10 Testes para Duas Amostras

Objetivos do Aprendizado

Neste capítulo, você aprenderá:

- As médias aritméticas de duas populações independentes
- As médias aritméticas de duas populações relacionadas
- As proporções de duas populações independentes
- As variâncias de duas populações independentes

UTILIZANDO A ESTATÍSTICA

@ BLK Alimentos

O tipo de disposição de mercadorias utilizado em um supermercado afeta as vendas dos produtos? Como gerente regional de vendas da BLK Alimentos, você deseja comparar o volume de vendas de refrigerantes da marca BLK sabor cola quando o produto é exposto em uma localização normal de prateleira com o volume de vendas quando o produto é exposto em posições especiais de ponta de corredor. Para testar a eficácia de exposições em pontas de corredor, você seleciona 20 filiais da cadeia de supermercados BLK que apresentem, todas elas, volumes de vendas similares em todas as seções da loja. Depois disso, você designa, aleatoriamente, 10 dentre as 20 filiais para a amostra 1 e 10 filiais para a amostra 2. Os gerentes das 10 filiais na amostra 1 posicionam o refrigerante da marca BLK em prateleiras em localizações regulares, juntamente com outros tipos de refrigerante. As 10 lojas na amostra 2 utilizam a exposição em locais especiais em pontas de corredor, destinados a promoções. Ao final de uma semana, são registradas as vendas do refrigerante da marca BLK. Como você consegue determinar se as vendas de refrigerantes da marca BLK com a exposição promocional de ponta de corredor são as mesmas que as ocorridas quando o refrigerante é exposto em uma prateleira com localização regular? Como você seria capaz de decidir se a variabilidade das vendas de refrigerantes da marca BLK, de loja para loja, é a mesma no que diz respeito aos dois tipos de exposição do produto? Como você poderia utilizar as respostas para essas perguntas de modo a incrementar as vendas dos refrigerantes da marca BLK?

O teste de hipóteses proporciona uma abordagem *confirmatória* para a análise de dados. No Capítulo 9, você aprendeu uma variedade de procedimentos de testes de hipóteses habitualmente utilizados que se relacionam a uma única amostra de dados extraída de uma única população. Neste capítulo, você aprenderá a estender o teste de hipóteses para **testes de duas amostras** que comparam estatísticas oriundas de amostras de dados extraídas de duas populações. Um desses tipos de teste para o cenário da BLK Alimentos seria: "A média aritmética das vendas semanais dos refrigerantes da marca BLK, quando é utilizada a exposição em localização regular de prateleiras (uma população), é igual à média aritmética das vendas semanais de refrigerantes da marca BLK quando é utilizada a exposição especial de ponta de corredor (uma segunda população)?"

10.1 Comparando as Médias Aritméticas de Duas Populações Independentes

Nas Seções 8.1 e 9.1, você aprendeu que em quase todos os casos você não conheceria o desvio-padrão da população no que diz respeito à população em estudo. Do mesmo modo, quando extrai uma amostra aleatória de cada uma de duas populações independentes, você quase sempre não conhece o desvio-padrão de nenhuma das populações. Além disso, você precisa saber se pode pressupor que as variâncias nas duas populações são iguais, porque o método que você utiliza para comparar as médias aritméticas de cada uma das populações depende do fato de você poder ou não pressupor que as variâncias das duas populações são iguais.

Teste *t* de Variância Agrupada para a Diferença entre Duas Médias Aritméticas

Se você pressupõe que as amostras aleatórias são selecionadas independentemente a partir de duas populações e que as populações são normalmente distribuídas nos moldes de uma distribuição normal e possuem variâncias iguais, você pode utilizar o **teste *t* de variância agrupada** para determinar se existe uma diferença significativa entre as médias aritméticas de duas populações. Se as populações não forem normalmente distribuídas, o teste *t* de variância agrupada pode, ainda assim, ser utilizado se os tamanhos de amostras forem grandes o suficiente (geralmente ≥ 30 para cada uma das amostras[1]).

Utilizando subscritos para diferenciar entre a média aritmética da primeira população, μ_1 e a média aritmética da segunda população, μ_2, a hipótese nula de nenhuma diferença nas médias aritméticas de duas populações independentes pode ser enumerada como

$$H_0: \mu_1 = \mu_2 \quad \text{ou} \quad \mu_1 - \mu_2 = 0$$

e a hipótese alternativa de que as médias aritméticas não são iguais pode ser enumerada como

$$H_1: \mu_1 \neq \mu_2 \quad \text{ou} \quad \mu_1 - \mu_2 \neq 0$$

Para testar a hipótese nula, utilize a estatística do teste t_{ESTAT} para variância agrupada ilustrado na Equação (10.1). O teste *t* de variância agrupada herdou seu nome do fato de que a estatística do teste agrupa, ou combina, as variâncias das duas amostras S_1^2 e S_2^2 de modo a calcular S_p^2, a melhor estimativa da variância comum a ambas as populações, sob o pressuposto de que as variâncias das duas populações são iguais.[2]

[1] Reveja a discussão da Seção 7.4 sobre o Teorema do Limite Central na Seção 7.4 do Capítulo 7 para compreender mais sobre tamanhos de amostras "grandes o suficiente".

[2] Quando os dois tamanhos de amostra são iguais (ou seja, $n_1 = n_2$), a equação para a variância agrupada pode ser simplificada para

$$S_p^2 = \frac{S_1^2 + S_2^2}{2}$$

TESTE *t* DE VARIÂNCIA AGRUPADA PARA A DIFERENÇA ENTRE DUAS MÉDIAS ARITMÉTICAS

$$t_{ESTAT} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad (10.1)$$

em que

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)}$$

e

S_p^2 = variância agrupada

$\bar{X}_1$ = média aritmética da amostra extraída da população 1

S_1^2 = variância da amostra extraída da população 1

n_1 = tamanho da amostra extraída da população 1

$\bar{X}_2$ = média aritmética da amostra extraída da população 2

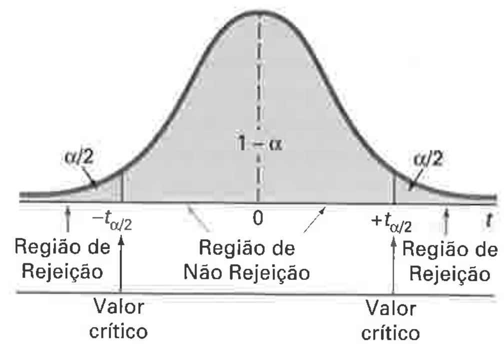
S_2^2 = variância da amostra extraída da população 2

n_2 = tamanho da amostra extraída da população 2

A estatística do teste t_{ESTAT} segue uma distribuição *t* com $n_1 + n_2 - 2$ graus de liberdade.

Para um determinado nível de significância, α, em um teste bicaudal, você rejeita a hipótese nula caso a estatística do teste t_{ESTAT} calculada seja maior do que o valor crítico da cauda superior da distribuição *t*, ou caso a estatística do teste t_{ESTAT} calculada seja menor do que o valor crítico da cauda inferior da distribuição *t*. A Figura 10.1 apresenta as regiões de rejeição.

FIGURA 10.1
Regiões de rejeição e de não rejeição para o teste *t* de variância agrupada para a diferença entre as médias aritméticas (teste bicaudal)



Em um teste unicaudal em que a região de rejeição está posicionada na cauda inferior, você rejeita a hipótese nula se a estatística do teste t_{ESTAT} calculada for menor do que o valor crítico da cauda inferior da distribuição *t*. Em um teste unicaudal no qual a região de rejeição está posicionada na cauda superior, você rejeita a hipótese nula se a estatística do teste t_{ESTAT} calculada for maior do que o valor crítico da cauda superior da distribuição *t*.

Para demonstrar a utilização do teste *t* de variância agrupada, retorne ao cenário Utilizando a Estatística no início deste capítulo. Você define o objetivo da BLK Alimentos como sendo determinar se a média aritmética das vendas semanais dos refrigerantes da marca BLK é a mesma quando é utilizada a localização regular de prateleiras e quando é utilizada uma exposição em ponta de corredor. Existem duas populações de interesse. A primeira população corresponde ao conjunto de todas as vendas semanais possíveis de refrigerantes da marca BLK se todos os supermercados BLK utilizarem a exposição dos produtos em prateleiras regulares. A segunda população corresponde ao conjunto de todas as vendas semanais possíveis de refrigerantes da marca BLK se todos os supermercados BLK utilizarem a modalidade de exposição dos produtos em pontas de corredor. Você coleta os dados extraídos de uma amostra de 10 supermercados BLK aos quais tenha sido designada a exposição de produtos em prateleiras regulares e uma segunda amostra de 10 supermercados BLK aos quais tenha sido designada a exposição de produtos em pontas de corredor. Você organiza e guarda os resultados no arquivo Cola. A Tabela 10.1 contém as vendas dos refrigerantes BLK (em número de embalagens) correspondentes às duas amostras.

TABELA 10.1
Comparando as Vendas Semanais dos Refrigerantes da Marca BLK, Sabor cola, partir de Duas Diferentes Localizações para Exposição do Produto (em Número de Embalagens)

Local de Exposição do Produto									
Normal					Ponta de Corredor				
22	34	52	62	30	52	71	76	54	67
40	64	84	56	59	83	66	90	77	84

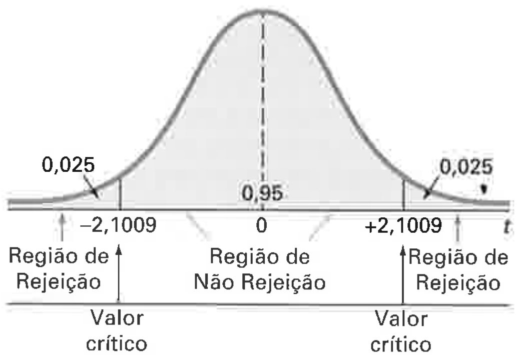
A hipótese nula e a hipótese alternativa são:

$$H_0: \mu_1 = \mu_2 \quad \text{ou} \quad \mu_1 - \mu_2 = 0$$
$$H_1: \mu_1 \neq \mu_2 \quad \text{ou} \quad \mu_1 - \mu_2 \neq 0$$

Considerando que as amostras sejam extraídas de populações normais com variâncias iguais, você pode utilizar o teste t de variância agrupada. A estatística do teste t_{ESTAT} segue uma distribuição t com $10 + 10 - 2 = 18$ graus de liberdade. Utilizando o nível de significância $\alpha = 0{,}05$, você divide a região de rejeição entre as duas caudas para esse teste bicaudal (ou seja, duas partes iguais, com 0,025 cada uma). A Tabela E.3 demonstra que os valores críticos para esse teste bicaudal correspondem a $+2{,}1009$ e $-2{,}1009$. Conforme ilustrado na Figura 10.2, a regra de decisão é:

$$\text{Rejeitar } H_0 \text{ se } t_{ESTAT} > +2{,}1009$$
$$\text{ou se } t_{ESTAT} < -2{,}1009;$$
$$\text{caso contrário, não rejeitar } H_0.$$

FIGURA 10.2
Teste de hipóteses bicaudal para a diferença entre as médias aritméticas, no nível de significância de 0,05, com 18 graus de liberdade



Com base na Figura 10.3, a estatística t_{ESTAT} calculada para esse teste é $-3{,}0446$, e o valor-p é 0,0070.

FIGURA 10.3
Planilha para o teste t de variância agrupada para os dois locais de exposição do refrigerante sabor cola da marca BLK

A Figura 10.3 exibe a ***planilha CÁLCULO*** *da* ***pasta de trabalho T variância agrupada****. Crie essa planilha utilizando as instruções da Seção GE10.1.*

	A	B	C
1	Teste t de Variância Agrupada para a Diferença entre Duas Médias Aritméticas		
2	(pressupõe iguais variâncias de população)		
3	**Dados**		
4	Hipótese da Diferença	0	
5	Nível de Significância	0,05	
6	**Amostra da População 1**		
7	Tamanho da Amostra	10	=CONT.NÚM(CÓPIADADOS!$A:$A)
8	Média Aritmética da Amostra	50,3	=MÉDIA(CÓPIADADOS!$A:$A)
9	Desvio-padrão da Amostra	18,7264	=DESVPAD(CÓPIADADOS!$A:$A)
10	**Amostra da População 2**		
11	Tamanho da Amostra	10	=CONT.NÚM(CÓPIADADOS!$B:$B)
12	Média Aritmética da Amostra	72	=MÉDIA(CÓPIADADOS!$B:$B)
13	Desvio-padrão da Amostra	12,5433	=DESVPAD(CÓPIADADOS!$B:$B)
14			
15	Cálculos Intermediários		
16	Graus de Liberdade da Amostra da População 1	9	=B7 - 1
17	Graus de Liberdade da Amostra da População 2	9	=B11 - 1
18	Total dos Graus de Liberdade	18	=B16 + B17
19	Variância Agrupada	254,0056	=((B16 * B9^2) + (B17 * B13^2))/B18
20	Erro-padrão	7,1275	=RAIZ(B19 * (1/B7 + 1/B11))
21	Diferença entre Médias das Amostras	-21,7	=B8 - B12
22	Estatística do Teste t	-3,0446	=(B21 - B4)/B20
23			
24	**Teste Bicaudal**		
25	Valor Crítico Inferior	-2,1009	=-INVT(B5, B18)
26	Valor Crítico Superior	2,1009	=INVT(B5, B18)
27	Valor p	0,0070	=DISTT(ABS(B22), B18, 2)
28	Rejeitar a hipótese nula		=SE(B27 < B5, "Rejeitar a hipótese nula", "Não rejeitar a hipótese nula")

Utilizando a Equação (10.1) e as estatísticas descritivas apresentadas na Figura 10.3,

$$t_{ESTAT} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$$

em que

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)}$$
$$= \frac{9(18{,}7264)^2 + 9(12{,}5433)^2}{9 + 9} = 254{,}0056$$

Portanto,

$$t_{ESTAT} = \frac{(50{,}3 - 72{,}0) - 0{,}0}{\sqrt{254{,}0056\left(\dfrac{1}{10} + \dfrac{1}{10}\right)}} = \frac{-21{,}7}{\sqrt{50{,}801}} = -3{,}0446$$

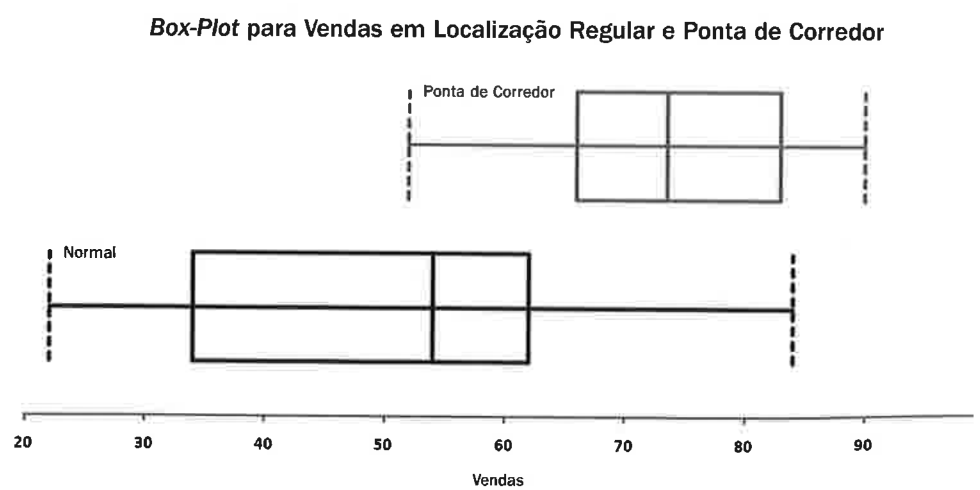
Você rejeita a hipótese nula uma vez que $t_{ESTAT} = -3{,}0446 < -2{,}1009$ e o valor-p é 0,0070. Em outras palavras, a probabilidade de que $t_{ESTAT} > 3{,}0446$ ou $t_{ESTAT} < -3{,}0446$ é igual a 0,0070. Como o valor-p é menor do que $\alpha = 0{,}05$, existem evidências suficientes para rejeitar a hipótese nula. Você pode concluir que as médias das vendas são diferentes para a localização em prateleiras regulares e para a localização em pontas de corredor. Com base nesses resultados, as vendas são mais baixas para a localização regular (em comparação com a localização em ponta de corredor).

Ao testar a diferença entre as médias aritméticas, você pressupõe que as populações são normalmente distribuídas, com variâncias iguais. Para situações em que, as duas populações apresenta variâncias iguais, o teste t de variância agrupada é **robusto** (ou seja, não sensível) em relação a afastamentos moderados da premissa da normalidade, contanto que os tamanhos das amostras sejam grandes. Nesses tipos de situações, você pode utilizar o teste t de variância agrupada sem sérios efeitos sobre a sua eficácia. Por outro lado, caso você não possa adotar o pressuposto de que ambas as populações são normalmente distribuídas, você tem duas opções. Você pode utilizar um procedimento não paramétrico, tal como o teste da soma de classificações de Wilcoxon (abordado na Seção 12.5), que não depende do pressuposto da normalidade para as duas populações, ou pode utilizar uma transformação normalizadora (veja a referência 5) em cada um dos resultados e, depois, utilizar o teste t de variância agrupada.

Para verificar o pressuposto da normalidade em cada uma das duas populações, construa o *box-plot* correspondente às vendas nas duas localizações para fins de exposição do produto, ilustrado na Figura 10.4. Para essas duas pequenas amostras, parece existir um afastamento apenas um pouco moderado em relação à normalidade, de modo que o pressuposto da normalidade necessário para o teste t não é seriamente violado.

FIGURA 10.4
Box-plot para as vendas em duas localizações para exposição do produto

Crie box-plots utilizando as instruções na Seção GE3.3.



O Exemplo 10.1 fornece uma outra aplicação para o teste t para variância agrupada.

EXEMPLO 10.1

Testando a Diferença entre as Médias Aritméticas Correspondentes aos Tempos de Entrega

Você e alguns amigos decidiram testar a validade de um anúncio feito por uma pizzaria local, que afirma entregar pizza nos dormitórios mais rápido do que uma filial de uma cadeia nacional de pizzarias. Tanto a pizzaria local quanto a da cadeia nacional estão localizadas do outro lado da rua do *campus* de sua universidade. Você define a variável de interesse como o tempo de entrega, em minutos, desde o tempo em que a pizza é encomendada até o momento em que é entregue. Você coleta os dados encomendando 10 pizzas da pizzaria local e 10 pizzas da filial da cadeia nacional, todas em diferentes horários. Você organiza e armazena os dados em **TempoPizza**. A Tabela 10.2 mostra os tempos correspondentes à entrega.

TABELA 10.2
Tempos de Entrega (em minutos) da Pizzaria Local e da Filial da Cadeia Nacional

Local		Cadeia Nacional	
16,8	18,1	22,0	19,5
11,7	14,1	15,2	17,0
15,6	21,8	18,7	19,5
16,7	13,9	15,6	16,5
17,5	20,8	20,8	24,0

No nível de significância de 0,05, existem evidências de que a média aritmética do tempo de entrega da pizzaria local é menor do que o tempo de entrega da filial da cadeia nacional de pizzarias?

SOLUÇÃO Tendo em vista que você deseja saber se a média aritmética é *mais baixa* para a pizzaria local do que para a filial da cadeia nacional de pizzarias, você tem um teste unicaudal com as seguintes hipóteses nula e alternativa:

H_0: $\mu_1 \geq \mu_2$ (A média aritmética do tempo de entrega da pizzaria local é igual ou maior do que a média aritmética do tempo de entrega da filial da cadeia nacional de pizzarias.)

H_1: $\mu_1 < \mu_2$ (A média aritmética do tempo de entrega da pizzaria local é menor do que a média aritmética do tempo de entrega da filial da cadeia nacional de pizzarias.)

