de $f(y)$ para y entre $\mu + z\sigma$ e ∞ é igual a probabilidade da cauda apresentada na Tabela A.)

*4.57 A fórmula do erro padrão $\sigma_{\bar{y}} = \sigma/\sqrt{n}$ trata o tamanho da população N como infinitamente grande relativo ao tamanho da amostra n. A fórmula $\sigma_{\bar{y}}$ para uma população finita de tamanho N é:

$$\sigma_{\bar{y}} = \sqrt{\frac{N-n}{N-1}}\left(\frac{\sigma}{\sqrt{n}}\right).$$

O termo $\sqrt{(N-n)/(N-1)}$ é chamado de **correção de população finita**.

(a) Quando $n = 300$ estudantes são selecionados de um corpo docente de uma faculdade de tamanho $N = 30000$, mostre que $\sigma_{\bar{y}} = 0.995\sigma/\sqrt{n}$. (Em geral, n é bem menor quando comparado a N; assim a correção, na maioria das vezes, tem pouca influência prática.)

(b) Se $n = N$ (isto é, amostramos toda a população), mostre que $\sigma_{\bar{y}} = 0$. Em outras palavras, não ocorre um erro amostral porque $\bar{y} = \mu$.

(c) Para $n = 1$, explique por que a distribuição amostral de $\bar{y}$ e seu erro padrão são idênticos à distribuição da população e seu desvio padrão.

☑ NOTAS

1 *Journey to Work*, editado em 2004 pela Agência do Censo dos Estados Unidos.
2 www.cnn.com/ELECTION/2006.
3 *British Medical Journal*. V. 307, 1993, p. 234.

☑ 5 INFERÊNCIA ESTATÍSTICA: ESTIMAÇÃO

Este capítulo mostra como usar os dados amostrais para estimar os parâmetros da população. Com variáveis quantitativas estimamos a média da população. Um estudo que trata de assuntos do sistema de saúde, por exemplo, pode estimar os parâmetros da população como a quantia média de dinheiro gasta em medicamentos prescritos durante o último ano e o número médio de visitas ao médico. Com variáveis categóricas, estimamos as proporções da população para as categorias. O estudo do sistema de saúde pode estimar as proporções das pessoas que (têm, não têm) seguro de saúde e as proporções que (estão satisfeitas, não estão satisfeitas) com seu plano de saúde.

Inicialmente aprenderemos sobre dois tipos de estimativas dos parâmetros. Após, nas Seções 5.2 e 5.3, as aplicaremos às médias e proporções da população. A Seção 5.4 encontra o tamanho da amostra necessário para alcançar a precisão desejada da estimativa. A Seção 5.5 discute a estimativa da mediana e de outros parâmetros.

5.1 ESTIMAÇÃO POR PONTO E INTERVALAR

Existem dois tipos de estimativas de parâmetros:

- Uma **estimativa por ponto** é um único número que é a melhor avaliação do parâmetro.
- Uma **estimativa intervalar** é um conjunto de números em torno da estimativa por ponto dentro do qual o valor do parâmetro deve estar.

Por exemplo, uma PSG perguntou: "Você acredita em vida após a morte?". Para 1958 sujeitos amostrados, a estimativa por ponto para a proporção de todos os norte-americanos que iriam responder *sim* é igual a 0,73. Uma estimativa intervalar prevê que a proporção da população que respondeu *sim* está entre 0,71 e 0,75. Isto é, ela prevê que a estimativa por ponto de 0,73 está dentro de *uma margem de erro* de 0,02 do valor real. Portanto, uma estimativa intervalar nos ajuda a avaliar a precisão provável de uma estimativa por ponto.

O termo *estimativa*, apenas, é geralmente usado como uma abreviação de *estimativa por ponto*. O termo *estimador* se refere a um tipo específico de estatística para estimar um parâmetro e *estimativa* se refere ao seu valor para uma amostra específica. Por exemplo, a proporção amostral é um estimador da proporção populacional. O valor 0,73 é a estimativa para a proporção da população que acredita em vida após a morte.

Estimativa por ponto

Qualquer parâmetro em particular tem muitos estimadores possíveis. Para uma população normalmente distribuída, por exemplo, o centro é a média e a mediana, uma vez que ela é simétrica. Assim, com dados amostrais, dois estimadores possíveis do centro da população são a média e a mediana da amostra.

As estimativas são as inferências estatísticas mais comuns relatadas pelos meios de comunicação de massa. Por exemplo, um levantamento de dados da Gallup de janeiro de 2007 verificou que 36% do público norte-americano aprovava o desempenho do presidente George W. Bush no seu cargo. Essa é uma estimativa em vez de um parâmetro porque foi baseada nas entrevistas de uma amostra de aproximadamente 1000 pessoas em vez de em toda a população.

Estimadores por ponto eficientes e não tendenciosos

Um bom estimador tem uma distribuição amostral que: (1) está centrada em torno do parâmetro e (2) tem um erro padrão tão pequeno quanto possível.

Um estimador é **não tendencioso** se sua distribuição amostral está centrada em torno do parâmetro. Especificamente, o parâmetro é a média da distribuição amostral. Da Seção 4.4, para amostras aleatórias, a média da distribuição amostral da média da amostra $\bar{y}$ é igual à média da população μ. Assim, $\bar{y}$ é um estimador não tendencioso da média da população μ. A Figura 5.1 ilustra a situação. Para qualquer amostra em particular, a média da amostra pode subestimar μ ou pode superestimá-la. Se, contudo, a média da amostra fosse encontrada repetidamente com diferentes amostras, as superestimativas tenderiam a contrabalançar as subestimativas.

Ao contrário, um estimador **tendencioso** tende, em média, a subestimar ou a superestimar o parâmetro. Por exemplo, o intervalo ou amplitude da amostra é tipicamente menor do que o intervalo populacional e não pode ser maior, porque o mínimo e o máximo da amostra não podem ser mais extremos do que o mínimo e o máximo da população. Portanto, o intervalo amostral tende a subestimar o intervalo populacional. Ele é um estimador tendencioso da amplitude da população.

Suponha que a distribuição da população é assimétrica à direita, como mostra a Figura 5.1, e você quer estimar a média dessa população. Se você está preocupado com o efeito dos valores atípicos, você pode decidir estimar a média utilizando a mediana amostral em vez da média. Entretanto, a mediana da população é menor do que a média da população, nesse caso, e a mediana amostral tende a ser menor do que a média da população μ. Portanto, a mediana amostral é um estimador tendencioso de μ, em média, subestimando μ. É melhor usar a média amostral, talvez calculá-la após eliminar os valores extremos, se eles tiverem uma influência indevida ou refletirem erros no registro dos dados.

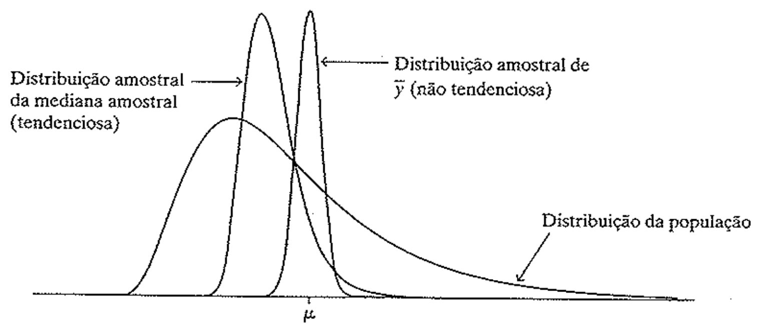


Figura 5.1 Distribuições amostrais de dois estimadores da média populacional para uma distribuição populacional assimétrica.

A segunda propriedade desejável para um estimador é que ele tenha um erro padrão relativamente pequeno. Um estimador que tem um erro padrão menor do que outros estimadores do mesmo parâmetro é dito ser **eficiente**. Um estimador eficiente está mais próximo, em média, do parâmetro do que outros estimadores. Por exemplo, quando uma população tem uma distribuição normal, o erro padrão da mediana amostral é 25% maior do que o erro padrão da média amostral. A média amostral tende a estar mais próxima do que a mediana do centro da população. A média da amostra, nesse caso, é um estimador mais eficiente do que a mediana.

Em resumo, um bom estimador de um parâmetro é *não tendencioso*, ou quase isto, e *eficiente*. Os métodos estatísticos utilizam estimadores que possuem essas propriedades.

Estimadores da média, do desvio padrão e da proporção

É comum, mas não necessário, usar o valor amostral análogo de um parâmetro como seu estimador. Por exemplo, para estimar a proporção populacional, a proporção amostral é um estimador não tendencioso e eficiente. Para estimar a média da população μ, a média amostral $\bar{y}$ é não tendenciosa. Ela é eficiente para as distribuições populacionais mais comuns. Da mesma forma, usamos o desvio padrão da amostra s como um estimador do desvio padrão da população σ.

O sinal "^" sobre um símbolo do parâmetro é geralmente usado para representar uma estimativa daquele parâmetro. O sinal "^" é chamado de *circunflexo* e com frequência é lido como *chapéu*. Por exemplo, $\hat{\mu}$ é lido como *mi-chapéu*. Portanto, $\hat{\mu}$ representa uma estimativa da média da população μ.

Método de estimação de máxima verossimilhança*

As contribuições mais importantes para a ciência da estatística moderna foram feitas por um estatístico e geneticista britânico, Ronald A. Fisher (1890 – 1962). Enquanto trabalhava em uma estação de pesquisa agrícola ao norte de Londres, ele desenvolveu muito da teoria da estimativa por ponto, assim como a metodologia do projeto de experimentos e da análise de dados.

Para uma estimativa por ponto, Fisher advogava a **estimativa de máxima verossimilhança**. Essa estimativa é o valor do parâmetro que é mais consistente com os dados observados no seguinte sentido: se o parâmetro é igual àquele número (isto é, o valor da estimativa), os dados observados teriam maior chance de ocorrer do que se o parâmetro fosse igual a qualquer outro número. Por exemplo, um levantamento de dados recente de aproximadamente 1000 norte-americanos adultos relataram que a estimativa de máxima verossimilhança da proporção populacional que acredita em astrologia é de 0,37. Portanto, a amostra observada seria mais provável de acontecer se a proporção da população fosse 0,37 do que se ela fosse igual a qualquer outro valor possível.

Para muitas distribuições da população, como a normal, a estimativa de máxima verossimilhança da média da população é a média amostral. As principais estimativas por ponto apresentadas neste livro são, sob certas hipóteses populacionais, estimativas de máxima verossimilhança. Fisher mostrou que, para amostras grandes, estimadores de máxima verossimilhança apresentam três propriedades desejáveis:

- Eles são eficientes para amostras relativamente grandes: outros estimadores não têm erros padrão menores e não tendem a estar próximos do parâmetro.
- Eles têm pouca, ou nenhuma, tendenciosidade, com a tendenciosidade di-

minuindo à medida que o tamanho da amostra aumenta.
- Eles têm distribuições amostrais aproximadamente normais.

Intervalo de confiança é estimativa por ponto ± margem de erro

Para ser bem claro, uma inferência sobre um parâmetro deve fornecer não somente uma estimativa por ponto mas também indicar quão próximo a estimativa é provável de estar do valor do parâmetro. Por exemplo, desde 1988 a cada ano a Florida Poll (Pesquisa da Flórida) conduzida pela Universidade Internacional da Flórida (www.fiu.edu/orgs/ipor/ffp) pergunta para aproximadamente 1200 moradores do estado se relações sexuais entre dois adultos do mesmo sexo é errado. O percentual que disse que sempre é errado diminuiu de 78% em 1988 para 54% em 2006. Quão precisas são essas estimativas? Dentro de 2%? Dentro de 5%? Dentro de 10%?

A informação sobre a precisão de uma estimativa por ponto determina a amplitude de uma *estimativa intervalar* para o parâmetro, que consiste em um intervalo de valores em torno da estimativa por ponto. O intervalo é construído de forma a tentar conter o parâmetro com uma probabilidade fixada e próxima a 1. Visto que a estimativa intervalar pode conter o parâmetro com certo grau de confiança, eles são denominados de **intervalos de confiança**.

> ☑ **Intervalo de confiança**
>
> Um **intervalo de confiança** para um parâmetro é um intervalo de números que, se acredita, contenha o parâmetro com determinada probabilidade. A probabilidade de que o intervalo construído em torno da estimativa pontual contenha o parâmetro é denominada de **nível de confiança**. Ele é um número escolhido e está próximo a 1, como 0,95 ou 0,99.

A chave para construir um intervalo de confiança é a distribuição amostral do estimador por ponto. Frequentemente, a distribuição amostral é aproximadamente normal. A distribuição normal, então, determina a probabilidade de que o estimador esteja a certa distância do parâmetro. Com a probabilidade aproximada de 0,95, o estimador estará dentro de dois erros padrão. É quase certo que ele estará dentro de três erros padrão. Quanto menor o erro padrão, mais preciso o estimador tende a ser.

Na prática, com frequência a distribuição amostral é aproximadamente normal. Então, para construir um intervalo de confiança, somamos e subtraímos da estimativa do ponto um múltiplo (um escore-z) do seu erro padrão. Esse múltiplo do erro padrão é a **margem de erro**. Isto é:

> Um intervalo de confiança tem a forma: **estimativa por ponto ± margem de erro.**

Para construir um intervalo tendo "95% de confiança", tomamos a estimativa por ponto e somamos e subtraímos a margem de erro que é a aproximadamente igual a dois erros padrão. Veremos os detalhes nas próximas duas seções.

5.2 INTERVALO DE CONFIANÇA PARA UMA PROPORÇÃO

Para dados categóricos, uma observação ocorre em uma de um conjunto de categorias. Esse tipo de medida ocorre quando a variável é nominal, como o candidato de preferência (Democrata, Republicano, Independente) ou ordinal, como a opinião sobre o gasto do governo (aumenta, mantém o mesmo, diminui). Ela também ocorre quando variáveis inerentemente contínuas são mensuradas com escalas categóricas, como quando a renda anual tem categorias ($0 – $24999, $25000 – $49999, $50000 – $74999, pelo menos $75000).

Para resumir dados categóricos, registramos as *proporções* (ou *percentuais*) das observações nas categorias. Por exemplo, um estudo pode fornecer uma estimativa por ponto ou intervalo da:

- Proporção de norte-americanos que têm seguro de saúde.
- Proporção de canadenses que são a favor da independência de Quebec.
- Proporção de australianos que estão desempregados.

A proporção amostral e seu erro padrão

Considere π a proporção populacional.[1] Portanto, π está entre 0 e 1. Sua estimativa por ponto é a *proporção da amostra*. Representamos a proporção amostral por $\hat{\pi}$, visto que ela estima π.

Lembre que a proporção amostral é uma média quando consideramos $y = 1$ para uma observação na categoria e interesse e $y = 0$, caso contrário. (Veja a discussão sobre a Tabela 3.6 na página 62 e o Exemplo 4.7 seguinte, na página 108.) Da mesma forma, a proporção populacional π é a média μ da distribuição da probabilidade tendo as probabilidades

$$P(1) = \pi \text{ e } P(0) = 1 - \pi.$$

O desvio padrão dessa distribuição da probabilidade é $\sigma = \sqrt{\pi(1-\pi)}$. (O Exercício 4.55 derivou essa fórmula.) Visto que a fórmula do erro padrão da média amostral é igual a $\sigma_{\bar{y}} = \sigma/\sqrt{n}$, o erro padrão $\sigma_{\hat{\pi}}$ da proporção amostral $\hat{\pi}$ é

$$\sigma_{\hat{\pi}} = \sigma/\sqrt{n} = \sqrt{\frac{\pi(1-\pi)}{n}}.$$

À medida que o tamanho da amostra aumenta, o erro padrão fica menor. A proporção amostral, então, tende a estar mais próxima da proporção populacional.

Intervalo de confiança para uma proporção com grandes amostras

Visto que a proporção amostral $\hat{\pi}$ é uma média amostral, o Teorema Central do Limite se aplica: para amostras aleatórias grandes, a distribuição amostral de $\hat{\pi}$ é aproximadamente normal em relação ao parâmetro π que ele estima. A Figura 5.2 ilustra o fato.

Lembre que 95% de uma distribuição normal estão dentro de dois desvios padrão da média, ou mais precisamente, 1,96 desvios padrão. Nós acabamos de ver que o erro padrão da proporção amostral é $\sigma_{\hat{\pi}} = \sqrt{\pi(1-\pi)/n}$. Assim, com a probabilidade de 0,95, $\hat{\pi}$ está dentro de $1{,}96\sigma_{\hat{\pi}}$ unidades do parâmetro π, isto é, entre $\pi - 1{,}96\sigma_{\hat{\pi}}$ e $\pi + 1{,}96\sigma_{\hat{\pi}}$, como mostra a Figura 5.2.

