

DEFINIÇÃO

A análise da variância (ANOVA) é um método para testar a igualdade de três ou mais médias populacionais, baseado na análise de variâncias amostrais.

Utiliza-se a ANOVA em aplicações tais como:

- Quando três grupos diferentes de pessoas (fumantes; não-fumantes expostos aos efeitos do fumo ambiental; e não-fumantes não expostos a essa contingência) são avaliados quanto à cotinina (um medidor da nicotina), podemos testar se eles têm o mesmo nível.
- Quando Kaplan, Princeton Review e Collins Test Prep ministram cursos preparatórios diferentes para os testes SAT, podemos testar se há diferenças nos escores SAT médios para as três populações de estudantes provenientes dos três cursos distintos.

Por que nos preocuparmos com um novo processo, quando podemos testar a igualdade de duas médias com os métodos apresentados no Capítulo 8, onde estabelecemos testes para H_0: $\mu_1 = \mu_2$? Por exemplo, se pretendemos usar os dados amostrais da Tabela 11-1 para testar a afirmação (em nível de 0,05) de que as quatro populações têm a mesma média, por que não as separamos em grupos de duas e testamos H_0: $\mu_1 = \mu_2$, H_0: $\mu_2 = \mu_3$ e assim por diante? Esta abordagem (considerar dois de cada vez) exigiria seis testes de hipóteses diferentes, de modo que o grau de confiança ficaria reduzido a $0{,}95^6$ (ou 0,735). De modo geral, à medida que aumentamos o número de testes individuais de significância, aumentamos a possibilidade de encontrar uma diferença apenas casual. O risco de um erro tipo I — encontrar uma diferença em um dos pares, quando tal diferença realmente não existe — é excessivamente elevado. O método da análise da variância permite-nos evitar este perigo (rejeitar uma hipótese nula verdadeira) pela utilização de um teste de igualdade de várias médias.

A Distribuição *F*

Os métodos de ANOVA utilizam a distribuição *F*, introduzida na Seção 8-4. Ali vimos que a distribuição *F* apresenta as seguintes propriedades importantes (veja Figura 11-2):

1. A distribuição *F* não é simétrica; é assimétrica à direita.
2. Os valores de *F* podem ser 0 ou positivos, mas nunca negativos.
3. Há uma distribuição *F* diferente para cada par de graus de liberdade (do numerador e do denominador).

Os valores críticos de *F* encontram-se na Tabela A-5.

A análise da variância (ANOVA) se baseia na comparação de duas estimativas diferentes da variância comum às diferentes populações. Na Seção 11-2 descreveremos essas estimativas (a *variância entre amostras* e a *variância dentro das amostras*). Utiliza-se a expressão *um critério* porque os dados amostrais são separados em grupos segundo uma característica, ou fator. Por exemplo, os tempos relacionados na Tabela 11-1 são separados em quatro grupos diferentes de acordo com a característica (ou fator) do ano em que foram observados. Na Seção 11-3 introduz-

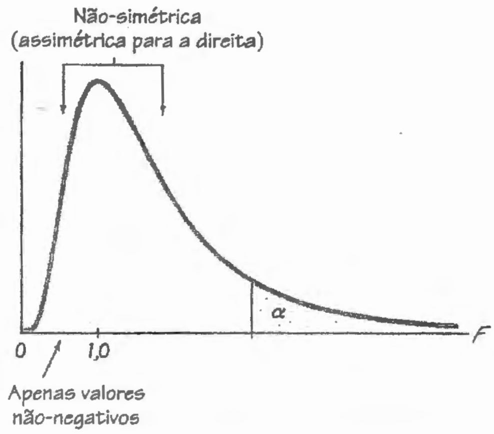


Fig. 11-2 A distribuição *F*.
Há uma distribuição *F* diferente para cada par de graus de liberdade do numerador e do denominador.

se a análise da variância de dois critérios, que permite comparar populações separadas em categorias com base em duas características (ou fatores). Por exemplo, poderíamos classificar as durações dos filmes por um dos dois critérios seguintes: (1) sua classificação por estrelas (1, 2, 3 ou 4 estrelas) e (2) a avaliação do espectador (G, PG, PG-13, R).

Como os procedimentos utilizados neste capítulo exigem cálculos complicados, recomendamos o uso de programas de computador, como STATDISK ou Minitab, juntamente com a calculadora TI-83.

11-2 ANOVA de Um Critério

À vista da complexidade dos cálculos em jogo, recomendamos a seguinte abordagem a esta seção:

1. Desenvolver uma perfeita compreensão de como interpretar o painel de um computador que relacione resultados da análise da variância.
2. Procurar entender a lógica do processo, focalizando cálculos que se apliquem a um exemplo em que as amostras tenham todas o mesmo número de valores.
3. Compreender a natureza da SQ (soma de quadrados) e do QM (quadrado médio), e seu papel na determinação da estatística de teste *F*, mas recorrer a programas estatísticos para achar esses valores.

Nesta seção vamos considerar testes de hipóteses de que três ou mais médias populacionais sejam todas iguais, como em H_0: $\mu_1 = \mu_2 = \mu_3$.

Suposições Valem as seguintes suposições quando testamos a hipótese de que três ou mais amostras provêm de populações com a mesma média:

1. As populações têm distribuições normais.
2. As populações têm a mesma variância σ^2 (ou o mesmo desvio-padrão σ).
3. As amostras são aleatórias e mutuamente independentes.
4. As diferentes amostras provêm de populações classificadas em apenas uma categoria.

As exigências de normalidade e igualdade de variâncias podem ser um tanto relaxadas, pois os métodos desta seção funcionam razoavelmente bem, a menos que a distribuição seja fortemente não-normal ou as variâncias populacionais sejam muito diferentes. O estatístico George E. P. Box, da Universidade de Wisconsin, mostrou que, desde que os tamanhos das amostras sejam iguais (ou quase iguais), a diferença entre as variâncias pode ser de tal ordem que a maior seja nove vezes a menor e, ainda assim, os resultados da ANOVA continuam a ser essencialmente confiáveis. (Se as amostras são independentes mas as distribuições fortemente não-normais, podemos utilizar o teste de Kruskal-Wallis apresentado na Seção 13-5.)

O método que utilizamos é chamado **análise da variância de um critério** (ou **critério único**) porque lançamos mão de uma única característica, ou critério, para categorizar as populações. Esta característica costuma chamar-se *tratamento*, ou *fator*.

DEFINIÇÃO

Um **tratamento** (ou **fator**) é uma característica que nos permite distinguir diferentes populações umas das outras.

Por exemplo, os tempos relacionados na Tabela 11-1 são dados amostrais extraídos de quatro populações diferentes que se distinguem segundo o tratamento (ou fator), que é o ano em que foram observados. Usa-se o termo *tratamento* porque as primeiras aplicações da análise da variância se referiam a experimentos agrícolas em que diferentes lotes de terra eram tratados com diferentes fertilizantes, tipos de semente, inseticidas etc.

Utilização de Computadores e Calculadoras para ANOVA de Um Critério

Como os cálculos exigidos pela análise da variância são extremamente complexos, quase sempre utilizamos programas de computador. Inicialmente, descrevemos a interpretação dos resultados obtidos por computador, para então considerar as bases lógicas do método de análise da variância.

EXEMPLO Com os dados amostrais da Tabela 11-1 e um nível de significância $\alpha = 0{,}05$, utilize STATDISK, ou Minitab, ou uma calculadora TI-83 para testar a afirmação de que as quatro amostras provêm de populações com a mesma média ($\mu_{1951} = \mu_{1985} = \mu_{1995} = \mu_{1996}$). Com STATDISK, selecione `Analysis` da barra do menu principal, em seguida `One-Way Analysis of Variance` e passe a introduzir os dados amostrais. Com Minitab, introduza primeiro os dados amostrais nas colunas C1, C2, C3, C4; em seguida selecione `Stat, ANOVA, ONEWAY (UNSTACKED)` e introduza `C1 C2 C3 C4` na caixa identificada como `Responses` (em colunas separadas). Com uma calculadora TI-83, introduza os dados como listas em L1, L2, L3 e L4, acione em seguida STAT, selecione TESTS e escolha a opção ANOVA.

SOLUÇÃO Mostram-se a seguir os resultados de STATDISK e Minitab. Um elemento-chave é o valor *P*, que é 0,00066 em STATDISK ou 0,001 em Minitab. Como o valor *P* não supera o nível de significância $\alpha = 0{,}05$, rejeitamos a hipótese nula de igualdade das médias. Há evidência suficiente para rejeitar a afirmação de que as quatro amostras provenham de populações com a mesma média. Parece que o tempo médio entre as erupções do Old Faithful está se modificando.

A parte inferior do quadro Minitab inclui também as médias e desvios-padrão amostrais individuais, e os gráficos das estimativas intervalares de 95% de confiança para cada uma das quatro médias populacionais. Esses intervalos de confiança são calculados pelos mesmos métodos usados no Capítulo 6, com a diferença que se toma um desvio-padrão combinado em lugar dos desvios-padrão amostrais individuais. (O último valor no quadro Minitab mostra que o desvio-padrão combinado é 12,22.)

APRESENTAÇÃO STATDISK

File Edit Analysis Data Help

Equal Length Samples

Total Num Values	48
Upper Deg Free	3
Lower Deg Free	44

SS(treatment)	3090.1
SS(error)	6571.4
SS(total)	9661.5
MS(treatment)	1030.0
MS(error)	149.35
MS(total)	205.56

Test Statistic, F	6.8967
Critical F	2.8165
P-Value	0.000660

Reject the Null Hypothesis

Data provides evidence that the sample means are unequal

APRESENTAÇÃO MINITAB

```
Analysis of Variance

Source     DF        SS        MS        F        p
Factor      3      3090      1030     6.90    0.001
Error      44      6571       149
Total      47      9661
                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev   ----+---------+---------+---------+--
C1         12     63.25      9.45   (------*------)
C2         12     74.25     14.17            (------*------)
C3         12     81.67     13.72                   (------*------)
C4         12     83.75     10.91                     (------*------)
                                    ----+---------+---------+---------+--
Pooled StDev= 12.22                     60        70        80        90
```

Fazendo passar os dados da Tabela 11-1 pela TI-83, o quadro ANOVA incluirá a estatística de teste $F = 6{,}896673726$ e o valor P 0,00066042992, assim como outros resultados relevantes.

Com um programa de computador, o processo de análise da variância é bastante fácil: Introduza os dados, acione o programa ANOVA, identifique o valor P e rejeite a hipótese nula de igualdade das médias se o valor P não superar o nível α de significância (em caso contrário, não rejeite a hipótese nula).