A Figura 10.5 ilustra em uma planilha os resultados do Microsoft Excel do teste t para variância agrupada para esses dados.

FIGURA 10.5
Planilha do teste t de variância agrupada para os dados relativos aos tempos de entrega de pizzas

Crie essa planilha utilizando as instruções na Seção GE10.1.

	A	B
1	Teste *t* de Variância Agrupada para a Diferença entre Duas Médias Aritméticas	
2	(pressupõe iguais variâncias de população)	
3	**Dados**	
4	Hipótese da Diferença	0
5	Nível de Significância	0,05
6	**Amostra da População 1**	
7	Tamanho da Amostra	10
8	Média Aritmética da Amostra	16,7
9	Desvio-padrão da Amostra	3,0955
10	**Amostra da População 2**	
11	Tamanho da Amostra	10
12	Média Aritmética da Amostra	18,88
13	Desvio-padrão da Amostra	2,8662
14		
15	Cálculos Intermediários	
16	Graus de Liberdade da Amostra da População 1	9
17	Graus de Liberdade da Amostra da População 2	9
18	Total dos Graus de Liberdade	18
19	Variância Agrupada	8,8987
20	Erro-padrão	1,3341
21	Diferença entre Médias das Amostras	-2,18
22	Estatística do Teste *t*	-1,6341
23		
24	**Teste da Cauda Inferior**	
25	Valor Crítico Inferior	-1,7341
26	Valor *p*	0,0598
27	Não rejeitar a hipótese nula	

Para ilustrar os cálculos, utilizando a Equação (10.1),

$$t_{ESTAT} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

em que

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)}$$

$$= \frac{9(3{,}0955)^2 + 9(2{,}8662)^2}{9 + 9} = 8{,}8987$$

Portanto,

$$t_{ESTAT} = \frac{(16{,}7 - 18{,}88) - 0{,}0}{\sqrt{8{,}8987\left(\frac{1}{10} + \frac{1}{10}\right)}} = \frac{-2{,}18}{\sqrt{1{,}7797}} = -1{,}6341$$

Você não rejeita a hipótese nula porque $t_{ESTAT} = -1{,}6341 > -1{,}7341$. O valor-$p$ (conforme calculado na planilha da Figura 10.5) é igual a 0,0598. Esse valor-p indica que a probabilidade de que $t_{ESTAT} < -1{,}6341$ é igual a 0,0598. Em outras palavras, se as médias aritméticas das populações forem iguais, a probabilidade de que a média aritmética da amostra para o tempo de entrega na pizzaria local seja pelo menos 2,18 minutos menor do que o tempo de entrega da filial da cadeia nacional é de 0,0598. Uma vez que o valor-p é maior do que $\alpha = 0{,}05$, não existem evidências suficientes para rejeitar a hipótese nula. Com base nesses resultados, não existem evidências suficientes para que a pizzaria local alegue em sua propaganda ter um tempo de entrega mais rápido.

Estimativa do Intervalo de Confiança para a Diferença entre Duas Médias Aritméticas

Em vez de, ou além de, testar a diferença entre as médias aritméticas de duas populações independentes, você pode utilizar a Equação (10.2) para desenvolver uma estimativa de um intervalo de confiança para a diferença entre duas médias aritméticas.

A planilha CÁLCULO da pasta de trabalho T Variância Agrupada calcula a estimativa de um intervalo de confiança para a diferença entre duas médias aritméticas.

ESTIMATIVA DO INTERVALO DE CONFIANÇA PARA A DIFERENÇA ENTRE DUAS MÉDIAS ARITMÉTICAS DE SUAS POPULAÇÕES INDEPENDENTES

$$(\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

ou

$$(\bar{X}_1 - \bar{X}_2) - t_{\alpha/2}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \leq \mu_1 - \mu_2 \leq (\bar{X}_1 - \bar{X}_2) + t_{\alpha/2}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \quad \textbf{(10.2)}$$

em que $t_{\alpha/2}$ corresponde ao valor crítico da distribuição t, com $n_1 + n_2 - 2$ graus de liberdade, para uma área de $\alpha/2$ na cauda superior.

Para as estatísticas das amostras pertinentes às duas localizações em corredor apresentadas na Figura 10.3, ao serem utilizados 95% de confiança e a Equação (10.2),

$$\bar{X}_1 = 50{,}3,\ n_1 = 10,\ \bar{X}_2 = 72,\ n_2 = 10,\ S_p^2 = 254{,}0056 \text{ e com } 10 + 10 - 2$$

$$= 18 \text{ graus de liberdade, } t_{0{,}025} = 2{,}1009$$

$$(50{,}3 - 72{,}0) \pm (2{,}1009)\sqrt{254{,}0056\left(\frac{1}{10} + \frac{1}{10}\right)}$$
$$-21{,}7 \pm (2{,}1009)(7{,}1275)$$
$$-21{,}7 \pm 14{,}97$$
$$-36{,}67 \leq \mu_1 - \mu_2 \leq -6{,}73$$

Por conseguinte, você está 95% confiante de que a diferença na média aritmética para as vendas entre o local de exposição de produtos em prateleiras regulares e em pontas de corredor encontra-se entre −36,67 embalagens de refrigerante sabor cola e −6,73 embalagens de refrigerante sabor cola. Em outras palavras, a localização especial de produtos em pontas de corredor vende, em média, 6,73 a 36,67 embalagens a mais do que a localização em prateleiras regulares. Partindo de uma perspectiva de teste de hipóteses, uma vez que o intervalo não inclui o zero, você rejeita a hipótese nula de que não existe nenhuma diferença entre as médias aritméticas das duas populações.

Teste *t* para a Diferença entre Duas Médias Aritméticas Pressupondo Variâncias Diferentes

Caso você não consiga adotar o pressuposto de que as duas populações independentes apresentam variâncias iguais, você não consegue agrupar as variâncias das duas amostras em um estimador comum, S_p^2 e, portanto, não pode usar o teste *t* de variância agrupada. Em vez disso, você utiliza o **teste *t* de variâncias separadas**, desenvolvido por Satterthwaite (veja a referência 4). Esse procedimento de teste utiliza uma série de cálculos que envolve o cálculo de duas variâncias de amostras separadas para calcular graus de liberdade para a estatística do teste *t*.

A Figura 10.6 apresenta uma solução por meio de uma planilha que realiza o teste *t* de variâncias separadas para os dados sobre localizações para a exposição de produtos.

FIGURA 10.6

Planilha do teste *t* de variâncias separadas para os dados sobre localizações para a exposição de produtos

	A	B		D	E	
1	Teste *t* de Variância Agrupada para a Diferença entre Duas Médias Aritméticas					
2	(pressupõe diferentes variâncias de população)					
3	Dados					
4	Hipótese da Diferença	0				
5	Nível de Significância	0,05				
6	Amostra da População 1					
7	Tamanho da Amostra	10	=CONT.NÚM(CÓPIADADOS!$A:$A)			
8	Média Aritmética da Amostra	50,3	=MÉDIA(CÓPIADADOS!$A:$A)			
9	Desvio-padrão da Amostra	18,7264	=DESVPAD(CÓPIADADOS!$A:$A)			
10	Amostra da População 2					
11	Tamanho da Amostra	10	=CONT.NÚM(CÓPIADADOS!$B:$B)			
12	Média Aritmética da Amostra	72	=MÉDIA(CÓPIADADOS!$B:$B)			
13	Desvio-padrão da Amostra	12,5433	=DESVPAD(CÓPIADADOS!$B:$B)			
14						
15	Cálculos Intermediários			Área de Cálculos		
16	Graus de Liberdade do Numerador	2580,7529	=(E18 + E19)^2	Variância da Amostra da Pop 1	350.6778	=B9^2
17	Graus de Liberdade do Denominador	164,1430	=(E18^2)/(B7 - 1) + (E19^2)/(B11 - 1)	Variância da Amostra da Pop 2	157.3333	=B13^2
18	Total dos Graus de Liberdade	15,7226	=B16/B17	Variância da Amostra/Tamanho da Amostra da Pop 1	35.0678	=E16/B7
19	Graus de Liberdade	15	=INV(B18)	Variância da Amostra/Tamanho da Amostra da Pop 2	15.7333	=E17/B11
20	Erro-padrão	7,1275	=RAIZ(E18 + E19)	Para testes unicaudais:		
21	Diferença entre Médias das Amostras	-21,7	=B8 - B12	Valor de DISTT	0.0041	=DISTT(ABS(B22), B19, 1)
22	Estatística do Teste *t* de Variâncias Separadas	-3,0446	=B21/B20	Valor de 1 - DISTT	0.9959	=1 - E21
23						
24	Teste Bicaudal					
25	Valor Crítico Inferior	-2,1314	=-INVT(B5, B19)			
26	Valor Crítico Superior	2,1314	=INVT(B5, B19)			
27	Valor-*p*	0,0082	=DISTT(ABS(B22), B19, 2)			
28	Rejeitar a hipótese nula		=SE(B27 < B5, "Rejeitar a hipótese nula", "Não rejeitar a hipótese nula")			

*A Figura 10.6 ilustra a **planilha CÁLCULO** da **pasta de trabalho T Variâncias Separadas**. Crie essa planilha utilizando as instruções na Seção GE10.1.*

Na Figura 10.6, a estatística do teste $t_{ESTAT} = 3{,}0446$ e o valor-*p* é $0{,}0082 < 0{,}05$. Por conseguinte, os resultados para o teste *t* de variâncias separadas são quase exatamente os mesmos que os do teste *t* de variância agrupada. O pressuposto da igualdade entre as variâncias das populações não teve nenhum efeito real sobre os resultados. Algumas vezes, no entanto, os resultados dos testes *t* para variância agrupada e dos testes *t* para variâncias separadas são conflitantes, pelo fato de ter sido violado o pressuposto de variâncias iguais. Assim, é importante que você avalie os pressupostos e utilize esses resultados como uma diretriz para selecionar apropriadamente um procedimento de teste. Na Seção 10.4, é utilizado o teste *F* para a proporcionalidade entre duas variâncias a fim de determinar se existem evidências de uma diferença entre as variâncias nas duas populações. Os resultados desse teste podem ajudar você a determinar qual dentre os testes *t* — variância agrupada ou variâncias separadas — é o mais apropriado.

PENSE SOBRE ISSO “Essa Chamada Pode Ser Gravada...”

Se você alguma vez já utilizou um telefone para um serviço de atendimento ao cliente, provavelmente já escutou pelo menos uma vez a mensagem que começa com “esta chamada pode ser gravada...” Na maior parte das vezes, a mensagem explica que a gravação e o monitoramento são para “garantir o padrão de qualidade”, mas será que as empresas realmente monitoram as chamadas dos clientes para melhorar o padrão de qualidade?

De uma de nossas ex-alunas, ficamos sabendo que uma determinada empresa financeira de grande porte realmente grava e monitora as ligações para garantir o padrão de qualidade. Foi pedido a essa aluna que desenvolvesse um programa de treinamento para melhoria no atendimento para uma central de atendimento a clientes que estava contratando pessoas para responderem a chamadas telefônicas dos clientes referentes a empréstimos pendentes de pagamento. Com o propósito de resposta e avaliação, ela planejou selecionar aleatoriamente chamadas telefônicas recebidas por cada um dos novos empregados e classificar cada empregado em relação a 10 aspectos da chamada, incluindo o fato de o empregado manter, ou não, um tom de voz agradável ao falar com o cliente.

Para Quem Você Vai Telefonar?

Nossa aluna apresentou seu plano a seu chefe, para fins de aprovação, mas ele, lembrando-se das palavras de um famoso estatístico, afirmou: “Confiamos em Deus; todos os outros devem apresentar dados.” Ou seja, ele queria provas de que seu novo programa de treinamento melhoraria o serviço de atendimento a clientes. Diante desse tipo de exigência, a que você recorreria? Ela telefonou para um de nós. “Oi, professor, você não adivinha o porquê de eu ter ligado. Eu trabalho para uma grande empresa, e no projeto em que estou atualmente envolvida tenho que colocar em prática um pouco da estatística que você nos ensinou! Você pode ajudar? A resposta foi “sim”, e em conjunto eles formularam o seguinte teste:

- Atribua, aleatoriamente, as 60 contratações mais recentes a dois programas de treinamento. Metade iria para o programa de treinamento preexistente, e metade seria treinada com o uso do novo programa.
- Ao final do primeiro mês, compare a média aritmética do resultado correspondente aos 30 empregados no novo programa de treinamento em relação à média aritmética do resultado para os 30 empregados do programa de treinamento preexistente.

Ela escutava, à medida que seu professor ia explicando: “O que você está tentando provar é que a média aritmética do resultado do novo programa de treinamento é mais alta do que a média aritmética do resultado do programa atual. Você pode formular a hipótese nula de que as médias aritméticas são iguais e verificar se consegue rejeitá-la em favor da hipótese alternativa de que a média aritmética do resultado do novo programa é mais alta.

“Ou, como você costumava dizer, ‘se o valor-*p* for pequeno, H_0 vai para o dreno!’ — sim, eu me lembro!”, respondeu ela. O professor deu uma boa risada e disse: “Sim, está correto. E se você conseguir rejeitar H_0, você terá a comprovação para apresentar a seu chefe.” Ela lhe agradeceu a ajuda e retornou ao trabalho, com a recém-conquistada confiança de que seria capaz de aplicar com sucesso o teste *t* que compara as médias aritméticas de duas populações independentes.

Problemas para a Seção 10.1

APRENDENDO O BÁSICO

10.1 Se você tem amostras de tamanho $n_1 = 12$ e $n_2 = 15$, ao realizar o teste *t* de variância agrupada, quantos graus de liberdade você tem?

10.2 Suponha que você tem uma amostra de tamanho $n_1 = 8$, com a média aritmética da amostra $\bar{X}_1 = 42$, e um desvio-padrão de amostra $S_1 = 4$, e que você tem uma amostra independente, de tamanho $n_2 = 15$, extraída de uma outra população, com uma média aritmética de amostra $\bar{X}_2 = 34$ e um desvio-padrão de amostra $S_2 = 5$.

a. Qual é o valor da estatística do teste t_{ESTAT} de variância agrupada, para testar H_0: $\mu_1 = \mu_2$?

b. Quando encontramos o valor crítico, $t_{\alpha/2}$, quantos graus de liberdade estão disponíveis?

c. Utilizando o nível de significância $\alpha = 0{,}01$, qual é o valor crítico para um teste unicaudal correspondente à hipótese H_0: $\mu_1 \leq \mu_2$, contra a hipótese alternativa, H_1: $\mu_1 > \mu_2$?

d. Qual é sua decisão estatística?

10.3 Que premissas em relação às duas populações são necessárias no Problema 10.2?

10.4 Com referência ao Problema 10.2, construa uma estimativa para o intervalo de confiança de 95% para a diferença entre μ_1 e μ_2 nas populações.

10.5 Com referência ao Problema 10.2, se $n_1 = 5$ e $n_2 = 4$, quantos graus de liberdade você tem?

10.6 Com referência ao Problema 10.2, se $n_1 = 5$ e $n_2 = 4$, no nível de significância de 0,01, há evidencia de que $\mu_1 > \mu_2$?

APLICANDO OS CONCEITOS

10.7 De acordo com um estudo recente, ao realizar compras virtuais de artigos de luxo, os homens gastam uma média aritmética de \$2.401, enquanto as mulheres gastam uma média aritmética de \$1.527 (dados extraídos de R. A, Smith, “Fashion Online: Retailers Tackle the Gender Gap”, *The Wall Street Journal*, 13 de março de 2008, pp. D1, D10). Suponha que o estudo tenha se baseado em uma amostra de 600 homens e 700 mulheres, e o desvio-padrão da quantia gasta foi $1.200 para os homens e $1.000 para as mulheres.

a. Declare a hipótese nula e a hipótese alternativa, se você deseja determinar se a média aritmética para a quantia gasta é mais alta para os homens do que para as mulheres.
b. No contexto desse estudo, qual é o significado do erro do Tipo I?
c. No contexto desse estudo, qual é o significado do erro do Tipo II?
d. No nível de significância de 0,01, existem evidências de que a média aritmética da quantia gasta é mais alta para os homens do que para as mulheres?

10.8 Um estudo recente ("Snack Ads Spur Children to Eat More", *The New York Times*, 20 de julho de 2009, p. B3) descobriu que crianças que assistiam a um desenho animado com propaganda de alimentos comiam, em média, 28,5 gramas de biscoitos Goldfish em comparação com uma média de 19,7 gramas de biscoitos Goldfish para crianças que assistiam a um desenho animado sem propaganda de alimentos. Embora houvesse 118 crianças no estudo, nem o tamanho da amostra em cada um dos grupos nem o desvio-padrão foram relatados. Suponha que houvesse 59 crianças em cada um dos grupos e que o desvio-padrão da amostra para as crianças que assistiam ao comercial de alimentos era 8,6 gramas e o desvio-padrão da amostra para as crianças que não assistiam ao comercial de alimentos era 7,9 gramas.

a. Pressupondo que as variâncias das populações sejam iguais e $\alpha = 0{,}05$, existem evidências de que a média aritmética da quantidade de biscoitos Goldfish ingeridos tenha sido significativamente maior para as crianças que assistiram ao comercial de alimentos?
b. Pressupondo que as variâncias das populações sejam iguais, construa uma estimativa de intervalo de confiança de 95% para a diferença entre a média aritmética da quantidade de biscoitos Goldfish ingeridos pelas crianças que assistiram ao comercial de alimentos e a média aritmética para as crianças que não assistiram ao comercial de alimentos.
c. Compare os resultados de (a) e (b) e argumente.

10.9 Um problema relacionado a uma linha telefônica que impede que o cliente receba ou faça ligações é desagradável tanto para o cliente como para a companhia telefônica. O arquivo **Telefone** contém amostras relativas a 20 problemas relatados, a duas diferentes centrais telefônicas de uma mesma companhia telefônica e o tempo necessário para solucionar esses problemas (em minutos) nas linhas dos clientes:

Tempo para Solucionar Problemas na Central Telefônica I (minutos)

1,48	1,75	0,78	2,85	0,52	1,60	4,15	3,97	1,48	3,10
1,02	0,53	0,93	1,60	0,80	1,05	6,32	3,93	5,45	0,97

Tempo para Solucionar Problemas na Central Telefônica II (minutos)

7,55	3,75	0,10	1,10	0,60	0,52	3,30	2,10	0,58	4,02
3,75	0,65	1,92	0,60	1,53	4,23	0,08	1,48	1,65	0,72

a. Supondo que as variâncias das populações de ambas as centrais sejam iguais, existem evidências de uma diferença entre as médias aritméticas dos tempos de espera para as duas centrais? (Utilize $\alpha = 0{,}05$.)
b. Determine o valor-p em (a) e interprete o seu significado.
c. Que outro pressuposto é necessário em (a)?
d. Supondo que as variâncias das populações de ambas as centrais sejam iguais, construa e interprete uma estimativa para o intervalo de confiança de 95% para a diferença entre as médias aritméticas das populações nas duas centrais telefônicas.

AUTO-AVALIAÇÃO **10.10** A Computer Anxiety Rating Scale (CARS) mede o nível de ansiedade dos indivíduos causada por computadores, em uma escala que vai de 20 (nenhuma ansiedade) até 100 (maior nível de ansiedade). Pesquisadores da Universidade de Miami administraram a CARS a 172 alunos de escolas de negócios. Um dos objetivos do estudo era determinar se existe uma diferença no nível de ansiedade sofrido por alunos de escolas de negócios dos sexos feminino e masculino. Eles descobriram o seguinte:

	Homens	Mulheres
$\bar{X}$	40,26	36,85
S	13,35	9,42
n	100	72

Fonte: *Dados extraídos de T. Broome e D. Havelka, "Determination of Computer Anxiety in Business Students"*, The Review of Business Information Systems, *Spring 2002, 6(2), pp. 9-16.*

a. No nível de significância de 0,05, existem evidências de uma diferença na média aritmética do nível de ansiedade sofrido por alunos de escolas de negócios dos sexos feminino e masculino?
b. Determine o valor-p em (a) e interprete o seu significado.
c. Que pressupostos você precisa adotar sobre as duas populações para justificar o uso do teste t?