Uma vez que a amostra é selecionada, se $\hat{\pi}$ está dentro de $1{,}96\sigma_{\hat{\pi}}$ unidades de π, então o intervalo entre $\hat{\pi} - 1{,}96\sigma_{\hat{\pi}}$ a $\hat{\pi} + 1{,}96\sigma_{\hat{\pi}}$ contém π. Veja a linha 1 da Figura 5.2. Em outras palavras, com a probabilidade de 0,95, um valor $\hat{\pi}$ ocorre de tal forma que o intervalo $\hat{\pi} \pm 1{,}96\sigma_{\hat{\pi}}$ contém a proporção populacional π.

Por outro lado, a probabilidade é de 0,05 de que $\hat{\pi}$ *não* esteja dentro de $1{,}96\sigma_{\hat{\pi}}$ de π. Se isto acontecer, então o intervalo de $\hat{\pi} - 1{,}96\sigma_{\hat{\pi}}$ a $\hat{\pi} + 1{,}96\sigma_{\hat{\pi}}$ *não* contém π (veja a Figura 5.2, linha 2). Portanto, a probabilidade é 0,05 de que $\hat{\pi}$ é tal que $\hat{\pi} \pm 1{,}96\sigma_{\hat{\pi}}$ *não* contém π.

O intervalo $\hat{\pi} \pm 1{,}96\sigma_{\hat{\pi}}$ é uma estimativa intervalar para π com um nível de confiança de 0,95. Ele é chamado de **intervalo de 95% de confiança**. Na prática, o valor do erro padrão $\sigma_{\hat{\pi}} = \sqrt{\pi(1-\pi)/n}$ não pode ser calculado porque ele depende do parâmetro desconhecido π. Assim, estimamos o erro padrão substituindo o parâmetro pela proporção amostral e a fórmula anterior fica:

$$ep = \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}.$$

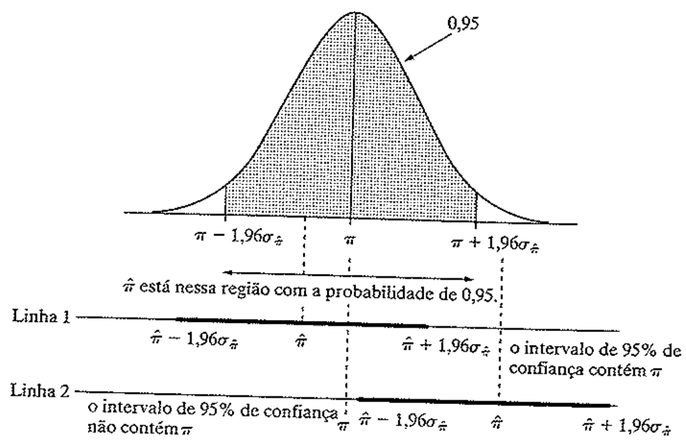


Figura 5.2 Distribuição amostral de $\hat{\pi}$ e possíveis intervalos de 95% de confiança para π.

Usamos o símbolo s para representar o desvio padrão da amostra que estima o desvio padrão da população σ. No restante deste livro, usaremos o símbolo ep para representar a estimativa amostral do erro padrão.

A fórmula do intervalo de confiança usa esse erro padrão estimado. Em resumo, o intervalo de 95% de confiança para π é dado por

$$\hat{\pi} \pm 1{,}96(ep) \quad \text{onde } ep = \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}.$$

EXEMPLO 5.1 Estimando a proporção de quem é a favor da restrição do aborto legalizado

A Florida Poll de 2006 (www.fiu.edu/orgs/ipor/ffp) conduzida pela Universidade Internacional da Flórida perguntou: "Em geral, você acha que é apropriado para o governo do estado fazer leis que restringem o acesso ao aborto?". De 1200 moradores da Flórida, aleatoriamente escolhidos, 396 disseram *sim* e 804 disseram *não*. Vamos estimar a proporção populacional que iria responder *sim* a esta pergunta.

Considere π o valor que representa a proporção populacional de adultos da Flórida que responderam *sim*. De $n = 1200$ respondentes, 396 disseram *sim*, então $\hat{\pi} = 396/1200 = 0{,}33$. Então, $1 - \hat{\pi} = 0{,}67$. Isto é, 33% da amostra disseram *sim* e 67% disseram *não*.

O erro padrão estimado da proporção amostral $\hat{\pi}$ é igual a

$$ep = \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}} = \sqrt{\frac{(0{,}33)(0{,}67)}{1200}} = \sqrt{0{,}000184} = 0{,}0136.$$

Um intervalo de 95% de confiança para π é

$$\hat{\pi} \pm 1{,}96(ep) = 0{,}33 \pm 1{,}96(0{,}0136) = 0{,}33 \pm 0{,}03, \text{ ou } (0{,}30; 036).$$

O percentual da população que apoia a restrição ao acesso ao aborto parece ser pelo menos 30%, mas não mais do que 36%. Todos os números no intervalo de confiança (0,30, 0,36) estão abaixo de 0,50. Portanto, aparentemente menos do que metade da população adulta da Flórida apoia a restrição ao acesso ao aborto. ■

Os resultados de tais levantamentos de dados variam grandemente dependendo da formulação da pergunta e de onde a pesquisa foi conduzida. Por exemplo, quando a PSG de 2006 perguntou se as mulheres grávidas deveriam ser capazes de obter um aborto legal se assim o desejassem por qualquer razão (variável "ABANY"), 1155 disseram *não* e 784 disseram *sim*. Você pode verificar que o intervalo de 95% de confiança para a proporção populacional que disse *não* é igual a (0,57, 0,62).

EXEMPLO 5.2 Estimando a proporção "contra" da proporção "a favor"

Na Florida Poll, para estimar a proporção populacional que apoia a restrição ao acesso ao aborto, obtivemos $ep = 0{,}0136$ para a estimativa por ponto $\hat{\pi} = 0{,}33$. Da mesma forma, o erro padrão estimado para $1 - \hat{\pi} = 0{,}67$, a proporção dos eleitores que dizem *não* à restrição ao acesso ao aborto, é

$$ep = \sqrt{(1-\hat{\pi})\hat{\pi}/n} = \sqrt{(0{,}67)(0{,}33)/1200} = 0{,}0136.$$

Assim, as duas proporções têm o mesmo ep.

Um intervalo de 95% de confiança para a proporção populacional de respostas *não* à restrição ao acesso ao aborto é

$$0{,}67 \pm 1{,}96(0{,}0136) = 0{,}67 \pm 0{,}03 \text{ ou } (0{,}64; 0{,}70).$$

Agora, $0{,}64 = 1 - 0{,}36$ e $0{,}70 = 1 - 0{,}30$, onde (0,3; 0,36) é o intervalo de 95% de confiança para π. Portanto, inferências para a proporção $1 - \pi$ seguem diretamente daquelas para a proporção π subtraindo de 1,0 cada ponto final do intervalo de confiança. ■

Se você construir um intervalo de confiança usando uma calculadora, não arredonde enquanto estiver calculando ou sua resposta pode ser afetada, mas arredonde quando você chegar à resposta final. Da mesma forma, quando apresentar os resultados de um *software*, você deve usar os dois ou três primeiros dígitos significativos. Escreva o intervalo de confiança como (0,30; 0,36) ou (0,303; 0,357) em vez de (0,303395; 0,356605). A precisão extra do *software* fornece cálculos precisos para encontrar o ep e o intervalo de confiança. Entretanto, os dígitos extras são apenas distração nos relatórios e não são úteis. Eles não nos dizem nada extra, na prática, sobre a proporção populacional.

Controlando o nível de confiança

Com um intervalo de 0,95 de confiança, isto é, "95% de confiança", existe uma probabilidade de 0,05 de que o método produza um intervalo de confiança que *não* contenha o valor do parâmetro. Em algumas aplicações, uma chance de 5% de uma inferência incorreta é inaceitável. Para aumentar a chance de uma inferência correta usamos um nível de confiança maior, como 0,99.

EXEMPLO 5.3 Encontrando um intervalo de 99% de confiança

Utilizando os dados do Exemplo 5.1, vamos encontrar um intervalo de 99% de confiança para a proporção populacional dos que são a favor de leis que restringem o acesso ao aborto. Agora, 99% de uma distribuição normal ocorrem dentro de 2,58 desvios padrão da média. Assim, a probabilidade será de 0,99 de que a proporção amostral $\hat{\pi}$ estará dentro de 2,58 erros padrão da proporção populacional π. Um intervalo de 99% de confiança para π é obtido por $\hat{\pi} \pm 2{,}58(ep)$.

No Exemplo 5.1, a proporção amostral era 0,33 com $ep = 0{,}0136$. Assim, o intervalo de 99% de confiança será:

$$\hat{\pi} \pm 2{,}58(ep) = 0{,}33 \pm 2{,}58(0{,}0136) = 0{,}33 \pm 0{,}04, \text{ ou } (0{,}29; 0{,}37).$$

Comparado com o intervalo de 95% de confiança de (0,30, 0,36), esta estimativa é menos precisa pois o intervalo tem uma amplitude um pouco maior. Para termos mais certeza de capturarmos o parâmetro, devemos sacrificar a precisão da estimativa usando um intervalo maior. ■

A forma geral para o intervalo de confiança para a proporção populacional π é

$$\hat{\pi} \pm z(ep), \quad \text{com } ep = \sqrt{\hat{\pi}(1 - \hat{\pi})/n},$$

onde z depende do nível de confiança. Quanto mais alto o nível de confiança, maior a chance de que o intervalo contenha o parâmetro. Níveis de confiança altos são usados na prática, de modo que a chance de erro é pequena. O intervalo de confiança mais comum é o de 0,95, com 0,99 sendo usado quando for mais crucial não cometer um erro.

O valor-z multiplicado pelo ep é a *margem de erro*. Com uma confiança maior, o intervalo de confiança é maior porque o escore-z na margem de erro é maior, por exemplo, $z = 1{,}96$ para 95% de confiança e $z = 2{,}58$ para 99% de confiança.

Por que não estabelecemos nada menos do que 100% de confiança? Para estarmos 100% certos de uma inferência correta, o intervalo deve conter todos os valores possíveis para π. Um intervalo de 100% de confiança para a proporção populacional a favor da limitação ao acesso ao aborto vai de 0,0 a 1,0. Isso não é útil. Na prática, estabelecemos menos do que a perfeição para estimarmos mais precisamente o valor do parâmetro. Na formação de um intervalo de confiança, temos um compromisso entre a confiança desejada de que a inferência esteja correta e a precisão requerida da estimação. À medida que uma fica melhor, a outra fica pior. É por isso que você não irá ver, normalmente, um intervalo de 99,9999% de confiança. Seria pouco preciso estimar onde o parâmetro da população está, pois o valor-z é 4,9.

Tamanhos de amostra maiores dão intervalos menores

Esperamos ser capazes de estimar a proporção populacional π mais precisamente com uma amostra de tamanho maior. A margem de erro é $z(ep)$, onde $ep = \sqrt{\hat{\pi}(1 - \hat{\pi})/n}$. Quanto maior o valor de n, menor a margem de erro e, portanto, menor o intervalo.

Para ilustrar, suponha que $\hat{\pi} = 0{,}33$, no Exemplo 5.1, para estimar a proporção de quem é a favor da legalização do aborto, tivesse como base uma amostra de $n = 300$, isto é, somente um quarto do tamanho real da amostra de $n = 1200$. Então, o erro padrão estimado de

$$ep = \sqrt{\hat{\pi}(1 - \hat{\pi})/n}$$
$$= \sqrt{(0{,}33)(0{,}67)/300} = 0{,}027$$

é duas vezes maior do que o ep do Exemplo 5.1. O intervalo de 95% de confiança resultante é:

$$\hat{\pi} \pm 1{,}96(ep)$$
$$= 0{,}33 \pm 1{,}96(0{,}027)$$
$$= 0{,}33 \pm 0{,}067.$$

Assim, ele tem o dobro da amplitude do intervalo de confiança formado quando a amostra utilizada foi a de tamanho $n = 1200$.

Visto que a margem de erro é inversamente proporcional à raiz quadrada de n e visto que $\sqrt{4n} = 2\sqrt{n}$, o tamanho da amostra deve *quadruplicar* para *dobrar* a precisão (isto é, metade da amplitude). A Seção 5.4 mostra como encontrar o tamanho da amostra necessário para alcançar determinada precisão.

> ☑ **A amplitude de um intervalo de confiança**
>
> 1. Aumenta à medida que o nível de confiança aumenta.
> 2. Diminui à medida que o tamanho da amostra aumenta.

Essas propriedades se aplicam a todos os intervalos de confiança, não apenas àquele para uma proporção.

Probabilidade de erro = 1 – nível de confiança

A probabilidade de que um método de estimação de um intervalo gere um intervalo de confiança que *não* contenha o parâmetro é chamado de **probabilidade de erro**. Isso é igual a 1 menos o nível de confiança. Para um nível de confiança de 0,95, a probabilidade de erro é igual a $1 - 0{,}95 = 0{,}05$. A letra grega α (alfa) representa essa probabilidade de erro e, então, $1 - \alpha$ é o nível de confiança. Para uma probabilidade de erro de $\alpha = 0{,}05$, o nível de confiança é igual a $1 - \alpha = 0{,}95$.

O valor-z para o intervalo de confiança é tal que a probabilidade é α que $\hat{\pi}$ está *mais do que* z erros padrão de π. O valor-z corresponde a probabilidade total de α nas duas caudas de uma distribuição normal ou $\alpha/2$ (metade da probabilidade de erro) em cada cauda. Por exemplo, para um intervalo de 95% de confiança, $\alpha = 0{,}05$ e o escore-z é aquele com a probabilidade $\alpha/2 = 0{,}05/2 = 0{,}025$ em cada cauda. Isto é $z = 1{,}96$.

Nível de confiança é a proporção correta a longo prazo*

O nível de confiança para um intervalo de confiança descreve como é o desempenho do método quando usado repetidamente com várias amostras aleatórias diferentes. A proporção populacional desconhecida π é um número fixo. Um intervalo de confiança construído de qualquer amostra em particular contém ou não contém π. Se repetidamente selecionarmos amostras aleatórias daquele tamanho e a cada vez construirmos um intervalo de 95% de confiança, então, a longo prazo, aproximadamente 95% dos intervalos conterão π. Isso acontece porque aproximadamente 95% das proporções amostrais estariam dentro de $1{,}96(ep)$ de π, como está $\hat{\pi}$ na linha 1 da Figura 5.2 (página 136). Dizer que um intervalo em particular contém π com "95% de confiança" significa que, *a longo prazo*, 95% de tais intervalos conterão π. Isto é, em 95% das vezes a inferência estará correta.

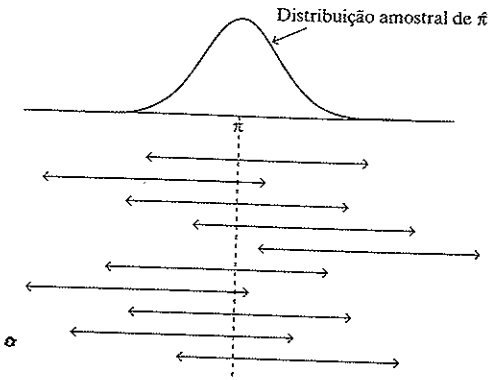
A Figura 5.3 mostra os resultados da seleção de dez amostras separadas e o cálculo da proporção amostral para cada uma e um intervalo de 95% de confiança para a proporção populacional em cada caso. Os intervalos de confiança se alternam porque $\hat{\pi}$ varia de amostra para amostra. Entretanto, nove dos dez intervalos contêm a proporção populacional π. Em média, somente aproximadamente uma em 20 vezes o intervalo de 95% de confiança não contém o parâmetro da população.

Na prática, selecionamos somente *uma* amostra de tamanho n e construímos *um* intervalo de confiança usando as observações dessa amostra. Não sabemos se esse intervalo de confiança verdadeiramente contém π. Nossa confiança no intervalo é baseada nas propriedades de longo prazo do procedimento. Podemos controlar, pela nossa escolha do nível de confiança, a chance de que o intervalo contenha π. Se uma probabilidade de erro de 0,05 nos deixa nervosos, podemos, então, formar um intervalo de 99% de confiança, para o qual o método erra apenas em 1% das vezes.

Tamanho da amostra necessário para a validade do método

Na prática, a probabilidade de que o intervalo de confiança contenha π é *aproximadamente* igual ao nível de confiança escolhido. A aproximação é melhor para amostras grandes.

* N. de T. T.: Por "longo prazo" entenda-se "após um grande número de repetições".



☑ **Figura 5.3** Dez Intervalos de 95% de confiança para a proporção populacional π. A longo prazo, somente 5% dos intervalos não conterão π.