Fundamentos Lógicos

O método da análise da variância se baseia neste conceito fundamental: Com a suposição de que as populações tenham todas a mesma variância σ^2, estimamos seu valor comum utilizando duas abordagens diferentes. A estatística de teste F é a razão dessas duas estimativas, de modo que um valor de F significativamente *grande* (localizado muito à direita do gráfico da distribuição F) constitui evidência contra a igualdade das médias populacionais. As duas abordagens para estimar o valor comum de σ^2 são:

1. A **variância entre amostras** (também chamada **variação devida ao tratamento**) é uma estimativa da variância populacional comum σ^2 que se baseia na variabilidade entre as *médias* amostrais.
2. A **variância dentro das amostras** (também chamada **variação devida ao erro**) é uma estimativa da variância populacional comum σ^2 baseada nas *variâncias* amostrais.

Estatística de Teste para ANOVA de Um Critério

$$F = \frac{\text{variância entre amostras}}{\text{variância dentro de amostras}}$$

O numerador mede a variação entre as médias amostrais. A estimativa da variância no denominador depende somente das variâncias amostrais e não é afetada pelas diferenças entre as médias amostrais. Conseqüentemente, as médias amostrais que apresentam valores próximos uns dos outros resultam em uma estatística de teste F próxima de 1, e concluímos que não há diferença significativa entre as médias amostrais. Mas se o valor de F é excessivamente *grande*, rejeitamos a afirmação de igualdade de médias. (As expressões vagas "próximo de 1" e "excessivamente grande" tornam-se objetivas com a adoção de um valor crítico específico, que estabelece claramente a diferença entre uma estatística de teste F que está na região crítica e uma que não está.) Como os valores excessivamente grandes de F refletem médias desiguais, o teste é unilateral à direita.

Resistência à Pesquisa

As pesquisas baseadas em amostras relativamente pequenas podem ser bastante precisas, desde que a amostra seja aleatória ou representativa da população. Todavia, as crescentes taxas de recusa às pesquisas estão tornando mais difícil a obtenção de amostras aleatórias. O Council of American Survey Research Organizations relatou que, em ano recente, 38% dos consumidores recusaram-se a responder pesquisas. O gerente de uma empresa de pesquisa de mercado disse: "Todo o mundo teme a auto-seleção e se preocupa com o fato de as generalizações feitas se basearem apenas naqueles que cooperam." Os resultados do mercado multibilionário da indústria de pesquisa afetam os produtos que adquirimos, os programas de televisão a que assistimos e muitos outros aspectos de nossas vidas.

Cálculos com Tamanhos Iguais de Amostras

Se os conjuntos de dados têm todos o mesmo tamanho amostral (como na Tabela 11-1), os cálculos necessários não são excepcionalmente difíceis. Inicialmente, determine a variância entre amostras calculando $ns_{\bar{x}}^2$, onde $s_{\bar{x}}^2$ é a variância das médias amostrais. Isto é, considere as médias amostrais como um conjunto ordinário de valores e calcule a variância. (Pelo teorema central do limite, $\sigma_{\bar{x}} = \sigma/\sqrt{n}$ pode ser resolvido em relação a σ, dando $\sigma = \sqrt{n} \cdot \sigma_{\bar{x}}$, de forma que podemos estimar σ^2 por $ns_{\bar{x}}^2$. Por exemplo, as médias amostrais na Tabela 11-1 são 63,3, 74,3, 81,7 e 83,8. Aqueles quatro valores têm um desvio-padrão $s_{\bar{x}}$ = 9,26116, de modo que

$$\text{variância entre amostras} = ns_{\bar{x}}^2 = 12(9{,}26116)^2 = 1029{,}23$$

Em seguida, estime a variância dentro das amostras calculando s_p^2, que é a variância combinada que se obtém achando-se a média das variâncias amostrais. Os desvios-padrão amostrais na Tabela 11-1 são 9,4, 14,2, 13,7 e 10,9, e assim

$$\text{variância dentro de amostras} = s_p^2$$

$$= \frac{9{,}4^2 + 14{,}2^2 + 13{,}7^2 + 10{,}9^2}{4} = 149{,}125$$

Finalmente, calcule a estatística de teste F como se segue:

$$F = \frac{\text{variância entre amostra}}{\text{variância dentro de amostras}} = \frac{1029{,}23}{149{,}125} = 6{,}9018$$

Tomando maior número de decimais, obteríamos a estatística de teste mais precisa F = 6,8967.

Obtém-se o valor crítico de F admitindo-se um teste unilateral à direita, porque grandes valores de F correspondem a diferenças significativas entre médias. Com k amostras, cada uma com n valores, calculam-se como se segue os números de graus de liberdade.

Graus de Liberdade com k Amostras de Mesmo Tamanho n
graus de liberdade do numerador = $k - 1$ graus de liberdade do denominador = $k(n - 1)$

Para os dados amostrais da Tabela 11-1, $k = 4$ e $n = 12$, de modo que os graus de liberdade são 3 para o numerador e 4(11) = 44 para o denominador. Com $\alpha = 0{,}05$, 3 graus de liberdade para o numerador e 44 para o denominador, o valor crítico é $F = 2{,}8387$. (A Tabela A-5 não inclui 44 graus de liberdade para o denominador, o que nos leva a tomar o valor mais próximo, que corresponde a 40 graus de liberdade.) Note que STATDISK e Minitab

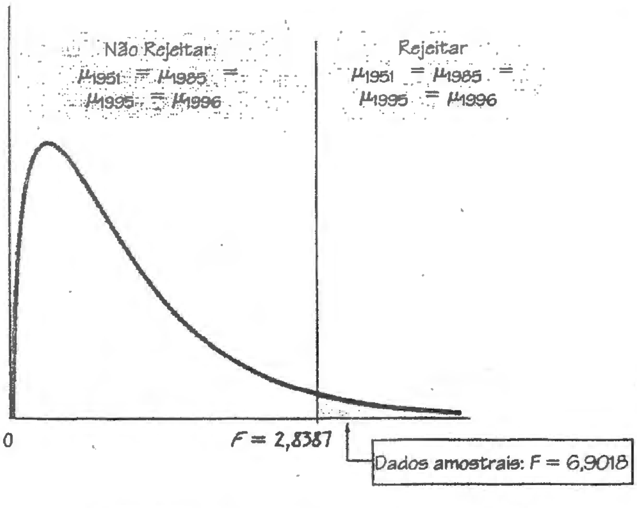


Fig. 11-3 Distribuição F para os dados do Old Faithful.

dão os números de graus de liberdade juntamente com o valor calculado da estatística de teste F. Veja também a Figura 11-3, que mostra a estatística de teste e o valor crítico. Com base nestes resultados, rejeitamos a hipótese nula de médias iguais. Há evidência suficiente para rejeitarmos a afirmação de que as quatro amostras provêm de populações com a mesma média.

Cálculos com Amostras de Tamanhos Diferentes

Conquanto os cálculos necessários para os casos de amostras de mesmo tamanho sejam razoáveis, no caso de amostras de tamanhos diferentes eles se tornam realmente complicados. Aplica-se o mesmo raciocínio básico porque calculamos a estatística de teste F que é a razão de duas estimativas diferentes da variação populacional comum σ^2, mas essas estimativas envolvem medidas ponderadas que levam em conta os tamanhos das amostras, como se vê a seguir.

$$F = \frac{\text{variância entre amostras}}{\text{variância dentro de amostras}} = \frac{\left[\dfrac{\Sigma n_i(\bar{x}_i - \bar{\bar{x}})^2}{k - 1}\right]}{\left[\dfrac{\Sigma(n_i - 1)s_i^2}{\Sigma(n_i - 1)}\right]}$$

onde $\bar{\bar{x}}$ = média de todos os valores amostrais combinados
k = número de médias populacionais que estão sendo comparadas
n_i = número de valores na i^{ma} amostra
N = número total de valores em todas as amostras combinadas
$\bar{x}_i$ = média dos valores na i^{ma} amostra
s_i^2 = variância dos valores na i^{ma} amostra

Note que o numerador nesta estatística de teste é, na verdade, uma forma da fórmula

$$s^2 = \frac{\Sigma(x - \bar{x})^2}{n - 1}$$

da variância dada no Capítulo 2. Inclui-se o fator n_i a fim de que amostras maiores tenham maior peso. O denominador da estatística de teste nada mais é do que a média das variâncias amostrais, mas é uma média ponderada com os pesos baseados nos tamanhos das amostras.

Como o cálculo desta estatística de teste pode levar a grandes erros de arredondamento, os diversos programas adotam uma expressão diferente (porém equivalente) que envolve a notação SQ (soma de quadrados) e QM (quadrado médio). Embora a notação e as componentes que se seguem sejam complicadas, a idéia básica é a mesma: A estatística de teste F é uma razão cujo numerador reflete a variação *entre* as médias amostrais e cujo denominador reflete a variação *dentro* das amostras. Se as populações têm médias iguais, a razão F tende a ficar próxima de 1, mas se as médias populacionais não são iguais, a razão F tende a ser significativamente maior do que 1. Identificam-se a seguir as componentes-chave de nosso método de ANOVA.

SQ(total) ou soma total de quadrados: é uma medida da variação total (em torno de $\bar{\bar{x}}$) em todos os dados amostrais combinados.
Fórmula 11-1 SQ(total) = $\sum (x - \bar{\bar{x}})^2$

SQ(total) pode decompor-se em SQ(tratamento) e SQ(erro), como se segue:

SQ(tratamento) é uma medida da variação entre as médias amostrais. [Em ANOVA de um critério, SQ(tratamento) é chamada às vezes SQ(fator). E como é uma medida de variabilidade *entre* as médias amostrais, é também conhecida como SQ(entre grupos) ou SQ(entre amostras).]

Fórmula 11-2

$$\text{SQ(tratamento)} = n_1(\bar{x}_1 - \bar{\bar{x}})^2 + n_2(\bar{x}_2 - \bar{\bar{x}})^2 + \cdots + n_k(\bar{x}_k - \bar{\bar{x}})^2 = \Sigma\, n_i(\bar{x}_i - \bar{\bar{x}})^2$$

Se as médias populacionais ($\mu_1, \mu_2, \ldots \mu_k$) são iguais, então as médias amostrais $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k$ tenderão a concentrar-se em torno de $\bar{\bar{x}}$. O resultado será um valor relativamente pequeno de SQ(tratamento). Entretanto, se as médias populacionais não são todas iguais, então pelo menos um dos valores $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k$ tenderá a distanciar-se dos outros e também de $\bar{\bar{x}}$. O resultado será um valor relativamente grande de SQ(tratamento).