10.11 Câmeras fotográficas digitais dominam o mercado de máquinas fotográficas do tipo focalizar e disparar. Uma das características importantes de uma câmera é o tempo de vida útil da bateria, medido pelo número de fotos tiradas até que a bateria precise ser recarregada. O arquivo **CâmerasDigitais** contém o tempo de vida útil de 29 câmeras subcompactas e 16 câmeras compactas (dados extraídos de "Digital Cameras", *Consumer Reports*, julho de 2009, pp. 28-29).

a. Supondo que as variâncias das populações de ambos os tipos de câmeras digitais são iguais, existem evidências de uma diferença na média aritmética da vida útil da bateria entre os dois tipos de câmeras digitais ($\alpha = 0{,}05$)?
b. Determine o valor-p em (a) e interprete o seu significado.
c. Supondo que as variâncias das populações de ambos os tipos de câmeras digitais são iguais, construa e interprete uma estimativa para o intervalo de confiança de 95% para a diferença entre as médias aritméticas das populações de vida útil de baterias para os dois tipos de câmeras digitais.

10.12 Um banco com uma agência localizada em um bairro comercial de uma cidade desenvolveu um processo de aperfeiçoamento para atendimento aos clientes no horário de pico do almoço, do meio-dia até as 13 horas. O tempo de espera em minutos (definido operacionalmente como o tempo decorrido desde o momento em que o cliente entra na fila até que seja atendido no caixa) precisa ser diminuído, de modo a aumentar a satisfação dos clientes. É selecionada uma amostra aleatória de 15 clientes, e os resultados se apresentam como se segue (e armazenados em **Banco1**):

4,21	5,55	3,02	5,13	4,77	2,34	3,54	3,20
4,50	6,10	0,38	5,12	6,46	6,19	3,79	

Suponha que uma outra agência bancária, localizada em um bairro residencial, esteja também preocupada com o horário de pico do almoço, do meio-dia até as 13 horas. É selecionada uma amostra aleatória de 15 clientes, e os resultados se apresentam como se segue (e armazenados em **Banco2**):

9,66	5,90	8,02	5,79	8,73	3,82	8,01	8,35
10,49	6,68	5,64	4,08	6,17	9,91	5,47	

a. Pressupondo que as variâncias das populações de ambas as agências bancárias sejam iguais, existem evidências de uma diferença nas médias aritméticas dos tempos de espera entre as duas agências? (Utilize $\alpha = 0{,}05$.)
b. Determine o valor-p em (a) e interprete o seu significado.
c. Além de variâncias iguais, que outro pressuposto é necessário em (a)?
d. Construa e interprete uma estimativa para o intervalo de confiança de 95% para a diferença entre as médias aritméticas das populações nas duas agências.

10.13 Repita o Problema 10.12(a) considerando que as variâncias das populações nas duas agências não são iguais. Compare os resultados com os do Problema 10.12(a).

10.14 Na gravação em entalhe, um desenho, ou figura, é gravado na superfície de um metal resistente ou de uma pedra. Suponha que seja projetado um experimento com o objetivo de comparar diferenças em termos da média aritmética da solidez da superfície das placas de aço utilizadas na gravação em entalhe (medida em números de identação), com base em duas diferentes condições de superfície — não tratada e ligeiramente polida com folha de esmeril. No experimento, 40 placas de aço são distribuídas aleatoriamente — 20 que não são tratadas e 20 que são ligeiramente polidas. Os resultados do experimento (em **Entalhe**) são os seguintes:

Não Tratadas		Tratadas	
164,368	177,135	158,239	150,226
159,018	163,903	138,216	155,620
153,871	167,802	168,006	151,233
165,096	160,818	149,654	158,653
157,184	167,433	145,456	151,204
154,496	163,538	168,178	150,869
160,920	164,525	154,321	161,657
164,917	171,230	162,763	157,016
169,091	174,964	161,020	156,670
175,276	166,311	167,706	147,920

a. Supondo que as variâncias das populações para ambas as condições são iguais, existem evidências de uma diferença nas médias aritméticas correspondentes à solidez da superfície entre placas de aço não tratadas e placas de aço tratadas? (Utilize $\alpha = 0{,}05$.)
b. Determine o valor-p em (a) e interprete o seu significado.
c. Além da igualdade entre as variâncias, qual outro pressuposto se faz necessário em (a)?
d. Construa e interprete uma estimativa para o intervalo de confiança de 95% da diferença entre as médias aritméticas das populações de placas de aço tratadas e placas de aço não tratadas.

10.15 Repita o Problema 10.14(a), pressupondo que as variâncias das populações extraídas de placas de aço não tratadas e de placas de aço tratadas não são iguais. Compare os resultados com os do Problema 10.14(a).

10.16 Crianças pequenas usam telefones celulares? Aparentemente sim, de acordo com um estudo recente (dados extraídos de A. Ross, "Message to Santa: Kids Want a Phone", *Palm Beach Post*, 16 de dezembro de 2008, pp. 1A, 4A), que afirma que usuários de telefones celulares abaixo de 12 anos de idade fazem uma média de 137 chamadas telefônicas por mês, comparadas a 231 chamadas por mês de usuários de telefones celulares de 13 a 17 anos de idade. Não foram informados os tamanhos das amostras. Suponha que os resultados tenham se baseado em amostras com 50 usuários de telefones celulares em cada um dos grupos e que o desvio-padrão da amostra para os usuários de telefones celulares abaixo de 12 anos de idade tenha sido de 51,7 chamadas por mês, enquanto o desvio-padrão da amostra para os usuários de telefones celulares com 13 a 17 anos de idade foi de 67,6 chamadas por mês.

a. Supondo que as variâncias das populações de usuários de telefones celulares sejam iguais, existem evidências de uma diferença na média aritmética do uso de telefones celulares para usuários abaixo de 12 anos de idade e usuários de telefones celulares com 13 a 17 anos de idade? (Utilize um nível de significância de 0,05.)
b. Além da igualdade entre as variâncias, qual outro pressuposto se faz necessário em (a)?

10.17 Avaliação não destrutiva é um método utilizado para descrever as propriedades de componentes ou materiais sem que seja causada alteração física permanente nas unidades. Isso inclui a determinação de propriedades de materiais e a classificação de falhas em termos de tamanho, formato, tipo e localização. Esse método é bastante efetivo para detectar falhas na superfície e para caracterizar propriedades da superfície de materiais que são condutores de eletricidade. Foram coletados dados que classificaram cada um dos componentes como tendo ou não uma falha, com base em inspeções manuais ou por julgamento do operador, e os dados informaram também os tamanhos de fendas no material. Os componentes classificados como isentos de falhas possuíam uma média aritmética menor para o tamanho de fenda do que os componentes classificados como portadores de falhas? Os resultados, em termos do tamanho da fenda (em polegadas), estão no arquivo **Fenda** (dados extraídos de B. D. Olin e W. Q. Meeker, "Applications of Statistical Methods to Nondestructive Evaluation", *Technometrics*, 38, 1996, p. 101).

a. Supondo que as variâncias das populações sejam iguais, existem evidências de que a média aritmética dos tamanhos de fendas é menor para os materiais considerados isentos de falhas do que para espécimes considerados portadores de falhas? (Utilize $\alpha = 0{,}05$.)
a. Repita (a), pressupondo que as variâncias das populações não são iguais.
a. Compare os resultados de (a) e (b).

10.2 Comparando as Médias Aritméticas de Duas Populações Inter-Relacionadas

Os procedimentos de testes de hipóteses examinados na Seção 10.1 possibilitam que você faça comparações e examine diferenças nas médias aritméticas de duas populações *independentes*. Nesta seção, você aprenderá um procedimento para analisar a diferença entre as médias aritméticas de duas populações quando coleta dados de amostras extraídas de populações que são relacionadas uma à outra — ou seja, quando os resultados da primeira população *não* são independentes dos resultados da segunda população.

Existem dois casos que envolvem dados correlacionados entre populações. Você pode adotar medições repetidas a partir do mesmo conjunto de itens ou indivíduos, ou você pode combinar os itens ou indivíduos de acordo com alguma característica. Em qualquer uma dessas situações, você está interessado na *diferença entre os dois valores inter-relacionados*, e não nos *valores individuais* propriamente ditos.

Quando adota **medições repetidas** nos mesmos itens ou indivíduos, você supõe que os mesmos itens ou indivíduos se comportarão de modo semelhante se forem tratados de modo semelhante. Seu objetivo é demonstrar que quaisquer diferenças entre duas medições dos mesmos itens ou indivíduos são decorrentes de diferentes condições de tratamento. Por exemplo, ao realizar um experimento relacionado a um teste de degustação comparando duas bebidas, você pode utilizar cada uma das pessoas na amostra como seu próprio controle, de modo que você possa ter *medições repetidas* em relação ao mesmo indivíduo.

A segunda situação que envolve dados correlacionados entre populações é quando você tem **amostras combinadas**. Nesse caso, os itens ou indivíduos são colocados em pares, conjuntamente, de acordo com alguma característica de interesse. Por exemplo, ao fazer um teste de comercialização de um produto utilizando duas estratégias diferentes de publicidade, uma amostra de testes de mercado pode ser *combinada* com base no tamanho da população do teste de mercado e/ou variáveis demográficas. Ao levar em conta as diferenças em relação ao tamanho da população do teste de mercado e/ou variáveis demográficas, você passa a estar mais bem capacitado para mensurar os efeitos decorrentes de duas diferentes campanhas de publicidade.

Independentemente de você ter amostras combinadas ou medições repetidas, o objetivo é estudar as diferenças entre duas medições reduzindo o efeito da variabilidade que é decorrente dos itens ou indivíduos propriamente ditos. A Tabela 10.3 mostra as diferenças nos valores individuais para duas populações inter-relacionadas. Para ler essa tabela, sejam $X_{11}, X_{12},..., X_{1n}$ os n valores extraídos de uma amostra. Além disso, sejam $X_{21}, X_{22},..., X_{2n}$ os n valores combinados extraídos de uma segunda amostra, ou as n medições repetidas correspondentes extraídas da amostra inicial. Depois disso, $D_1, D_2,..., D_n$ representarão o conjunto correspondente de *n resultados* de diferenças, de modo tal que:

$$D_1 = X_{11} - X_{21}, D_2 = X_{12} - X_{22}, \ldots, \text{e } D_n = X_{1n} - X_{2n}.$$

Para testar a diferença entre as médias aritméticas de duas populações inter-relacionadas, você trata os resultados das diferenças, cada D_i, como valores extraídos de uma única amostra.

TABELA 10.3
Determinando a Diferença entre Duas Amostras Inter-Relacionadas

	Amostra		
Valor	**1**	**2**	**Diferença**
1	X_{11}	X_{21}	$D_1 = X_{11} - X_{21}$
2	X_{12}	X_{22}	$D_2 = X_{12} - X_{22}$
.	.	.	.
.	.	.	.
.	.	.	.
i	X_{1i}	X_{2i}	$D_i = X_{1i} - X_{2i}$
.	.	.	.
.	.	.	.
.	.	.	.
n	X_{1n}	X_{2n}	$D_n = X_{1n} - X_{2n}$

Teste t em Pares

Se você adota o pressuposto de que os resultados das diferenças foram selecionados de maneira aleatória e independente a partir de uma população normalmente distribuída, você pode utilizar o **teste t em pares para a média aritmética da diferença** entre populações relacionadas a fim de determinar se existe alguma diferença significativa entre as médias aritméticas das populações. De maneira análoga ao teste t para uma amostra, desenvolvido na Seção 9.2 [veja a Equação (9.2) na Seção 9.2], a estatística do teste t aqui desenvolvida segue a distribuição t, com $n - 1$ graus de liberdade. Embora você necessariamente deva pressupor que a população é distribuída nos moldes de uma distribuição normal, e desde que o tamanho da amostra não seja demasiadamente pequeno e a população não seja extremamente assimétrica, você pode utilizar o teste t em pares.

Para testar a hipótese nula de que não existe nenhuma diferença entre as médias aritméticas de duas populações inter-relacionadas:

$$H_0: \mu_D = 0 \text{ (em que } \mu_D = \mu_1 - \mu_2)$$

contra a hipótese alternativa de que as médias aritméticas não são iguais:

$$H_1: \mu_D \neq 0$$

você calcula a estatística do teste t_{ESTAT} utilizando a Equação (10.3).

TESTE t EM PARES PARA A MÉDIA ARITMÉTICA DA DIFERENÇA

$$t_{ESTAT} = \frac{\overline{D} - \mu_D}{\frac{S_D}{\sqrt{n}}} \quad \textbf{(10.3)}$$

em que

μ_D = média aritmética da diferença formulada na hipótese

$$\overline{D} = \frac{\sum_{i=1}^{n} D_i}{n}$$

$$S_D = \sqrt{\frac{\sum_{i=1}^{n}(D_i - \overline{D})^2}{n - 1}}$$

A estatística do teste t_{ESTAT} segue uma distribuição t, com $n - 1$ graus de liberdade.

Para um teste bicaudal com um determinado nível de significância, α, você rejeita a hipótese nula caso a estatística do teste t_{ESTAT} calculada seja maior do que o valor crítico da cauda superior, $t_{\alpha/2}$, a partir da distribuição t, ou caso a estatística do teste t_{ESTAT} calculada seja menor do que o valor crítico da cauda inferior, $-t_{\alpha/2}$, da distribuição t. A regra de decisão é

Rejeitar H_0 se $t_{ESTAT} > t_{\alpha/2}$

ou se $t_{STAT} < -t_{\alpha/2}$;

caso contrário, não rejeitar H_0.

O exemplo a seguir ilustra o uso do teste t para a média aritmética da diferença. A Automobile Association of America (AAA) conduziu um teste de milhagem para comparar a milhagem de combustível em testes de direção da vida real realizados por membros da AAA e os resultados de testes de direção realizados de acordo com os padrões estabelecidos pelo governo (J. Healey, "Fuel Economy Calculations to Be Altered, *USA Today*, 11 de janeiro de 2006, p. 1B).

Qual é a melhor maneira de projetar um experimento para comparar a milhagem de combustível de testes de direção da vida real realizados por membros da AAA e os resultados de testes de direção realizados em conformidade com os padrões estabelecidos pelo governo? Uma das abordagens seria tomar duas amostras independentes e, então, utilizar os testes de hipóteses discutidos na Seção 10.1. Nessa abordagem, você utilizaria um conjunto de automóveis para testar a direção da vida real feita pelos membros da AAA. Em seguida, você utilizaria um segundo conjunto de automóveis diferentes para testar os resultados de testes de direção realizados em conformidade com os padrões estabelecidos pelo governo.

Entretanto, uma vez que o primeiro conjunto de automóveis para testar a direção na vida real, feita pelos membros da AAA, pode resultar em uma milhagem de combustível mais baixa ou mais alta do que o segundo conjunto de automóveis, essa não é uma boa abordagem. Uma abordagem melhor seria utilizar o experimento com medições repetidas. Nesse tipo de experimento, você utiliza um único conjunto de automóveis. Para cada um dos automóveis, você conduz um teste de direção da vida real feito por um membro da AAA e um teste de direção realizado em conformidade com os padrões estabelecidos pelo governo. Ao mensurar as duas milhagens de combustível para os mesmos automóveis, você consegue reduzir a variabilidade em termos das milhagens de combustível comparadas com o que ocorreria caso você utilizasse dois conjuntos independentes de automóveis. Essa abordagem se concentra nas diferenças entre testes de direção da vida real realizados por um membro da AAA e o teste de direção realizado em conformidade com os padrões estabelecidos pelo governo.

Você decide coletar os dados por meio da condução de um experimento com uma amostra de $n = 9$ automóveis. Você organiza e armazena os dados em **AAAMilhagem**. A Tabela 10.4 mostra os resultados.

TABELA 10.4
Medições Repetidas para a Milhagem de Combustível para Testes de Direção da Vida Real Realizados por Membros da AAA e para Testes de Direção Realizados em Conformidade com os Padrões do Governo

Modelo	Membros	Governo	Diferença (D_i)
Ford F-150 2005	14,3	16,8	−2,5
Chevrolet Silverado 2005	15,0	17,8	−2,8
Honda Accord LX 2002	27,8	26,2	+1,6
Honda Civic 2002	27,9	33,2	−5,3
Honda Civic Hybrid 2004	48,8	47,6	+1,2
Ford Explorer 2002	16,8	18,3	−1,5
Toyota Camry 2005	23,7	28,5	−4,8
Toyota Corolla 2003	32,8	33,1	−0,3
Toyota Prius 2005	37,3	44,0	−6,7
			−21,1

Seu objetivo é determinar se existe alguma diferença entre a média aritmética da milhagem de combustível entre testes de direção da vida real realizados por membros da AAA e testes de direção na cidade e em autoestradas feitos conforme os padrões do governo. Em outras palavras, existem evidências de que a média aritmética da milhagem de combustível seja diferente em relação aos dois tipos de testes de direção? Consequentemente, a hipótese nula e a hipótese alternativa são

H_0: $\mu_D = 0$. (Não existe nenhuma diferença entre a média aritmética da milhagem de combustível para testes de direção da vida real realizados por membros da AAA e a média aritmética de testes de direção realizados em conformidade com os padrões do governo.)

H_1: $\mu_D \neq 0$. (Existe uma diferença entre a média aritmética da milhagem de combustível para testes de direção da vida real realizados por membros da AAA e a média aritmética de testes de direção realizados em conformidade com os padrões do governo.)

Ao escolher um nível de significância de $\alpha = 0{,}05$ e pressupor que as diferenças são distribuídas nos moldes de uma distribuição normal, você utiliza o teste t em pares [(Equação 10.3)]. Para uma amostra com $n = 9$ automóveis, existem $n - 1 = 8$ graus de liberdade. Utilizando a Tabela E.3, a regra de decisão é

$$\text{Rejeitar } H_0 \text{ se } t_{ESTAT} > t_{0,025} = 2{,}3060$$

$$\text{ou se } t_{ESTAT} < -t_{0,025} = -2{,}3060;$$

$$\text{caso contrário, não rejeitar } H_0.$$

Para as $n = 9$ diferenças (veja a Tabela 10.4), a média aritmética das diferenças na amostra é

$$\overline{D} = \frac{\sum_{i=1}^{n} D_i}{n} = \frac{-21{,}1}{9} = -2{,}3444$$

e

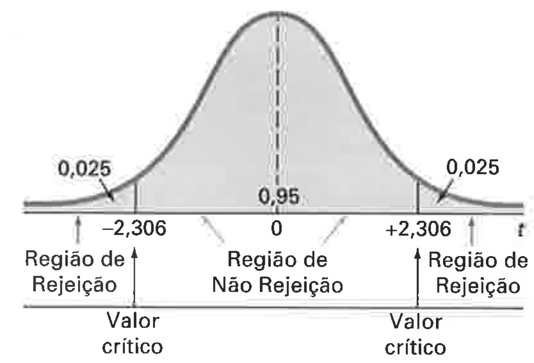
$$S_D = \sqrt{\frac{\sum_{i=1}^{n}(D_i - \overline{D})^2}{n-1}} = 2{,}8936$$

Com base na Equação (10.3),

$$t_{ESTAT} = \frac{\overline{D} - \mu_D}{\frac{S_D}{\sqrt{n}}} = \frac{-2{,}3444 - 0}{\frac{2{,}8936}{\sqrt{9}}} = -2{,}4307$$

Uma vez que $t_{ESTAT} = -2{,}4307$ é menor do que $-2{,}3060$, você rejeita a hipótese nula, H_0 (veja a Figura 10.7). Existem evidências de uma diferença na média aritmética da milhagem de combustível entre os testes de direção da vida real realizados por membros da AAA e os testes de direção realizados em conformidade com os padrões do governo. Os testes de direção da vida real resultam em uma média aritmética mais baixa para a milhagem de combustível.