À medida que n aumenta, a distribuição amostral de $\hat{\pi}$ se aproxima cada vez da normal, pelo Teorema Central do Limite. Isto é o que nos permite usar os escores-*z* da distribuição normal para encontrar a margem de erro. Também, à medida que n aumenta, o erro padrão *estimado* $ep = \sqrt{\hat{\pi}(1-\hat{\pi})/n}$ fica mais próximo do erro padrão *verdadeiro* $\sigma_{\hat{\pi}} = \sqrt{\pi(1-\pi)/n}$.

Por esta razão, a fórmula do intervalo de confiança se aplica a amostras aleatórias *grandes*. Quão grande é "grande"? Uma norma geral declara que você deve ter pelo menos 15 observações tanto na categoria de interesse quanto fora dela. Isso é verdadeiro para a maioria dos estudos em Ciências Sociais. No Exemplo 5.1, as freqüências nas duas categorias eram de 396 e 804 respectivamente, portanto os requisitos do tamanho da amostra foram facilmente satisfeitos. A Seção 5.4 e o Exercício 5.77 mostram métodos que funcionam bem quando a norma não for satisfeita.

Aqui está um resumo do intervalo de confiança para uma proporção:

☑ **Intervalo de confiança para a proporção populacional π**

Para uma amostra aleatória, um intervalo de confiança para a proporção populacional π baseada em uma proporção amostral $\hat{\pi}$ é

$$\hat{\pi} \pm z(ep), \text{ que é } \hat{\pi} \pm z\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}.$$

O valor-*z* é tal que a probabilidade sob uma curva normal dentro de *z* erros padrão da média é igual ao nível de confiança. Para intervalos de 95% e 99% de confiança, *z* é igual a 1,96 e 2,58. O tamanho da amostra *n* deve ser suficientemente grande para que pelo menos 15 observações estejam na categoria sendo estimada e pelo menos 15 nas demais.

5.3 INTERVALO DE CONFIANÇA PARA UMA MÉDIA

Aprendemos como construir um intervalo de confiança para uma proporção populacional para dados categóricos. Agora aprenderemos como construir um intervalo de confiança para a média populacional para dados quantitativos.

Erro padrão estimado para a margem de erro

Como o intervalo de confiança para uma proporção, o intervalo de confiança para uma média tem a forma

Estimativa por ponto ± Margem de erro,

onde a margem de erro é um múltiplo do erro padrão. A estimativa por ponto da média populacional μ é a média amostral $\bar{y}$. Para amostras aleatórias grandes, pelo Teorema Central do Limite, a distribuição amostral de $\bar{y}$ é aproximadamente normal. Assim, para amostras grandes podemos, novamente, encontrar uma margem de erro multiplicando um escore-*z* da distribuição normal pelo erro padrão.

Da Seção 4.5, o erro padrão da média amostral é igual a

$$\sigma_{\bar{y}} = \frac{\sigma}{\sqrt{n}},$$

onde σ é o desvio padrão da população. Como no caso do erro padrão da proporção amostral, ele depende de um parâmetro desconhecido, neste caso o σ. Na prática, estimamos σ pelo desvio padrão da amostra s. Assim, os intervalos de confiança usam o erro padrão *estimado*:

$$ep = s/\sqrt{n}.$$

EXEMPLO 5.4 Estimando o número médio de parceiros sexuais

Algumas Pesquisas Sociais Gerais (PSGs) perguntaram aos respondentes quantos parceiros sexuais eles tiveram desde os seus 18 anos. Em 2006, quando perguntadas sobre quantos parceiros sexuais masculinos tiveram, as 231 mulheres da amostra com idades entre 20 e 29 relataram uma média de 4,96. Uma saída de computador resume os resultados dessa variável da PSG (representada por "NUMMEN"):

Variável	n	Média	Desv. Padrão	Média EP	IC de 95%
NUMMEN	231	4,96	6,81	0,45	(4,1; 5,8)

Como um *software* relata o erro padrão? Como ele é interpretado e, também, o intervalo de confiança mostrado?

O desvio padrão da amostra é $s = 6{,}81$. O tamanho da amostra é $n = 231$. Assim, o erro padrão estimado da média amostral é

$$ep = s/\sqrt{n} = 6{,}81/\sqrt{231} = 0{,}45.$$

Em várias amostras aleatórias de 231 mulheres neste grupo etário, o número médio de parceiros sexuais masculinos varia de amostra para amostra com um desvio padrão de aproximadamente 0,45.

O intervalo de 95% de confiança (4,1; 5,8) apresentado é um intervalo estimado de μ, o número médio de parceiros sexuais masculinos desde os 18 anos para a população de mulheres adultas dos Estados Unidos com idade entre 20 e 29 anos. Podemos estar 95% confiantes de que esse intervalo contem μ. A estimativa por ponto de μ é 5,0 e a estimativa intervalar prevê que μ não é menor do que 4,1 e nem maior do que 5,8.

Este exemplo destaca algumas coisas para serem lembradas quando fizermos análises estatísticas. Primeiro, a média amostral de 5,0 e o desvio padrão de 6,8 sugerem que a distribuição da variável "NUMMEN" é altamente assimétrica à direita. A média pode ser ilusória como uma medida do centro. Na verdade, uma análise de toda a distribuição de "NUMMEN" em 2006 (no *site* do PSG) revela que a resposta mediana foi de 3, talvez um representante mais útil. Vale a pena também observar que a moda foi 1, com 23% da amostra.

Segundo, a margem de erro em intervalos de confiança se refere apenas ao erro amostral. Outros erros potenciais incluem aqueles devido a não resposta ou erro de mensuração (mentindo ou dando uma resposta imprecisa). Se tais erros não são insignificantes, a margem de erro é, na verdade, maior do que a relatada pelo *software* usando as fórmulas estatísticas.

Finalmente, como foi mencionado nos Capítulos 2 e 4, a PSG usa um delineamento multiestágio que incorpora a amostragem por conglomerados.[2] Para esse delineamento, as estimativas não são tão precisas quanto uma amostragem aleatória simples iria fornecer. Por simplicidade, neste livro, estamos agindo como se a PSG fosse uma amostra aleatória simples. ■

Como um *software* encontrou a margem de erro para o intervalo de confiança no exemplo anterior? Assim como com a proporção, para um intervalo de 95% de confiança isto é aproximadamente duas vezes o erro padrão estimado. Encontraremos a seguir a margem de erro precisa multiplicando o *ep* pelo escore que é muito similar ao escore-*z*, exceto se *n* for bem pequeno.

A distribuição *t*

Iremos, agora, aprender sobre o intervalo de confiança que serve para *qualquer* tamanho de amostra aleatória. Para conseguir essa generalidade, ele tem a desvantagem de assumir que a distribuição populacional é normal. Neste caso, a distribuição amostral de $\bar{y}$ é normal mesmo para tamanhos de amostras pequenos. (O painel direito da Figura 4.14 na página 114, que mostrou as distribuições amostrais para várias distribuições populacionais, ilustrou esse fato.)

Suponha que conhecemos o erro padrão exato da média amostral, $\sigma_{\bar{y}} = \sigma/\sqrt{n}$. Então, com a suposição adicional de que a população é normal, com todo *n* podemos usar a fórmula:

$$\bar{y} \pm z\sigma_{\bar{y}}, \text{ que é } \bar{y} \pm z\sigma/\sqrt{n},$$

por exemplo, com $z = 1{,}96$ para uma confiança de 95%. Na prática, não conhecemos o desvio padrão da *população* σ, portanto não conhecemos o erro padrão *exato*. Substituir o desvio padrão da população σ pelo da *amostra* s obtendo o erro padrão *estimado*, $ep = s/\sqrt{n}$, introduz um erro adicional. Esse erro pode ser de tamanho considerável quando *n* é pequeno. Para lidar com esse erro, temos que substituir o escore-*z* por um escore um pouco maior, chamado de escore-*t*. O intervalo de confiança, então, é um pouco mais amplo. O escore-*t* é como o escore-*z*, mas ele vem de uma distribuição com a forma de sino que é um pouco mais dispersa do que uma distribuição normal. Essa distribuição é chamada de **distribuição *t***.

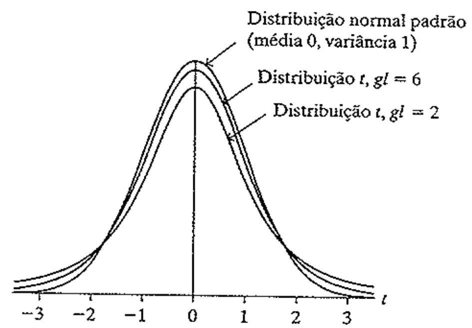
Propriedades da distribuição *t*

Aqui estão as propriedades principais da distribuição *t*:

- A distribuição *t* tem a forma de sino e é simétrica em relação à média de 0.
- O desvio padrão é um pouco maior do que 1. O valor preciso depende do que é chamado de **graus de liberdade**, representado por *gl*. A distribuição *t* tem uma dispersão levemente diferente para cada valor distinto do *gl*, e escores-*t* diferentes se aplicam para cada valor do *gl*.
- Para inferências da média populacional, os graus de liberdade são iguais a $gl = n - 1$, um a menos do que o tamanho da amostra utilizada.
- A distribuição *t* tem caudas mais grossas e é mais dispersa do que a distribuição normal padrão. A Figura 5.4 ilustra o fato. Quando *gl* é aproximadamente 30 ou mais, as duas distribuições são quase idênticas.
- Um escore-*t* multiplicado pelo erro padrão estimado dá a margem de erro para um intervalo de confiança para a média.

A Tabela B, no final do livro, lista os escores-*t* para várias probabilidades da cauda direita. A Tabela 5.1 é uma parte daquela tabela. A coluna rotulada $t_{0,025}$ tem uma probabilidade de 0,025 na cauda direita e uma probabilidade nas duas caudas de 0,05. Este é o escore-*t* que é usado em intervalos de 95% de confiança.

Para ilustrar, quando o tamanho da amostra é 29, os graus de liberdade são $gl = n - 1 = 28$. Com $gl = 28$, veja que $t_{0,025} = 2{,}048$. Isto significa que 2,5% da distribuição *t* estão na cauda direita acima de 2,048. Por simetria, 2,5% estão na cauda esquerda abaixo de $-t_{0,025} = -2{,}048$. Veja a Figura 5.5. Quando $gl = 28$, a probabilidade é igual a 0,95 entre $-2{,}048$ e $2{,}048$. Estes são os escores-*t* para um intervalo de 95% de confiança quando $n = 29$. O intervalo de confiança é $\bar{y} \pm 2{,}048(ep)$.



☑ **Figura 5.4** A distribuição *t* comparada à distribuição normal padrão. A *t* fica mais próxima à normal à medida que os graus de liberdade (*gl*) aumentam, e as duas distribuições são praticamente idênticas quando $gl > 30$.

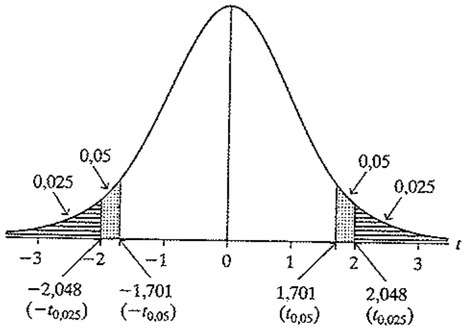
☑ **Tabela 5.1** Parte da Tabela B mostrando os escores-*t*. Os escores têm probabilidades na cauda direita de 0,100; 0,050; 0,025; 0,010; 0,005 e 0,001

	Nível de confiança					
	80%	90%	95%	98%	99%	99,8%
gl	$t_{0,100}$	$t_{0,050}$	$t_{0,025}$	$t_{0,010}$	$t_{0,005}$	$t_{0,001}$
1	3,078	6,314	12,706	31,821	63,657	318,3
10	1,372	1,812	2,228	2,764	3,169	4,144
28	1,313	1,701	2,048	2,467	2,763	3,408
30	1,310	1,697	2,042	2,457	2,750	3,385
100	1,290	1,660	1,984	2,364	2,626	3,174
Infinito	1,282	1,645	1,960	2,326	2,576	3,090

Escores-*t* no intervalo de confiança para uma média

Os intervalos de confiança para uma média se parecem com aqueles para as proporções, exceto que eles usam a distribuição *t* em vez da normal padrão.

> ☑ **Intervalo de confiança para a média populacional μ**
>
> Para uma amostra aleatória retirada de uma distribuição populacional normal, um intervalo de 95% de confiança para μ, quando o desvio padrão da população é desconhecido é
>
> $$\bar{y} \pm t_{0,025}(ep), \text{ com } ep = s/\sqrt{n},$$
>
> onde $gl = n - 1$ para o escore-*t*.



☑ **Figura 5.5** A distribuição t com gl = 28.

Como o intervalo de confiança para uma proporção, esse intervalo de confiança tem uma margem de erro que é um valor multiplicado pelo erro padrão estimado. A diferença principal é a substituição do escore-t pelo escore-z. O método t também faz a suposição de uma distribuição populacional normal. Isto é muito relevante para amostras pequenas. Na prática, a distribuição populacional pode não estar próxima da normal e discutiremos a importância desta suposição mais tarde nesta seção.

EXEMPLO 5.5 Estimando a mudança média de peso para meninas anoréxicas

Este exemplo vem de um estudo experimental que comparou vários tratamentos para jovens meninas que sofriam de anorexia, um distúrbio alimentar. Para cada jovem, o peso era mensurado antes e depois de um período fixo de tratamento. A variável de interesse era a mudança de peso, isto é, o peso ao final do estudo menos o peso no início do estudo. A mudança no peso era positiva se a jovem ganhasse peso e negativa se ela perdesse peso. Os tratamentos foram delineados para ajudar no ganho de peso. As mudanças do peso para as 29 meninas submetidas ao tratamento cognitivo-comportamental foram[3]

1,7; 0,7; −0,1; −0,7; −3,5; 14,9; 3,5; 17,1; −7,6; 1,6;
11,7; 6,1; 1,1; −4,0; 20,9; −9,1; 2,1; 1,4; −0,3; −3,7;
−1,4; −0,8; 2,4; 12,6; 1,9; 3,9; 0,1; 15,4; −0,7.

O *software* usado para analisar os dados apresenta um resumo dos resultados:

```
-----------------------------------------------
           Número                        EP da
Variável   de Casos   Média    DP        média
MUDANÇA    29          3,01    7,31      1,36
-----------------------------------------------
```

Portanto, para as n = 29 meninas que receberam o tratamento, a mudança do peso médio foi $\bar{y}$ = 3,01 libras com um desvio padrão (DP) de s = 7,31. A média amostral tinha um erro padrão estimado $ep = s/\sqrt{n} = 7{,}31/\sqrt{29} = 1{,}36$ (relatado como EP da Média).

Considere μ a representação da mudança média populacional no peso para o tratamento cognitivo-comportamental, para a população representada por esta amostra. Se este tratamento tem um efeito benéfico, então μ é positivo. Visto que n = 29, $gl = n - 1 = 28$. Para um intervalo de 95% de confiança usamos $t_{0,025} = 2{,}048$. O intervalo de 95% de confiança é, então:

$$\bar{y} \pm t_{0,025}\,(ep) = 3{,}01 \pm 2{,}048(1{,}36) = 3{,}0 \pm 2{,}8 \text{ ou } (0{,}2;\ 5{,}8).$$

Com 95% de confiança, inferimos que este intervalo contém a mudança média do peso da população. Torna-se visível que a mudança média do peso é positiva, mas pode ser bem pequena em termos práticos. Entretanto, este estudo experimental usou uma amostra voluntária porque não é possível identificar e amostrar aleatoriamente uma população de meninas anoréxicas. Por causa disto, as inferências são tentativas, e "95% de confiança" nos resultados podem ser excessivamente otimistas. Os resultados são mais convincentes se os pesquisadores podem argumentar que a amostra era representativa da população. O estudo empregou a aleatorização designando as meninas a três terapias (somente uma é considerada aqui), o que traz mais confiança para as análises comparando as terapias conduzidas mais tarde no livro.

Outra advertência sobre nossa conclusão é apresentada na Figura 5.6 no histograma que o *software* mostra para os dados. Ele revela que os dados são assimétricos à direita. A suposição sobre a normalidade da população pode ser violada – veremos mais sobre isso depois. O peso mediano mudou apenas 1,4 libras, algo menos do que a média de 3,0 por causa da assimetria à direita. A mediana amostral é outra indicação de que o tamanho do efeito pode ser menor. ■

Efeito do nível de confiança e do tamanho da amostra

Usamos a distribuição t para encontrar o intervalo de 95% de confiança. Outros níveis de confiança usam a mesma fórmula com um escore-t diferente.

Considere α representando a probabilidade de erro. Isto é a probabilidade de que o método gere um intervalo de confiança que não contenha μ. O intervalo de confiança usa o escore-t com uma probabilidade em cada cauda de $\alpha/2$. Para um intervalo de 99% de confiança, por exemplo, $\alpha = 1 - 0{,}99 = 0{,}01$, assim $\alpha/2 = 0{,}005$ e o escore-t apropriado é $t_{0,005}$ para o valor do gl especificado.