SQ(erro) é uma soma de quadrados que representa a variabilidade que supomos seja comum a todas as populações em consideração.

Fórmula 11-3

$$\text{SQ(erro)} = (n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \cdots + (n_k - 1)s_k^2 = \Sigma(n_i - 1)s_i^2$$

Como SQ(erro) é uma medida da variância dentro dos grupos, é designada às vezes por SQ(dentro dos grupos) ou SQ(dentro das amostras). Consideradas as expressões precedentes para SQ(total), SQ(tratamento) e SQ(erro), a relação seguinte é sempre válida:

Fórmula 11-4 SQ(total) = SQ(tratamento) + SQ(erro)

SQ(tratamento) e SQ(erro) são ambas somas de quadrados, e se dividirmos cada uma delas pelo correspondente número de graus de liberdade, obteremos quadrados médios, definidos como se segue:

QM(tratamento) é o quadrado médio para tratamento, obtido como se segue:

Fórmula 11-5 $\text{QM(tratamento)} = \dfrac{\text{SQ(tratamento)}}{k-1}$

QM(erro) é o quadrado médio para erro, e se obtém como se segue:

Fórmula 11-6 $\text{QM(erro)} = \dfrac{\text{SQ(erro)}}{N-k}$

QM(total) é o quadrado médio para a variação total:

Fórmula 11-7 $\text{QM(total)} = \dfrac{\text{SQ(total)}}{N-1}$

Com os valores SQ e QM determinamos a estatística de teste F que se aplica quando as amostras não têm todas o mesmo tamanho.

Estatística de Teste para ANOVA com Tamanhos Diferentes de Amostras

Ao testarmos a hipótese nula H_0: $\mu_1 = \mu_2 = \ldots = \mu_k$ contra a hipótese alternativa (essas médias *não* são todas iguais), a estatística de teste

Fórmula 11-8 $F = \dfrac{\text{SQ(tratamento)}}{\text{SQ(erro)}}$

tem distribuição F (quando a hipótese nula é verdadeira) com os graus de liberdade dados por

graus de liberdade do numerador $= k - 1$
graus de liberdade do denominador $= N - k$

Esta estatística de teste é a mesma dada anteriormente, e sua interpretação é também a mesma. O denominador depende somente das variâncias amostrais que medem a variação dentro dos tratamentos, a qual não é afetada pelas diferenças entre as médias amostrais. Em contraste, o numerador depende efetivamente das diferenças entre as médias amostrais. Se essas diferenças são muito grandes, o mesmo ocorrerá com o numerador, de forma que F também será excessivamente grande. Como conseqüência, valores muito grandes de F sugerem médias desiguais, e o teste ANOVA é, assim, assimétrico à direita.

As tabelas constituem um dispositivo conveniente para resumir os resultados fundamentais dos cálculos de ANOVA, e a Tabela 11-2 apresenta um formato comumente usado em resultados de computador. (Veja o quadro Minitab precedente.) Os valores da Tabela 11-2 resultam dos dados da Tabela 11-1.

Aplicamos os métodos de ANOVA para concluir que há evidência suficiente para rejeitar (ou não rejeitar) a afirmação de que as médias populacionais são iguais, mas não podemos concluir que determinada média seja diferente das outras. Com os dados da Tabela 11-1, rejeitamos a igualdade das médias, e os diagramas em caixa da Figura 11-1 sugerem que os tempos em 1951 têm menor média, mas não identificaríamos as médias que são diferentes. Há vários outros testes que podem ser usados para tais identificações. Os processos para identificar especificamente as médias que são diferentes são chamados **processos de comparação múltipla**. A comparação de intervalos de confiança, o teste de Scheffé, o teste estendido de Tukey e o teste de Bonferroni são quatro processos comuns de comparação múltipla em geral incluídos em textos mais avançados.

Por mais eficientes e confiáveis que sejam os programas de computador, eles não têm valor algum se não compreendermos

TABELA 11-2 Tabela de ANOVA para Dados do Old Faithful

Fonte de Variação	Soma de Quadrados SQ	Graus de Liberdade	Quadrado Médio QM	Estatística de Teste F
Tratamentos	3090,06250	3	1030,02083	6,8967
Erro	6571,41667	44	149,350379	
Total	9661,47917	47		

os conceitos fundamentais. Deve ficar entendido que os métodos desta seção são utilizados para testar a afirmação de que várias amostras provêm de populações com a mesma média. Esses métodos exigem populações distribuídas normalmente com a mesma variância, e as amostras devem ser independentes. Rejeitamos ou deixamos de rejeitar a hipótese nula de médias iguais analisando essas duas estimativas da variância: a variância entre amostras e a variância dentro das amostras. QM(tratamento) é uma estimativa da variação entre amostras, e QM(erro) é uma estimativa da variação dentro das amostras. Se QM(tratamento) é significativamente maior do que QM(erro), rejeitamos a afirmação de igualdade de médias; em caso contrário, não rejeitamos tal afirmação.

11-2 Exercícios A: Habilidades e Conceitos Básicos

1. Abaixo apresentamos os resultados da análise de variância dados por Minitab, correspondentes a três amostras com cinco valores cada uma. As três amostras correspondem a avaliações das reações do consumidor a comerciais de TV de três empresas de publicidade diferentes. Suponha que queiramos usar o nível de 0,05 de significância para testar a hipótese nula de que as três empresas diferentes produziram comerciais com a mesma avaliação média.
   a. Identifique a estatística de teste.
   b. Ache o valor crítico na Tabela A-5.
   c. Identifique o valor *P*.
   d. Com base nos resultados precedentes, devemos rejeitar ou não a hipótese nula?
   e. Escreva um relatório sucinto que resuma com clareza a conclusão adequada.

```
Source    DF      SS      MS      F       P
Factor     2    3.20    1.60   0.49   0.627
Error     12   39.51    3,29
Total     14   42.71
```

2. Refaça o Exercício 1, supondo que os resultados Minitab sejam:

```
Source    DF      SS      MS      F       P
Factor     2   20.68   10.34   6.68   0.011
Error     12   18.59    1.55
Total     14   39.27
```

3. Diferentes grupos etários têm temperaturas médias distintas? A tabela a seguir relaciona as temperaturas do corpo de cinco indivíduos escolhidos aleatoriamente de cada um de três grupos etários diferentes.
   a. Determine a variância *entre* amostras calculando $ns_{\bar{x}}^2$. (Recorde que $s_{\bar{x}}^2$ é a variância das três médias amostrais.)
   b. Ache a variância *dentro* das amostras calculando a variância combinada s_p^2, que é a média das variâncias amostrais.
   c. Com os resultados das partes (a) e (b), calcule a estatística *F*.
   d. Ache o valor crítico *F*. (Recorde que, com iguais tamanhos de amostra, os graus de liberdade são $k - 1$ para o numerador e $k(n - 1)$ para o denominador.)
   e. Com base nos resultados precedentes, o que se pode concluir sobre a afirmação de que os três grupos etários tenham a mesma temperatura média?

Temperaturas do Corpo (°F) Classificadas por Idade

18-20	21-29	30 e mais
98,0	99,6	98,6
98,4	98,2	98,6
97,7	99,0	97,0
98,5	98,2	97,5
97,1	97,9	97,3
$n_1 = 5$	$n_2 = 5$	$n_3 = 5$
$\bar{x}_1 = 97{,}940$	$\bar{x}_2 = 98{,}580$	$\bar{x}_3 = 97.800$
$s_1 = 0{,}568$	$s_2 = 0{,}701$	$s_3 = 0{,}752$

Com base em dados dos Drs. Philip Mackowiak, Steven Wasserman e Myron Levine, da Universidade de Maryland.

4. A City Resource Recovery Company (CRRC) coleta os resíduos descartados pelas residências de uma região. Esses resíduos devem ser separados nas categorias de metal, papel, plástico e vidro. Ao planejar o equipamento necessário para coletar e processar o lixo, a CRRC recorre aos dados resumidos no Conjunto de Dados 1 do Apêndice B (fornecidos pelo Projeto do Lixo da Universidade do Arizona). A lista que se segue resume os resultados (pesos em libras). Com um nível de 0,05 de significância, teste a afirmação de que as quatro populações específicas de lixo tenham a mesma média. Com base nos resultados, acha que essas quatro categorias exigem os mesmos recursos de coleta e processamento? Há alguma categoria que pareça constituir uma parcela particularmente grande do problema de resíduos?

Metal	Papel	Plástico	Vidro
$n = 62$	$n = 62$	$n = 62$	$n = 62$
$\bar{x} = 2{,}218$	$\bar{x} = 9{,}428$	$\bar{x} = 1{,}911$	$\bar{x} = 3{,}752$
$s = 1{,}091$	$s = 4{,}168$	$s = 1{,}065$	$s = 3{,}108$

Para os Exercícios 5-8, utilize os dados da tabelaa seguir. Os dados selecionados aleatoriamente representam 30 casas diferentes vendidas no condado de Dutchess, Nova York. As zonas (1, 4, 7) correspondem a diferentes regiões geográficas do condado. Os valores PV são os preços de venda em milhares de dólares. Os valores AR são as áreas das residências em centenas de pés quadrados. Os valores Acres são os tamanhos dos lotes em acres, e os Impostos representam os impostos anuais em milhares de dólares. Por exemplo, a primeira casa está na zona 1; foi vendida por $147.000; tem área de 2000 pés quadrados; está em um lote de 0,50 acres; e o imposto anual é de $1900.