FIGURA 10.7
Teste t bicaudal, em pares, no nível de significância de 0,05, com 8 graus de liberdade



A Figura 10.8 apresenta uma planilha de solução para esse exemplo que compara a estatística do teste e o valor-p. Uma vez que o valor-$p = 0{,}0412 < \alpha = 0{,}05$, você rejeita H_0. O valor-p indica que, se os dois tipos de testes de direção apresentarem a mesma média aritmética em termos de milhagem de combustível, a probabilidade de que um determinado tipo de teste de direção venha a ter uma média aritmética de amostra igual a 2,3444 milhas por galão a menos do que o outro tipo é de 0,0412. Como essa probabilidade é menor do que $\alpha = 0{,}05$, você conclui que a hipótese alternativa é verdadeira.

FIGURA 10.8
Planilha para o teste t em pares, para os dados relacionados à milhagem de automóveis

	A	B	
1	Teste *t* em Pares		
2			
3	**Dados**		
4	Hipótese da Diferença de Média	0	
5	Nível de Significância	0,05	
6			
7	Cálculos Intermediários		
8	Tamanho da Amostra	9	=CONT.NÚM(PtCalcs!$A:$A)
9	D-Barra	-2,3444	=MÉDIA(PtCalcs!$C:$C)
10	Graus de Liberdade	8	=B8 - 1
11	S_D	2,8936	=RAIZ(SOMA(PtCalcs!$D:$D)/B10)
12	Erro-padrão	0,9645	=B11/RAIZ(B8)
13	Estatística do Teste *t*	-2,4307	=(B9 - B4)/B12
14			
15	**Teste Bicaudal**		
16	Valor Crítico Inferior	-2,3060	=-INVT(B5, B10)
17	Valor Crítico Superior	2,3060	=INVT(B5, B10)
18	Valor-*p*	0,0412	=DISTT(ABS(B13), B10, 2)
19	Rejeitar a hipótese nula		=SE(B28 < B5, D27, D28)

Não ilustrado
Célula D27: Rejeitar a hipótese nula
Célula D28: Não rejeitar a hipótese nula

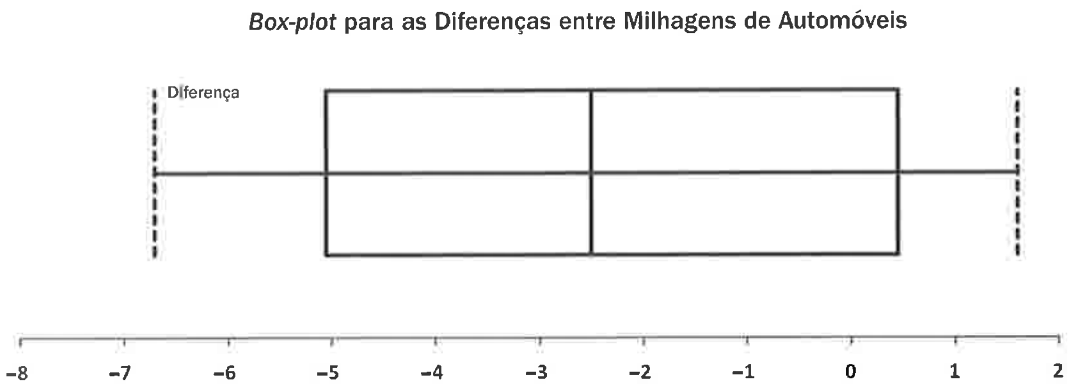
	A	B	C	D	E	F	G
1	Membros	Governo	D_i	(D_i - DBarra^2)		DBarra	-2,3444
2	14,3	16,8	-2,5	0,0242			
3	15	17,8	-2,8	0,2075			
4	27,8	26,2	1,6	15,5586			
5	27,9	33,2	-5,3	8,7353			
6	48,8	47,6	1,2	12,5631			
7	16,8	18,3	-1,5	0,7131			
8	23,7	28,5	-4,8	6,0298			
9	32,8	33,1	-0,3	4,1798			
10	37,3	44	-6,7	18,9709			

*A Figura 10.8 exibe a **planilha CÁLCULO** (à esquerda) e a **planilha PtCalcs** (à direita) da **pasta de trabalho T em Pares**. A planilha CÁLCULO utiliza valores calculados em PtCalcs.*

Crie essas planilhas utilizando as instruções da Seção GE10.2.

Para avaliar a validade do pressuposto da normalidade, você constrói um box-plot das diferenças, conforme ilustrado na Figura 10.9.

FIGURA 10.9
Box-plot para os dados sobre a milhagem de automóveis



O box-plot da Figura 10.9 apresenta uma simetria aproximada. Assim os dados não contradizem muito o pressuposto subjacente da normalidade. Caso um box-plot, um histograma ou gráfico da probabilidade normal revelem que o pressuposto de normalidade subjacente na população foi seriamente violado, então o teste t não é apropriado, sobretudo se o tamanho da amostra for pequeno. Se você acreditar que o teste t não é apropriado, pode utilizar um procedimento *não paramétrico* que não exija o pressuposto restritivo de normalidade subjacente (veja as referências 1 e 2) ou realizar uma transformação de dados (veja a referência 5) e, a partir de então, verificar novamente os pressupostos que determinam se você deve ou não utilizar o teste t.

EXEMPLO 10.2

Teste T em Pares para os Tempos de Entrega de Pizzas

Lembre-se, com base no Exemplo 10.1, na Seção 10.1, no qual uma pizzaria local, situada do outro lado da rua do *campus* universitário onde você estuda, diz em sua propaganda que entrega pizzas nos dormitórios mais rápido do que a filial local de uma cadeia nacional de pizzarias. Para determinar se essa propaganda é válida, você e alguns amigos decidiram encomendar 10 pizzas da pizzaria local e 10 pizzas da filial da cadeia nacional. Na realidade, cada vez que você encomendou uma pizza da pizzaria local, seus amigos, ao mesmo tempo, encomendaram uma pizza da filial da cadeia nacional. Consequentemente, você tem amostras combinadas. Para cada uma das 10 vezes em que foram encomendadas pizzas, você tem uma medição para a pizzaria local e uma para a cadeia nacional de pizzarias. No nível de significância de 0,05, a média aritmética do tempo de entrega da pizzaria local é menor do que a média aritmética do tempo de entrega da filial da cadeia nacional de pizzarias?

SOLUÇÃO Utilize o teste t em pares para analisar os dados na Tabela 10.5 (armazenados em **TempoPizza**). A Figura 10.10 mostra uma planilha com os resultados do teste t em pares para os dados sobre entrega de pizzas.

TABELA 10.5
Tempos de Entrega da Pizzaria Local e da Pizzaria da Cadeia Nacional

Tempo	Local	Cadeia Nacional	Diferença
1	16,8	22,0	−5,2
2	11,7	15,2	−3,5
3	15,6	18,7	−3,1
4	16,7	15,6	1,1
5	17,5	20,8	−3,3
6	18,1	19,5	−1,4
7	14,1	17,0	−2,9
8	21,8	19,5	2,3
9	13,9	16,5	−2,6
10	20,8	24,0	−3,2
			−21,8

FIGURA 10.10
Planilha para o teste t em pares para os dados sobre entrega de pizzas

A Figura 10.10 contém uma planilha baseada na planilha CÁLCULO_INFERIOR da pasta de trabalho T em Pares. Crie planilhas para testes unicaudais utilizando as instruções da Seção GE10.2.

	A	B	
1	Teste *t* em Pares		
2			
3	**Dados**		
4	Hipótese da Diferença de Média	0	
5	Nível de Significância	0,05	
6			
7	Cálculos Intermediários		
8	Tamanho da Amostra	10	=CONT.NÚM(PtCalcs!$A:$A)
9	D-Barra	-2,1800	=MÉDIA(PtCalcs!$C:$C)
10	Graus de Liberdade	9	=B8 - 1
11	S_D	2,2641	=RAIZ(SOMA(PtCalcs!$D:$D)/B10)
12	Erro-padrão	0,7160	=B11/RAIZ(B8)
13	Estatística do Teste *t*	-3,0448	=(B9 - B4)/B12
14			
15	**Teste da Cauda Inferior**		
16	Valor Crítico Inferior	-1,8331	=-INVT(2 * B5, B10)
17	Valor-*p*	0,0070	=SE(B13, E22, E24)
18	Rejeitar a hipótese nula		=SE(B17 < B5, D27, D28)

Não ilustrado
Célula E23: =DISTT(ABS(B13), B10, 1)
Célula E24: =1 - E23
Célula D27: Rejeitar a hipótese nula
Célula D28: Não rejeitar a hipótese nula
e a planilha PtCalcs para este problema

A hipótese nula e a hipótese alternativa são:

H_0: $\mu_D \geq 0$ (A média aritmética do tempo de entrega da pizzaria local é maior ou igual à média aritmética do tempo de entrega da filial da cadeia nacional de pizzarias.)

H_1: $\mu_D < 0$ (A média aritmética do tempo de entrega da pizzaria local é menor do que a média aritmética do tempo de entrega da filial da cadeia nacional de pizzarias.)

Ao escolher um nível de significância de $\alpha = 0{,}05$, e pressupondo que as diferenças são distribuídas segundo uma distribuição normal, você utiliza o teste t em pares [Equação (10.3)]. Para uma amostra de $n = 10$ tempos de entrega, existem $n - 1 = 9$ graus de liberdade. Utilizando a Tabela E.3, a regra de decisão seria

$$\text{Rejeitar } H_0 \text{ se } t_{ESTAT} < -t_{0,05} = -1{,}8331;$$

$$\text{caso contrário, não rejeitar } H_0.$$

Para ilustrar os cálculos, para $n = 10$ diferenças (veja a Tabela 10.5), a média aritmética das diferenças na amostra é

$$\overline{D} = \frac{\sum_{i=1}^{n} D_i}{n} = \frac{-21{,}8}{10} = -2{,}18$$

e o desvio-padrão das diferenças na amostra é

$$S_D = \sqrt{\frac{\sum_{i=1}^{n}(D_i - \overline{D})^2}{n - 1}} = 2{,}2641$$

Com base na Equação (10.3),

$$t_{ESTAT} = \frac{\overline{D} - \mu_D}{\frac{S_D}{\sqrt{n}}} = \frac{-2{,}18 - 0}{\frac{2{,}2641}{\sqrt{10}}} = -3{,}0448$$

Como $t_{ESTAT} = -3{,}0448$ é menor do que $-1{,}8331$, você rejeita a hipótese nula, H_0 (o valor-p é $0{,}0070 < 0{,}05$). Existem evidências de que a média aritmética do tempo de entrega é mais baixa para a pizzaria local do que para a filial da cadeia nacional de pizzarias.

Essa conclusão é diferente daquela a que você chegou no Exemplo 10.1, quando utilizou o teste t de variância agrupada para esses mesmos dados. Ao parear os tempos de entrega, você é capaz de se concentrar nas diferenças entre os dois serviços de entrega de pizzas, e não na variabilidade criada pelo fato de pedir pizzas em diferentes horários no dia. O teste t em pares é um procedimento estatístico mais eficaz e que tem maior capacidade de detectar a diferença entre os dois serviços de entrega de pizzas, uma vez que você está controlando a hora do dia em que elas foram pedidas.

Estimativa do Intervalo de Confiança para a Média Aritmética da Diferença

Em vez de, ou além de, testar a diferença entre as médias aritméticas de duas populações inter-relacionadas, você pode utilizar a Equação (10.4) para construir uma estimativa para o intervalo de confiança da média aritmética da diferença.

ESTIMATIVA DO INTERVALO DE CONFIANÇA PARA A MÉDIA ARITMÉTICA DA DIFERENÇA

$$\bar{D} \pm t_{\alpha/2}\frac{S_D}{\sqrt{n}}$$

ou

$$\bar{D} - t_{\alpha/2}\frac{S_D}{\sqrt{n}} \leq \mu_D \leq \bar{D} + t_{\alpha/2}\frac{S_D}{\sqrt{n}} \qquad (10.4)$$

em que $t_{\alpha/2}$ é o valor crítico da distribuição t, com $n - 1$ graus de liberdade, para uma área de $\alpha/2$ na cauda superior.

Retorne ao exemplo que compara a milhagem de combustível gerada por testes de direção com base na vida real e com base nos padrões do governo. Utilizando a Equação (10.4), $\bar{D} = -2{,}3444$, $S_D = 2{,}8936$, $n = 9$ e $t_{\alpha/2} = 2{,}306$ (para 95% de confiança e $n - 1 = 8$ graus de liberdade),

$$-2{,}3444 \pm (2{,}306)\frac{2{,}8936}{\sqrt{9}}$$

$$-2{,}3444 \pm 2{,}2242$$

$$-4{,}5686 \leq \mu_D \leq -0{,}1202$$

Assim, com 95% de confiança, a média aritmética da diferença de milhagem de combustível entre testes de direção da vida real feitos por um membro da AAA e testes de direção na cidade realizados em conformidade com os padrões do governo está entre $-4{,}5686$ e $-0{,}1202$ milhas por galão. Como a estimativa do intervalo contém somente valores menores do que zero, você pode concluir que existe uma diferença nas médias aritméticas das populações. A média aritmética das milhas por galão para os testes de direção da vida real feitos por um membro da AAA é menor do que a média aritmética das milhas por galão para testes de direção realizados em conformidade com os padrões do governo.

Problemas para a Seção 10.2

APRENDENDO O BÁSICO

10.18 Um projeto experimental para um teste t em pares apresenta 20 pares de gêmeos idênticos. Quantos graus de liberdade existem nesse teste t?

10.19 Quinze voluntários são recrutados para participar de um experimento. É feita uma medição (tal como a medição da pressão sanguínea) antes de ser solicitado a cada um dos voluntários que leia uma passagem particularmente desagradável de um livro e depois de cada um deles ter lido essa passagem do livro. Na análise dos dados coletados a partir desse experimento, quantos graus de liberdade existem no teste?

APLICANDO OS CONCEITOS

AUTO-AVALIAÇÃO **10.20** Nove especialistas classificaram duas marcas de café colombiano em um experimento sobre testes de degustação. Uma classificação em uma escala de sete pontos (1 = extremamente desagradável, 7 = extremamente agradável) é atribuída a cada uma de quatro características: gosto, aroma, intensidade e acidez. Os dados a seguir (armazenados em **Café**) apresentam as classificações resumidas — acumuladas ao longo de todas as quatro características.

	MARCA	
ESPECIALISTA	**A**	**B**
C.C.	24	26
S.E.	27	27
E.G.	19	22
B.L.	24	27
C.M.	22	25
C.N.	26	27
G.N.	27	26
R.M.	25	27
P.V.	22	23

a. No nível de significância de 0,05, existem evidências de alguma diferença, em termos da média aritmética para as classificações resumidas, entre as duas marcas?
b. Que pressuposto em relação à distribuição da população é necessário para realizar esse teste?
c. Determine o valor-p em (a) e interprete o seu significado.
d. Construa e interprete uma estimativa para o intervalo de confiança de 95% para a diferença da média aritmética das classificações resumidas para as duas marcas.

10.21 Em ambientes industriais, geralmente existem métodos alternativos para mensurar variáveis de interesse. Os dados no arquivo **Medição** (codificado de modo a manter a confidencialidade) representam medições realizadas em série coletadas por um analista durante o processo de produção e de um laboratório analítico (dados extraídos de M. Leitnaker, "Comparing Measurement Processes: In-line Versus Analytical Measurements", *Quality Engineering*, 13, 2000-2001, pp. 293-298).
a. No nível de significância de 0,05, existem evidências de uma diferença nas médias aritméticas para as medições realizadas em série e para as medições realizadas no laboratório analítico?
b. Qual pressuposto em relação à distribuição da população é necessário para que esse teste seja realizado?
c. Utilize um método gráfico para avaliar a validade do pressuposto adotado em (a).
d. Construa e interprete uma estimativa para o intervalo de confiança de 95% para a diferença nas médias aritméticas das medições realizadas em série e das medições realizadas no laboratório analítico.

10.22 Os estudantes conseguem economizar dinheiro adquirindo seus livros didáticos na Amazon.com? Para investigar essa possibilidade, foi selecionada uma amostra aleatória de 19 livros didáticos utilizados durante o verão de 2009 na Universidade de Miami. Foram registrados os preços desses livros em uma livraria local e na Amazon.com. Os preços dos livros didáticos estão armazenados no arquivo **PreçosLivros**:
a. No nível de significância de 0,01, existem evidências de uma diferença entre a média aritmética dos preços dos livros didáticos na livraria local e na Amazon.com?
b. Qual pressuposto em relação à distribuição da população é necessário para que esse teste seja realizado?
c. Construa uma estimativa para o intervalo de confiança de 99% para a média aritmética da diferença em termos de preço. Interprete esse intervalo.
d. Compare os resultados de (a) e (c).

10.23 Em tempos difíceis da economia, revistas e outros meios de comunicação têm problemas com a venda de espaço para propaganda. Por conseguinte, um dos indicadores de uma economia fraca é a redução no número de páginas de revistas dedicadas a propaganda. O arquivo **PagProp** contém o número de páginas dedicadas a propagandas, em maio de 2008 e maio de 2009, de 12 revistas masculinas (extraído de W. Levith, "Magazine Monitor", *Mediaweek*, 20 de abril de 2009, p. 53)
a. No nível de significância de 0,05, existem evidências de que a média aritmética do número de páginas de revistas dedicadas a propaganda tenha sido mais alta em maio de 2008 do que em maio de 2009?
b. Qual pressuposto em relação à distribuição da população é necessário para que esse teste seja realizado?
c. Utilize um método gráfico para avaliar a validade do pressuposto adotado em (b).
d. Construa e interprete uma estimativa do intervalo de confiança de 95% para a diferença no número médio de páginas dedicadas a propagandas nas revistas masculinas entre maio de 2008 e maio de 2009.

10.24 O mieloma múltiplo, uma neoplasia maligna que se origina da medula óssea, é caracterizado pelo aumento da formação de artérias (angiogênese) na medula óssea, que constitui um fator prognóstico decisivo para a sobrevivência. Um dos métodos de tratamento utilizado no mieloma múltiplo é o transplante de células-tronco, com as células-tronco do próprio paciente. Os dados a seguir (armazenados em **Mieloma**) representam a densidade das microartérias na medula óssea de pacientes que apresentaram resposta completa ao transplante de células-tronco (medida por exames de sangue e de urina). As medições foram feitas imediatamente antes do transplante das células-tronco e no momento em que foi determinada a resposta completa:

Paciente	Antes	Depois
1	158	284
2	189	214
3	202	101
4	353	227
5	416	290
6	426	176
7	441	290

Fonte: *Dados extraídos de S. V. Rajkumar, R. Fonseca, T. E. Witzig, M. A. Gertz e P.R. Greipp, "Bone Marrow Angiogenesis in Patients Achieving Complete Response After Stem Cell Transplantation for Multiple Myeloma"*, Leukemia, *1999, 13, pp. 469-472.*

a. No nível de significância de 0,05, existem evidências de que a média aritmética da densidade das microartérias na medula óssea é mais alta antes do transplante de células-tronco do que depois?
b. Interprete o significado do valor-p em (a).
c. Construa e interprete a estimativa do intervalo de confiança de 95% para a média aritmética da diferença da densidade das microartérias antes e depois do transplante de células-tronco.
d. Qual pressuposto em relação à distribuição da população é necessário para que o teste em (a) seja realizado?

10.25 Ao longo do ano passado, a vice-presidente de recursos humanos de um grande centro médico conduziu uma série de seminários com duração de três meses visando fazer crescer a motivação e o desempenho dos empregados. Para verificar a eficácia dos seminários, ela selecionou uma amostra aleatória de 35 empregados, dos arquivos de pessoal da empresa. Ela coletou os resultados das avaliações de desempenho antes e depois de os empregados terem frequentado o seminário, e armazenou os resultados em pares para todos os 35 empregados no arquivo **Desempenho**. Os resultados sob a forma de planilha para essa amostra são ilustrados a seguir. Expresse suas descobertas e conclusões em um relatório para a vice-presidente de recursos humanos.