Para ficarmos mais seguros na estimação da mudança média do peso da população para o estudo da anorexia do Exemplo 5.5, poderíamos usar um intervalo de 99% de confiança. Visto que gl = 28 quando n = 29, o escore-t é $t_{0,005} = 2{,}763$. O erro padrão não muda. O intervalo de 99% de confiança é, então:

$$\bar{y} \pm 2{,}763(ep) = 3{,}01 \pm 2{,}763(1{,}36),$$
$$\text{que é } (-0{,}7;\ 6{,}8)$$

O intervalo de confiança é mais amplo do que o intervalo de 95% (0,2; 5,8). Esse é o custo de se ter mais confiança. O intervalo de 99% de confiança contém o 0. Isso nos diz que é plausível, a um nível de 99%

```
Frequência
   +                   *****
10 +                   *****
   +          *****    *****
 8 +          *****    *****
   +          *****    *****
 6 +          *****    *****
   +          *****    *****
 4 +          *****    *****              *****
   +          *****    *****              *****
 2 + *****    *****    *****    *****     *****
   + *****    *****    *****    *****     *****    *****
---------------------------------------------------------- Mudança_peso
   (-12,-6)  (-6,0)    (0,6)    (6,12)   (12,18)  (18,24)
```

☑ **Figura 5.6** Histograma dos valores da mudança do peso para o estudo de anorexia.

de confiança, que a mudança na média da população seja 0, isto é, que a terapia pode não resultar em alterações do peso médio.

Como a amplitude do intervalo de confiança para uma proporção, a amplitude de um intervalo de confiança para uma média também depende do tamanho da amostra n. Tamanhos da amostra maiores resultam em intervalos menores.

Robustez para violações da suposição de normalidade

As suposições para a construção de um intervalo de confiança para uma média são (1) que a aleatorização é usada para a coleta da amostra e (2) que a distribuição da população é normal. Sob a suposição da normalidade, a distribuição amostral de $\bar{y}$ é normal mesmo para pequenas amostras. Da mesma forma, o escore-z que mensura os erros padrão que $\bar{y}$ está de μ tem, então, uma distribuição normal padrão. Na prática, quando usamos o erro padrão *estimado* $ep = s/\sqrt{n}$ (em vez do verdadeiro, $\sigma/\sqrt{n}$), o número de ep que $\bar{y}$ está de μ tem uma distribuição t.

O intervalo de confiança no Exemplo 5.5 estimou a mudança média do peso para as meninas anoréxicas. O histograma dos dados da mudança de peso exibido acima não é preciso quando n é tão pequeno quanto naquele exemplo ($n = 29$), mas ele mostrou evidência de assimetria. Em geral, a suposição de população normal parece inadequada porque muitas variáveis nas Ciências Sociais têm distribuições que estão longe de serem normais.

Um método estatístico é dito ser **robusto** com respeito a uma suposição em particular se ele tem um desempenho adequado mesmo quando aquela suposição for violada. Os estatísticos têm mostrado que o intervalo de confiança para uma média usando a distribuição t é robusto contra violações da suposição de normalidade. Mesmo se a população não for normal, os intervalos de confiança baseados na distribuição t ainda funcionam bem, especialmente quando n excede aproximadamente 15. À medida que o tamanho da amostra fica maior, a suposição de normalidade se torna menos importante por causa do Teorema Central do Limite. A distribuição amostral da média amostral tem, então, a forma de sino mesmo quando a distribuição populacional não tem. A probabilidade real de que um intervalo de 95% de confiança forneça a inferência correta está próxima a 0,95 e fica mais próxima à medida que n aumenta.

Uma situação importante onde o método não funciona bem é quando os dados são extremamente assimétricos ou contêm valores atípicos extremos. Parte disso é devido ao efeito no próprio método, mas outra parte é porque a média, nesse caso, pode não ser um representante adequado do centro dos dados. Por essa razão, o intervalo de confiança sobre o número de parceiros sexuais relatado no Exemplo 5.4 (página 141) tem uma utilidade limitada.

Na prática, é raro que as suposições sejam totalmente satisfeitas. Portanto, saber se um método estatístico é robusto em relação à violação de determinada suposição é importante. O método dos intervalos de confiança t *não* é robusto a violações na aleatorização. O modelo t, como todos os métodos de estatística inferencial, tem validade questionável se o método para a produção de dados não utiliza a aleatorização.

A normal padrão é a distribuição t com infinitos graus de liberdade

Veja a tabela dos escores-t (Tabela B), parte dela foi exibida na Tabela 5.1. À medida que o gl aumenta, você move a tabela para baixo. O escore-t diminui e fica cada vez mais próximo do escore-z para a distribuição normal padrão. Isso significa que a distribuição t se torna menos dispersa e mais similar na aparência à distribuição normal padrão à medida que o gl aumenta. Você pode pensar na distribuição normal padrão como uma distribuição t com infinitos graus de liberdade.

Por exemplo, quando o gl aumenta de 1 para 100 na Tabela 5.1, o escore-t $t_{0,025}$ com a probabilidade unicaudal à direita igual a 0,025 diminui de 12,706 para 1,984. O escore-z com a mesma probabilidade na cauda direita para a distribuição normal padrão é $z = 1{,}96$. Os escores-t não estão impressos para $gl > 100$, mas eles estão próximos aos escores-z. A última linha da Tabela 5.1 e da Tabela B (*gl infinito*) lista os escores-z para vários níveis de confiança.

Você pode obter os escores-t para qualquer valor do gl usando um *software* e muitas calculadoras, assim você não está restrito ao uso da Tabela B. Para valores do gl maiores do que os exibidos na Tabela B (acima de 100), você pode usar um escore-z para aproximar o escore-t. Para um intervalo de 95% de confiança você irá, então, usar

$$\bar{y} \pm 1{,}96(ep) \text{ em vez de } \bar{y} \pm t_{0,025}(ep).$$

Você não irá obter *exatamente* o mesmo resultado que um *software* iria dar, mas ele será próximo o suficiente para finalidades práticas. Por exemplo, para obter o intervalo de confiança para o número médio de parceiros sexuais no Exemplo 5.4, para o qual a amostra da PSG tinha $n = 231$, o *software* usa o escore-t para $gl = 231 - 1 = 230$, que é 1,97. Isto está muito próximo ao escore-z de 1,96 da distribuição normal padrão.

Por que a distribuição t se parece mais com a distribuição normal padrão à medida que n (e, portanto, o gl) aumenta? Porque s é cada vez mais preciso como uma estimativa por ponto de σ na aproximação do erro padrão verdadeiro $\sigma/\sqrt{n}$ por $ep = s/\sqrt{n}$, à medida que n aumenta. O erro amostral adicional para amostras pequenas resulta na distribuição t sendo mais dispersa do que a normal padrão.

A distribuição t recentemente celebrou seu 100º aniversário. Ela foi descoberta em 1908 pelo estatístico e químico inglês W. S. Gosset. Naquela época, Gosset era empregado da Cervejaria Guinness em Dublin, Irlanda, projetando experimentos relacionados à seleção, cultivo e tratamento da cevada e lúpulo para o processo de fermentação da cerveja. Devido à política da empresa que proibia a publicação de segredos comerciais, Gosset usou o pseudônimo *Student* em artigos que ele escreveu sobre a sua descoberta. A distribuição t ficou conhecida como o t de *Student*, o nome que ainda é hoje é usado algumas vezes. Os intervalos de confiança, entretanto, só foram introduzidos por uma série de artigos de Jerzy Neyman e Egon Pearson em 1928.

Uma advertência sobre o uso de *software*

Os exemplos nesta seção usaram resultados de *softwares* estatísticos para nos ajudar a analisar os dados. Faremos isto de modo crescente nos capítulos futuros à medida que cobrirmos métodos que requeiram muita computação. Você deve usar o *software* em alguns exercícios para ter uma ideia de como os pesquisadores analisam, na prática, os dados. Qualquer *software* tem um grande número de opções e é muito fácil de usá-lo de maneira imprópria. Só porque os resultados aparecem em uma janela de saída não significa que são os corretos ou que as suposições foram suficientemente satisfeitas para fazer aquela análise.

Quando você começa a usar um *software* para um dado método, sugerimos que você primeiro o use para o exemplo daquele método neste livro. Observe se você obtém os mesmos resultados, como uma forma de verificar se você está usando corretamente o *software*.

5.4 ESCOLHA DO TAMANHO DA AMOSTRA

As organizações de pesquisa de opinião pública, como a Gallup, coletam amostras que normalmente contêm aproximadamente mil sujeitos. Isto é grande o suficiente para uma estimativa da proporção amostral ter uma margem de erro de aproximadamente 0,03. À primeira vista, parece surpreendente que uma amostra desse tamanho de uma população de talvez muitos milhões seja adequada para prever resultados de eleições, resumos de opiniões sobre assuntos controversos, mostrar tamanhos relativos de audiência de televisão e assim por diante.

Lembre que a margem de erro para um intervalo de confiança depende do *erro padrão* da estimativa por ponto. Portanto, a base para este poder inferencial está nas fórmulas para os erros padrão. Contanto que a amostragem seja executada propriamente, estimativas boas irão resultar de amostras relativamente pequenas, não importando o tamanho da população (na verdade, os métodos realmente tratam o tamanho da população como infinito; veja o Exercício 4.57). As organizações de pesquisa de opinião pública usam métodos amostrais que são mais complexos do que amostras aleatórias simples, geralmente envolvendo aglomeração e/ou estratificação. Entretanto, os erros padrão sob seus planos amostrais são aproximadamente bem razoáveis também para fórmulas utilizando amostras aleatórias simples ou para inflar tais fórmulas por determinado fator (como 25%) para refletir o efeito do plano amostral.

Antes de iniciar a coleta dos dados, a maioria dos estudos tenta determinar o tamanho da amostra que irá fornecer certo grau de precisão para a estimativa. Uma medida relevante é o valor de n para o qual um intervalo de confiança para o parâmetro tem uma margem de erro igual a um valor especificado. Esta seção mostra como eles fazem isto. A solução resulta em encontrar o tamanho da amostra da seguinte forma:

- A *margem de erro* depende diretamente do *erro padrão* para a distribuição amostral da estimativa por ponto.
- O *erro padrão* depende, ele próprio, do *tamanho da amostra*.

O tamanho da amostra para estimar proporções

Para determinar o tamanho da amostra, devemos decidir sobre a margem de erro desejada. Em alguns estudos, uma estimativa altamente precisa não é tão importante como em outros. Uma pesquisa de boca de urna em uma eleição próxima requer uma estimativa precisa para prever o vencedor. Se, por outro lado, o objetivo for estimar a proporção de residentes de Siracusa, Nova Iorque, que têm seguro de saúde, uma margem de erro maior pode ser aceitável. Assim, devemos decidir primeiro se a margem de erro deve ser aproximadamente 0,03 (três pontos percentuais), 0,05 ou outra qualquer.

Devemos especificar, também, a *probabilidade* com a qual a margem de erro é alcançada. Por exemplo, podemos decidir que a margem de erro, quando estimamos uma proporção da população, não deva exceder a um erro de 0,04, com 0,95 de probabilidade. Essa probabilidade deve ser declarada, visto que com qualquer tamanho da amostra o erro não é maior do que 0,04 com *alguma* probabilidade, talvez uma pequena probabilidade.

O próximo exemplo ilustra a determinação do tamanho da amostra para estimar a proporção populacional.

EXEMPLO 5.6 Tamanho da amostra para um levantamento de dados de pais solteiros

Uma cientista social queria estimar a proporção de crianças em idade escolar de Boston que vivem com somente um dos pais. Como seu relatório seria publicado, ela queria uma estimativa razoavelmente precisa. Entretanto, já que seu financiamento era

limitado, a cientista social não queria coletar uma amostra maior do que a necessária. Ela decidiu usar um tamanho de amostra com uma probabilidade de 0,95 e um erro não superior a 0,04. Portanto, ela precisava determinar n tal que um intervalo de 95% de confiança para π fosse igual a $\hat{\pi} \pm 0{,}04$.

Visto que a distribuição da proporção amostral $\hat{\pi}$ é aproximadamente normal, $\hat{\pi}$ está dentro de 1,96 erros padrão de π com probabilidade de 0,95. Portanto, se o tamanho da amostra é tal que 1,96 erros padrão são iguais a 0,04, então com a probabilidade de 0,95, $\hat{\pi}$ está dentro de 0,04 unidades de π. Veja a Figura 5.7.

Lembre que o erro padrão verdadeiro é $\sigma_{\hat{\pi}} = \sqrt{\pi(1 - \pi)/n}$. Como encontramos o valor de n que fornece um valor de $\sigma_{\hat{\pi}}$ para o qual $0{,}04 = 1{,}96\sigma_{\hat{\pi}}$? Devemos resolver para n a expressão:

$$0{,}04 = 1{,}96\sqrt{\frac{\pi(1 - \pi)}{n}}.$$

Multiplicando ambos os lados da expressão por $\sqrt{n}$ e dividindo ambos os lados por 0,04, temos:

$$\sqrt{n} = \frac{1{,}96\sqrt{\pi(1 - \pi)}}{0{,}04}.$$

Elevando ao quadrado ambos os lados, obtemos o resultado:

$$n = \frac{(1{,}96)^2\pi(1 - \pi)}{(0{,}04)^2}.$$

Agora, nos deparamos com um problema. Queremos determinar n para o propósito de estimar a proporção da população π, mas a fórmula requer o conhecimento do valor de π. Isso ocorre porque a dispersão da distribuição amostral depende de π. A distribuição é menos dispersa e é mais fácil de estimar π, se π está próximo a 0 ou 1 do que se estiver próximo a 0,50. Visto que π é desconhecido, devemos substituí-lo por um valor estimado (suposto) nessa equação e então determinar n.

O maior valor possível para $\pi(1 - \pi)$ ocorre quando $\pi = 0{,}50$. Nesse caso, $\pi(1 - \pi) = 0{,}25$. Na verdade, $\pi(1 - \pi)$ é muito próximo a 0,25 a não ser que π esteja bem longe de 0,50. Por exemplo, $\pi(1 - \pi) = 0{,}24$ quando $\pi = 0{,}40$ ou $\pi = 0{,}60$ e $\pi(1 - \pi) = 0{,}21$ quando $\pi = 0{,}70$ ou $\pi = 0{,}30$. Portanto, outra abordagem possível é substituir $\pi = 0{,}50$ na equação acima. Isso gera:

$$n = \frac{(1{,}96)^2\pi(1 - \pi)}{(0{,}04)^2} = \frac{(1{,}96)^2(0{,}50)(0{,}50)}{(0{,}04)^2} = 600.$$

Essa abordagem assegura que com um intervalo de 0,95 de confiança a margem

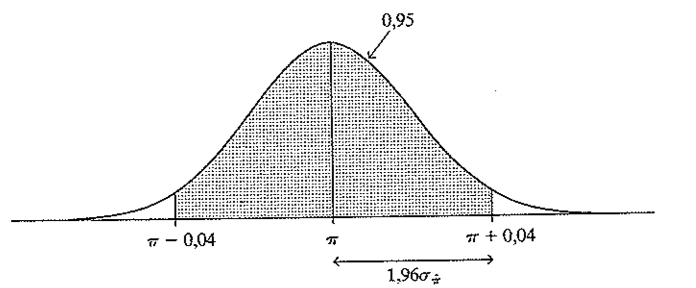


Figura 5.7 Distribuição amostral de $\hat{\pi}$ com erro de estimação não maior do que 0,04, com uma probabilidade de 0,95.

de erro não irá exceder a 0,04, não importando qual o valor de π. ■

Obtendo n estabelecendo $\pi = 0{,}50$ é a abordagem "segura". Mas este valor de n é excessivamente grande se π não estiver próximo a 0,50. Suponha que, baseada em outros estudos, a cientista social acredita que π não é maior do que 0,25. Então, o tamanho da amostra adequado será:

$$n = \frac{(1{,}96)^2\pi(1-\pi)}{(0{,}04)^2}$$

$$= \frac{(1{,}96)^2(0{,}25)(0{,}75)}{(0{,}04)^2} = 450.$$

Um tamanho da amostra de 600 é maior do que o necessário. Com ele, a margem de erro para um intervalo de 95% de confiança seria menor do que 0,04.

Fórmula do tamanho da amostra para estimar proporções

Forneceremos, a seguir, uma fórmula geral para determinar o tamanho da amostra. Considere M a representação da margem de erro desejada. A fórmula também usa um escore-z geral (no lugar de 1,96) determinado pela probabilidade com a qual o erro não é maior do que M.