5. Com o nível de 0,05 de significância, teste se as medias dos preços de venda são as mesmas para as zonas 1, 4 e 7. Alguma das zonas parece ter casas com preços de venda mais elevados?
6. Com o nível de 0,05 de significância, teste a afirmação de que as zonas 1, 4 e 7 têm mesma média de área residencial. Alguma das zonas aparenta ter casas maiores?
7. Com o nível de 0,05 de significância, teste a afirmação de que o tamanho médio dos lotes (em acres) é o mesmo para as zonas 1, 4 e 7. Alguma zona aparenta ter lotes maiores?
8. Com o nível de 0,05 de significância, teste a afirmação de que a média de impostos é a mesma para as zonas 1, 4 e 7. Alguma zona aparenta ter impostos mais elevados?
9. Foram feitos testes de inflamabilidade em roupas de dormir infantis. Utilizou-se o Vertical Semirestrained Test, em que são incineradas peças de roupa sob condições controladas. Terminada a queima, mediu-se e registrou-se a extensão da porção incinerada. Os resultados são dados a seguir, para o mesmo tecido testado em diferentes laboratórios. Como foi utilizado o mesmo tecido, os di-

Zona	PV	AR	Acres	Impostos
1	147	20	0,50	1,9
1	160	18	1,00	2,4
1	128	27	1,05	1,5
1	162	17	0,42	1,6
1	135	18	0,84	1,6
1	132	13	0,33	1,5
1	181	24	0,90	1,7
1	138	15	0,83	2,2
1	145	17	2,00	1,6
1	165	16	0,78	1,4
4	160	18	0,55	2,8
4	140	20	0,46	1,8
4	173	19	0,94	3,2
4	113	12	0,29	2,1
4	85	9	0,26	1,4
4	120	18	0,33	2,1
4	285	28	1,70	4,2
4	117	10	0,50	1,7
4	133	15	0,43	1,8
4	119	12	0,25	1,6
7	215	21	3,04	2,7
7	127	16	1,09	1,9
7	98	14	0,23	1,3
7	147	23	1,00	1,7
7	184	17	6,20	2,2
7	109	17	0,46	2,0
7	169	20	3,20	2,2
7	110	14	0,77	1,6
7	68	12	1,40	2,5
7	160	18	4,00	1,8

versos laboratórios deveriam obter os mesmos resultados. Isto ocorreu realmente? Com o nível de 0,05 de significância, teste a afirmação de que os diferentes laboratórios têm a mesma média populacional. (Os dados foram fornecidos por Minitab, Inc.)

Laboratório				
1	2	3	4	5
2,9	2,7	3,3	3,3	4,1
3,1	3,4	3,3	3,2	4,1
3,1	3,6	3,5	3,4	3,7
3,7	3,2	3,5	2,7	4,2
3,1	4,0	2,8	2,7	3,1
4,2	4,1	2,8	3,3	3,5
3,7	3,8	3,2	2,9	2,8
3,9	3,8	2,8	3,2	
3,1	4,3	3,8	2,9	
3,0	3,4	3,5		
2,9	3,3			

10. Os valores da tabela abaixo são larguras máximas de crânios de homens egípcios de diferentes épocas (com base em dados de *Ancient Races of the Thebaid*, por Thomson and Randall-Maciver). Modificações na forma da cabeça através dos tempos sugerem uma miscigenação com populações imigrantes. Com o nível de 0,05 de significância, teste a igualdade das médias para as diferentes épocas.

4000 A.C.	1850 A.C.	150 A.D.
131	129	128
138	134	138
125	136	136
129	137	139
132	137	141
135	129	142
132	136	137
134	138	145
138	134	137

11. Consulte o Conjunto de Dados 5 do Apêndice B. Utilizando o fator mês isoladamente, teste a afirmação de que Abril, Maio, Junho, Julho, Agosto e Setembro têm as mesmas concentrações médias de ozônio nos estados sob amostragem.
12. Consulte o Conjunto de Dados 6 do Apêndice B referente aos depósitos de nitrato conseqüentes da chuva ácida. Utilizando o fator ano isoladamente, teste a afirmação de que os anos de 1980 a 1990 têm a mesma quantidade média.
13. Consulte o Conjunto de Dados 11 do Apêndice B. Com o nível de 0,05 de significância, teste a afirmação de que o peso médio dos M&Ms é o mesmo para cada uma das seis diferentes populações (cores). Se a intenção de Mars, Inc., é produzir confeitos cujas populações (cores) tenham a mesma média, os resultados sugerem que a companhia tenha algum problema que exija ação corretiva?
14. Consulte o Conjunto de Dados 9 do Apêndice B. Com o nível de 0,01 de significância, teste a afirmação de que o nível médio de cotinina é diferente para estes três grupos: não-fumantes não expostos ao fumo ambiental, não-fumantes expostos ao fumo ambiental, e fumantes. Que sugerem os resultados quanto ao fumo "de segunda mão"?

11-2 Exercícios B: Além do Básico

15. Refaça o Exercício 3 após aumentar de 2° cada temperatura relacionada na faixa etária 18-20. Compare esses resultados com os obtidos no Exercício 3.
16. Refazendo-se o Exercício 3 após transformar os valores da temperatura da escala Fahrenheit para a escala Celsius, os resultados sofrem alteração? De modo geral, qual é a influência da escala utilizada sobre a estatística de teste F da análise da variância?
17. Como são afetados os resultados da análise da variância em cada um dos seguintes casos?
    a. Adiciona-se a mesma constante a cada valor amostral.
    b. Cada valor amostral é multiplicado pela mesma constante.
    c. Modifica-se a ordem das amostras.
18. Extraem-se aleatoriamente cinco amostras independentes de tamanho 50, de populações que têm distribuição normal com variâncias iguais. Desejamos testar a afirmação $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$.
    a. Se utilizássemos apenas os métodos do Capítulo 8, deveríamos testar as igualdades individuais $\mu_1 = \mu_2$, $\mu_1 = \mu_3$, ..., $\mu_4 = \mu_5$. Quantos pares poderíamos formar com as cinco médias?
    b. Suponha que, para cada teste de igualdade entre duas médias, haja uma probabilidade de 0,95 de não cometer um erro tipo I. Testando-se a igualdade de todos os pares possíveis, qual é a probabilidade de não cometer nenhum erro tipo I? (Embora os testes não sejam realmente independentes, suponha que o sejam.)
    c. Se aplicamos a análise da variância para testar a afirmação de que $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$, com o nível de 0,05 de significância, qual é a probabilidade de não cometer um erro tipo I?
    d. Compare os resultados das partes (b) e (c). Qual é a melhor abordagem, no sentido de nos dar maior chance de não cometer um erro tipo I?
19. Neste exercício, o leitor verificará que, no caso de dois conjuntos de dados amostrais, o teste t para amostras independentes e o método ANOVA desta seção são equivalentes. Consulte os dados sobre imóveis dos Exercícios 5-8, excluindo os dados referentes à zona 7.
    a. Com o nível de 0,05 de significância, teste se as zonas 1 e 4 têm o mesmo preço de venda médio. (Admita que ambas as zonas tenham a mesma variância.)
    b. Com o nível de 0,05 e com o método ANOVA desta seção, teste a afirmação da parte (a).

c. Verifique que os quadrados da estatística de teste t e do valor crítico da parte (a) são iguais à estatística de teste F e ao valor crítico da parte (b).

20. Complete a tabela de ANOVA a seguir, sabendo que há três amostras de tamanhos 5, 7 e 7, respectivamente.

Fonte de Variação	Soma de Quadrados SQ	Graus de Liberdade	Quadrado Médio QM	Estatística de Teste
Tratamentos	?	?	?	
Erro	100,00	?	?	$F = ?$
Total	123,45	?		

11-3 ANOVA de Dois Critérios

Na Seção 11-2 ilustramos o uso da análise da variância para decidir se três ou mais populações tinham a mesma média. Ali utilizamos processos conhecidos como análise da variância de um critério (ou de um fator) porque os dados estavam categorizados em grupos de acordo com um *único* fator (ou tratamento). Recorde que um fator, ou tratamento, é uma propriedade que serve de base para categorizar diferentes grupos de dados. Por exemplo, na Tabela 11-3 os comprimentos de 30 filmes são divididos em três categorias, de acordo com o fator único "número de estrelas" — um filme regular recebe 2,0 ou 2,5 estrelas, um bom filme recebe 3,0 ou 3,5 estrelas e um filme excelente recebe 4,0 estrelas. Como as três amostras da Tabela 11-3 são categorizadas segundo um único fator (número de estrelas), podemos aplicar os métodos da Seção 11-2 para testar a afirmação $\mu_1 = \mu_2 = \mu_3$. Segue a apresentação Minitab para os dados da Tabela 11-3.

Pela apresentação Minitab, vemos que a estatística de teste é

$$F = \frac{\text{QM(tratamento)}}{\text{QM(erro)}} 5\ 5\ 1{,}17$$

O valor P de 0,325 indica que, com o nível de 0,05 de significância, não rejeitamos a igualdade de médias. Com base em nossos dados amostrais, parece que as três categorias "regular", "bom" e "excelente" têm a mesma duração média.

A análise da variância de um critério para a Tabela 11-3 envolve um único fator — a classificação por estrelas. Se estivéssemos pesquisando as durações de filmes, uma análise mais complexa poderia incluir fatores adicionais como natureza do filme (aventura, drama, policial, comédia etc.), o estúdio que produziu o filme, o orçamento da filmagem, os atores e atrizes que tomaram parte, o diretor e o fundo musical. A **análise da variância de dois critérios** envolve dois fatores, como no caso da Tabela 11-4: (1) a classificação do filme, por número de estrelas, em regular, bom ou excelente, e (2) a classificação pela Motion Picture Association of America (MPAA), que divide os filmes em uma categoria R, ou outra categoria G, PG ou PG-13. Com a classificação por estrelas e a classificação MPAA, particionamos os dados em seis categorias, conforme a Tabela 11-4.

Tais categorias costumam ser chamadas *células*, de modo que a Tabela 11-4 tem seis células contendo cada uma cinco valores. As Tabelas 11-3 e 11-4 incluem os mesmos números, mas a Tabela 11-3 classifica os dados pelo fator único "número de estrelas", enquanto a Tabela 11-4 os classifica segundo os dois fatores "número de estrelas" e "classificação da MPAA".

Como já abordamos a análise da variância de um critério para o caso de *um* fator (classificação de filmes por número de estre-

APRESENTAÇÃO MINITAB PARA UM FATOR DE CLASSIFICAÇÃO POR ESTRELAS

```
Analysis of Variance
Source    DF      SS      MS        F      p
Factor     2    1582     791     1.17  0.325
Error     27   18244     676
Total     29   19826
                                 Individual 95% CIs for Mean
                                 Based on Pooled StDev
Level      N    Mean   StDev   ---------+---------+---------+------
C1        10  108.80   18.74    (-----------*----------)
C2        10  107.30   14.81   (-----------*----------)
C3        10  123.40   38.16              (----------*-----------)
                               ---------+---------+---------+------
Pooled StDev =  25.99                  105       120       135
```

TABELA 11-3 Durações (em min) de Filmes Categorizados por Número de Estrelas

Regular (2,0-2,5 Estrelas)	98	100	123	92	99	110	114	96	101	155
Bom (3,0-3,5 Estrelas)	93	94	94	105	111	115	133	106	93	129
Excelente (4,0 Estrelas)	103	193	168	88	121	72	120	106	104	159

TABELA 11-4 Duração (em min.) de Filmes Categorizados por Dois Fatores: Estrela e MPAA (G/PG/PG-13, R)

	Regular 2,0-2,5 Estrelas	Bom 3,0-3,5 Estrelas	Excelente 4,0 Estrelas
Classificação MPAA: G/PG/PG-13	98	93	103
	100	94	193
	123	94	168
	92	105	88
	99	111	121
Classificação MPAA: R	110	115	72
	114	133	120
	96	106	106
	101	93	104
	155	129	159

las), parece razoável que passemos diretamente a outra ANOVA de um critério para a classificação MPAA. Infelizmente, essa abordagem desperdiça informação e ignora totalmente qualquer efêito de uma interação entre dois fatores.