	A	B
1	*Diferença*	
2		
3	Média Aritmética	-5,2571
4	Erro-padrão	1,9478
5	Mediana	-5
6	Modo	-10
7	Desvio-padrão	11,5232
8	Variância da Amostra	132,7849
9	Curtose	1,1038
10	Assimetria	0,1103
11	Amplitude	61
12	Mínimo	-34
13	Máximo	27
14	Soma	-184
15	Contagem	35

10.26 Os dados no arquivo **Concreto1** representam a força de compressão, em milhares de libras por polegada quadrada (psi), de 40 amostras de concreto, extraídas dois e sete dias depois da aplicação.

Fonte: *Dados extraídos de O. Carrillo-Gamboa e R. F. Gunst, "Measurement-Error-Model Collinearities"*, Technometrics, *34, 1992, pp. 454-464.*

a. No nível de significância de 0,01, existem evidências de que a média aritmética da resistência é mais baixa em dois dias do que em sete dias?

b. Qual pressuposto em relação à distribuição da população é necessário para que esse teste seja realizado?

c. Encontre o valor-p em (a) e interprete seu significado.

	A	B
1	Teste *t* em Pares	
2		
3	**Dados**	
4	Hipótese da Diferença de Média	0
5	Nível de Significância	0,05
6		
7	Cálculos Intermediários	
8	Tamanho da Amostra	35
9	D-Barra	-5,2571
10	Graus de Liberdade	34
11	S_D	11,5232
12	Erro-padrão	1,9478
13	Estatística do Teste *t*	-2,6990
14		
15	**Teste Bicaudal**	
16	Valor Crítico Inferior	-2,0322
17	Valor Crítico Superior	2,0322
18	Valor-*p*	0,0108
19	Rejeitar a hipótese nula	
20		
21	**Teste da Cauda Inferior**	
22	Valor Crítico Inferior	-1,6909
23	Valor-*p*	0,0054
24	Rejeitar a hipótese nula	

10.3 Comparando as Proporções de Duas Populações Independentes

Frequentemente você precisa realizar comparações e analisar diferenças entre as proporções de duas populações. Você pode realizar um teste para a diferença entre duas proporções selecionadas de amostras independentes utilizando dois métodos diferentes. Esta seção apresenta um procedimento cuja estatística do teste Z_{ESTAT} é aproximada por uma distribuição normal padronizada. Na Seção 12.1, é desenvolvido um procedimento cuja estatística do teste, χ^2_{ESTAT}, é aproximada por uma distribuição qui-quadrada. Como você poderá verificar ao ler essa seção, os resultados desses dois testes são equivalentes.

Teste Z para a Diferença entre Duas Proporções

Ao avaliar as diferenças entre as proporções de duas populações, você pode utilizar um **teste Z para a diferença entre duas proporções**. A estatística para o teste Z_{ESTAT} baseia-se na diferença entre as proporções de duas amostras ($p_1 - p_2$). Essa estatística de teste, apresentada na Equação (10.5), segue, aproximadamente, uma distribuição normal padronizada para tamanhos de amostras suficientemente grandes.

TESTE *Z* PARA A DIFERENÇA ENTRE DUAS PROPORÇÕES

$$Z_{ESTAT} = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \qquad \textbf{(10.5)}$$

com

$$\bar{p} = \frac{X_1 + X_2}{n_1 + n_2} \qquad p_1 = \frac{X_1}{n_1} \qquad p_2 = \frac{X_2}{n_2}$$

em que

p_1 = proporção de itens de interesse na amostra 1

X_1 = número de itens de interesse na amostra 1

n_1 = tamanho da amostra para a amostra 1

π_1 = proporção de itens de interesse na população 1

p_2 = proporção de itens de interesse na amostra 2

X_2 = número de itens de interesse na amostra 2

n_2 = tamanho da amostra para a amostra 2

π_2 = proporção de itens de interesse na população 2

$\bar{p}$ = estimativa agrupada para a proporção de itens de interesse na população

A estatística do teste Z_{ESTAT} segue aproximadamente uma distribuição normal padronizada

Sob a égide da hipótese nula, você pressupõe que as proporções das duas populações são iguais ($\pi_1 = \pi_2$). Como a estimativa agrupada para a proporção da população é baseada na hipótese nula, você combina, ou agrupa, as proporções das duas amostras para calcular $\bar{p}$, uma estimativa geral para a proporção comum da população. Essa estimativa é igual ao número de itens de interesse nas duas amostras combinadas ($X_1 + X_2$) dividido pelo tamanho total da amostra a partir dos dois grupos de amostras ($n_1 + n_2$).

Conforme demonstrado na tabela a seguir, você pode utilizar esse teste Z para a diferença entre proporções de populações a fim de determinar se existe alguma diferença na proporção de itens de interesse nos dois grupos (teste bicaudal) ou se um dos grupos apresenta uma proporção mais elevada de itens de interesse do que o outro grupo (teste unicaudal):

Teste Bicaudal	Teste Unicaudal	Teste Unicaudal
$H_0: \pi_1 = \pi_2$	$H_0: \pi_1 \geq \pi_2$	$H_0: \pi_1 \leq \pi_2$
$H_1: \pi_1 \neq \pi_2$	$H_1: \pi_1 < \pi_2$	$H_1: \pi_1 > \pi_2$

em que

π_1 = proporção de itens de interesse na população 1

π_2 = proporção de itens de interesse na população 2

Para testar a hipótese nula de que não existe nenhuma diferença entre as proporções de duas populações independentes:

$$H_0: \pi_1 = \pi_2$$

contra a hipótese alternativa de que as proporções para as duas população não as mesmas:

$$H_1: \pi_1 \neq \pi_2$$

utilize a estatística do teste Z_{ESTAT}, fornecida pela Equação (10.5). Para um determinado nível de significância, α, você rejeita a hipótese nula caso a estatística do teste Z_{ESTAT} calculada seja maior do que o valor crítico da cauda superior oriundo da distribuição normal padronizada, ou se a estatística do teste Z_{ESTAT} calculada for menor do que o valor crítico da cauda inferior, a partir da distribuição normal padronizada.

Para ilustrar a utilização do teste Z para a igualdade entre duas proporções, suponha que você seja o gerente da T.C. Resort Properties, um grupo de hotéis da categoria de cinco estrelas, localizados em duas ilhas balneárias. Em uma das ilhas, a T.C. Resort Properties possui dois hotéis, o Beachcomber e o Windsurfer. Você definiu como objetivo estratégico da empresa melhorar a taxa de retorno de hóspedes nos hotéis Beachcomber e Windsurfer. No questionário preenchido pelos hóspedes na hora em que deixam o hotel, uma das perguntas era se ele retornaria ao hotel. Respostas a essa e a outras perguntas foram coletadas de 227 hóspedes no Beachcomber e 262 hóspedes no Windsurfer. Os resultados para a pergunta em questão indicaram que 163 dos 227 hóspedes do Beachcomber responderam que sim, retornariam ao hotel, enquanto 154 dentre 262 hóspedes do Windsurfer responderam que sim, retornariam ao hotel. No nível de significância de 0,05, existem evidências de uma diferença significativa em termos da satisfação dos hóspedes (medida com base na probabilidade de retorno ao hotel) entre os dois hotéis?

A hipótese nula e a hipótese alternativa são

$$H_0: \pi_1 = \pi_2 \quad \text{ou} \quad \pi_1 - \pi_2 = 0$$
$$H_1: \pi_1 \neq \pi_2 \quad \text{ou} \quad \pi_1 - \pi_2 \neq 0$$

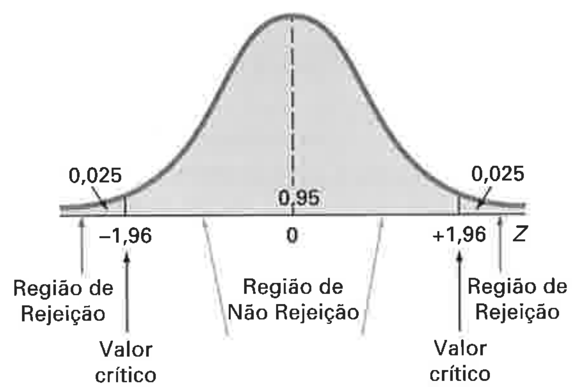
Utilizando o nível de significância de 0,05, os valores críticos são $-1{,}96$ e $+1{,}96$ (veja a Figura 10.11), e a regra de decisão é

Rejeitar H_0 se $Z_{ESTAT} < -1{,}96$

ou se $Z_{ESTAT} > +1{,}96$;

caso contrário, não rejeitar H_0.

FIGURA 10.11
Regiões de rejeição e de não rejeição, ao ser testada a hipótese para a diferença entre duas proporções, no nível de significância de 0,05



Utilizando a Equação (10.5),

$$Z_{ESTAT} = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

em que

$$p_1 = \frac{X_1}{n_1} = \frac{163}{227} = 0{,}7181 \quad p_2 = \frac{X_2}{n_2} = \frac{154}{262} = 0{,}5878$$

e

$$\bar{p} = \frac{X_1 + X_2}{n_1 + n_2} = \frac{163 + 154}{227 + 262} = \frac{317}{489} = 0{,}6483$$

de modo tal que

$$Z_{ESTAT} = \frac{(0{,}7181 - 0{,}5878) - (0)}{\sqrt{0{,}6483(1 - 0{,}6483)\left(\frac{1}{227} + \frac{1}{262}\right)}}$$
$$= \frac{0{,}1303}{\sqrt{(0{,}228)(0{,}0082)}}$$
$$= \frac{0{,}1303}{\sqrt{0{,}00187}}$$
$$= \frac{0{,}1303}{0{,}0432} = +3{,}0088$$

Utilizando o nível de significância de 0,05, você rejeita a hipótese nula, tendo em vista que $Z_{ESTAT} = +3{,}0088 > +1{,}96$. O valor-$p$ é 0,0026 (calculado utilizando-se a Tabela E.2 ou com base na planilha da Figura 10.12) e indica que, se a hipótese nula for verdadeira, a probabilidade de que uma estatística de teste Z_{ESTAT} seja menor do que $-3{,}0088$ é de 0,0013 e, de modo semelhante, a probabilidade de que uma estatística de teste Z_{ESTAT} seja maior do que $+3{,}0088$ é de 0,0013. Assim, no que diz respeito a esse teste bicaudal, o valor-p corresponde a $0{,}0013 + 0{,}0013 = 0{,}0026$. Tendo em vista que $0{,}0026 < \alpha = 0{,}05$, você rejeita a hipótese nula. Existem evidências para que se conclua que os dois hotéis são significativamente diferentes no que se refere à satisfação de seus hóspedes; uma maior proporção de hóspedes está propensa a retornar ao Beachcomber do que ao Windsurfer.

FIGURA 10.12
Teste Z para a diferença entre duas proporções, para o problema que trata da satisfação de hóspedes de hotéis

A Figura 10.12 apresenta a planilha CÁLCULO da pasta de trabalho Z Duas Proporções. Crie essa planilha utilizando as instruções da Seção GE10.3.

	A	B	
1	Teste Z para Diferenças em Duas Proporções		
2			
3	**Dados**		
4	Hipótese da Diferença	0	
5	Nível de Significância	0,05	
6	**Grupo 1**		
7	Número de Itens de Interesse	163	
8	Tamanho da Amostra	227	
9	**Grupo 2**		
10	Número de Itens de Interesse	154	
11	Tamanho da Amostra	262	
12			
13	Cálculos Intermediários		
14	Proporção do Grupo 1	0,7181	=B7/B8
15	Proporção do Grupo 2	0,5878	=B10/B11
16	Diferença nas Duas Proporções	0,1303	=B14 - B15
17	Média da Proporção	0,6483	=(B7 + B10)/(B8 + B11)
18	Estatística do Teste Z	3,0088	=(B16 - B4)/RAIZ(B17 * (1 - B17 * (1/B8 + 1/B11))
19			
20	**Teste Bicaudal**		
21	Valor Crítico Inferior	-1,9600	=INV.NORMP(B5/2)
22	Valor Crítico Superior	1,9600	=INV.NORMP(1 - B5/2)
23	Valor-p	0,0026	=2 * (1 - DISTT(ABS(B18)))
24	Rejeitar a hipótese nula		=SE(B23 < B5, "Rejeitar a hipótese nula", "Não rejeitar a hipótese nula")

EXEMPLO 10.3

Testando a Diferença entre Duas Proporções

A tecnologia vem acarretando um crescimento na quantidade de trabalhadores extremados, que trabalham 60 horas ou mais por semana. Uma das razões citadas pelos empregados sobre o porquê de trabalharem tantas horas é o fato de eles amarem seu trabalho por ele ser estimulante/desafiador/proporcionar um fluxo acelerado de adrenalina (dados extraídos de S. Armour, "Hi, I'm Joan and I'm a Workaholic", *USA Today*, 23 de maio de 2007, pp. 1B, 2B). Suponha que a pesquisa realizada junto a 1.564 *workaholics* tenha incluído 786 homens e 778 mulheres e que 707 homens e 638 mulheres tenham declarado amar seu trabalho por ele ser estimulante/desafiador/proporcionar um fluxo acelerado de adrenalina. No nível de significância de 0,05, você deseja determinar se a proporção de homens *workaholic* e que amam seu trabalho por ele ser estimulante/desafiador/proporcionar um fluxo acelerado de adrenalina é maior do que a proporção de mulheres.

SOLUÇÃO Tendo em vista que você quer saber se existem evidências de que a proporção de homens *workaholics* que amam seu trabalho por ele ser estimulante/desafiador/proporcionar um fluxo acelerado de adrenalina é *maior* do que a proporção de mulheres, você tem um teste unicaudal. A hipótese nula e a hipótese alternativa são

H_0: $\pi_1 \leq \pi_2$ (A proporção de homens *workaholics* que amam seu trabalho por ele ser estimulante/desafiador/proporcionar um fluxo acelerado de adrenalina é menor ou igual à proporção de mulheres.)

H_1: $\pi_1 > \pi_2$ (A proporção de homens *workaholics* que amam seu trabalho por ele ser estimulante/desafiador/proporcionar um fluxo acelerado de adrenalina é maior do que a proporção de mulheres.)

Utilizando o nível de significância de 0,05, para o teste unicaudal na cauda superior, o valor crítico é +1,645. A regra de decisão é

Rejeitar H_0 se $Z_{ESTAT} > +1{,}645$;

caso contrário, não rejeitar H_0.

Utilizando a Equação (10.5),

$$Z_{ESTAT} = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

em que

$$p_1 = \frac{X_1}{n_1} = \frac{707}{786} = 0{,}8995 \quad p_2 = \frac{X_2}{n_2} = \frac{638}{778} = 0{,}8201$$

e

$$\bar{p} = \frac{X_1 + X_2}{n_1 + n_2} = \frac{707 + 638}{786 + 778} = \frac{1.345}{1.564} = 0{,}8600$$

de modo que

$$Z_{ESTAT} = \frac{(0{,}8995 - 0{,}8201) - (0)}{\sqrt{0{,}86(1 - 0{,}86)\left(\frac{1}{786} + \frac{1}{778}\right)}}$$

$$= \frac{0{,}0794}{\sqrt{(0{,}1204)(0{,}0025575)}}$$

$$= \frac{0{,}0794}{\sqrt{0{,}0003079}}$$

$$= \frac{0{,}0794}{0{,}017547} = +4{,}53$$

Utilizando o nível de significância de 0,05, você rejeita a hipótese nula, tendo em vista que $Z_{ESTAT} = +4{,}53 > +1{,}645$. O valor-*p* é aproximadamente 0,0000. Por conseguinte, se a hipótese nula for verdadeira, a probabilidade de que uma estatística de teste Z_{ESTAT} seja maior do que +4,53 é de aproximadamente 0,0000 (que é menor do que $\alpha = 0{,}05$). Você conclui que existem evidências de que a proporção de homens *workaholics* que amam seu trabalho por ele ser estimulante/desafiador/proporcionar um fluxo acelerado de adrenalina é maior do que a proporção de mulheres *workaholics* que amam seu trabalho.

Estimativa do Intervalo de Confiança para a Diferença entre Duas Proporções

Em vez de, ou além de, testar a diferença entre as proporções de duas populações independentes, você pode construir uma estimativa para o intervalo de confiança da diferença entre duas proporções utilizando a Equação (10.6).

*A planilha **CÁLCULO** da pasta de trabalho **Z Duas Proporções** calcula a estimativa de um intervalo de confiança para a diferença entre duas proporções nas colunas D e E (não ilustradas na Figura 10.12).*

ESTIMATIVA DO INTERVALO DE CONFIANÇA PARA A DIFERENÇA ENTRE DUAS PROPORÇÕES

$$(p_1 - p_2) \pm Z_{\alpha/2}\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

ou

$$(p_1 - p_2) - Z_{\alpha/2}\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} \leq (\pi_1 - \pi_2)$$

$$\leq (p_1 - p_2) + Z_{\alpha/2}\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} \quad \textbf{(10.6)}$$

Para construir uma estimativa do intervalo de confiança de 95% para a diferença da população entre a proporção de hóspedes que retornariam ao Beachcomber e que retornariam ao Windsurfer, você utiliza os resultados da Figura 10.12:

$$p_1 = \frac{X_1}{n_1} = \frac{163}{227} = 0{,}7181 \quad p_2 = \frac{X_2}{n_2} = \frac{154}{262} = 0{,}5878$$

Utilizando a Equação (10.6),

$$(0{,}7181 - 0{,}5878) \pm (1{,}96)\sqrt{\frac{0{,}7181(1 - 0{,}7181)}{227} + \frac{0{,}5878(1 - 0{,}5878)}{262}}$$

$$0{,}1303 \pm (1{,}96)(0{,}0426)$$

$$0{,}1303 \pm 0{,}0835$$

$$0{,}0468 \leq (\pi_1 - \pi_2) \leq 0{,}2138$$

Assim, você tem 95% de confiança de que a diferença entre a proporção da população de hóspedes que retornariam ao Beachcomber e ao Windsurfer está entre 0,0468 e 0,2138. Em termos percentuais, a diferença está entre 4,68% e 21,38%. A satisfação dos hóspedes é mais alta no Beachcomber do que no Windsurfer.

Problemas para a Seção 10.3

APRENDENDO O BÁSICO

10.27 Sejam $n_1 = 100$, $X_1 = 50$, $n_2 = 100$ e $X_2 = 30$.

a. No nível de significância de 0,05, existem evidências de uma diferença significativa entre as proporções das duas populações?

b. Construa uma estimativa do intervalo de confiança de 95% para a diferença entre as proporções das duas populações.

10.28 Sejam $n_1 = 100$, $X_1 = 45$, $n_2 = 50$ e $X_2 = 25$.

a. No nível de significância de 0,01, existem evidências de uma diferença significativa entre as proporções das duas populações?

b. Construa uma estimativa do intervalo de confiança de 99% da diferença entre as proporções das duas populações.

APLICANDO OS CONCEITOS

10.29 Foi realizada uma pesquisa junto a 500 consumidores, em uma grande área metropolitana, com o objetivo de determinar várias informações em relação ao comportamento de consumo. Entre as perguntas feitas estava: "Você gosta de comprar roupas?" De um total de 240 homens, 136 responderam que sim. De um total de 260 mulheres, 224 responderam que sim.

a. Existem evidências de uma diferença significativa entre homens e mulheres no que diz respeito à proporção deles que gosta de comprar roupas, no nível de significância de 0,01?

b. Encontre o valor-*p* em (a) e interprete o seu significado.

c. Construa e interprete uma estimativa para o intervalo de confiança de 99% referente à diferença entre as proporções de homens e mulheres que gostam de comprar roupas.
d. Quais seriam suas respostas de (a) a (c) se 206 homens gostassem de comprar roupas?