> ☑ **Tamanho da amostra para estimar a proporção da população π**
>
> O tamanho da amostra aleatória n tendo a margem de erro M necessária para estimar π a proporção populacional é:
>
> $$n = \pi(1-\pi)\left(\frac{z}{M}\right)^2.$$
>
> O escore-z é o escore para um intervalo de confiança com o nível de confiança desejado, tal que $z = 1{,}96$ para o nível 0,95. Você deve supor um valor para π ou empregar a abordagem segura utilizando $\pi = 0{,}50$.

Para ilustrar, suponha que o estudo sobre as crianças de pais solteiros queria estimar a proporção populacional dentro de 0,08 com uma probabilidade de pelo menos 0,95. Então, a margem de erro é igual a $M = 0{,}08$ e $z = 1{,}96$. O tamanho da amostra requerido usando a abordagem segura é:

$$n = \pi(1-\pi)\left(\frac{z}{M}\right)^2$$

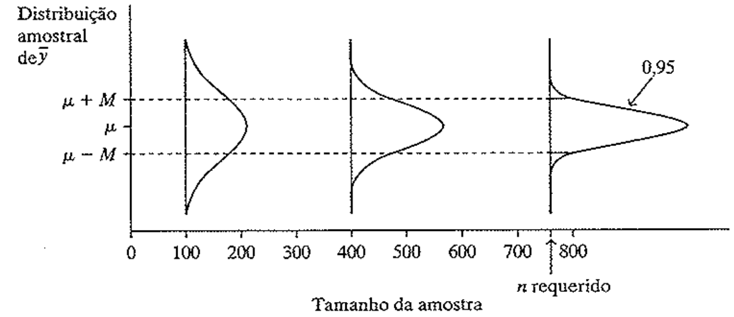
$$= (0{,}50)(0{,}50)\left(\frac{1{,}96}{0{,}08}\right)^2 = 150.$$

Esse tamanho da amostra de 150 é um quarto do tamanho da amostra de 600 necessário para garantir uma margem de erro não maior do que $M = 0{,}04$. Reduzir a margem de erro pela metade (1/2) requer que o tamanho da amostra seja quadruplicado.

Tamanho da amostra para estimar médias

Um resultado análogo é válido para estimar a média populacional μ. Queremos determinar quão grande deve ser n para que a distribuição amostral de $\bar{y}$ tenha margem de erro M. A Figura 5.8 ilustra a situação. Ela mostra como a distribuição amostral fica mais estreita à medida que n aumenta até que, para o n desejado, 95% está dentro da margem de erro especificada. Uma derivação usando uma amostra maior da distribuição amostral normal de $\bar{y}$ gera o seguinte resultado:

> ☑ **Tamanho da amostra para estimar a média populacional μ**
>
> O tamanho da amostra aleatória n tendo a margem de erro M necessária para estimar μ, utilizando a média amostral $\bar{y}$ é:
>
> $$n = \sigma^2\left(\frac{z}{M}\right)^2.$$
>
> O escore-z é o utilizado no intervalo de confiança com o nível de confiança desejado, tal que $z = 1{,}96$ para o nível 0,95. Você precisa estimar o desvio padrão da população σ.



☑ **Figura 5.8** Determinando n tal que $\bar{y}$ tenha a probabilidade de 0,95 de estar dentro da margem de erro de M unidades da média populacional μ.

Quanto maior a dispersão da distribuição populacional, como foi medida pelo seu desvio padrão σ, maior o tamanho da amostra necessário para estar dentro de certa margem de erro. Se os sujeitos tiverem pouca variação (isto é, se σ for pequeno), precisamos de menos dados do que se eles fossem altamente heterogêneos. Na prática, σ é desconhecido. Precisamos substituí-lo por uma estimativa, talvez baseada em um estudo anterior.

Uma pequena complicação surge visto que não conhecemos σ, para a inferência usamos, na verdade, a distribuição t em vez da normal padrão. Mas, se não conhecemos n, também não conhecemos os graus de liberdade e o escore-t. Vimos na Tabela B, entretanto, que a não ser que o gl seja pequeno, o escore-t está próximo do escore-z. Assim, nós não iremos nos preocupar com esta complicação. A aproximação de substituir um escore-t desconhecido na fórmula do tamanho da amostra por um escore-z é geralmente bem menor do que aquela envolvida em ter de usar uma estimativa para σ.

EXEMPLO 5.7 Estimando o grau educacional médio dos nativos norte-americanos

Um estudo é planejado para estudar nativos norte-americanos idosos. As variáveis a serem estudadas incluem o nível educacional. Que tamanho da amostra é necessário para estimar o número médio de anos de escolaridade, com um erro de 1 ano e uma probabilidade de 0,99?

Suponha que o estudo não tenha informação prévia sobre o desvio padrão do nível educacional dos nativos norte-americanos. Como uma estimativa, talvez todos os valores dessa variável estejam dentro de um intervalo de aproximadamente 15 anos, como entre 5 e 20 anos. Se essa distribuição é aproximadamente normal, então, visto que o intervalo entre $\mu - 3\sigma$ para $\mu + 3\sigma$ contém aproximadamente toda a distribuição normal, o intervalo de 15 seria igual a 6σ. Então, $15/6 = 2{,}5$ seria uma estimativa para σ.

Agora, para 99% de confiança, a probabilidade de erro é 0,01. O escore-z é o escore com um probabilidade unicaudal de $0{,}01/2 = 0{,}005$, e ele é igual a 2,58. Visto que a margem de erro desejada é $M = 1$ ano, o tamanho da amostra necessário é:

$$n = \sigma^2\left(\frac{z}{M}\right)^2$$

$$= (2{,}5)^2\left(\frac{2{,}58}{1}\right)^2 = 42.$$

Uma abordagem mais cautelosa utilizaria um valor maior para σ. Por exemplo, se o intervalo de 5 a 20 inclui somente aproximadamente 95% dos valores educacionais, poderíamos tratar isto como um intervalo de $\mu - 2\sigma$ a $\mu + 2\sigma$ e estabelecer $15 = 4\sigma$. Então, $\sigma = 15/4 = 3{,}75$ e $n = (3{,}75)^2(2{,}58/1)^2 = 94$. Se σ é realmente menor do que 3,75, a margem de erro de um intervalo de 99% de confiança com $n = 94$ observações será ainda menor do que 1. ■

Essas fórmulas do tamanho da amostra se aplicam à amostragem aleatória simples e sistemática. Amostras por conglomerados e amostras multiestágio complexas devem geralmente ser maiores para alcançar a mesma precisão, enquanto que as amostras estratificadas podem ser frequentemente menores. Em tais casos, você deve procurar orientação de um consultor estatístico.

Outras considerações na determinação do tamanho da amostra

Vimos que o tamanho da amostra necessário depende da *precisão* e da *confiança* desejadas. A precisão se refere à margem de erro. A confiança se refere à probabilidade de que o intervalo de confiança irá conter o parâmetro. Também vimos que o tamanho da amostra depende da *variabilidade* da população. Para estimar médias, o tamanho da amostra necessário aumenta à medida que σ aumenta. Na maioria das pesquisas sociais, amostras grandes (1000 ou mais) são necessárias, mas para populações homogêneas (por exemplo, residentes de asilos), amostras menores são geralmente adequadas devido à reduzida variabilidade da população.

De um ponto de vista prático, outras considerações também afetam o tamanho da amostra. Uma questão é a *complexidade da análise* planejada. Quanto mais complexa for a análise, como por exemplo, mais variáveis analisadas simultaneamente, maior o tamanho da amostra necessária. Para analisar uma única variável usando uma média, uma amostra relativamente pequena pode ser adequada. As comparações planejadas de vários grupos usando métodos multivariados complexos, entretanto, requerem uma amostra maior. Como ilustração, o Exemplo 5.7 mostrou que podemos ser capazes de estimar a média do grau educacional obtido muito bem com somente 42 pessoas. Mas, se também quisermos comparar a média para vários grupos étnicos e raciais e estudar como a média depende de outras variáveis como gênero, renda e educação dos pais e tamanho da comunidade, uma amostra maior seria necessária.

Outra consideração tem relação com o tempo, dinheiro e *recursos*. As amostras maiores são mais caras e consomem mais tempo, pois podem requerer mais recursos do que os disponíveis. Por exemplo, as fórmulas do tamanho da amostra podem sugerir que 1000 casos forneçam a precisão desejada. Talvez você tenha recursos para reunir somente 400. Você deveria continuar com a amostra menor e sacrificar a precisão ou confiança ou você deveria desistir a menos que obtenha recursos adicionais? Você pode ter que responder perguntas como: "É realmente crucial estudar todos os grupos ou eu posso reduzir a amostra focando em dois grupos apenas?".

Em resumo, nenhuma fórmula simples poderá sempre dar um tamanho da amostra apropriado. Apesar de o tamanho da amostra ser um assunto importante, sua escolha depende de recursos e da análise planejada, o que requer um julgamento cuidadoso. Uma advertência final: se o estudo for conduzido deficientemente, se os dados nunca são obtidos para um percentual substancial da amostra, se alguns sujeitos mentem ou se algumas informações são registradas incorretamente pelo coletor dos dados ou pelo analista estatístico, então a real probabilidade e a precisão da margem de erro especificadas poderão ser bem menores do que o pretendido. Quando alguém alega ter alcançado certa precisão e confiança, seja cético, a não ser que o estudo esteja substancialmente livre de tais problemas.

E se você tiver somente uma amostra pequena?*

Algumas vezes, por causa de fatores financeiros e éticos, não é possível ter uma amostra tão grande quanto gostaríamos. Se n é pequeno, como isso afeta a validade dos métodos do intervalo de confiança? Os métodos t para a média podem ser usados com qualquer n. Quando n é pequeno, entretanto, você deve ser cuidadoso na procura de valores atípicos ou grandes desvios da normalidade (como os sugeridos por dados altamente assimétricos). Eles podem afetar os resultados e a validade de usar a média como um representante do conjunto.

Lembre que a fórmula do intervalo de confiança para uma proporção requer pelo menos 15 observações de cada tipo. De outra forma, a distribuição amostral da proporção não precisa estar próxima à normal e a estimativa $ep = \sqrt{\hat{\pi}(1 - \pi)/n}$ do erro padrão verdadeiro $\sqrt{\pi(1 - \pi)/n}$ pode ser deficiente. Como resultado, o intervalo de confiança será deficiente, como mostra o próximo exemplo.

EXEMPLO 5.8 Qual a proporção dos estudantes que são vegetarianos?

Para um projeto de aula, uma estudante amostrou aleatoriamente 20 colegas na Universidade da Flórida para estimar a proporção, naquela universidade, dos que eram vegetarianos. Dos 20 alunos que ela amostrou, nenhum era vegetariano. Considere π a proporção de vegetarianos da população na universidade. A proporção amostral era $\hat{\pi} = 0/20 = 0{,}0$.

Quando $\hat{\pi} = 0{,}0$, então $ep = \sqrt{\hat{\pi}(1 - \hat{\pi})/n} = \sqrt{(0{,}0)(1{,}0)/20} = 0{,}0$. O intervalo de 95% de confiança para a proporção da população de vegetarianos é

$$\hat{\pi} \pm 1{,}96(ep) = 0{,}0 \pm 1{,}96(0{,}0),$$
$$\text{que é } 0{,}0 \pm 0{,}0 \text{ ou } (0{,}0; 0{,}0).$$

A estudante conclui que ela poderia estar 95% confiante de que π está entre 0 e 0. Mas esta fórmula do intervalo de confiança é válida somente se a amostra tiver pelo menos 15 vegetarianos e pelo menos 15 não vegetarianos. A amostra não tinha pelo menos 15 vegetarianos, portanto o método não é apropriado. ■

Para amostras pequenas, a fórmula do intervalo de confiança é ainda válida se a usarmos depois de acrescentar 4 observações artificiais, 2 de cada tipo. O tamanho da amostra de $n = 20$ no Exemplo 5.8 tinha 0 vegetarianos e 20 não vegetarianos. Podemos aplicar a fórmula do intervalo de confiança com $0 + 2 = 2$ vegetarianos e $20 + 2 = 22$ não vegetarianos. O valor do tamanho da amostra para a fórmula é então $n = 24$. Aplicando a fórmula, obtemos

$$\hat{\pi} = 2/24 = 0{,}083,\ ep = \sqrt{\hat{\pi}(1 - \hat{\pi})/n}$$
$$= \sqrt{(0{,}083)(0{,}917)/24} = 0{,}056.$$

O intervalo de 95% de confiança resultante é, então:

$$\hat{\pi} \pm 1{,}96(ep), \text{ que é } 0{,}083 \pm 1{,}96(0{,}056)$$
$$\text{ou } (-0{,}03; 0{,}19).$$

Uma proporção não pode ser negativa, assim relatamos o intervalo como (0,0; 0,19). Podemos estar 95% confiantes de que a proporção de vegetarianos na Universidade da Flórida não é maior do que 0,19.

Por que acrescentamos 2 às contagens dos dois tipos? A razão é que aquele intervalo de confiança então aproxima um baseado em um método mais complexo (descrito no Exercício 5.77) que não requer a estimativa do erro padrão.[4]

5.5 INTERVALOS DE CONFIANÇA PARA MÉDIAS E OUTROS PARÂMETROS*

Focamos, até agora, na estimação de médias e proporções. O Capítulo 3 mostrou, entretanto, que outras estatísticas são também úteis para descrever dados. Essas outras estatísticas também têm distribuições amostrais. Para amostras aleatórias grandes, suas distribuições amostrais são também aproximadamente normais e são a base dos intervalos de confiança de medidas populacionais. Vamos ilustrar nesta seção o caso da mediana.

Ineficiência da mediana amostral para dados normais

Quando a distribuição populacional é normal e a amostra é aleatória, o erro padrão da mediana amostral tem uma fórmula similar àquela da média amostral. O erro padrão é igual a $1{,}25\sigma/\sqrt{n}$.

A mediana populacional para uma distribuição normal é igual à média da população μ. Assim, a mediana amostral e a média amostral são ambas estimativas por pontos do mesmo número. A mediana amostral não é tão eficiente quanto a média amostral porque seu erro padrão é 25% maior. Quando a distribuição populacional for aproximadamente normal, a média amostral é uma estimativa melhor do centro da distribuição. Esta é uma das razões de a média ser mais comumente usada do que a mediana na inferência estatística.

Quando a distribuição populacional for altamente assimétrica, a mediana populacional é geralmente um resumo mais útil do que a média populacional. Usamos a mediana amostral para estimar a mediana populacional. Entretanto, a fórmula do erro padrão $1{,}25\sigma/\sqrt{n}$ é válida somente quando a distribuição populacional for aproximadamente normal.

Intervalo de confiança com grandes amostras para a mediana

O intervalo de confiança para a mediana discutido a seguir é válido para amostras grandes (n pelo menos aproximadamente 20-30). Ele não requer suposição sobre a distribuição populacional a não ser que ela seja essencialmente contínua. Sua lógica utiliza ideias deste capítulo.

Por definição, a probabilidade π de que uma observação aleatoriamente selecionada esteja abaixo da mediana é de 0,50. Assim, para uma amostra aleatória de tamanho n, a proporção amostral $\hat{\pi}$ que está abaixo da mediana tem média de 0,50 e erro padrão $\sigma_{\hat{\pi}} = \sqrt{\pi(1-\pi)/n} = \sqrt{0{,}50(0{,}50)/n} = 0{,}50/\sqrt{n}$. Em particular, a probabilidade é aproximadamente 0,95 de que a proporção amostral das observações que estão abaixo da mediana esteja entre dois erros padrão ou $1/\sqrt{n}$ de 0,50. O *número* amostral das observações que estão abaixo da mediana é n vezes a proporção amostral. Assim, a probabilidade é aproximadamente 0,95 de que o número de observações que estão abaixo da mediana esteja dentro de $n(1/\sqrt{n}) = \sqrt{n}$ da metade da amostra, e o número de observações que estão *acima* da mediana esteja dentro de $n(1/\sqrt{n}) = \sqrt{n}$ da metade da amostra.

Agora, para uma amostra ordenada de tamanho n, a mediana é a medida do meio, que tem um índice $(n+1)/2$. A observação com índice $(n+1)/2 - \sqrt{n}$ é o extremo inferior de um intervalo de 95% de confiança para a mediana. A observação com índice $(n+1)/2 + \sqrt{n}$ é o extremo superior.

EXEMPLO 5.9 Estimando o tempo mediano na prateleira em uma biblioteca

Uma bibliotecária da Universidade da Flórida queria estimar várias características dos livros de uma das coleções especiais da universidade. Entre as perguntas de interesse estavam: "Quão velho é um

típico livro da coleção?" e "Quanto tempo faz que um típico livro da coleção foi retirado?". Suspeitamos que a distribuição de tais variáveis possa ser muito assimétrica à direita. Portanto, usamos a mediana para descrever o centro.