DEFINIÇÃO

Há uma **interação** entre dois fatores se o efeito de um dos fatores se modifica conforme a categoria do outro fator.

Ao aplicarmos ANOVA aos dados da Tabela 11-4, consideraremos o efeito de uma *interação* entre classificações por estrelas e classificações da MPAA, assim como os efeitos das classificações por estrelas e das classificações da MPAA sobre a duração do filme. Como os cálculos são bastante complicados, *supomos o uso de um programa de computador.* (A calculadora TI-83 pode ser usada para a análise da variância de um critério, mas não está programada para os métodos desta seção.) A apresentação Minitab a seguir resulta dos dados da Tabela 11-4. Para utilizar o Minitab, introduza os 30 valores amostrais na Coluna C1, introduza os números de linha correspondentes (1 ou 2) na coluna C2 para identificar a classificação MPAA e introduza os números correspondentes de colunas (1 ou 2 ou 3) em C3 para identificar a classificação por estrelas. Do menu principal, selecione `Stat`, em seguida `ANOVA`, em seguida `Two-Way`, e passe a introduzir C1 para `Response`; introduza C2 para `row factor`, e C3 para `column factor`; depois clique em OK.

A apresentação Minitab inclui componentes de SQ(soma de quadrados) análogas às descritas na Seção 11-2. Como as circunstâncias da Seção 11-2 envolviam apenas um fator, utilizamos SQ(tratamento) como medida da variação devida às diferentes categorias de tratamento, e SQ(erro) como medida da variação devida ao erro amostral. Vamos agora usar SQ(ESTRELA) como medida da variação entre as médias das classificações por estrelas. Usaremos SQ(MPAA) como medida da variação entre as médias de MPAA. Continuaremos usando SQ(erro) como medida da variação devida ao erro amostral. Da mesma forma, usaremos QM(ESTRELA) e QM((MPAA) para os dois diferentes quadrados médios; continuaremos a usar QM(erro) como antes. Outrossim, usaremos gl(ESTRELA) e gl(MPAA) para os dois graus de liberdade diferentes.

A Tabela 11-5 compara as análises dos dados nas Tabelas 11-3 e 11-4. Como as Tabelas 11-3 e 11-4 têm os mesmos dados nas categorias de classificação por estrelas, os valores calculados de SQ(ESTRELA), QM(ESTRELA) e gl(ESTRELA) são também os mesmos em ambos os casos. Note, entretanto, que quando passamos do caso de um fator (Tabela 11-3) para o caso de dois fatores (Tabela 11-4), particionando os dados de acordo com o segundo fator — classificação de MPAA —, o valor de SQ(erro) é particionado em SQ(MPAA), SQ(interação) e SQ(erro). (Pode-se ver que 18.244 = 32 + 2256 + 15.956.) Outrossim, gl(erro) é particionado em gl(MPAA), gl(interação)

APRESENTAÇÃO MINITAB

```
Two-Way Analysis of Variance

Analysis of Variance for C1

Source          DF        SS       MS
MPAA             1        32       32
STAR             2      1582      791
Interaction      2      2256     1128
Error           24     15956      665
Total           29     19826
```

TABELA 11-5 Comparação de Componentes Utilizadas em ANOVA de Um Critério e ANOVA Dois Critérios

ANOVA Um Critério	ANOVA Dois Critérios
Dados amostrais: Tabela 11-3	Dados amostrais: Tabela 11-4
Um fator: classificação ESTRELA	Dois fatores: classificação ESTRELA e classificação MPAA
SQ(ESTRELA) = 1582	mesmo
QM(ESTRELA) = 791	mesmo
gl(ESTRELA) = 2	mesmo
SQ(total) = 19.826	mesmo
gl(total) = 29	mesmo
	SQ(MPAA) = 32
SQ(erro) = 18.244	SQ(interação) = 2256
	SQ(erro) = 15.956
	gl(MPAA) = 1
gl(erro) = 27	gl(interação) = 2
	gl(erro) = 24
	QM(MPAA) = 32
QM(erro) = 676	QM(interação) = 1128
	QM(erro) = 665

e gl(erro), o que se pode verificar facilmente notando que 27 = 1 + 2 + 24. Não se aplica um tal particionamento a QM(erro); o valor de QM(erro) = 676 não é igual a 32 + 1128 + 665.

Ao fazer uma análise da variância de dois critérios, consideramos três efeitos:

1. O efeito devido à *interação* entre os dois fatores — classificação por estrelas e classificação MPAA
2. O efeito devido ao fator *linha* (classificação MPAA)
3. O efeito devido ao fator *coluna* (classificação por estrelas)

Os comentários que se seguem resumem o processo básico da análise de variância de dois critérios. É, na verdade, análogo aos processos apresentados na Seção 11-2. Formulamos conclusões sobre igualdade de médias analisando duas estimativas de variância, e a estatística de teste F é a razão dessas duas estimativas. Um valor significativamente grande de F indica que há uma diferença de médias estatisticamente significativa. A Figura 11-4 resume o processo para ANOVA de dois critérios.

Processo para ANOVA de Dois Critérios

Passo 1: Na análise da variância de dois critérios, começamos testando a hipótese nula, de que não há interação entre os dois fatores. Utilizando Minitab para os dados da Tabela 11-4, obtemos a seguinte estatística de teste:

$$F = \frac{\text{QM(interação)}}{\text{QM(erro)}} = \frac{1128}{665} = 1{,}6962$$

Com gl(interação) = 2 e gl(erro) = 24, obtemos um valor crítico $F = 3{,}4028$ (admitindo o nível de 0,05 de significância). Como a estatística de teste não excede o valor crítico, não rejeitamos a hipótese nula de não haver interação entre os dois fatores. Não há evidência para concluirmos que a duração do filme seja afetada por uma interação entre classificação por estrelas e classificação MPAA.

Passo 2: Se rejeitamos efetivamente a hipótese nula, de nenhuma interação entre fatores, então devemos parar por aqui, não prosseguindo com os testes adicionais. (Se há uma interação entre fatores, não devemos considerar os efeitos de qualquer deles sem considerar os efeitos do outro.)

Se não rejeitamos a hipótese nula de nenhuma interação entre fatores, então devemos testar as duas hipóteses seguintes:

H_0: Não há efeitos devidos ao fator linha (isto é, as médias de linhas são iguais).
H_0: Não há efeitos devidos ao fator coluna (isto é, as médias das colunas são iguais).

No Passo 1, deixamos de rejeitar a hipótese nula, de nenhuma interação entre fatores; passamos, assim, aos dois próximos testes de hipóteses identificados no Passo 2.

Para o fator linha,

$$F = \frac{\text{QM(MPAA)}}{\text{QM(erro)}} = \frac{32}{665} = 0{,}0481$$

Este valor não é significativo, porque o valor crítico, baseado em gl(MPPAA) = 1, gl(erro) = 24 e um nível de 0,05 de significância, é $F = 4{,}2597$. Não rejeitamos a hipótese nula, de nenhum efeito da classificação MPAA. A classificação MPAA não parece ter qualquer efeito sobre a duração do filme.

Para o fator coluna,

$$F = \frac{\text{QM(ESTRELA)}}{\text{QM(erro)}} = \frac{791}{665} = 1{,}1895$$

Este valor não é significativo, pois o valor crítico, baseado em gl(ESTRELA) = 2, gl(erro) = 24, e um nível de 0,05 de significância, é $F = 3{,}4082$. Não rejeitamos a hipótese nula, de nenhum efeito devido à classificação por estrelas. A classificação de um filme por estrelas não parece ter qualquer efeito sobre sua duração.

Caso Especial: Uma Observação por Célula e Nenhuma Interação

A Tabela 11-4 contém cinco observações por célula. Se nossos dados amostrais consistem em apenas uma observação por célula, perdemos QM(interação), SQ(interação) e gl(interação) porque esses valores se baseiam em variâncias amostrais calculadas para cada célula individual. Se há apenas uma observação por célula, não há variação dentro das células individuais, não sendo possível calcular aquelas variâncias amostrais. Eis como proceder quando há apenas uma observação por célula: *Se parece razoável admitir (com base no conhecimento das circunstâncias) que não há interação entre os dois fatores, adotar esta suposição e proceder como anteriormente para testar separadamente as duas hipóteses seguintes:*

H_0: Não há efeitos devidos ao fator linha.
H_0: Não há efeitos devidos ao fator coluna.

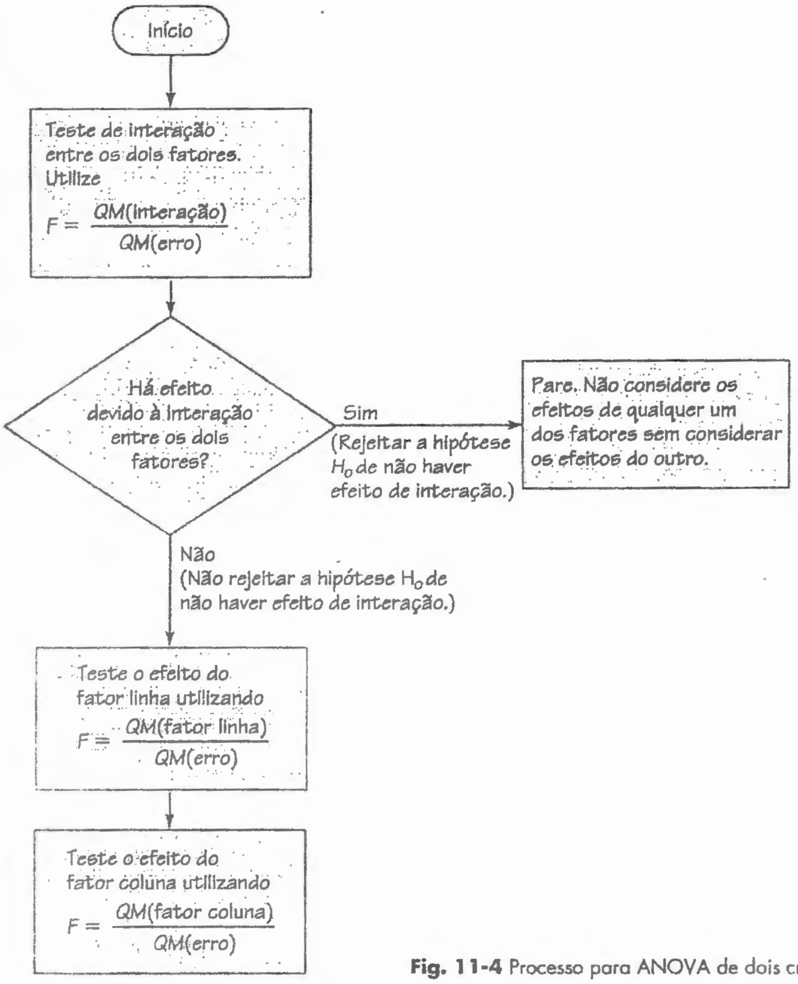


Fig. 11-4 Processo para ANOVA de dois critérios.