10.30 Um estudo patrocinado pelo Massachusetts Institute of Technology testou a noção de que até mesmo quando se trata de adoçantes em comprimidos, algumas pessoas acreditam que o mais caro tem melhores resultados do que o mais barato. Pesquisadores dividiram, aleatoriamente, em dois grupos, 82 voluntários remunerados, todos eles saudáveis. Todos os voluntários imaginavam que estavam testando um novo analgésico. A um grupo foi informado que cada comprimido do analgésico que usariam custava \$2,50, enquanto ao outro grupo foi dito que cada comprimido custaria somente 10 centavos. Na realidade, as pílulas que todos ingeririam eram simplesmente de comprimidos contendo adoçante. Foi dado um leve choque elétrico no pulso dos voluntários. Depois, foi dado aos voluntários um comprimido de adoçante e pouco tempo depois eles levaram um novo choque. Dos voluntários que tomaram a pílula cara, 35 de 41 afirmaram sentir menos dor depois do comprimido. Dos voluntários que tomaram a pílula barata, 25 dos 41 disseram sentir menos dor depois de tomarem o comprimido (dados extraídos de R. Rubin, "Placebo Study Tests 'Costlier Is Better' Notion", **www.usatoday.com**, 5 de março de 2008).

a. Estabeleça a hipótese nula e a hipótese alternativa para tentar provar que as pessoas acreditam que uma pílula cara tem melhores resultados do que uma pílula barata.
b. Conduza o teste de hipóteses definido em (a) utilizando o nível de confiança de 0,05.
c. O resultado de seu teste em (b) torna apropriada a declaração de que as pessoas acreditam que uma pílula cara apresenta melhores resultados do que uma pílula barata?

10.31 Algumas pessoas gostam de desfrutar da *expectativa* de um produto ou evento que está por vir ou acontecer e preferem pagar antecipadamente e postergar a data real do consumo/entrega. Em outros casos, as pessoas não desejam essa postergação. Um artigo no *Journal of Marketing Research* relatou um experimento no qual foi informado a 50 indivíduos que tinham acabado de adquirir um ingresso para um concerto, e 50 outros foram informados de que tinham acabado de adquirir um PDA (personal digital assistant – assistente digital pessoal). Foi então solicitado aos participantes que indicassem suas preferências em relação a ir a um concerto ou receber o PDA. Eles prefeririam aquela noite ou a noite seguinte, ou prefeririam esperar de duas a quatro semanas? Foi dito aos indivíduos que ignorassem suas restrições em termos de agenda, de modo tal que pudessem melhor mensurar sua predisposição de postergar o consumo/entrega de sua compra. A tabela a seguir apresenta resultados parciais do estudo:

	Concerto	PDA
Mesma noite ou noite seguinte	28	47
Duas a quatro semanas	22	3
Total	50	50

Fonte: *Dados adaptados de O. Amir e D. Ariely, "Decisions by Rules: The Case of Unwillingness to Pay for Beneficial Delays"*, Journal of Marketing Research, *fevereiro de 2007, Vol. XLIV, pp. 142-152.*

a. Qual proporção dos participantes prefeririam postergar a data do concerto?
b. Qual proporção dos participantes prefeririam postergar o recebimento do novo PDA?
c. Utilizando o nível de significância de 0,05, existem evidências de uma diferença significativa entre a proporção que preferiria postergar a data do concerto e a proporção que preferiria postergar o recebimento do novo PDA?

AUTO-AVALIAÇÃO **10.32** As pessoas em diferentes faixas etárias diferem em termos de suas respostas a mensagens de correio eletrônico? Uma pesquisa científica realizada pelo Center for the Digital Future da Universidade do Sul da Califórnia (dados extraídos de A. Mindlin, "Older E-mail Users Favor Fast Replies", *The New York Times*, 14 de julho de 2008, p. B3) relatou que 70,7% dos usuários com mais de 70 anos de idade acreditam que mensagens de correio eletrônico devem ser respondidas mais prontamente, comparados a 53,6% dos usuários de 12 a 50 anos de idade. Suponha que a pesquisa tenha se baseado em 1.000 usuários com mais de 70 anos de idade e 1.000 usuários entre 12 e 50 anos de idade.

a. No nível de significância de 0,01, existem evidências de uma diferença significativa entre as duas faixas etárias no que se refere a acreditar que as mensagens de correio eletrônico devem ser respondidas mais prontamente?
b. Encontre o valor-*p* em (a) e interprete o seu significado.

10.33 As mulheres têm maior aversão ao risco no mercado de ações? Foi feita a seguinte pergunta a uma amostra de homens e mulheres: "Se tanto o mercado de ações quanto uma determinada ação que você possua caírem 25% em três meses, você compraria mais ações enquanto o preço está baixo?" (dados extraídos de "Women Are More Risk Averse in the Stock Market", *USA Today*, 25 de setembro de 2006, p. 1C). De 965 mulheres, 338 afirmaram que sim. De 1.066 homens, 554 disseram que sim.

a. No nível de significância de 0,05, existem evidências de que a proporção de mulheres que comprariam mais ações enquanto o preço estivesse baixo é menor do que a proporção de homens?
b. Encontre o valor-*p* em (a) e interprete o seu significado.

10.34 Foi conduzido um experimento para estudar as opções feitas em termos da escolha de fundos mútuos. Foram apresentados a alunos da graduação e de MBA diferentes fundos integrantes do Índice S&P 500 que eram idênticos, exceto pelas taxas cobradas. Suponha que tenham sido selecionados 100 alunos da graduação e 100 alunos de MBA. Os resultados parciais são mostrados na tabela a seguir:

	GRUPO DE ALUNOS	
FUNDOS	**Graduação**	**MBA**
Fundo de custo mais elevado	27	18
Fundo de custo menos elevado	73	82

Fonte: *Dados adaptados de J. Choi, D. Laibson and B. Madrian, "Why Does the Law of One Practice Fail? An Experiment on Mutual Funds"*, **www.som.yale.edu/faculty/jjc83/fees.pdf**.

a. No nível de significância de 0,05, existem evidências de uma diferença entre os alunos de graduação e alunos de MBA, no que se refere à proporção que selecionou o fundo de custo mais elevado?
b. Encontre o valor-*p* em (a) e interprete o seu significado.

10.35 Os lugares nos quais as pessoas buscam notícias variam nas diferentes faixas etárias (dados extraídos de P. Johnson, "Young People Turn to the Web for News", *USA Today*, 23 de março de 2006, p. 9D.) Suponha que um estudo conduzido sobre esse tema tenha se baseado em 200 respondentes com idades entre 36 e 50 anos e 200 respondentes com mais de 50 anos de idade. Dentre os 200 respondentes entre 36 e 50 anos de idade, 82 buscavam as notícias principalmente em jornais impressos. Dentre os 200 entrevistados acima de 50 anos de idade, 104 buscavam as notícias principalmente em jornais impressos.

a. Existem evidências de uma diferença significativa na proporção de pessoas que buscam as notícias principalmente em jornais impressos entre os respondentes com 36 a 50 anos de idade e aqueles com mais de 50 anos de idade? (Utilize $\alpha = 0{,}05$.)
b. Determine o valor-*p* em (a) e interprete o seu significado.
c. Construa e interprete uma estimativa para o intervalo de confiança de 95% para a diferença na proporção da população de respondentes que buscam as notícias principalmente em jornais impressos entre pessoas com 36 a 50 anos de idade e pessoas com mais de 50 anos de idade.

10.4 Teste *F* para a Proporcionalidade entre Duas Variâncias

Com frequência, você precisa testar se duas populações independentes apresentam a mesma variabilidade. Ao testar variâncias, você consegue detectar diferenças na variabilidade em duas populações independentes. Uma importante razão para que seja testada a diferença entre as variâncias de duas populações é determinar se deve ser utilizado o teste *t* para variância agrupada (que pressupõe variâncias iguais) ou o teste *t* de variâncias separadas (que não pressupõe variâncias iguais), quando estamos comparando as médias aritméticas de duas populações independentes.

O teste *t* para a diferença entre as variâncias de duas populações independentes é baseado na proporcionalidade entre as variâncias das duas amostras. Se você considerar que cada uma das populações é distribuída segundo uma distribuição normal, então a fração S_1^2/S_2^2 segue a distribuição *F* (veja a Tabela E.5). Os valores críticos da **distribuição *F*** na Tabela E.5 dependem dos graus de liberdade nas duas amostras. Os graus de liberdade no numerador da fração correspondem à primeira amostra, e os graus de liberdade no denominador correspondem à segunda amostra. A primeira amostra, extraída da primeira população, é definida como a amostra que apresenta a *maior* variância de amostra. A segunda amostra, extraída da segunda população, é definida como a amostra com a *menor* variância de amostra. A Equação (10.7) define a **estatística do teste *F* para a proporcionalidade entre duas variâncias**.

ESTATÍSTICA DO TESTE *F* PARA TESTAR A PROPORCIONALIDADE ENTRE DUAS VARIÂNCIAS

A estatística do teste F_{ESTAT} é igual à variância da amostra 1 (a maior variância de amostra) dividida pela variância da amostra 2 (a menor variância de amostra).

$$F_{ESTAT} = \frac{S_1^2}{S_2^2} \qquad \textbf{(10.7)}$$

em que

S_1^2 = variância da amostra 1 (a maior variância de amostra)

S_2^2 = variância da amostra 2 (a menor variância de amostra)

n_1 = tamanho da amostra 1

n_2 = tamanho da amostra 2

$n_1 - 1$ = graus de liberdade da amostra 1 (ou seja, os graus de liberdade do numerador)

$n_2 - 1$ = graus de liberdade da amostra 2 (ou seja, os graus de liberdade do denominador)

A estatística do teste F_{ESTAT} segue uma distribuição *F*, com $n_1 - 1$ e $n_2 - 1$ graus de liberdade.

Para um determinado nível de significância, α, a fim de testar a hipótese nula de igualdade entre variâncias de populações:

$$H_0: \sigma_1^2 = \sigma_2^2$$

contra a hipótese alternativa de que as variâncias para as duas populações não são iguais:

$$H_1: \sigma_1^2 \neq \sigma_2^2$$

você rejeita a hipótese nula se a estatística do teste F_{ESTAT} calculada for maior do que o valor crítico da cauda superior, $F_{\alpha/2}$, a partir da distribuição F, com $n_1 - 1$ graus de liberdade no numerador e $n_2 - 1$ graus de liberdade no denominador. Portanto, a regra de decisão é

$$\text{Rejeitar } H_0 \text{ se } F_{ESTAT} > F_{\alpha/2};$$

$$\text{caso contrário, não rejeitar } H_0.$$

Para ilustrar como utilizar o teste F para determinar se as duas variâncias são iguais, retorne ao cenário da BLK Alimentos, no início do capítulo, que trata das vendas de refrigerantes sabor cola da marca BLK em duas diferentes localizações de exposição do produto. Para determinar se deve ser utilizado o teste t de variância agrupada ou o teste t de variâncias separadas, na Seção 10.1, você pode testar a igualdade entre as variâncias das duas populações. A hipótese nula e a hipótese alternativa são:

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_1: \sigma_1^2 \neq \sigma_2^2$$

Uma vez que você está definindo a amostra 1 como a que apresenta a maior variância de amostra, a região de rejeição na cauda superior da distribuição F contém $\alpha/2$. Utilizando o nível de significância $\alpha = 0{,}05$, a região de rejeição na cauda superior da distribuição F contém 0,025 da distribuição.

Pelo fato de existirem amostras de 10 lojas para cada uma das duas localizações de exposição do produto, existem $10 - 1 = 9$ graus de liberdade no numerador (a amostra com a maior variância) e também no denominador (a amostra com a menor variância). $F_{\alpha/2}$, o valor crítico da cauda superior da distribuição F, é encontrado diretamente na Tabela E.5, uma parcela da qual é apresentada na Tabela 10.6. Uma vez que existem 9 graus de liberdade no numerador e 9 graus de liberdade no denominador, você encontra o valor crítico da cauda superior, $F_{\alpha/2}$ procurando na coluna com a legenda 9 e a linha com a legenda 9. Consequentemente, o valor crítico da cauda superior para essa distribuição F é igual a 4,03. Portanto, a regra de decisão é

$$\text{Rejeitar } H_0 \text{ se } F_{ESTAT} > F_{0{,}025} = 4{,}03;$$

$$\text{caso contrário, não rejeitar } H_0.$$

TABELA 10.6

Encontrando o Valor Crítico da Cauda Superior de F com 9 e 9 Graus de Liberdade para uma Área da Cauda Superior de 0,025

Probabilidades Acumuladas = 0,975
Área da Cauda Superior = 0,025
Numerador gl_1

Denominador gl_2	**1**	**2**	**3**	**...**	**7**	**8**	**9**
1	647,80	799,50	864,20	...	948,20	956,70	963,30
2	38,51	39,00	39,17	...	39,36	39,37	39,39
3	17,44	16,04	15,44	...	14,62	14,54	14,47
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
7	8,07	6,54	5,89	...	4,99	4,90	4,82
8	7,57	6,06	5,42	...	4,53	4,43	4,36
9	7,21	5,71	5,08	...	4,20	4,10	4,03

Fonte: *Extraída da Tabela E.5.*

Utilizando a Equação (10.7) e os dados sobre vendas de refrigerantes sabor cola (veja a Tabela 10.1 na Seção 10.1),

$$S_1^2 = (18{,}7264)^2 = 350{,}6778 \qquad S_2^2 = (12{,}5433)^2 = 157{,}3333$$

de modo que

$$F_{ESTAT} = \frac{S_1^2}{S_2^2} = \frac{350{,}6778}{157{,}3333} = 2{,}2289$$

Uma vez que $F_{ESTAT} = 2{,}2289 < 4{,}03$, você não rejeita H_0. A Figura 10.13 mostra um planilha com os resultados para esse teste, incluindo o valor-p, 0,2482. Uma vez que $0{,}2482 > 0{,}05$, você conclui que não existe nenhuma diferença significativa na variabilidade das vendas de refrigerantes sabor cola nas duas localizações para exposição do produto.

FIGURA 10.13

Planilha para o teste F relativo aos dados sobre as vendas de refrigerantes sabor cola da marca BLK

Figura 10.13 exibe a planilha CÁLCULO da pasta de trabalho F Duas Variâncias. Crie essa planilha utilizando as instruções da Seção GE10.4.

	A	B	
1	Teste F para Diferenças em Duas Variâncias		
2			
3	**Dados**		
4	Nível de Significância	0,05	
5	**Amostra com Maior Variância**		
6	Tamanho da Amostra	10	=CONT.NÚM(CÓPIADADOS!$A:$A)
7	Variância da Amostra	350,6778	=VAR(CÓPIADADOS!$A:$A)
8	**Amostra com Menor Variância**		
9	Tamanho da Amostra	10	=CONT.NÚM(CÓPIADADOS!$B:$B)
10	Variância da Amostra	157,3333	=VAR(CÓPIADADOS!$B:$B)
11			
12	Cálculos Intermediários		
13	Estatística do Teste *F*	2,2289	=B7/B10
14	Graus de Liberdade da Amostra da População 1	9	=B6 - 1
15	Graus de Liberdade da Amostra da População 2	9	=B9 - 1
16			
17	**Teste Bicaudal**		
18	Valor Crítico Superior	4,0260	=INVF(B4/2, B14, B15)
19	Valor-*p*	0,2482	=2 * E17
20	Não rejeitar a hipótese nula		=SE(B19 < B4, "Rejeitar a hipótese nula", "Não rejeitar a hipótese nula")

Não ilustrado
Célula E17: DISTF(B13, B14, B15)

Ao testar a diferença entre duas variâncias utilizando o teste F descrito nesta seção, você pressupõe que cada uma das duas populações é distribuída conforme a distribuição normal. O teste F é bastante sensível ao pressuposto da normalidade. Caso box-plots ou gráficos da distribuição normal sugiram até mesmo um leve distanciamento da normalidade em relação a qualquer uma das duas populações, você não deve utilizar o teste F. Caso isso venha a ocorrer, você deve utilizar o teste de Levene (veja a Seção 11.1) ou uma abordagem não paramétrica (veja as referências 1 e 2).

Ao testar a igualdade entre variâncias como parte da avaliação sobre a validade do procedimento do teste t para variância agrupada, o teste F representa um teste bicaudal com $\alpha/2$ na cauda superior. No entanto, quando você está interessado na variabilidade em situações outras que não se refiram ao teste t de variância agrupada, o teste F é geralmente um teste unicaudal. O Exemplo 10.4 ilustra um teste unicaudal.

EXEMPLO 10.4

Um Teste Unicaudal para a Diferença entre Duas Variâncias

Remessas de carne, produtos derivados da carne e outros ingredientes são misturadas em várias linhas de abastecimento em uma fábrica de ração enlatada para cães. Gerentes de operações suspeitam que, embora a quantidade de carne abastecida por lata seja geralmente estável, a variabilidade nas latas abastecidas na linha de abastecimento A é maior do que a variabilidade da linha de abastecimento B. Os dados de uma amostra de latas contendo 8 onças se apresentam do seguinte modo:

	Linha de abastecimento *A*	**Linha de abastecimento *B***
$\bar{X}$	8,005	7,997
S	0,012	0,005
n	11	16

No nível de significância de 0,05, existem evidências de que a variância na linha de abastecimento A seja maior do que a variabilidade da linha de abastecimento B? Suponha que a população das quantidades abastecidas seja distribuída segundo uma distribuição normal.

SOLUÇÃO A hipótese nula e a hipótese alternativa são

$$H_0: \sigma_A^2 \leq \sigma_B^2$$

$$H_1: \sigma_A^2 > \sigma_B^2$$

A estatística do teste F_{ESTAT} é fornecida pela Equação (10.7):

$$F_{ESTAT} = \frac{S_1^2}{S_2^2}$$

Você utiliza a Tabela E.5 para encontrar o valor crítico superior da distribuição F. Com $n_1 - 1 = 11 - 1 = 10$ graus de liberdade no numerador, $n_2 - 1 = 16 - 1 = 15$ graus de liberdade no denominador e $\alpha = 0{,}05$, o valor crítico superior, $F_{0{,}05}$, é igual a 2,54.

A regra de decisão é

Rejeitar H_0 se $F_{ESTAT} > 2{,}54$;

caso contrário, não rejeitar H_0.

Com base na Equação (10.7),

$$F_{ESTAT} = \frac{S_1^2}{S_2^2} = \frac{(0{,}012)^2}{(0{,}005)^2} = 5{,}76$$

Uma vez que $F_{ESTAT} = 5{,}76 > 2{,}54$, você rejeita H_0. Utilizando um nível de significância de 0,05, você conclui que existem evidências de que a variância na linha de abastecimento A é maior do que a variância na linha de abastecimento B. Em outras palavras, a quantidade de ração nas latas abastecidas na linha de abastecimento A é mais variável do que a quantidade de ração nas latas abastecidas na linha de abastecimento B.

Problemas para a Seção 10.4

APRENDENDO O BÁSICO

10.36 Determine os valores críticos das caudas superior de F em cada um dos seguintes testes bicaudais:

a. $\alpha = 0{,}10$, $n_1 = 16$, $n_2 = 21$
b. $\alpha = 0{,}05$, $n_1 = 16$, $n_2 = 21$
c. $\alpha = 0{,}01$, $n_1 = 16$, $n_2 = 21$

10.37 Determine o valor crítico da cauda superior de F em cada um dos seguintes testes unicaudais:

a. $\alpha = 0{,}05$, $n_1 = 16$, $n_2 = 21$
b. $\alpha = 0{,}01$, $n_1 = 16$, $n_2 = 21$

10.38 As informações a seguir estão disponíveis para duas amostras selecionadas a partir de populações independentes distribuídas por meio de uma distribuição normal:

População A: $n_1 = 25$ $S_1^2 = 16$

População B: $n_2 = 25$ $S_2^2 = 25$

a. Que variância de amostra você coloca no numerador de F_{ESTAT}?
b. Qual é o valor de F_{ESTAT}?