A Tabela 5.2 mostra dados de P = número de anos desde a publicação do livro e C = número de anos desde que o livro foi retirado, para uma amostra aleatória sistemática de 54 livros da coleção. A Figura 5.9 mostra um diagrama de caixa e bigodes do SPSS para os valores P. Os cinco valores marcados com asteriscos representam os valores atípicos extremos que estão mais do que 3,0IIQ acima do quartil superior. A mediana amostral de 17 é mais representativa dos dados do que a média amostral de 22,6. Vamos construir um intervalo de 95% de confiança para a mediana da população da distribuição de P.

Para $n = 54$, os pontos finais de um intervalo de 95% de confiança têm índices:

$$\frac{n+1}{2} \pm \sqrt{n} = \frac{54+1}{2} \pm \sqrt{54}$$
$$= 27{,}5 \pm 7{,}3, \text{ ou } (20{,}2; 34{,}8).$$

O intervalo de confiança consiste nos 20º menores e nos 35º menores (20º maiores) valores da variável.

Para uma amostra pequena como esta, é mais simples ordenar os valores em um diagrama de caule e folhas. A Tabela 5.3 mostra parte deste diagrama para as 44 observações menores do total das 54 observações, dividindo o caule em dois. A 20ª observação menor é igual a 11, e a 35ª observação menor é igual a 19. O intervalo de 95% de confiança é igual a (11; 19). Podemos estar 95% confiantes de que o tempo mediano desde a publicação é de pelo menos 11 anos e não maior do que 19 anos. Para obter um intervalo menor, precisamos de um tamanho da amostra maior. ■

O *Bootstrap*

Para alguns parâmetros, não é possível escrever uma fórmula do intervalo de confiança que funcione bem, apesar da distribuição da população, do tamanho da amostra ou do método de amostragem. Para estes casos, uma recente invenção computacional chamada de *bootstrap* é útil. Esse método trata a distribuição amostral como se fosse a distribuição populacional verdadeira. Você amostra n observações (com reposição) desta distribuição, onde cada um dos n pontos de dados originais tem a probabilidade $1/n$ de ser selecionado para cada "nova" observação. Para esta "nova" amostra de tamanho n, você, então, constrói uma estimativa por ponto do parâmetro. Você repete

Tabela 5.2 Número de anos desde a publicação (P) e desde a retirada (C) para 54 livros

C	P	C	P	C	P	C	P	C	P
1	3	9	9	4	4	1	18	1	5
30	30	0	17	2	7	0	12	1	13
7	19	5	5	47	47	3	15	9	17
11	140	2	19	5	8	2	10	11	18
1	5	1	22	1	11	5	19	2	3
2	97	0	10	1	21	7	7	4	19
4	4	11	11	5	20	14	14	5	43
2	19	10	10	10	10	0	18	10	17
4	13	17	71	8	19	0	17	48	48
2	19	11	11	6	6	7	20	4	4
92	92	4	44	1	5	1	54		

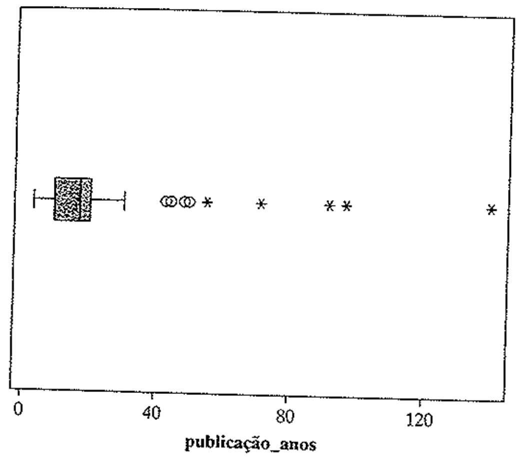


Figura 5.9 Diagrama de caixa e bigodes para o número de anos desde a publicação para uma amostra dos livros da biblioteca.

Tabela 5.3 Parte inferior do diagrama de caule e folhas para o número de anos desde a publicação. O diagrama não mostra a longa cauda direita da distribuição, que não é necessária para encontrar o intervalo de confiança para a mediana

Caule	Folhas
0	3 3 4 4 4
0	5 5 5 5 6 7 7 8 9
1	0 0 0 0 1 1 1 2 3 3 4
1	5 7 7 7 7 8 8 8 9 9 9 9 9 9 9
2	0 0 1 2

esse processo de amostragem um grande número de vezes, por exemplo, seleciona 1000 amostras separadas de tamanho n.

A distribuição amostral gerada dos valores da estimativa por pontos fornece informação sobre o parâmetro verdadeiro. Com o método do *percentil* de *boostrapping*, o intervalo de 95% de confiança para o parâmetro é 95% do conjunto central dos valores estimados. Estes são os que estão entre o 2,5º e o 97,5º percentil da distribuição amostral gerada. Este é um processo computacional intensivo, mas facilmente obtido com os computadores modernos.

5.6 RESUMO DO CAPÍTULO

Este capítulo apresentou métodos de estimação, focando na média populacional μ para variáveis quantitativas e na proporção populacional π para variáveis categóricas.

- Uma **estimativa por ponto** é a melhor suposição individual para o valor do parâmetro. As estimativas por ponto da média populacional μ, do desvio padrão σ e da proporção π são os valores amostrais $\bar{y}$, s e $\hat{\pi}$.
- Uma **estimativa intervalar,** chamada de **intervalo de confiança**, é um conjunto de números que se acredita conter o valor do parâmetro. Os intervalos de confiança para a média populacional μ e para a proporção populacional π têm a forma

Estimativa por pontos $\pm$ Margem de erro, com Margem de erro = escore $\times$ (ep),

onde ep é o erro padrão estimado. O escore multiplicado por ep é um escore-z da distribuição normal para os intervalos de confiança para proporções e um escore-t da distribuição t para intervalos de confiança para a média.

- A probabilidade de que o método gere um intervalo que contenha o parâmetro é chamado de **nível de confiança**. Ele é controlado pela escolha do escore-z ou t na margem de erro. Aumentar o nível de confiança exige o uso de um escore maior e, portanto, o ônus de obter um intervalo mais amplo.
- A **distribuição** t se parece com a distribuição normal padrão, tendo uma média de 0, mas sendo um pouco mais dispersa. Sua dispersão é determinada pelos **graus de liberdade**, que são iguais a $n - 1$ para a inferência sobre a média.
- A amplitude de um intervalo de confiança também depende do erro padrão da distribuição amostral da estimativa por pontos. Grandes amostras produzem erros padrão menores e intervalos de confiança mais estreitos e, portanto, estimativas mais precisas.

Os intervalos de confiança assumem amostragem aleatória. Para amostras grandes, eles não precisam de uma suposição sobre a distribuição populacional porque a distribuição amostral é aproximadamente normal pelo Teorema Central do Limite mesmo se a população for altamente não normal. A Tabela 5.4 resume os métodos de estimação.

A Tabela 5.4 mostra, também, as fórmulas para o tamanho da amostra necessário para alcançar a margem de erro M desejada. Você deve selecionar o nível de confiança, que determina o escore-z. Também deve obter uma estimativa para o desvio padrão da população σ para determinar o tamanho da amostra necessária para estimar a média populacional μ. Deve substituir uma estimativa para a proporção populacional π para determinar o tamanho da amostra para estimar π. Substituir π por 0,50 garante que o tamanho da amostra é grande o suficiente para dar a precisão e a confiança desejadas.

Tabela 5.4 Resumo dos métodos de estimação para médias e proporções

Parâmetro	Estimativa por ponto	Erro padrão estimado	Intervalo de confiança	Tamanho da amostra para estimar com erro máximo M
Média μ	$\bar{y}$	$ep = \frac{s}{\sqrt{n}}$	$\bar{y} \pm t(ep)$	$n = \sigma^2\left(\frac{z}{M}\right)^2$
Proporção π	$\hat{\pi}$	$ep = \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}$	$\hat{\pi} \pm z(ep)$	$n = \pi(1-\pi)\left(\frac{z}{M}\right)^2$

Nota: $z = 1{,}96$ para uma confiança de 95%; para uma probabilidade de erro α e um nível $(1-\alpha)$, o escore-z ou escore-t têm probabilidades unicaudais à direita de $\alpha/2$ (por exemplo, $\alpha/2 = 0{,}025$ para 95% de confiança).

EXERCÍCIOS

Praticando o básico

5.1 Das 577006 pessoas envolvidas em acidentes de carro na Flórida em um ano recente, 412878 estavam usando o cinto de segurança (*Fonte:* Florida Department of Highway Safety and Motor Vehicles). Encontre uma estimativa por ponto da proporção populacional dos motoristas que usam cintos de segurança na Flórida.

5.2 Em resposta à pergunta da PSG de 2006 sobre o número de horas gastas assistindo à televisão, as respostas de oito sujeitos que se identificaram como hindus foram 2, 3, 2, 1, 0, 1, 4, 3.

(a) Encontre uma estimativa por ponto da média populacional de horas gastas assistindo à televisão para hindus.

(b) A margem de erro para esta estimativa por ponto é 1,1. Explique o que isso representa.

5.3 Uma matéria da Associated Press (8 de março de 2006) sobre um levantamento de dados autorizado pela American Medical Association (Associação Médica Americana) de uma amostra aleatória de todo o país de 644 universitárias ou diplomadas com idades entre 17 e 35 anos estimou que uma proporção de 0,74 mulheres durante as férias de primavera usa a bebida como uma desculpa para um comportamento escandaloso, incluindo nudez em público e dança em mesas. Encontre o erro padrão desta estimativa e interprete.

5.4 Um levantamento de dados conduzido nacionalmente, em julho de 2006, pelo Pew Forum on Religion & Public Life perguntou se o sujeito era a favor da permissão de que casais homossexuais entrassem em sindicatos civis – acordos legais que dariam a eles muitos dos mesmos direitos de marido e mulher. Dos 2003 adultos entrevistados, 54% disseram *sim*, 42% disseram *não* e 4% não tinham opinião. Encontre o erro padrão estimado para a proporção amostral que respondeu *sim* e interprete-o.

5.5 Quando pesquisados em 2006 e perguntados se a Turquia deveria ser incluída na União Europeia se satisfizesse todas as condições determinadas por esta, o percentual que disse *sim* foi de 51% na Dinamarca ($n = 1008$) e 42% no Reino Unido ($n = 1312$). Para o resultado da Dinamarca, o relatório declarou que a margem de erro foi de mais ou menos de 3,1%. Explique como eles conseguiram esse resultado.

5.6 Uma pergunta (codificada "EQUALIZE") em uma PSG recente quis saber: "Você acha que deveria ser responsabilidade do governo reduzir as diferenças de renda entre ricos e pobres?". Aqueles que responderam *sim* incluíam 90 dos 142 sujeitos que se denominavam de "Democratas entusiastas" na identificação do partido político e 26 dos 102 que se denominavam "Republicanos entusiastas".

(a) Encontre estimativas por ponto da proporção populacional que iria responder *sim* em cada grupo.

(b) O intervalo de 95% de confiança para a proporção populacional de respostas *sim* é (0,55; 0,71) para Democratas entusiastas e (0,17; 0,34) para Republicanos entusiastas. Explique como interpretar os intervalos.

5.7 A PSG perguntou se um sujeito concordava ou discordava da seguinte afirmação: "É melhor para cada envolvido que o homem seja o empreendedor fora de casa e a mulher a que toma conta da casa e da família" (variável "FEFAM"). A proporção amostral dos que concordam foi de 0,66 em 1977 e de 0,36 em 2004 ($n = 883$).

(a) Mostre que o erro padrão estimado em 2004 foi de 0,016.

(b) Mostre que a margem de erro para um intervalo de 95% de confiança usando a estimativa de 2004 foi de 0,03. Interprete.

(c) Construa um intervalo de 95% de confiança para 2004 e interprete-o.

5.8 Uma PSG recente perguntou: "Se a mulher em uma família quer filhos, mas o marido decide que não quer filhos, é certo para o marido se recusar a ter filhos?". Dos 598 respondentes, 366 disseram *sim* e 232 disseram *não*. Mostre que o intervalo de 99% de confiança para a proporção populacional que dirá *sim* é de (0,56; 0,66).

5.9 Em 2006, a *Florida Poll* conduzida pela Universidade Internacional da Flórida perguntou se os recentes regulamentos sobre o meio ambiente são ou não muito rígidos. Dos 1200 respondentes, 229 disseram que eles eram muito rígidos. Encontre e interprete um intervalo de: **(a)** 95%, **(b)** 99% de confiança para um parâmetro relevante da época do levantamento de dados.

5.10 Quando uma PSG recente perguntou se o governo deveria impor leis rígidas para que a indústria cause menos danos ao meio ambiente, um intervalo de 95% de confiança para a proporção populacional que respondeu *sim* foi (0,87; 0,90). Um intervalo de 99% de confiança seria maior ou menor? Por quê?

5.11 Determine o escore-*z* necessário para encontrar um intervalo de confiança para uma proporção com nível de confiança de:

(a) 0,98

(b) 0,90

(c) 0,50

(d) 0,9973

5.12 Na PSG de 2006, foi perguntado aos respondentes se eles eram a favor ou contra a pena de morte para pessoas condenadas por assassinato. Um *software* mostrou os resultados:

x	n	Proporção amostral	IC de 95%
1885	2815	0,6696	(0,652;0,687)

Aqui, x se refere ao número de respondentes que eram a favor.

(a) Mostre como obter o valor apresentado em "proporção amostral".

(b) Você pode concluir que mais da metade de todos os norte-americanos adultos estão a favor? Por quê?

(c) Encontre um intervalo de 95% de confiança para a proporção de norte-americanos adultos que eram *contra* a pena de morte, a partir do intervalo de confiança mostrado para a proporção a favor.

5.13 A PSG perguntou aos respondentes: "Você acha que o uso de maconha deve ser legalizado ou não?". Veja os resultados, para todos os anos, entrando com as variáveis "GRASS" e "YEAR" em sda.berkeley.edu/GSS.

(a) Dos respondentes em 2004, que proporção disse que deve ser legalizado e que proporção disse que não deve ser legalizado?

(b) Existe evidência o suficiente para concluir se a maioria ou a minoria da população apoia a legalização? Explique o seu raciocínio.

(c) Descreva a tendência que você vê, desde aproximadamente 1986, da proporção dos que são a favor da legalização.

5.14 Quando a PSG de 2000 perguntou se os seres humanos se desenvolveram a partir de espécies primitivas de animais (variável "SCITEST4"), 53,8% dos 1095 respondentes responderam que isso era definitivamente ou provavelmente não verdadeiro. Encontre um intervalo de 99% de confiança para a proporção populacional correspondente e indique se você pode concluir que a maioria dos norte-americanos pensa desta forma.

5.15 Um estudo do US National Center for Health Statistics (Centro Nacional dos Estados Unidos para Estatísticas da Saúde) forneceu uma estimativa por ponto de 25,5% para o percentual de norte-americanos adultos que eram, no momento, fumantes. O tamanho da amostra foi de 42000. Assumindo que esta amostra tem a característica de uma amostra aleatória, construa e interprete um intervalo de 99% de confiança para a proporção da população que é fumante. (*Nota*: quando n é muito grande, mesmo intervalos de confiança com grandes níveis de confiança são estreitos.)

5.16 Na pesquisa de boca de urna discutida no capítulo anterior, dos 2705 eleitores amostrados, 56,5% disseram que votaram em Schwarzenegger. Existe evidência suficiente para prever o vencedor da eleição? Baseie sua decisão em um intervalo de 95% de confiança, declarando as suposições necessárias para esta decisão.

5.17 Para uma pesquisa de boca de urna de pessoas que votaram em uma eleição governamental, 40% votaram em Jones e 60% votaram em Smith. Assumindo que essa é uma amostra aleatória de todos os eleitores, construa um intervalo de 99% de confiança para a proporção de votos que Jones recebeu se o tamanho da amostra foi de: **(a)** 400, **(b)** 40. Em cada caso, indique se você estaria disposto a prever o vencedor. Explique como e por que o tamanho da amostra afeta a inferência.

5.18 Em 2003, a Harris Poll relatou os resultados de um levantamento de dados sobre crenças religiosas. Dos 2201 norte-americanos adultos pesquisados, 27% acreditavam em reencarnação. Tratando isto como uma amostra aleatória, um intervalo de 95% de confiança para a proporção da população de norte-americanos adultos que acreditam em reencarnação é (0,25; 0,29). Sem fazer nenhum cálculo, explique como o intervalo iria mudar se o tamanho da amostra tivesse sido somente um quarto do tamanho, $n = 550$.

5.19 Informe o escore-t que deve ser multiplicado pelo erro padrão para formar um:
- **(a)** intervalo de 95% de confiança com 5 observações.
- **(b)** intervalo de 95% de confiança com 15 observações.
- **(c)** intervalo de 95% de confiança com 25 observações.
- **(d)** intervalo de 95% de confiança com $gl = 25$.
- **(e)** intervalo de 99% de confiança com $gl = 25$.