A título de exemplo, suponha que temos apenas o primeiro escore em cada célula da Tabela 11-4. Note que as duas médias de linhas são 98,0 e 99,0. Esta diferença é significativa, sugerindo um efeito devido à classificação MPAA? As médias das colunas são 104,0, 94,0 e 87,5. Essas diferenças são significativas, sugerindo um efeito devido à classificação por estrelas? Se acreditamos que o número de estrelas que um filme recebe não tem qualquer relação com sua classificação como G, PG, PG-13 ou R, podemos admitir que não há interação entre classificação por estrelas e classificação MPAA. (Se achamos que há alguma interação, então o método aqui apresentado não pode ser aplicado.) Para ilustrar o método, vamos supor que não há qualquer interação. Apresentamos a seguir o quadro Minitab para os dados da Tabela 11-4, com apenas o primeiro valor de cada célula.

Inicialmente utilizamos os resultados do quadro Minitab para testar a hipótese nula de que não há qualquer efeito do fator linha (classificação MPAA). Obtemos a estatística de teste

APRESENTAÇÃO MINITAB

```
Two-Way Analysis of Variance

Analysis of Variance for C1

Source      DF        SS        MS
MPAA         1         2         2
STAR         2       363       181
Error        2       793       397
Total        5      1157
```

$$F = \frac{\text{QM(MPAA)}}{\text{QM(erro)}} = \frac{2}{397} = 0{,}0050$$

Com um nível de 0,05 de significância, obtemos o valor crítico $F = 18{,}513$. [Consulte a Tabela A-5 e note que o quadro Minitab dá os graus de liberdade (gl) 1 e 2 para o numerador e denominador.] Como a estatística de teste $F = 0{,}0050$ não excede o valor crítico $F = 18{,}513$, não rejeitamos a hipótese nula; a classificação MPAA não parece afetar a duração do filme.

Usamos a seguir os resultados do Minitab para testar a hipótese nula de que não há qualquer efeito devido ao fator coluna (classificação por estrelas). A estatística de teste é

$$F = \frac{\text{QM(ESTRELA)}}{\text{QM(erro)}} = \frac{181}{397} = 0{,}4559$$

que não excede o valor crítico $F = 19{,}000$. [Veja a Tabela A-5, com gl(STAR) = 2 e gl(erro) = 2.] Não rejeitamos a hipótese nula; a classificação por estrela não parece afetar a duração do filme. Novamente, salientamos que essas conclusões se baseiam em um número muito limitado de dados, e podem não ser válidas quando lidamos com maior número de dados.

Planejamento em Blocos Aleatorizados

Quando utilizamos a análise da variância de um critério (ou de um fator) e concluímos que as diferenças entre médias são significativas, não podemos deduzir que o fator em causa seja necessariamente responsável pelas diferenças. É possível que a variação de algum outro fator desconhecido seja a causa das diferenças. Uma forma de reduzir o efeito de fatores estranhos consiste em planejar o experimento de modo que ele seja **completamente aleatorizado**, isto é, que cada elemento tenha a mesma chance de pertencer às diferentes categorias, ou tratamentos. Outra maneira de reduzir o efeito de fatores estranhos é utilizar um **planejamento rigorosamente controlado**, em que os elementos são escolhidos cuidadosamente de forma que nenhum dos outros fatores tenha qualquer variabilidade. Isto é, escolher elementos que sejam os mesmos em todas as características, exceto o fator isolado que está sendo considerado.

Ainda uma outra maneira de controlar a variação estranha consiste em usar um **planejamento em blocos aleatorizados,** em que se obtém uma medida para cada tratamento e cada um de vários indivíduos que são emparelhados de acordo com características análogas. Um **bloco** é um grupo de indivíduos análogos. Veja o exemplo a seguir, em que utilizamos quatro carros diferentes para testar a quilometragem obtida por três tipos diferentes de gasolina (regular, extra, premium). Com um planejamento em blocos aleatorizados como forma de controlar as diferenças entre carros, usamos cada tipo de gasolina em cada um dos quatro carros, selecionando aleatoriamente a ordem em que se faz este experimento. Temos três tratamentos que correspondem aos três tipos de gasolina, e quatro blocos correspondentes aos quatro carros utilizados.

EXEMPLO Usando cada um dos três tipos de gasolina em cada um dos quatro carros diferentes, obtemos os resultados apresentados na Tabela 11-6. Aplique a análise da variância de dois critérios para testar a afirmação de que nem o tipo de gasolina nem a diferença entre os carros afetam a milhagem.

SOLUÇÃO O método aqui será o mesmo usado para o caso especial de uma observação por célula, mas não precisamos de conhecimento especial das circunstâncias para admitir que não há interação. Podemos supor a ausência de interação em razão da maneira como planejamos o experimento. A apresentação Minitab para os dados da Tabela 11-6 é como segue.

Considerando os efeitos de tratamento (linha), calculamos a estatística de teste

$$F = \frac{\text{QM(GAS)}}{\text{QM(erro)}} = \frac{23{,}58}{1{,}92} = 12{,}2813$$

Pelo quadro obtemos gl(GAS) = 2 e gl(erro) = 6;* podemos agora recorrer à Tabela A-5. Com o nível de 0,05 de signifi-

TABELA 11-6 Milhagem (mi/gal) para Diferentes Tipos de Gasolina

		Bloco			
		Carro 1	Carro 2	Carro 3	Carro 4
Tratamento	Combustível Regular	19	33	23	27
	Combustível Extra	19	34	26	29
	Combustível Premium	22	39	26	34

*gl = graus de liberdade (de DF = *degrees of freedom*). (N. do T.)

APRESENTAÇÃO MINITAB

```
Two-way Analysis of Variance

Analysis of Variance for C1

Source       DF          SS          MS
GAS           2       47.17       23.58
CAR           3      390.25      130.08
Error         6       11.50        1.92
Total        11      448.92
```

cância, o valor crítico é $F = 5{,}1433$. Como a estatística de teste, $F = 12{,}2813$, excede o valor crítico $F = 5{,}1433$, rejeitamos a hipótese nula. O tipo da gasolina parece influir na milhagem.

Para os efeitos de bloco (colunas), a estatística de teste é

$$F = \frac{\text{QM(CAR)}}{\text{QM(erro)}} = \frac{130{,}08}{1{,}92} = 67{,}7500$$

e o valor crítico é $F = 4{,}7571$. Rejeitamos a hipótese nula de igualdade das médias de blocos e concluímos que os diferentes carros têm diferentes valores de milhagem. (Permutando os papéis dos blocos e dos tratamentos mediante transposição da Tabela 11-6, obtemos os mesmos resultados.)

Nesta seção apresentamos rapidamente um importante ramo da estatística. Enfatizamos a interpretação das apresentações do computador, e simultaneamente omitimos os cálculos manuais e fórmulas, que são bastante extensos. Textos mais avançados abordam este tópico com muito mais detalhes; nosso intento aqui é dar uma visão um pouco mais geral da natureza da análise da variância de dois critérios.

11-3 Exercícios A: Habilidades e Conceitos Básicos

Nos Exercícios 1-4, use a apresentação Minitab, que corresponde aos dados do quadro a seguir

APRESENTAÇÃO MINITAB

```
Analysis of Variance for C1

Source        DF      SS      MS
MPAA           1      14      14
STAR           3    1049     350
Interaction    3    3790    1263
Error          8    5560     695
Total         15   10413
```

Durações (em min) de Filmes Classificados por Número de Estrelas e por MPAA (G/PG/PG-13, R)

	Fraco 0,0-1,5 Estrelas	Regular 2,0-2,5 Estrelas	Bom 3,0-3,5 Estrelas	Excelente 4,0 Estrelas
Classificação MPAA:	108	98	93	103
G/PG/PG-13	91	100	94	193
Classificação	105	110	115	72
MPAA: R	96	114	133	120

1. Identifique os valores indicados.
   a. QM(interação)
   b. QM(erro)
   c. QM(ESTRELA)
   d. QM(MPAA)
2. Determine a estatística de teste e o valor crítico para a hipótese nula — nenhuma interação entre classificação por estrelas e MPAA. Qual é sua conclusão?
3. Suponha que a duração de um filme não seja afetada por uma interação entre classificação por estrelas e classificação MPAA. Ache a estatística de teste e o valor crítico para a hipótese nula de que a classificação por estrela não tenha influência sobre a duração do filme. Qual é sua conclusão?
4. Suponha que a duração de um filme não seja afetada por uma interação entre classificação por estrelas e classificação MPAA. Determine a estatística de teste e o valor crítico para a hipótese nula de que a classificação MPAA não influi na duração do filme. Qual é sua conclusão?

Nos Exercícios 5 e 6, utilize somente o primeiro valor de cada uma das oito células da tabela usada para os Exercícios 1-4. Quando utilizamos apenas esses primeiros valores, os resultados de Minitab se apresentam como se segue.

APRESENTAÇÃO MINITAB

```
Two-Way Analysis of Variance

Analysis of Variance for C1

Source     DF      SS     MS
MPAA        1       0      0
STAR        3     459    153
Error       3     799    266
Total       7    1258
```

5. Admitindo que a interação entre classificação por estrelas e classificação MPAA não influi na duração do filme, teste a hipótese nula de que a classificação MPAA não influi na duração do filme. Identifique a estatística de teste e o valor crítico, e formule a conclusão. Considere o nível de 0,05 de significância.
6. Admitindo que a interação entre classificação por estrelas e classificação MPAA não influi na duração do filme, teste a hipótese nula de que a classificação por estrelas não influi na duração do filme. Identifique a estatística de teste e o valor crítico, e formule a conclusão. Considere o nível de 0,05 de significância.