10.39 As informações a seguir estão disponíveis para duas amostras selecionadas a partir de populações independentes distribuídas segundo uma distribuição normal:

População A: $n_1 = 25$ $S_1^2 = 161{,}9$

População B: $n_2 = 25$ $S_2^2 = 133{,}7$

Qual é o valor para F_{ESTAT} caso você esteja testando a hipótese nula de H_0: $\sigma_1^2 = \sigma_2^2$?

10.40 No Problema 10.39, quantos graus de liberdade existem no numerador e no denominador do teste F?

10.41 Nos Problemas 10.39 e 10.41, qual é o valor crítico superior de F se o nível de significância, α, é de 0,05 e a hipótese alternativa é de H_1: $\sigma_1^2 \neq \sigma_2^2$?

10.42 Nos Problemas 10.39 e 10.40, qual é a sua decisão estatística?

10.43 As informações a seguir estão disponíveis para duas amostras selecionadas de populações independentes mas bastante assimétricas à direita:

População A: $n_1 = 16$ $S_1^2 = 47{,}3$

População B: $n_2 = 13$ $S_2^2 = 36{,}4$

Você deve utilizar o teste F para testar a hipótese nula de igualdade entre variâncias? Discuta.

10.44 No Problema 10.43, considere que as duas amostras foram selecionadas de populações independentes e distribuídas segundo uma distribuição normal.

a. No nível de significância de 0,05, existem evidências de uma diferença entre σ_1^2 e σ_2^2?
b. Suponha que você deseje realizar um teste unicaudal. No nível de significância de 0,05, qual seria o valor crítico para a cauda superior de F para determinar se existem evidências de que $\sigma_1^2 > \sigma_2^2$? Qual é a sua decisão estatística?

APLICANDO OS CONCEITOS

10.45 Um professor do departamento de contabilidade de uma faculdade de negócios declara que existe muito mais variabilidade nos resultados de exames finais dos alunos que estão matriculados no curso de Introdução à Contabilidade e que não estão se especializando em contabilidade do que dos alunos que estão matriculados no curso de Introdução à Contabilidade e estão se especializando em contabilidade. Amostras aleatórias de 13 alunos que não estão se especializando em contabilidade e de 10 alunos que estão se especializando em contabilidade são extraídas da pauta de chamada do professor em sala de aula, e os seguintes resultados são calculados com base nos resultados dos exames finais:

Sem especialização em contabilidade: $n_1 = 13$ $S_1^2 = 210{,}2$

Com especialização em contabilidade: $n_2 = 10$ $S_2^2 = 36{,}5$

a. No nível de significância de 0,05, existem evidências que respaldem a declaração do professor?
b. Interprete o valor-p.
c. Qual pressuposto você precisa adotar em (a), no que concerne às duas populações, que venha a justificar a sua utilização do teste F?

AUTO-AVALIAÇÃO **10.46** A CARS – Computer Anxiety Rating Scale (Escala de Medição da Ansiedade Causada por Computadores) mede o nível de ansiedade de um indivíduo com computadores em uma escala que vai de 20 (nenhuma ansiedade) a 100 (nível mais alto de ansiedade). Pesquisadores na Universidade de Miami administraram a CARS a 172 estudantes da área de administração. Um dos objetivos do estudo era determinar se existe uma diferença entre o nível de ansiedade causada por computadores em estudantes dos sexos feminino e masculino. Eles descobriram o seguinte:

	Homens	Mulheres
$\bar{X}$	40,26	36,85
S	13,35	9,42
n	100	72

Fonte: *Dados extraídos de T. Broome e D. Havelka, "Determinants of Computer Anxiety in Business Students"*, The Review of Business Information Systems, *Spring 2002, 6(2), pp. 9-16.*

a. No nível de significância de 0,05, existem evidências de alguma diferença na variabilidade da ansiedade de homens e mulheres causada por computadores?
b. Interprete o valor-p.
c. Qual pressuposto você precisa adotar para as duas populações a fim de justificar a sua utilização do teste F?
d. Com base em (a) e (b), qual dentre os testes t definidos na Seção 10.1 você deve utilizar para testar se existe alguma diferença significativa na média aritmética da ansiedade de estudantes dos sexos masculino e feminino causada por computadores?

10.47 Um banco com uma agência em um bairro comercial de uma cidade desenvolveu um processo de aperfeiçoamento para atendimento aos clientes no horário de pico do almoço, do meio-dia até as 13 horas. O tempo de espera (definido operacionalmente como o tempo decorrido desde o momento em que o cliente entra na fila até que seja atendido no caixa), para todos os clientes, ao longo desse intervalo de tempo, precisa ser diminuído para aumentar a satisfação dos clientes. É selecionada uma amostra aleatória de 15 clientes (e armazenada em Banco1), e os resultados (em minutos) são os seguintes:

4,21 5,55 3,02 5,13 4,77 2,34 3,54 3,20
4,50 6,10 0,38 5,12 6,46 6,19 3,79

Suponha que outra agência bancária, localizada em uma área residencial, esteja também preocupada com o horário de pico do almoço, do meio-dia até as 13 horas. É selecionada uma amostra aleatória de 15 clientes (e armazenada em Banco2), e os resultados (em minutos) se apresentam da seguinte maneira:

9,66 5,90 8,02 5,79 8,73 3,82 8,01 8,35
10,49 6,68 5,64 4,08 6,17 9,91 5,47

a. Existem evidências de alguma diferença em termos da variabilidade do tempo de espera entre as duas agências? (Utilize $\alpha = 0{,}05$.)
b. Determine o valor-p em (a) e interprete o seu significado.
c. Qual pressuposto em relação à distribuição da população é necessário em (a)? O pressuposto é válido para esses dados?
d. Com base nos resultados de (a), é apropriado utilizar o teste t de variância agrupada para comparar as médias aritméticas das duas agências?

10.48 Câmeras fotográficas digitais dominam o mercado de máquinas fotográficas do tipo focalizar e disparar. Uma das características importantes de uma câmera é o tempo de vida útil da bateria, conforme medido pelo número de fotos tiradas até que a bateria precise ser recarregada. O arquivo CâmerasDigitais contém o tempo de vida útil de 29 câmeras subcompactas e 16 câmeras compactas (dados extraídos de "Digital Cameras", *Consumer Reports*, julho de 2009, pp. 28-29).

a. Existem evidências de alguma diferença em termos da variabilidade do tempo de vida útil entre os dois tipos de câmeras digitais? (Utilize $\alpha = 0{,}05$.)
b. Determine o valor-p em (a) e interprete o seu significado.
c. Qual pressuposto em relação à distribuição da população é necessário em (a)? O pressuposto é válido para esses dados?
d. Com base nos resultados de (a), qual dentre os testes t definidos na Seção 10.1 você deve utilizar para comparar as médias aritméticas da vida útil dos dois tipos de câmeras digitais?

10.49 Crianças pequenas usam telefones celulares? Aparentemente sim, de acordo com um estudo recente (dados extraídos de A. Ross, "Message to Santa: Kids Want a Phone", *Palm Beach Post*, 16 de dezembro de 2008, pp. 1A, 4A), que declarou que usuários de telefones celulares com menos de 12 anos de idade faziam uma média de 137 chamadas telefônicas por mês, comparadas a 231 chamadas por mês de usuários de telefones celulares com 13 a 17 anos de idade. Não foram informados os tamanhos das amostras. Suponha que os resultados tenham se baseado em amostras com 50 usuários de telefones celulares em cada um dos grupos e que o desvio-padrão da amostra para os usuários de telefones celulares abaixo de 12 anos de idade tenha sido de 51,7 chamadas por mês, enquanto o desvio-padrão da amostra para os usuários de telefones celulares com 13 a 17 anos de idade tenha sido de 67,6 chamadas por mês.

a. Utilizando um nível de significância de 0,05, existem evidências de uma diferença entre as variâncias do uso de telefones celulares entre usuários de telefones celulares com menos de 12 anos de idade e usuários de telefones celulares com 13 a 17 anos de idade?
b. Com base nos resultados em (a), qual dentre os testes t definidos na Seção 10.1 você deve utilizar para comparar as médias aritméticas dos dois grupos de usuários de telefones celulares? Discuta.

10.50 Existe alguma diferença na variação em termos do rendimento de diferentes tipos de investimento entre os bancos? Os dados a seguir (em RendimentoBanco) representam os rendimentos mais altos, nos Estados Unidos, de contas do mercado monetário e de certificados de depósito (CD) de cinco anos, com base em 17 de maio de 2009.

Contas do Mercado Monetário	CD de 5 Anos
2,25 2,20 2,12 2,03 2,02	3,70 3,66 3,65 3,50 3,50

Fonte: *Dados extraídos de* **www.bankrate.com**, *17 de maio de 2009.*

No nível de significância de 0,05, existem evidências de uma diferença em termos da variância nos rendimentos entre as contas do mercado monetário e os CDs de cinco anos? Considere que os rendimentos das populações são normalmente distribuídos.

UTILIZANDO A ESTATÍSTICA @ BLK Alimentos Revisitada

No cenário Utilizando a estatística, você era o gerente regional de vendas da BLK Alimentos. Você comparou o volume de vendas de refrigerantes sabor cola da marca BLK quando o produto é exposto em uma localização normal de prateleira com o volume de vendas quando o produto é exposto em localizações especiais de ponta de corredor. Foi realizado um experimento no qual 10 lojas utilizaram as prateleiras em localizações regulares e 10 lojas utilizaram a exposição em locais especiais em pontas de corredor. Utilizando um teste t para a diferença entre duas médias aritméticas, você pôde concluir que a média aritmética das vendas utilizando a exposição em locais especiais em pontas de corredor é mais alta do que a média aritmética das vendas com localização em prateleiras regulares. Um intervalo de confiança permitiu que você inferisse, com 95% de confiança, que a exposição em pontas de corredor vende, em média, 6,73 a 36,67 caixas a mais do que a localização em prateleiras regulares. Você realizou também o teste F para a diferença entre duas variâncias, para verificar se a variabilidade de loja para loja, nas vendas das lojas que utilizam a exposição em pontas de corredor, diferia da variabilidade de loja para loja nas vendas das lojas que utilizam a localização em prateleiras regulares. Você concluiu que não existia diferença significativa em termos da variabilidade das vendas de refrigerantes sabor cola em relação aos dois tipos de localização. Como gerente regional de vendas, sua próxima etapa para aumentar as vendas é convencer um maior número de lojas a utilizar a exposição especial em pontas de corredor.

RESUMO

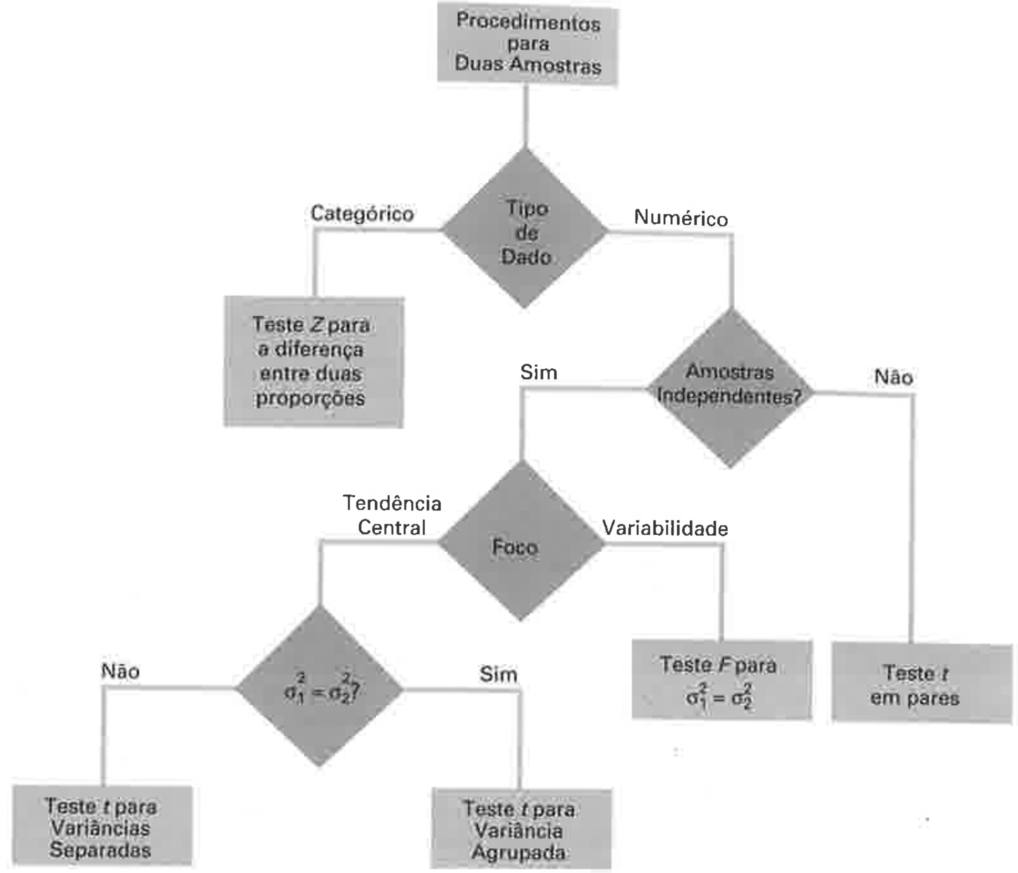
Neste capítulo, você conheceu uma variedade de testes para duas amostras. Em situações nas quais as amostras são independentes, você aprendeu a usar procedimentos de testes estatísticos para analisar possíveis diferenças entre médias aritméticas, variâncias e proporções. Além disso, você aprendeu um procedimento de teste que é frequentemente utilizado quando são analisadas diferenças entre as médias aritméticas de duas amostras inter-relacionadas. Lembre-se sempre de que você precisa selecionar o teste mais apropriado para um determinado conjunto de condições e investigar criteriosamente a validade dos pressupostos subjacentes a cada um dos procedimentos relacionados a testes de hipóteses.

A Tabela 10.7 apresenta uma lista de tópicos abordados neste capítulo. O roteiro na Figura 10.14 ilustra as etapas necessárias para determinar qual dos testes de hipóteses para duas amostras utilizar. São apresentadas a seguir as perguntas que você precisa levar em consideração.

TABELA 10.7
Resumo dos Tópicos do Capítulo 10

Tipo de Análise	Tipos de Dados: Numérico	Tipos de Dados: Categórico
Comparando duas populações	Testes t para a diferença entre as médias aritméticas de duas populações independentes (Seção 10.1)	Teste Z para a diferença entre duas proporções (Seção 10.3)
	Testes t em pares (Seção 10.2)	
	Teste F para a proporcionalidade entre duas variâncias (Seção 10.4)	

FIGURA 10.14
Roteiro para selecionar um teste de hipóteses para duas amostras



1. Que tipos de dados você tem em mãos? Caso esteja lidando com variáveis categóricas, utilize o teste Z para a diferença entre duas proporções. (Esse teste pressupõe amostras independentes.)
2. Caso tenha em mãos uma variável numérica, determine se você está lidando com amostras independentes ou com amostras inter-relacionadas. Caso esteja lidando com amostras inter-relacionadas, utilize o teste t em pares.
3. Se você tem em mãos amostras independentes, seu foco está na variabilidade ou na tendência central? Caso o foco esteja na variabilidade, utilize o teste F.
4. Caso o seu foco esteja na tendência central, determine se você pode pressupor que a variância dos dois grupos é igual. (Esse pressuposto pode ser testado com o uso do teste F.)
5. Caso você possa pressupor que os dois grupos apresentam variâncias iguais, utilize o teste t para variância agrupada. Caso você não possa pressupor que os dois grupos apresentam variâncias iguais, utilize o teste t para variâncias separadas.

EQUAÇÕES-CHAVE

Teste t de Variância agrupada para a Diferença entre Duas Médias Aritméticas

$$t_{ESTAT} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad (10.1)$$

Estimativa do Intervalo de Confiança para a Diferença entre as Médias Aritméticas de Duas Populações Independentes

$$(\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \quad (10.2)$$

ou

$$(\bar{X}_1 - \bar{X}_2) - t_{\alpha/2}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \le \mu_1 - \mu_2 \le (\bar{X}_1 - \bar{X}_2) + t_{\alpha/2}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

Teste t em Pares para a Média Aritmética da Diferença

$$t_{ESTAT} = \frac{\overline{D} - \mu_D}{\dfrac{S_D}{\sqrt{n}}} \tag{10.3}$$

Estimativa do Intervalo de Confiança para a Média Aritmética da Diferença

$$\overline{D} \pm t_{\alpha/2}\frac{S_D}{\sqrt{n}} \tag{10.4}$$

ou

$$\overline{D} - t_{\alpha/2}\frac{S_D}{\sqrt{n}} \leq \mu_D \leq \overline{D} + t_{\alpha/2}\frac{S_D}{\sqrt{n}}$$

Teste Z para a Diferença entre Duas Proporções

$$Z_{ESTAT} = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{\overline{p}(1 - \overline{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} \tag{10.5}$$

Estimativa do Intervalo de Confiança para a Diferença entre Duas Proporções

$$(p_1 - p_2) \pm Z_{\alpha/2}\sqrt{\left(\frac{p_1(1 - p_1)}{n_1} + \frac{p_2(1 - p_2)}{n_2}\right)} \tag{10.6}$$

ou

$$(p_1 - p_2) - Z_{\alpha/2}\sqrt{\frac{p_1(1 - p_1)}{n_1} + \frac{p_2(1 - p_2)}{n_1}} \leq (\pi_1 - \pi_2)$$
$$\leq (p_1 - p_2) + Z_{\alpha/2}\sqrt{\frac{p_1(1 - p_1)}{n_1} + \frac{p_2(1 - p_2)}{n_1}}$$

Teste F para a Proporcionalidade entre Duas Variâncias

$$F_{ESTAT} = \frac{S_1^2}{S_2^2} \tag{10.7}$$

TERMOS-CHAVE

amostras combinadas
distribuição F
medições repetidas
robusto
teste F para a proporcionalidade entre duas variâncias
teste para duas amostras
teste t em pares para a média aritmética da diferença
teste t para variância agrupada
teste t para variâncias separadas
teste Z para a diferença entre duas proporções

PROBLEMAS DE REVISÃO DO CAPÍTULO 10

AVALIANDO O SEU ENTENDIMENTO

10.51 Quais são alguns dos critérios utilizados na seleção de um determinado procedimento de teste de hipóteses?

10.52 Sob quais condições você utiliza o teste t de variância agrupada para examinar as possíveis diferenças entre as médias aritméticas de duas populações independentes?

10.53 Sob quais condições você deve utilizar o teste F para examinar as possíveis diferenças entre as variâncias de duas populações independentes?

10.54 Qual é a diferença entre duas populações independentes e duas populações inter-relacionadas?

10.55 Qual é a diferença entre medições repetidas e itens combinados?

10.56 Quando você tem em mãos duas populações independentes, explique as semelhanças e as diferenças entre o teste de hipóteses para a diferença entre as médias aritméticas e a estimativa do intervalo de confiança para a diferença entre as médias aritméticas.

10.57 Sob quais condições você deve utilizar o teste t em pares para a média aritmética da diferença entre duas populações inter-relacionadas?