5.20 Encontre e interprete o intervalo de 95% de confiança para μ, se $\bar{y} = 70$ e $s = 10$, baseado em um tamanho da amostra de:
- **(a)** 5
- **(b)** 20

5.21 A PSG de 2004 perguntou aos respondentes masculinos quantas parceiras femininas eles tiveram desde os seus 18 anos. A mediana = 6 e a moda = 1 (16,0% da amostra). Uma saída de computador resume os outros resultados:

Variável	n	Média	Desvio Padrão	EP da média	IC de 95,0%
NUMWOMEN	1007	24,745	52,554	1,656	(21,5; 28,0)

- **(a)** Mostre como o *software* obteve o erro padrão apresentado e interprete.
- **(b)** Interprete o intervalo de confiança calculado.
- **(c)** Apresente um fator que possa torná-lo cético sobre a utilidade desse intervalo de confiança.

5.22 Uma PSG perguntou: "Qual você acha que é o número ideal de filhos para uma família?". As 497 mulheres que responderam tiveram uma mediana de 2, uma média de 3,02 e um desvio padrão de 1,81.
- **(a)** Apresente uma estimativa por ponto da média populacional.
- **(b)** Encontre e interprete o erro padrão da média amostral.
- **(c)** O intervalo de 95% de confiança é (2,9; 3,2). Interprete.
- **(d)** É plausível que a média populacional seja igual a 2,0? Explique.

5.23 Considere o exercício anterior. Para os 397 homens da amostra, a média foi de 2,89 e o desvio padrão era 1,77.
- **(a)** Mostre que o erro padrão da média amostral é 0,089.
- **(b)** Mostre que o intervalo de 95% de confiança para a média populacional é (2,7; 3,1) e explique o que "95% de confiança" significa.

5.24 O Exemplo 5.5 (página 144) analisou dados de um estudo que comparou terapias para a anorexia. Para as 17 meninas que receberam a terapia familiar, as mudanças do peso durante o estudo foram:

11, 11, 6, 9, 14, −3,0, 7, 22, −5, −4, 13, 13, 9, 4, 6, 11.
- **(a)** Verifique que $\bar{y} = 7,29$ e $s = 7,18$ libras.
- **(b)** Verifique que o erro padrão da média amostral foi de 1,74.
- **(c)** Para usar a distribuição t, explique porque $gl = 16$ e um intervalo de 95% de confiança usa o escore t de 2,120.
- **(d)** Se μ representa a mudança média populacional de peso para esta terapia, verifique que o intervalo de 95% de confiança para μ é (3,6; 11,0). Interprete.

5.25 A PSG de 2004 perguntou: "Em um dia normal, aproximadamente quantas horas você assiste à televisão?". O *software* relata:

Variável	N	Média	EP da média	IC de 95,0%
TVHOURS	892	2,76	0,08	(2,60; 2,93)

O que está errado com a interpretação: "A longo prazo, 95% das vezes os sujeitos assistiram à televisão entre 2,60 e 2,93 horas por dia"? Informe a interpretação correta.

5.26 Em resposta à pergunta da PSG de 2006 sobre o número de horas diárias gastas assistindo à televisão, as respostas dos 15 sujeitos que se identificaram como budistas foram 0, 0, 0, 1, 1, 1, 2, 2, 2, 2, 2, 3, 4, 4, 5.
- **(a)** Estime a média, o desvio padrão e o erro padrão.
- **(b)** Construa um intervalo de 95% de confiança para a média populacional.
- **(c)** Especifique as suposições para o método. O que você pode dizer sobre a sua validade para esses dados?

5.27 A PSG perguntou as pessoas: "Há quanto tempo você mora na cidade ou comunidade onde você mora hoje?". As respostas de 1415 sujeitos desse levantamento de dados apresentaram uma mediana de 16 anos, uma média de 20,3 e um desvio padrão de 18,2.
- **(a)** Você acha que a distribuição populacional é normal? Por que sim ou por que não?
- **(b)** Baseado na sua resposta em (a), você pode construir um intervalo de 99% de confiança para a média populacional? Se não, explique porque não. Se você pode, construa-o e interprete-o.

5.28 Uma PSG recente perguntou: "Em quantos dias, nos últimos sete dias, você esteve triste?". As 816 mulheres que responderam tinham uma mediana de 1, média de 1,81 e desvio padrão de 1,98. Os 633 homens que responderam tinham uma mediana de 1, média de 1,42 e desvio padrão de 1,83.
- **(a)** Encontre um intervalo de 95% de confiança para a média populacional das mulheres. Interprete.
- **(b)** Explique porque os valores $\bar{y}$ e s sugerem que esta variável não tem uma distribuição normal. Isto causa um problema com o cálculo do intervalo de confiança em (a)? Explique.

5.29 A PSG de 2004 perguntou aos respondentes quantos parceiros sexuais eles tiveram nos últimos 12 meses. Um *software* apresentou os seguintes resultados:

Variável	N	Média	Desvio Padrão	EP da média	IC de 95,0%
Partners	2198	1,130	1,063	0,0227	(1,09; 1,18)

- **(a)** Interprete o intervalo de confiança apresentado.
- **(b)** Baseado nesses resultados, explique por que a distribuição é provavelmente assimétrica à direita. Explique por que a assimetria não causaria problema com a validade do intervalo de confiança, a não ser que existam valores atípicos.

5.30 Para a *Florida student survey*, o arquivo de dados mencionado no Exercício 1.11 (na página 25), um *software* apresenta os resultados para os respondentes sobre o número de vezes por semana que o sujeito lê um jornal:

```
Variável N  Média Desvio EP da  IC de
                  Padrão média  95,0%
News     50 4,1   3,0    0,387  (3,32;
                                 4,88)
```

(a) Interprete o intervalo de confiança obtido.
(b) Parece plausível que a distribuição populacional desta variável seja normal? Por quê?
(c) Explique as implicações do termo "robusto" com relação à suposição de normalidade para esta análise.

5.31 Uma PSG pede aos respondentes para externar sua opinião política em uma escala de sete pontos, onde 1 = extremamente liberal, 4 = moderado e 7 = extremamente conservador. Um pesquisador analisando os dados da PSG de 2004 obtém o seguinte resultado de um *software*:

```
Variável   n    Média Desvio EP da   IC de
                      Padrão média   95%
Polviews   1294 4,23  1,39   0,0387  (4,13,
                                      4,33)
```

(a) Mostre como construir o intervalo de confiança da outra informação fornecida.
(b) O intervalo de confiança seria maior ou menor (i) se você construísse um intervalo de 95% de confiança, (ii) se você encontrasse um intervalo de 99% de confiança somente para aqueles que se denominavam *fortemente Democratas* na identificação do partido político ("PARTYID"), para os quais a média era de 3,50 com um desvio padrão de 1,51?
(c) Que suposição você está fazendo sobre a escala de mensuração da ideologia política quando você usa a média amostral e o desvio padrão?

5.32 Em sda.berkeley.edu/GSS, considere as respostas à pergunta: "Em quantos dias, nos últimos sete dias, você se sentiu solitário?" (codificada "LONELY") no levantamento de dados no qual foi feita esta pergunta.
(a) Encontre uma estimativa por ponto da média populacional.
(b) Construa um intervalo de 95% de confiança e interprete.

5.33 Um estudo estima a renda média anual familiar para famílias que residem em habitações públicas de Chicago. Para uma amostra aleatória de 30 famílias, as rendas anuais (em centenas de dólares) são:

83	90	77	100	83	64	78	92	73	122
96	60	85	86	108	70	139	56	94	84
111	93	120	70	92	100	124	59	112	79.

(a) Construa um diagrama de caule e folhas da renda das famílias.
(b) Encontre e interprete estimativas por ponto de μ e σ, a média populacional e o desvio padrão.
(c) Construa e interprete um intervalo de 95% de confiança para μ.

5.34 Uma administradora de um hospital quer estimar a estadia média dos pacientes. Com base em uma amostra aleatória sistemática de 100 registros de pacientes do ano anterior, ela relata que: "A média amostral foi de 5,3. Em amostras aleatórias repetidas deste tamanho, a média amostral poderia estar dentro de 1,0 da verdadeira média aproximadamente 95% das vezes".
(a) Construa e interprete um intervalo de 95% de confiança para a média.
(b) A administradora decide que esse intervalo é muito grande e prefere um com somente metade da amplitude. Qual o tamanho da amostra necessário?

5.35 Para estimar a proporção de mortes do último ano no trânsito na Califórnia que estão relacionadas ao consumo de álcool, determine o tamanho da amostra necessário para que a estimativa seja precisa dentro de 0,06 com probabilidade de 0,90. Baseado em resultados de um estudo anterior essa proporção é aproximadamente 0,30.

5.36 Uma rede de televisão planeja prever o resultado de uma eleição entre Jacalyn Levin e Roberto Sanchez com uma pesquisa de boca de urna no dia da eleição. Os pesquisadores decidem usar uma amostra aleatória de tamanho tal que a margem de erro é de 0,04 para intervalos de 95% de confiança para as proporções populacionais.
(a) Que tamanho de amostra eles devem utilizar?
(b) Se acharem que a eleição será apertada, eles podem usar uma margem de erro de 0,02. Que tamanho deveria, nesse caso, ter a amostra? (Observe que reduzindo a margem de erro em 50% o valor de n quadruplica.)

5.37 Uma unidade de saúde pública quer amostrar os registros de mortes para o último ano em Toronto, para estimar as proporções das mortes que foram causadas por acidentes. Os funcionários da saúde querem uma estimativa com um erro de 0,02 e uma probabilidade de 0,95.
(a) Encontre o tamanho da amostra necessário se, baseado em estudos anteriores, os funcionários acreditam que esta proporção não excede a 0,10.
(b) Suponha que, ao determinar o tamanho da amostra necessário, os funcionários usam a abordagem segura que estabelece $\pi = 0,50$ na fórmula apropriada. Então, quantos registros precisam ser amostrados? Compare o resultado à resposta da parte (a) e observe a redução que ocorre no tamanho da amostra quando se faz uma estimativa para π.

5.38 Uma pesquisa no Canadá indicou que 48% dos canadenses são a favor da pena de morte (o Canadá não a tem). Um relatório da Amnesty International (Anistia Internacional) sobre isso e pesquisas relacionadas (www.amnesty.ca) não relatou o tamanho da amostra, mas declarou: "Pesquisas deste tamanho são consideradas precisas dentro de 2,5 pontos percentuais 95% das vezes". Aproximadamente de que tamanho era a amostra?

5.39 O relatório de junho de 2003 *Views of a Changing World* conduzido pelo Pew Global Attitudes Project (www.people-press.org) discutiu mudanças nas opiniões sobre os Estados Unidos por outros países. No maior país muçulmano, a Indonésia, uma pesquisa de opinião pública conduzida em maio de 2003, após o início da guerra contra o Iraque, relatou que 83% tinham uma opinião desfavorável sobre os Estados Unidos, comparada com 36% um ano antes. O resultado de 2003 foi relatado ter uma margem de erro de 3 pontos percentuais. Encontre o tamanho da amostra aproximado para o estudo.

5.40 É necessária uma estimativa do tamanho médio da área, em acres, das fazendas de Manitoba, Canadá. A estimativa deve ser correta dentro de 100 acres com a probabilidade de 0,95. Um estudo preliminar sugere que 500 acres é uma estimativa razoável para o desvio padrão do tamanho das fazendas.
(a) Quantas fazendas será necessário amostrar?
(b) Uma amostra aleatória é selecionada do tamanho encontrado em (a). A amostra tem um desvio padrão de 300 acres, em vez de 500. Qual é a margem de erro para um intervalo de 95% de confiança para a média de acres das fazendas?

5.41 Um cientista social planeja um estudo de sul-africanos adultos que residem em municípios na periferia de Cape Town, para investigar o nível educacional (o número de anos de educação completa) na comunidade negra. Muitos dos sujeitos potenciais do estudo foram forçados a deixar Cape Town em 1966 quando o governo aprovou uma lei proibindo negros de morar nas cidades. Sob o sistema de *apartheid*, as crianças negras da África do Sul não eram obrigadas a frequentar a escola, portanto alguns moradores têm pouca educação. Qual o tamanho da amostra necessário para que um intervalo de 95% de confiança para o nível educacional tenha uma margem de erro igual a 1 ano? Não existe informação sobre o desvio padrão do nível educacional, mas os pesquisadores esperam que todos os valores estejam entre 0 e 18 anos.

5.42 Que tamanho da amostra é necessário para estimar a renda média anual correta de nativos norte-americanos dentro de \$1000 com uma probabilidade de 0,99? Suponha que não exista informação anterior sobre o desvio padrão da renda anual dos nativos norte-americanos, mas estimamos que aproximadamente 95% dos seus rendimentos estão entre \$6000 e \$50000 e que esta distribuição dos rendimentos tem a forma aproximada de um monte.

5.43 Uma antropóloga quer estimar a proporção de crianças em uma tribo das Filipinas que morreram antes de atingir a idade adulta. Para famílias que ela sabia que tinham filhos nascidos entre 1980 e 1985, 3 das 30 crianças morreram antes de atingir a idade adulta. Você pode usar a fórmula usual das grandes amostras para construir um intervalo de 95% de confiança para a proporção populacional? Por que ou por que não? Construa um intervalo de confiança apropriado e interprete.

5.44 Você amostra aleatoriamente cinco estudantes da sua faculdade para estimar a proporção de estudantes que gosta de tofu. Nenhum dos cinco estudantes diz que gosta de tofu.
- **(a)** Encontre a proporção amostral dos que gostam de tofu e seu erro padrão. A interpretação usual do *ep* faz sentido?
- **(b)** Por que não é apropriado usar a fórmula do intervalo de confiança normal (da Seção 5.1) para estes dados? Use uma abordagem mais apropriada e interprete.

5.45 Considerando o Exercício 5.33, construa um intervalo de 95% de confiança para a renda média anual dos residentes de habitações públicas. Interprete.

5.46 Considere o Exemplo 5.9 (página 154). Construa um intervalo de 95% de confiança para o tempo médio desde que um livro foi retirado da biblioteca. Interprete.

Conceitos e aplicações

5.47 Você pode usar um *applet* para gerar repetidamente amostras aleatórias e construir intervalos de confiança para ilustrar seu comportamento quando usado por muitas amostras. Tente isso, vá para www.grupoa.com.br e use o *applet* dos *intervalos de confiança para uma proporção*. No menu, estabeleça o valor da proporção populacional (rotulado de *p*) em 0,50 e estabeleça o tamanho da amostra em 200. Clique em [*Simulate*]. O *software* irá gerar 100 amostras de tamanho 200 cada. Para cada amostra ele exibe os intervalos de 95% e 99% de confiança para a proporção populacional e destaca os intervalos de 95% que não contêm o valor do parâmetro. Ele também conta o número dos intervalos que contêm o valor do parâmetro e o número que não contém.
- **(a)** Na sua simulação, que percentual dos 100 intervalos de 95% de confiança gerou os que realmente continham o valor do parâmetro? Quantos teriam o parâmetro esperado?
- **(b)** Para ter uma ideia do que acontece "a longo prazo", clique em [*Simulate*] mais 49 vezes (50 vezes no total). Você terá, então, formado 50 × 100 = 5000 intervalos separados de 95% de confiança. Que percentual realmente continha o valor real do parâmetro? Você deverá ver que perto de 95% dos intervalos de confiança contêm o parâmetro verdadeiro.

5.48 Considere o exercício anterior. Usando o *applet* dos *intervalos de confiança para uma proporção*, vamos verificar que o intervalo de confiança para uma proporção pode funcionar com deficiência com amostras pequenas. Estabeleça n = 10 e $\pi = 0{,}90$. Clique em [*Simulate*] para gerar 100 amostras aleatórias, cada uma de tamanho 10, formando intervalos de confiança para π para cada uma.
- **(a)** Quantos intervalos não contêm o valor verdadeiro, $\pi = 0{,}90$? Quantos você esperaria que não contivessem o valor verdadeiro? O que isso sugere? (Observe que muitos dos intervalos contêm somente o valor 1,0, que acontece quando $\hat{\pi} = 1{,}0$.)
- **(b)** Para ver que isto não é uma casualidade, clique agora em [*Simulate*] mais 49 vezes, assim você terá um total de 5000 intervalos de confiança. Que percentual contém $\pi = 0{,}90$? (*Nota:* para cada intervalo formado, o número de *fracassos* é menor do que 15, assim a fórmula da amostra grande não é adequada.)
- **(c)** Usando o *applet* da *distribuição amostral* do mesmo *site*, selecione a distribuição populacional *binária*. Use o seu *mouse* para clicar na primeira barra e mude a proporção do 1 na população para 0,90. (Este é o valor do parâmetro π.) Especifique N = 1 tamanho da amostra aleatória n = 10. Clique em [*Sample*] e irá gerar o tamanho da amostra 10 e encontre a proporção amostral e represente-a em um histograma de proporções amostrais. Continue clicando em [*Sample*] 100 vezes, assim você gerou proporções amostrais para 100 amostras de tamanho 10 cada. Olhe para a distribuição amostral empírica dos valores da proporção amostral. Ela tem a forma de sino e é simétrica? Use isto para ajudar a explicar por que o intervalo de confiança tem um desempenho deficiente nesse caso.