Os Exercícios 7 e 8 se referem aos dados amostrais do quadro a seguir e à apresentação Minitab correspondente. Os valores da tabela são os números de vigas de apoio fabricadas por quatro operadores diferentes em cada uma de três máquinas distintas. Nos resultados Minitab, os operadores são representados por C2 e as máquinas são representadas por C3.

APRESENTAÇÃO MINITAB

```
Analysis of Variance for C1

Source     DF       SS       MS
C2          3    59.58    19.86
C3          2    93.17    46.58
Error       6    48.17     8.03
Total      11   200.92
```

		Máquina		
		1	2	3
Operador	1	66	74	67
	2	58	67	68
	3	65	71	65
	4	60	64	66

7. Com um nível de 0,05 de significância, teste a afirmação de que os quatro operadores têm a mesma média de produção. Identifique a estatística de teste e o valor crítico, e formule sua conclusão.
8. Com o nível de 0,05 de significância, teste a afirmação de que a escolha da máquina não influi no resultado da produção. Identifique a estatística de teste e o valor crítico, e formule a conclusão. (Recorra às instruções iniciais.)
9. Consulte o conjunto de Dados 8 do Apêndice B e construa uma tabela com as taxas de pulsação categorizadas segundo os fatores sexo e condição de fumante. Selecione nove valores para cada célula e teste a hipótese nula de que não há interação entre sexo e fumo.
10. Com os mesmos dados coletados para o Exercício 9, suponha que as taxas de pulsação não sejam afetadas por uma interação entre sexo e fumo, e teste a hipótese nula de que o sexo não influi na taxa de pulsação.
11. Com os mesmos dados coletados para o Exercício 9, suponha que a taxa de pulsação não seja afetada por uma interação entre sexo e fumo, e teste a hipótese nula de que o fumo não influi na taxa de pulsação.

11-3 Exercícios B: Além do Básico

12. Utilize um programa estatístico, como Minitab ou SPSS/PC, que produza resultados para a análise da variância de dois critérios. Inicialmente introduza os dados da tabela usada para os Exercícios 1-4 e verifique que os resultados são apresentados nesta seção. Transponha então a tabela de modo que a classificação MPAA seja o fator coluna e a classificação por estrelas seja o fator linha. Obtenha a apresentação de computador para a tabela transposta e compare os resultados com os apresentados previamente.
13. Considere os dados da tabela usada para os Exercícios 1-4 e subtraia 10 de cada valor. Utilizando um programa estatístico com capacidade para análise da variância de dois critérios, determine os efeitos da subtração de 10 de cada valor.
14. Considere os dados da tabela usada para os Exercícios 1-4 e multiplique cada valor por 10. Utilizando um programa estatístico com capacidade para análise da variância de dois critérios, determine os efeitos da multiplicação de cada valor por 10.
15. Analisando a Tabela 11-4, concluímos que a duração de um filme não é afetada por uma interação entre classificação por estrelas e classificação MPAA; não é afetada tampouco pela classificação por estrelas, nem pela classificação MPAA.
    a. Modifique os dados da tabela de modo que haja um efeito da interação entre classificação por estrelas e classificação MPAA.
    b. Modifique os dados da tabela de modo que não haja efeito da interação entre classificação por estrelas e classificação MPAA, nem efeito da classificação por estrelas, mas haja efeito da classificação MPAA.
    c. Modifique os dados da tabela de modo que não haja efeito da interação entre classificação por estrelas e classificação MPAA, nem efeito da classificação MPAA, mas haja efeito da classificação por estrelas.

Vocabulário

análise da variância (ANOVA)
análise da variância de um critério
análise da variância de um fator
tratamento
fator
análise da variância de dois critérios
interação
planejamento completamente aleatorizado
variância entre amostras
variação devida ao tratamento
variância dentro de amostras
variação devida ao erro
processos de comparação múltipla
planejamento rigorosamente controlado
planejamento em blocos aleatorizados
bloco

Revisão

Neste capítulo aplicamos a análise da variância (ou ANOVA) para testar a igualdade de médias populacionais. Este método exige (1) populações distribuídas normalmente, (2) populações com o mesmo desvio-padrão (ou mesma variância), e (3) amostras aleatórias mutuamente independentes.

Na Seção 11-2 consideramos a análise da variância de um critério, caracterizada por dados amostrais categorizados segundo um único fator. São características-chave da análise da variância de um critério:

- A estatística de teste F se baseia na razão de duas estimativas diferentes da variância populacional comum σ^2, conforme abaixo:

$$F = \frac{\text{variância entre amostras}}{\text{variância dentro de amostras}} = \frac{\text{QM(tratamento)}}{\text{QM(erro)}}$$

- Os valores críticos de F encontram-se na Tabela A-5 e os graus de liberdade (gl) são:

gl para o numerador $= k - 1$ (onde k = número de amostras)
gl para o denominador $= N - k$ (onde N = número total de valores em todas as amostras combinadas)

Na Seção 11-3 abordamos a análise da variância de dois critérios, caracterizada por dados categorizados segundo dois fatores diferentes. A Figura 11-4 resume o método de análise da variância de dois critérios. Consideramos também a análise da variância de dois critérios com uma observação por célula e análise da variância de dois critérios com planejamento em blocos aleatorizados. Esses casos utilizam os mesmos métodos de ANOVA, mas em um experimento com blocos aleatorizados procuramos controlar a variação estranha considerando os blocos como um dos fatores.

Em razão da complexidade dos cálculos exigidos em todo este capítulo, enfatizamos a interpretação de apresentações de computador.

Exercícios de Revisão

1. O Associated Insurance Institute patrocina estudos dos efeitos da bebida sobre a direção de automóveis. Em um desses estudos, selecionaram-se aleatoriamente 3 grupos de homens adultos para um experimento destinado a medir seus níveis de álcool no sangue após consumir 5 drinques. Os componentes do grupo A foram testados após 1 hora, os do grupo B foram testados após 2 horas, e os do grupo C, após 4 horas. Os resultados constam da tabela a seguir; dá-se também a apresentação Minitab. Com um nível de 0,05 de significância, teste a alegação de que os 3 grupos têm o mesmo nível médio.

A	B	C
0,11	0,08	0,04
0,10	0,09	0,04
0,09	0,07	0,05
0,09	0,07	0,05
0,10	0,06	0,06
		0,04
		0,05

```
Analysis of Variance

Source   DF        SS          MS        F      p
Factor    2  0.0076571   0.0038286   46.90  0.000
Error    14  0.0011429   0.0000816
Total    15  0.0088000
```

2. A relação a seguir dá os preços de venda (em milhares de dólares) de casas situadas em Long Beach Island (New Jersey). Esperam-se preços diferentes, conforme as diversas localizações. Os dados amostrais apóiam a hipótese de diferença entre os preços médios de venda? Tome o nível de significância de 0,05.

À beira-mar:	235	395	547	469	369	279
Frente para o mar:	538	446	435	639	499	399
Na baía:	199	219	239	309	399	190
Frente para a baía:	695	389	489	489	599	549

3. Fez-se um teste de consumo de combustível (em mi/gal) em quatro carros diferentes, de 4 cilindros, após terem rodado em condições idênticas de rodovia; os resultados constam da tabela e da apresentação Minitab a seguir. Com o nível de 0,05 de significância, teste se o consumo de combustível é ou não afetado por uma *interação* entre tamanho do motor e tipo de transmissão.

```
ANALYSIS OF VARIANCE MPG

SOURCE        DF      SS     MS
TRANS          1    40.3   40.3
SIZE           2    43.2   21.6
INTERACTION    2     1.2    0.6
ERROR          6    68.0   11.3
TOTAL         11   152.7
```

Consumo de Combustível (mi/gal) em Rodovias para Diferentes Carros Compactos de 4 Cilindros

	Tamanho do Motor (litros)		
	1,5	2,2	2,5
Transmissão Automática	31,32	28,26	31,23
Transmissão Manual	33,36	33,30	27,34

4. Considere os mesmos dados do Exercício 3 e suponha que o consumo de combustível não seja afetado por uma interação entre tamanho do motor e tipo de transmissão. Considere um nível de 0,05 de significância para testar a afirmação de que o consumo de combustível não é afetado pelo tamanho do motor.
5. Considere os mesmos dados do Exercício 3 e suponha que o consumo de combustível não seja afetado por uma interação entre tamanho do motor e tipo de transmissão. Considere um nível de 0,05 de significância para testar a afirmação de que o consumo de combustível não é afetado pelo tipo de transmissão.

Exercícios Cumulativos de Revisão

1. A Rocky Mountain Brewing Company planeja lançar uma campanha publicitária de vulto. Três agências de propaganda prepararam provas de um comercial visando obter um contrato de $2 milhões. Os comerciais foram testados junto a consumidores selecionados aleatoriamente, medindo-se as respectivas reações; os resultados constam da tabela a seguir. (Escores mais elevados indicam maior reação positiva aos comerciais.)
   a. Construa um diagrama em caixa para cada uma das três amostras. Use a mesma escala a fim de que os diagramas possam ser comparados. Os diagramas revelam alguma diferença sensível?
   b. Ache a média e o desvio-padrão de cada um dos três conjuntos de dados amostrais.
   c. Com os métodos da Seção 8-5, teste a afirmação de que a população de valores de Barnum tem média igual à da população de valores de Solomon & Ford. Trabalhe com o nível de 0,05 de significância.
   d. Para cada uma das três amostras, construa uma estimativa intervalar de 95% de confiança para a média populacional μ. Os resultados sugerem alguma diferença sensível?
   e. Considere um nível de 0,05 de significância e teste a afirmação de que as três populações têm o mesmo escore médio de reação. Se o leitor fosse responsável pela propaganda dessa cervejaria, que companhia escolheria, com base nesses resultados? Por quê?