5.49 Considere o arquivo de dados *Student survey* (Exercício 1.11 na página 25). Usando um *software*, construa e interprete um intervalo de 95% de confiança para **(a)** o número mensal de horas gastas assistindo à televisão, **(b)** a proporção que acredita na vida após a morte. Interprete.

5.50 Considere o arquivo de dados criado no Exercício 1.12 (página 26). Para as variáveis escolhidas pelo seu professor, proponha uma questão de pesquisa e conduza análises estatísticas inferenciais usando métodos de estimação básicos. Resuma e interprete as suas descobertas e explique como você poderia usá-las para responder à questão de pesquisa.

5.51 Em 2006 a PSG perguntou sobre o número de horas por semana gastas na internet, excluindo *e-mail* (variável "WWWHR"). Formule uma pergunta de pesquisa que você poderia fazer sobre essa variável resposta e uma variável explicativa relevante. Vá a sda.berkeley.edu/GSS e analise os dados. Prepare um pequeno relatório resumindo sua análise e respondendo à pergunta proposta.

5.52 Uma PSG recente perguntou a respondentes casados: "Você vivia com seu marido/esposa antes do casamento?". As respostas foram 57 *sim*, 115 *não* para aqueles que se denominavam politicamente liberais; e 45 *sim*, 238 *não* para aqueles que se denominavam politicamente conservadores. Analise estes dados identificando a variável resposta e a variável explicativa. Resuma sua análise em um relatório de não mais do que 300 palavras.

5.53 Quando os sujeitos de uma PSG recente foram perguntados se concordavam com as seguintes afirmações, as contagens (*sim, não*) sob várias condições foram as seguintes:

- As mulheres deveriam administrar as suas casas e deixar a administração do país aos homens: (275, 1556)
- É melhor para todos os envolvidos que o homem seja o trabalhador fora de casa e a mulher cuide da casa e da família: (627, 1208)
- Uma criança de idade pré-escolar é provável que sofra se sua mãe trabalha: (776, 1054)

Analise estes dados. Prepare um relatório de uma página, declarando as suas suposições, mostrando os resultados da descrição e inferência. Resuma suas conclusões.

5.54 As observações sobre assistir à televisão para os sete muçulmanos da PSG de um ano recente foram 0, 0, 2, 2, 2, 4, 6. Um intervalo de 95% de confiança para a média populacional é (0,3; 4,3). Suponha que a observação 6 do sétimo sujeito foi registrada de forma incorreta como 60. O que teria sido obtido para o intervalo de 95% de confiança? Compare ao intervalo (0,3; 4,3). Como isto o adverte sobre os efeitos potenciais de valores atípicos em intervalos de confiança para médias?

5.55
- **(a)** Explique o que significa um estimador ser não tendencioso.
- **(b)** Explique por que a amplitude da amostra é uma estimativa tendenciosa da amplitude da população. (*Dica:* como o mínimo e máximo da amostra se comparam ao mínimo e

máximo da população? Explique por que o intervalo da amostra é tipicamente menor do que o intervalo da população e não pode ser maior.)

5.56 Qual é o propósito de formar um intervalo de confiança para um parâmetro? O que você pode aprender dele que você não pode aprender de uma estimativa por ponto de um parâmetro?

5.57 Uma estimativa intervalar para uma média é mais informativa do que uma estimativa por ponto, porque com uma estimativa intervalar você pode descobrir a estimativa por ponto, mas com apenas uma estimativa por ponto não se tem ideia da amplitude da estimativa intervalar.

(a) Explique por que esta afirmação é correta ilustrando com o uso do intervalo de 95% de confiança igual a (4,0; 5,6) para o número médio de encontros tidos, no mês anterior, pelas mulheres em uma faculdade em particular.

(b) O intervalo de confiança em (a) usou um tamanho da amostra de 50. Quais foram a média e o desvio padrão amostral?

5.58 Explique por que os intervalos de confiança são maiores com

(a) níveis de confiança maiores.

(b) tamanhos de amostra menores.

5.59 Por que seria incomum ver um

(a) intervalo de 99,9999%?

(b) intervalo de 25% de confiança?

5.60 Dê um exemplo de um estudo em que seria importante ter

(a) um alto grau de confiança.

(b) um alto grau de precisão.

5.61 Como a heterogeneidade da população afeta o tamanho da amostra necessário para estimar a média populacional? Ilustre com um exemplo.

5.62 Explique o raciocínio por trás da seguinte afirmação: "Os estudos sobre populações mais dispersas requerem tamanhos de amostras maiores". Ilustre para o problema da estimativa da renda média para todos os médicos nos Estados Unidos comparado à estimativa da renda média para todos os novos empregados dos restaurantes do McDonald's nos Estados Unidos.

5.63 Você gostaria de encontrar a proporção de leis aprovadas pelo Congresso que foram vetadas pelo presidente na última sessão do Congresso. Após verificar os registros do Congresso, você vê que, para a população de todas as 40 leis aprovadas, 2 foram vetadas. Faz sentido construir um intervalo de confiança usando estes dados? Explique. (*Dica:* identifique a amostra e a população.)

5.64 A publicação, em 2006, *Attitudes towards European Union Enlargement* (Atitudes em relação ao aumento da União Europeia) dos estados do Eurobarometer declara:

> Os leitores são lembrados de que os resultados do levantamento de dados são *estimativas*, a precisão delas depende do tamanho da amostra e do percentual observado. Com amostras de aproximadamente 1000 entrevistas, os percentuais reais variam dentro dos seguintes limites de confiança:

Observado	Limites
10% ou 90%	± 1,9
20%, 80%	± 2,5
30%, 70%	± 2,7
40%, 60%	± 3,0
50%	± 3,1

(a) Explique como eles obtiveram 3,0 pontos para 40% ou 60%.

(b) Explique por que a margem de erro difere para os diferentes percentuais observados.

(c) Explique por que a precisão é a mesma para um percentual em particular e para 100 menos aquele valor (por exemplo, para 40 e 60%).

(d) Explique por que é mais difícil estimar a proporção populacional quando ela está próxima a 0,50 do que quando ela está próxima a 0 ou 1.

5.65 Para usar o intervalo de confiança para grandes amostras para π, você usou pelo menos 15 resultados de cada tipo. Mostre que o menor valor de n para o qual o método pode ser usado é (a) 30 quando $\hat{\pi} = 0{,}50$, (b) 50 quando $\hat{\pi} = 0{,}30$, (c) 150 quando $\hat{\pi} = 0{,}10$. Isto é, o n geral deve aumentar à medida que $\hat{\pi}$ se aproxima de 0 ou 1. (Quando a proporção verdadeira está próxima de 0 ou 1, a distribuição amostral pode ser altamente assimétrica a não ser que n seja bem grande.)

Selecione a melhor resposta nos Exercícios 5.66 a 5.69.

5.66 A razão pela qual usamos um escore-z de uma distribuição normal na construção de um intervalo de confiança para grandes amostras para uma proporção é que:

(a) Para amostras aleatórias grandes, a distribuição amostral da proporção amostral é aproximadamente normal.

(b) A distribuição populacional é normal.

(c) Para amostras aleatórias grandes a distribuição dos dados é aproximadamente normal.

(d) Se em dúvida sobre a distribuição populacional, é mais seguro assumir que ela é uma distribuição normal.

5.67 Aumentar o nível de confiança causa sobre a amplitude do intervalo de confiança:

(a) um aumento.

(b) uma diminuição.

(c) nenhuma alteração.

5.68 Mantendo tudo constante, quadruplicar o tamanho da amostra causa sobre a largura do intervalo de confiança:

(a) duplicação.

(b) divisão pela metade.

(c) redução pela quarta parte.

(d) nenhuma alteração.

5.69 Baseado nas respostas de 1467 sujeitos da Pesquisa Social Geral, um intervalo de 95% de confiança para o número médio de amigos próximos é igual a (6,8; 8,0). Qual das seguintes interpretações é ou são as mais corretas?

(a) Podemos estar 95% confiantes de que $\bar{y}$ está entre 6,8 e 8,0.

(b) Podemos estar 95% confiantes de que μ está entre 6,8 e 8,0.

(c) Noventa e cinco porcento dos valores de y = número de amigos próximos (para esta amostra) estão entre 6,8 e 8,0.

(d) Se amostras aleatórias do tamanho de 1467 forem selecionadas repetidamente, então 95% das vezes $\bar{y}$ estaria entre 6,8 e 8,0.

(e) Se amostras aleatórias do tamanho de 1467 forem selecionadas repetidamente, então, a longo prazo, 95% dos intervalos de confiança iriam conter o valor verdadeiro de μ.

5.70 Uma amostra aleatória de 50 registros gera um intervalo de 95% de confiança para a idade média do primeiro casamento de mulheres de certo condado com idades entre 21,5 a 23,0 anos. Explique o que está errado com cada uma das seguintes interpretações desse intervalo.

(a) Se amostras aleatórias de 50 registros forem repetidamente selecionadas, então 95% das vezes a amostra da idade média do primeiro casamento para mulheres do condado estaria entre 21,5 e 23,0 anos.

(b) Noventa e cinco porcento das idades do primeiro casamento para mulheres do condado estão entre 21,5 e 23,0 anos.

(c) Podemos estar 95% confiantes de que $\bar{y}$ está entre 21,5 e 23,0 anos.

(d) Se, repetidamente, amostrarmos toda a população, então 95% das vezes a média da população estaria entre 21,5 e 23,5 anos.

5.71 Considere o exercício anterior. Forneça uma interpretação apropriada.

***5.72** Para uma amostra aleatória de n sujeitos, explique por que é aproximadamente 95% provável que a proporção amostral tenha erro não maior do que $1/\sqrt{n}$ na estimativa da proporção populacional. (*Dica:* para mostrar esta "Regra $1/\sqrt{n}$", encontre dois erros padrão quando $\pi = 0{,}50$ e explique como isto é comparado a dois erros padrão a outros valores de π.) Usando este resultado, mostre que $n = 1/M^2$ é um tamanho da amostra seguro para estimar uma proporção com erro M com 95% de confiança.

*5.73 Você conhece a média amostral de n observações. Uma vez que você conhece $(n - 1)$ das observações, mostre que você pode encontrar a remanescente. Em outras palavras, para um dado valor de $\bar{y}$, os valores de $(n - 1)$ observações determinam a remanescente. Ao obter escores de uma variável quantitativa, ter $(n - 1)$ graus de liberdade significa que apenas aquela quantidade de observações é independente.

*5.74 Encontre o erro padrão da proporção amostral quando $\pi = 0$ ou $\pi = 1$. O que isto reflete?

*5.75 Seja π a probabilidade de que um eleitor aleatoriamente selecionado prefira o candidato Republicano. Você amostra 2 pessoas e nenhuma delas prefere o Republicano. Encontre a estimativa por ponto de π. Esta estimativa parece lógica? Por quê? (A estimativa *bayesiana* é uma estimativa alternativa que utiliza uma abordagem *subjetiva*, combinando os dados da amostra com opinião prévia sobre π antes de ver os dados. Por exemplo, se você achava que π era igualmente provável de estar entre 0 e 1, a estimativa bayesiana adiciona duas observações, uma de cada tipo, assim gerando a estimativa ¼.)

*5.76 Para encorajar os sujeitos a dar respostas sobre questões delicadas, o método de **resposta aleatorizada** é geralmente usado. É solicitado ao sujeito jogar uma moeda em segredo. Se for cara, o sujeito joga a moeda uma vez mais e relata o resultado, cara ou coroa. Se, ao contrário, o primeiro arremesso for coroa, o sujeito relata, então, a resposta a uma questão delicada; por exemplo, relata a resposta *cara* se a verdadeira resposta for *sim* e relata a resposta *coroa* se a verdadeira resposta for *não*. Considere π a representação da probabilidade verdadeira da resposta *sim* à questão delicada.

(a) Explique por que os números na Tabela 5.5 são as probabilidades dos quatro resultados possíveis.

(b) Considere p a representação da proporção amostral de sujeitos que relataram *cara* para a segunda resposta. Explique por que $\hat{\pi} = 2p - 0,5$ estima π.

(c) Usando esta abordagem, 200 sujeitos foram perguntados se nunca tinham fraudado intencionalmente seu imposto de renda. Relate a estimativa de π se o número de caras relatadas é igual a (i) 50, (ii) 70, (iii) 100, (iv) 150.

☑ Tabela 5.5

	Segunda resposta	
Primeira moeda	**Cara**	**Coroa**
Cara	0,25	0,25
Coroa	$\pi/2$	$(1 - \pi)/2$

*5.77 Para construir um intervalo de confiança para amostras grandes para a proporção π, não é necessário substituir $\hat{\pi}$ para o valor desconhecido de π na fórmula para o erro padrão de $\hat{\pi}$. Um método menos aproximado encontra os extremos de um intervalo de 95% de confiança determinando os valores de π que estão a 1,96 erros padrão da proporção amostral, solucionando para π a equação

$$|\hat{\pi} - \pi| = 1{,}96\sqrt{\frac{\pi(1 - \pi)}{n}}.$$

Para o Exemplo 5.8 (página 153) com nenhum vegetariano em uma amostra de tamanho 20, substitua $\hat{\pi}$ e n nesta equação e mostre que a equação é satisfeita em $\pi = 0$ e em $\pi = 0{,}161$. Portanto, o intervalo de confiança é (0; 0,161).

☑ NOTAS

1 Aqui, π *não* é a constante matemática, 3,1415...
2 Veja sda.berkeley.edu/D3/GSS06/Doc/gs06.htm.
3 Cortesia do Prof. Brian Everitt, Instituto de Psiquiatria, Londres.
4 Veja o artigo de A. Agresti e B. Coull (que propôs esse intervalo de confiança), *American Statistician*, v. 52, p. 119-26, 1998.

☑ 6 INFERÊNCIA ESTATÍSTICA: TESTES DE SIGNIFICÂNCIA

O objetivo de muitos estudos é verificar se os dados concordam com certas previsões. As previsões geralmente resultam da teoria que leva à pesquisa. Essas previsões são *hipóteses* sobre a população em estudo.

> ☑ **Hipótese**
>
> Na estatística, uma **hipótese** é uma afirmação sobre a população. Ela é geralmente uma previsão na qual um parâmetro que descreve uma característica de uma variável assume um valor numérico particular ou está em certo intervalo de valores.

Exemplos de hipóteses são os seguintes: "Para prestadores de serviço, a renda média é a mesma tanto para mulheres quanto para homens", "Não existe diferença em termos probabilísticos entre Democratas e Republicanos em relação ao voto seguindo a liderança do seu partido" e "A metade ou mais dos adultos canadenses está satisfeita com seu serviço nacional de saúde".

Um **teste de significância** usa dados para resumir a evidência sobre uma hipótese, comparando as estimativas por pontos dos parâmetros aos valores previstos pela hipótese. O seguinte exemplo ilustra os conceitos por trás dos testes de significância.

EXEMPLO 6.1 Testando a tendenciosidade de gênero na seleção de gerentes

Uma grande rede de supermercados na Flórida selecionou alguns empregados para receber treinamento para a gerência. Um grupo de funcionários femininos alegou que os homens foram selecionados em uma taxa desproporcionalmente alta para tal treinamento. A empresa negou esta alegação.[1] Uma alegação similar de tendenciosidade de gênero foi feita sobre as promoções e salários para mulheres que trabalham na Wal-Mart.[2] Como as funcionárias poderiam respaldar estatisticamente sua declaração?

Suponha que o grupo potencial de funcionários para uma seleção para o treinamento para gerente é formado meio a meio por homens e mulheres. Portanto, a alegação de não tendenciosidade de gênero é uma hipótese. Ela afirma que, mantendo as demais condições constantes, a cada escolha, a probabilidade de selecionar uma mulher é igual a ½ e a probabilidade de selecionar um homem é igual a ½. Se os funcionários realmente são selecionados aleatoriamente para o treinamento para gerente em termos de gênero, aproximadamente metade dos funcionários selecionados deveria ser de mulheres e metade deveria ser de homens. A alegação das mulheres é uma hipótese alternativa de que a probabilidade de selecionar um homem é maior do que ½.

Suponha que nove dos 10 funcionários escolhidos para o treinamento para gerente foram homens. Podemos ficar inclinados a acreditar na alegação das mulheres. Entretanto, devemos analisar se es-