Barnum Advertising Co	Solomon & Ford Advertising	Diaz and Florio Advertising
52 68 75 40	69 73 82 59	42 73 69 53
77 63 55 72	66.84 75 70	57 61 73 74

2. Nos Estados Unidos, os pesos de recém-nascidos têm distribuição normal com média de 7,54 lb (3,42 kg) e desvio-padrão de 1,09 lb (0,49 kg) (com base em dados de "Birth Weight and Perinatal Mortality", por Wilcox, Skjaerven, Buekens e Kiely, *Journal of the American Medical Association*, Vol. 273, N.º 9).
   a. Selecionado aleatoriamente um recém-nascido, qual é a probabilidade de ele ou ela pesar mais de 8,0 lb?
   b. Selecionados aleatoriamente 16 recém-nascidos, qual é a probabilidade de seu peso médio ser superior a 8,00 lb?
   c. Qual é a probabilidade de que cada um dos próximos três recém-nascidos tenha peso superior a 7,54 lb?

Projeto para Computador

A tabela a seguir dá o número de anos vividos por presidentes, papas e monarcas ingleses após sua posse, eleição ou coroação. Com auxílio de diagramas em caixa e análise da variância, determine se há diferença entre os tempos de sobrevivência dos três grupos. Faça a análise da variância com STATDISK, Minitab ou outro programa estatístico. Obtenha cópias das apresentações do computador e redija suas observações e conclusões.

Presidentes		Papas		Reis e Rainhas		Presidentes		Papas		Reis e Rainhas	
Washington	10	Alex VIII	2	James II	17	Hayes	16	Pius X	17		
J. Adams	29	Innoc XII	9	Mary II	6	Garfield	0	Pius XII	19		
Jefferson	26	Clem XI	21	William III	13	Arthur	7	John XXIII	5		
Madison	28	Innoc XIII	3	Anne	12	Cleveland	24	Paul VI	15		
Monroe	15	Ben XIII	6	George I	13	Harrison	12	John Paul I	0		
J. Q. Adams	23	Clem XII	10	George II	33	McKinley	4				
Jackson	17	Ben XIV	18	George III	59	T. Roosevelt	18				
Van Buren	25	Clem XIII	11	George IV	10	Taft	21				
Harrison	0	Clem XIV	6	William IV	7	Wilson	11				
Tyler	20	Pius VI	25	Victoria	63	Harding	2				
Polk	4	Pius VII	23	Edward VII	9	Coolidge	9				
Taylor	1	Leo XII	6	George V	25	Hoover	36				
Fillmore	24	Pius VIII	2	Edward VIII	36	F. Roosevelt	12				
Pierce	16	Greg XVI	15	George VI	15	Truman	28				
Buchanan	12	Pius IX	32			Kennedy	3				
Lincoln	4	Leo XIII	25			Eisenhower	16				
A. Johnson	10	Pius X	11			L. Johnson	9				
Grant	17	Ben XV	8			Nixon	25				

Com base em *Computer-Interactive Data Analysis*, por Lunn e McNeil, John Wiley & Sons.

DOS DADOS À DECISÃO

Terá Sido o Experimento Planejado Corretamente?

Planejou-se um experimento para comparar os efeitos da perda de peso com os efeitos da prática de exercícios aeróbicos sobre os fatores de risco de doenças em artéria coronária, para homens sadios porém de vida sedentária, obesos e de meia-idade ou mais velhos. (Veja "Effects of Weight Loss vs. Aerobic Exercise Training on Risk Factors for Coronary Diseases in Helthy, Obese, Middle-aged and Older Men," por Katzel, Bleecker *et al.*, *Journal of the American Medical Association,* Vol. 274, N.º 24.) O estudo envolveu 170 homens divididos em três grupos: (1) 73 homens foram destinados a perder peso, (2) 71 homens foram destinados a exercícios aeróbicos e (3) 26 homens ficaram em um grupo de controle, com a recomendação de não perder peso nem modificar o nível de sua atividade física. Em tal estudo, é importante partir de grupos que sejam análogos em características "básicas" de idade, peso, massa do corpo, gordura etc. Um relatório sobre o estudo afirmou que "Não houve diferenças significativas entre os três grupos no que toca qualquer das características básicas." Os pesos médios dos três grupos amostrais são 94,3 kg, 93,9 kg e 88,3 kg. Essas diferenças são significativas? Há evidência suficiente para rejeitar a afirmação de que não há diferença significativa entre os três grupos? Analise os dados da lista a seguir, formule uma conclusão e escreva um relatório indicando os resultados.

Pesos (em kg) de homens do grupo de perda de peso:

95	82	94	87	84	111	98	101	85	117	111	86	114	85	98	102	97
92	76	95	88	107	89	81	93	77	91	106	97	107	78	101	88	92
101	110	96	83	92	106	107	105	110	102	105	84	93	69	96	70	80
108	87	86	93	78	98	86	93	101	94	85	107	74	101	105	93	89
90	103	103	80	116												

Pesos (em kg) dos homens do grupo de exercícios aeróbicos:

90	101	89	76	93	110	87	108	109	105	83	98	100	82	71	67	126
114	102	78	101	79	97	100	106	84	80	100	103	80	71	99	108	84
78	92	92	101	107	95	106	84	111	108	109	79	116	98	77	110	93
95	85	86	104	93	92	86	93	74	74	89	85	91	104	97	97	80
104	95	109														

Pesonms (em kg) dos homens do grupo de controle:

91	71	74	86	97	78	81	82	78	106	79	110	91	93	87	72	95
88	83	96	98	77	84	95	103	102								

ATIVIDADES EM GRUPO

1. *Atividade em Classe*: Divida a turma em grupos de cinco ou seis estudantes. Esta atividade é um confronto de tempos de reação para determinar que grupo é mais rápido, se é que existe um grupo "mais rápido". Utilize o mesmo cronometrador de reação incluído nas Atividades em Grupo do Capítulo 5. Teste e regitre o tempo de reação utilizando a mão principal de cada elemento do grupo. (Apenas um teste por pessoa.) Cada grupo deve calcular os valores de n, $\bar{x}$ e s e registrar essas estatísticas resumo no quadro juntamente com a relação original de tempos de reação. Depois de todos os grupos terem reportado seus resultados, identifique o grupo com o tempo médio de reação mais rápido. Mas esse grupo é realmente o mais rápido? Recorra a ANOVA para determinar se as médias são significativamente diferentes. Se não o forem, não haverá um "ganhador" real.
2. *Atividade Extraclasse*: O World Almanac and Book of Facts contém uma seção intitulada "Noted Personalities" com subseções dedicadas a arquitetos, artistas, administradores, cartunistas, cientistas sociais, líderes militares, filósofos, líderes políticos, cientistas, escritores, compositores, anfitriões e outros. Planeje e faça um estudo observacional que comece com a extração de amostras de grupos selecionados e continue com uma comparação da duração de vida de pessoas de diferentes categorias. Algum grupo em particular parece ter duração de vida diferente da dos outros grupos? Pode explicar essas diferenças?

entrevista

Lester Curtin

Chefe da Equipe de Métodos Estatísticos do Office of Research and Methodology, National Center for Health Statistics, Centers for Disease Control.

Lester Curtin é doutor em estatística e atua temporariamente na Division of Vital Statistics. Atua normalmente como Chefe do Statistical Methods Staff no National Center for Health Statistics, que faz parte dos Centers for Disease Control.

Quais as atribuições do Centro Nacional de Estatística da Saúde?

A finalidade da agência, definida pelo Congresso, é coletar dados sobre a saúde da nação. Coletamos estatísticas vitais, inclusive nascimentos, mortes, casamentos e divórcios. Fazemos também pesquisas nacionais em larga escala, incluindo a National Health Interview Survey, que é uma pesquisa por entrevistas pessoais com cerca de 120.000 indivíduos cada ano. Realizamos a National Health and Nutrition Examination Survey, que envolve centros móveis de exame que percorrem o país coletando dados sobre os habitantes. Obtemos medidas biológicas em lugar de meras respostas de entrevistas, sujeitas a erros de mensuração. Para determinar o nível de cálcio, além de perguntar ao entrevistado o que ele comeu, extraímos também amostras de sangue para obter níveis de cálcio. Nossas estatísticas servem para monitorar a saúde da nação e para determinar as modificações em curso.

Pode citar um exemplo de como seus dados são utilizados?

O National Health and Nutrition Exam Survey vinha sendo feito mais ou menos na ocasião em que o Congresso proibiu o uso do chumbo na gasolina. Verificou-se então um movimento pela volta da adição do chumbo à gasolina. Como os pesquisadores estavam permanentemente em campo, pudemos fazer uma análise temporal, e constatamos uma queda dos níveis de chumbo nas pessoas. Esta conclusão foi combatida pelos técnicos da indústria de gasolina. O governo também tinha seus técnicos, e os dois grupos prosseguiram na disputa. Utilizamos dados do National Health ond Nutrition Exom Survey para testemunhar, perante o Congresso, em favor da exclusão do chumba da gasolina. A validade de nossos dados foi reconhecida e teve grande influência na lei que proibiu o chumbo na gasolina.

Às vezes é curioso lidar com estatísticas sanitárias. Medimos a cotinina para avaliar os efeitos do fumo passivo. Utilizamos dados sobre despesas monetárias para o modelamento matemático dos custos usados para a reforma de programa de saúde proposto pelo presidente. Pesquisamos o que poderia fazer diferença em termos de hábitos dietéticos sobre o futuro da saúde.

Em seu campo de atividade, o emprego da estatística está aumentando, diminuindo ou permanece estacionário?

O National Health and Nutrition Exam Survey tinha, em certa ocasião, o ponto de vista de que nosso grupo devia apenas coletar e difundir os dados. Em anos recentes, passamos a nos preocupar muito mais com os aspectos analíticos dos dados. Estamos progredindo mais ainda com a utilização de métodos estatísticos sofisticados.

Que métodos estatísticos utiliza?

Fazemos praticamente tudo: trabalho de planejamento de pesquisa, controle de qualidade para codificar e editar os dados; estudamos o aspecto cognitivo do planejamento de pesquisa e como o povo pensa ao responder uma questão. Podemos formular uma pergunta de duas ou três maneiras diferentes, para verificar o impacto na resposta do entrevistado. Utilizamos intervalos de confiança, teste de hipóteses, regressão, análise de séries temporais, modelamento ARIMA, regressão logística sobre fatores de risco e análise da variância. O National Health and Nutrition Exam Survey emprega cerca de 200 estatísticos e outros 50 funcionários com conhecimentos estatísticos quantitativos gerais. Temos funcionários com graus de bacharel, mestre e Ph.D.

Como consegue convencer as pessoas da validade de seus dados?

Adotamos a política básica de não apresentar uma estimativa a não ser que apresentemos também alguma medida do erro amostral. Outrossim, ao publicarmos um relatório do National Center for Health Statistics e dizermos que uma estatística é diferente de outra, temos como norma fundamentar nossa afirmação com um teste estatístico.