APLICANDO OS CONCEITOS

10.58 Um estudo comparou os preços de CDs de música de lojas virtuais na web e lojas tradicionais do comércio varejista [dados extraídos de L. Zoonky e S. Gosain, "A Longitudinal Price Comparison for Music CDs in Electronic and Brick-and-Mortar Markets: Pricing Strategies in Emergent Electronic Commerce", *Journal of Business Strategies*, primavera de 2002, 19(1), pp. 55-72]. Antes de coletar os dados, os pesquisadores definiram criteriosamente várias hipóteses relacionadas à pesquisa, incluindo as seguintes:

1. A dispersão de preços na Internet é mais baixa do que a dispersão de preços no mercado tradicional.
2. Os preços nos mercados eletrônicos são mais baixos do que os preços nos mercados físicos.

a. Considere a hipótese 1 da pesquisa. Escreva a hipótese nula e a hipótese alternativa em termos dos parâmetros da população. Defina criteriosamente os parâmetros da população utilizados.
b. Defina um erro do Tipo I e um erro do Tipo II para a hipótese em (a).
c. Que tipo de teste estatístico você deve utilizar?
d. Quais pressupostos são necessários para que seja realizado o teste que você selecionou?
e. Repita (a) a (d) para a hipótese 2 da pesquisa.

10.59 Um estudo conduzido em março de 2009 descobriu que aproximadamente metade dos adultos norte-americanos confiava mais no governo norte-americano do que nas empresas norte-americanas para solucionar os problemas econômicos dos Estados Unidos. Entretanto, quando a população é subdividida por afiliação a partidos políticos, os resultados são bastante diferentes. O estudo mostrou que 72% dos democratas confiavam mais no governo e somente 29% dos republicanos confiavam mais no governo. Suponha que você esteja encarregado de atualizar o estudo. Você extrairá uma amostra nacional de democratas e uma amostra nacional de republicanos e, depois, tentará utilizar os resultados para mostrar evidências estatísticas de que a proporção de democratas que confia mais no governo do que nas empresas é maior do que a proporção de republicanos que confia mais no governo do que nas empresas.

a. Quais são a hipótese nula e a hipótese alternativa?
b. Qual seria um erro do Tipo I nesse estudo?
c. Qual seria um erro do Tipo II nesse estudo?

10.60 A ASQ – American Society for Quality – conduziu uma pesquisa salarial junto a todos os seus membros. Os membros da ASQ atuam em todas as áreas de instituições de manufatura e de serviços, com um tema comum que é o interesse na qualidade. Dois nomes de funções associados a salários elevados são gerente e mestre faixa-preta. (Na Seção 17.7, você aprenderá que um mestre faixa-preta é a pessoa que assume uma função de liderança e treinamento em uma iniciativa de melhoria da qualidade no processo Seis Sigma). Estatísticas descritivas com relação aos salários desses dois nomes de função são apresentadas na tabela a seguir:

Nome do Cargo	Tamanho da Amostra	Média Aritmética	Desvio-padrão
Gerente	2.228	85.551	24.109
Mestre faixa-preta	134	113.385	24.738

Fonte: *Dados extraídos de I. E. Allen, "Salary Survey: Seeing Green",* Quality Progress, *dezembro de 2008, pp. 20-53.*

a. Utilizando um nível de significância de 0,05, existe uma diferença em termos da variabilidade de salários entre gerentes e mestres faixa-preta?
b. Com base nos resultados em (a), qual teste t definido na Seção 10.1 é apropriado para comparar a média aritmética dos salários?
c. Utilizando um nível de significância de 0,05, existe uma diferença entre a média aritmética do salário dos gerentes e a média aritmética do salário dos mestres faixa-preta?

10.61 Estudantes dos sexos masculino e feminino estudam a mesma quantidade de horas por semana? Em 2007, 58 alunos do primeiro ano de uma faculdade de administração foram entrevistados em uma grande universidade que possui mais de 1.000 alunos no primeiro ano de administração. O arquivo **TempoEstudo** contém o gênero e o número de horas gastas com estudo em uma semana típica, para os alunos entrevistados.

a. Em um nível de significância de 0,05, existe uma diferença na variância do tempo de estudo para estudantes do sexo masculino e estudantes do sexo feminino?
b. Utilizando os resultados em (a), qual teste t é apropriado para comparar a média aritmética do tempo de estudo para estudantes do sexo masculino e estudantes do sexo feminino?
c. No nível de significância de 0,05, conduza o teste selecionado em (b).
d. Escreva um resumo sucinto de suas descobertas.

10.62 Dois professores queriam saber como alunos das duas universidades em que lecionavam se comparavam em termos da capacidade de utilizar planilhas do Excel em cursos de sistemas de informação no nível de graduação (dados extraídos de H. Howe and M. G. Simkin, "Factors Affecting the Ability to Detect Spreadsheet Errors", *Decision Sciences Journal of Innovative Education*, janeiro de 2006, pp.101-122.) Foi realizada também uma comparação das características demográficas dos alunos. Uma das escolas é uma universidade estadual no oeste dos Estados Unidos, enquanto a outra é uma universidade estadual no leste dos Estados Unidos. A tabela a seguir contém informações relativas às idades dos alunos:

Faculdade	Tamanho da Amostra	Média Aritmética da Idade	Desvio-padrão
Oeste	93	23,28	6,29
Leste	135	21,16	1,32

a. Utilizando um nível de significância de 0,01, existem evidências de alguma diferença nas variâncias da idade dos alunos na universidade no oeste e na universidade no leste dos EUA?
b. Discuta as implicações práticas do teste realizado em (a). Aborde, especificamente, o impacto que variâncias iguais (ou diferentes) na idade exercem sobre o ensino no curso de sistemas de informação no nível de graduação.
c. Para testar a diferença na média aritmética da idade de alunos, seria mais apropriado utilizar o teste t de variância agrupada ou o teste t de variâncias separadas?

A tabela a seguir contém informações relativas aos anos de prática dos alunos no uso de planilhas:

Faculdade	Tamanho da Amostra	Média Aritmética dos Anos	Desvio-padrão
Oeste	93	2,6	2,4
Leste	135	4,0	2,1

d. Utilizando um nível de significância de 0,01, existem evidências de alguma diferença entre as variâncias nos anos de prática no uso de planilhas dos alunos na universidade no oeste e na universidade no leste dos EUA?
e. Com base nos resultados de (d), utilize o teste mais apropriado para determinar, no nível de significância de 0,01, se existem evidências de uma diferença em termos da média aritmética dos anos de prática no uso de planilhas dos alunos na universidade no oeste e na universidade no leste dos EUA.

10.63 O arquivo de dados **CustoRest** contém as cotações para comida, decoração, serviço e preço por pessoa de uma amostra de 50 restaurantes localizados em uma área urbana e 50 localizados em um subúrbio dos Estados Unidos. Analise integralmente as diferenças entre restaurantes da área urbana e dos subúrbios no que diz respeito às variáveis de cotação de comida, decoração, serviços e custo por pessoa, utilizando $\alpha = 0{,}05$.

Fonte: *Dados extraídos de* Zagat Survey 2008: New York City Restaurants *e* Zagat Survey 2007-2008: Long Island Restaurants.

10.64 Uma professora de informática está interessada em estudar a quantidade de tempo necessária para que alunos inscritos no curso de introdução à informática escrevam e executem um programa em Visual Basic. A professora contrata você para analisar os resultados a seguir (em minutos) extraídos de uma amostra aleatória de nove alunos (os dados estão armazenados no arquivo VB):

10 13 9 15 12 13 11 13 12

a. No nível de significância de 0,05, existem evidências de que a média aritmética da população correspondente à quantidade de tempo seja maior do que 10 minutos? O que você irá informar à professora?

b. Suponha que a professora de informática, ao verificar seus resultados, perceba que o quarto aluno precisou de 51 minutos, em vez dos 15 minutos registrados, para escrever e executar o programa em Visual Basic. No nível de significância de 0,05, analise novamente a pergunta apresentada em (a), utilizando os dados revisados. O que você dirá agora à professora?

c. A professora está surpresa diante desses resultados paradoxais, e solicita a você uma explicação para a diferença em suas descobertas em (a) e (b). Discuta.

d. Alguns dias mais tarde, a professora telefona para dizer a você que o dilema está completamente solucionado. O número original de 15 (o quarto valor de dados) estava correto, e, por conseguinte, as suas descobertas em (a) estão sendo utilizadas no artigo que ela está escrevendo para um periódico de informática. Agora, ela quer contratar você para comparar os resultados daquele grupo de alunos de introdução à informática com os resultados de uma amostra de 11 alunos graduados em informática, para determinar se existem evidências de que os alunos graduados em informática são capazes de escrever um programa em Visual Basic em menos tempo do que os alunos do curso introdutório. Para os alunos já formados, a média aritmética da amostra é de 8,5 minutos, e o desvio-padrão da amostra é de 2,0 minutos. No nível de significância de 0,05, analise completamente esses dados. O que você informará à professora?

e. Alguns dias depois, a professora telefona novamente para dizer a você que um revisor de seu artigo quer que ela inclua o valor-p para o resultado "correto" em (a). Além disso, a professora questiona você sobre o problema relacionado a variâncias diferentes, sobre o qual o revisor quer que ela discuta no artigo. Com suas próprias palavras, discuta o conceito do valor-p e descreva o problema relacionado a variâncias diferentes. Determine então o valor-p em (a) e analise se o problema relacionado a variâncias diferentes teve algum significado no estudo realizado pela professora.

10.65 Um artigo (dados extraídos de A. Jennings.,"What's Good for a Business Can Be Hard on Friends", *The New York Times*, 4 de agosto de 2007, pp. C1–C2) relatou que, de acordo com uma pesquisa de opinião, a média aritmética da quantidade de chamadas de telefones celulares por mês era de 290 para pessoas entre 18 e 24 anos de idade e de 194 para pessoas entre 45 e 54 anos de idade, enquanto a média aritmética para a quantidade de mensagens de texto por mês era de 290 para pessoas entre 18 e 24 anos de idade e de 57 para pessoas entre 45 e 54 anos de idade. Suponha que a pesquisa de opinião tenha se baseado em uma amostra de 100 pessoas entre 18 e 24 anos de idade e 100 pessoas entre 45 e 54 anos de idade e que o desvio-padrão para a quantidade de chamadas para telefones celulares por mês seja de 100 para pessoas entre 18 e 24 anos de idade e de 90 para pessoas entre 45 e 54 anos de idade, enquanto o desvio-padrão para a quantidade de mensagens de texto por mês seja de 90 para pessoas entre 18 e 24 anos de idade e de 77 para pessoas entre 45 e 54 anos de idade.

Utilizando um nível de significância de 0,05,

a. Existem evidências de uma diferença nas variâncias para a quantidade de chamadas de telefones celulares por mês de pessoas entre 18 e 24 anos de idade e de pessoas entre 45 e 54 anos de idade?

b. Existem evidências de uma diferença na média aritmética para a quantidade de chamadas de telefones celulares por mês de pessoas entre 18 e 24 anos de idade e de pessoas entre 45 e 54 anos de idade?

c. Construa e interprete uma estimativa do intervalo de confiança de 95% para a diferença na média aritmética da quantidade de chamadas de telefones celulares por mês de pessoas entre 18 e 24 anos de idade e de pessoas entre 45 e 54 anos de idade.

d. Existem evidências de uma diferença nas variâncias do número de mensagens de texto por mês, para pessoas entre 18 e 24 anos de idade e para pessoas entre 45 e 54 anos de idade?

e. Existem evidências de uma diferença nas médias aritméticas do número de mensagens de texto por mês para pessoas entre 18 e 24 anos de idade e de pessoas entre 45 e 54 anos de idade.

f. Construa e interprete uma estimativa para o intervalo de confiança de 95% para a diferença na média aritmética do número de mensagens de texto por mês de pessoas entre 18 e 24 anos de idade e de pessoas entre 45 e 54 anos de idade.

g. Com base nos resultados de (a) a (f), a que conclusões você pode chegar com relação à utilização de telefones celulares e de mensagens de texto de pessoas entre 18 e 24 anos de idade e de pessoas entre 45 e 54 anos de idade?

10.66 A extensão da vida útil (em horas) de uma amostra de 40 lâmpadas de 100 watts produzidas pelo fabricante A e de uma amostra de 40 lâmpadas de 100 watts produzidas pelo fabricante B encontra-se no arquivo Lâmpadas. Analise completamente as diferenças entre a extensão da vida útil das lâmpadas produzidas pelos dois fabricantes. (Utilize $\alpha = 0{,}05$.)

10.67 O gerente de um hotel está preocupado em aumentar a taxa de retorno dos hóspedes do hotel. Um dos aspectos que influencia a primeira impressão dos hóspedes está relacionado ao tempo para entregar a bagagem do hóspede no quarto depois do *check-in* no hotel. Em um determinado dia, foi selecionada uma amostra aleatória de 20 entregas na Ala A do hotel e uma amostra aleatória de 20 entregas na Ala B do hotel. Os resultados estão armazenados no arquivo Bagagem. Analise os dados e determine se existe alguma diferença na média aritmética para o tempo de entrega nas duas alas do hotel. (Utilize $\alpha = 0{,}05$.)

10.68 Muitas pessoas estão economizando em razão das condições recentes da economia, mas existe alguma diferença entre homens e mulheres no que diz respeito a o que fizeram para economizar? Uma pesquisa recente (dados extraídos de "He Spends, She Doesn't", *Consumer Reports*, dezembro de 2008, p. 17) relatou os seguintes resultados. Suponha que a pesquisa se baseou em 100 homens e 100 mulheres.

O que Eles Fizeram?	Homens	Mulheres
Gastaram menos com diversão e comendo fora	49%	62%
Reduziram gastos com cartões de crédito	49	57
Planejaram reduzir gastos nas férias	36	63
Colocaram mais dinheiro na poupança	39	40
Adiaram um projeto de reforma da casa	26	43
Desistiram de uma grande compra para a casa	24	39
Cancelaram ou adiaram um período de férias	27	36
Desistiram da compra de um carro novo	27	34
Cancelaram uma consulta médica ou um procedimento médico	20	27

Para *cada tipo de redução em termos de gasto*, determine se existe uma diferença entre homens e mulheres quanto a o que fariam para economizar.

10.69 O produtor das placas de asfalto das marcas Boston e Vermont sabe que o peso do produto é um fator importante na percepção de qualidade do consumidor. Mais do que isso, o peso representa o montante de matéria-prima que está sendo utilizado e, consequentemente, é muito importante para a empresa sob o ponto de vista de custos. O último estágio da linha de produção embala as placas antes que elas sejam colocadas em paletes de madeira. Uma vez preenchido (um palete, para a maioria das marcas, contém 16 pés de placas), o palete é pesado e a medição é registrada. O arquivo Palete contém o peso (em libras) gerado por uma amostra com 368 paletes de placas da marca Boston e 330 paletes de placas da marca Vermont. Analise completamente as diferenças nos pesos das placas Boston e Vermont, utilizando $\alpha = 0{,}05$.

10.70 O produtor das placas de asfalto das marcas Boston e Vermont fornece a seus clientes uma garantia de 20 anos para a maior parte de seus produtos. Para determinar se uma placa durará todo o período da garantia, são conduzidos, na área de produção, testes de aceleração de vida útil. O teste de aceleração de vida útil expõe a placa, em um ambiente de laboratório, às condições de desgaste às quais ela estaria sujeita ao longo de uma vida útil de utilização normal, por meio de um experimento que precisa apenas de alguns poucos minutos para ser conduzido. Nesse teste, uma placa é repetidamente esfregada com uma escova, por um curto período de tempo, e os grãos da placa removidos por meio da escovação são pesados (em gramas). Espera-se que as placas que sofrem a perda de poucas quantidades de grãos durem mais, mediante utilização normal, do que as placas que sofrem a perda de grandes quantidades de grãos. Nesse tipo de situação, uma placa deve sofrer a perda de uma quantidade de grãos não superior a 0,8 gramas, caso seja esperado que ela dure toda a extensão do período da garantia. O arquivo Grão contém uma amostra com 170 medições realizadas nas placas Boston e 140 medições realizadas nas placas Vermont. Analise completamente as diferenças na perda de grãos das placas Boston e Vermont, utilizando $\alpha = 0{,}05$.

EXERCÍCIOS DE REDAÇÃO DE RELATÓRIOS

10.71 Com referência aos resultados dos Problemas 10.69 e 10.70 relativos ao peso e à perda de grãos das placas Boston e Vermont, escreva um relatório que sintetize as suas conclusões.

PROJETOS DE GRUPO

O arquivo Fundos de Títulos contém informações relativas a oito variáveis extraídas de uma amostra de 180 fundos mútuos.

Tipo — Tipos de títulos que compõem o fundo mútuo (do governo de médio prazo e corporativos de curto prazo)
Ativos — Em milhões de dólares
Comissões — Taxas cobradas sobre vendas (não ou sim)
Proporção de despesas — Proporção entre despesas incorridas e ativos líquidos, em percentagens
Retorno 2008 — Retorno de 12 meses em 2008
Retorno de 3 anos — Retorno ano a ano, 2006-2008
Retorno de 5 anos — Retorno ano a ano, 2004-2000
Risco — Fator de risco de perdas do fundo mútuo (abaixo da média, médio ou acima da média)

10.72 Analise completamente as diferenças entre os fundos mútuos que não cobram taxas e os fundos que cobram taxas em termos do retorno em 2008, do retorno de 3 anos, do retorno de 5 anos e da proporção de despesas. Escreva um relatório sintetizando suas descobertas.

10.73 Analise completamente as diferenças entre os fundos de títulos do governo de médio prazo e corporativos de curto prazo, intermediário em termos do retorno em 2008, do retorno de 3 anos, do retorno de 5 anos e da proporção de despesas. Escreva um relatório sintetizando suas descobertas.

BANCO DE DADOS DE PESQUISA REALIZADA JUNTO A ALUNOS

10.74 O Problema 2.117, nos Problemas de Revisão do Capítulo 1, descreve uma pesquisa realizada junto a 50 alunos da graduação (dados armazenados em PesquisaGrad). Para esses dados,

a. no nível de significância de 0,05, existem evidências de uma diferença entre homens e mulheres, em termos da média final, para o salário inicial esperado, o salário esperado em 5 anos, a idade e o gasto com livros didáticos e o material escolar?

b. no nível de significância de 0,05, existem evidências de uma diferença entre os alunos que pretendem fazer um curso de pós-graduação e os alunos que não pretendem fazer um curso de pós-graduação, em termos da média final, para o salário inicial esperado, o salário esperado em 5 anos, a idade e o gasto com livros didáticos e o material escolar?

10.75 O Problema 2.117, nos Problemas de Revisão do Capítulo 1, descreve uma pesquisa realizada junto a 50 alunos da graduação (dados armazenados em PesquisaGrad).

a. Selecione uma amostra de 50 alunos da graduação em sua faculdade e conduza uma pesquisa semelhante para eles.

b. Para os dados coletados em (a), repita (a) e (b) no Problema 10.74.

c. Compare os resultados de (b) com os do Problema 10.74.

10.76 O Problema 2.119, nos Problemas de Revisão do Capítulo 1, descreve uma pesquisa realizada junto a 40 alunos de MBA (dados armazenados em **PesquisaPosGrad**). Em relação a esses dados, no nível de significância de 0,05, existem evidências de alguma diferença entre homens e mulheres, em termos de idade, média final antes da graduação, média final da graduação, resultado do GMAT, salário esperado logo após a graduação, salário esperado em cinco anos e gasto com livros didáticos e material escolar?

10.77 O Problema 2.119, nos Problemas de Revisão do Capítulo 1, descreve uma pesquisa realizada junto a 40 alunos de MBA (dados armazenados em **PesquisaPosGrad**.

a. Selecione uma amostra de 40 alunos de MBA em sua faculdade e conduza uma pesquisa semelhante com eles.

b. Para os dados coletados em (a), repita (a) no Problema 10.76.

c. Compare os resultados de (b) com os do Problema 10.76.

ADMINISTRANDO O *SPRINGVILLE HERALD*

A equipe do departamento de marketing está encarregada de melhorar o processo de telemarketing para aumentar o número de vendas de assinaturas com entrega domiciliar. Depois de diversas sessões de debates, ficou claro que, quanto mais longo o intervalo de tempo no qual o atendente conversasse com o respondente, maior seria a possibilidade de que o atendente vendesse uma assinatura com entrega domiciliar. Consequentemente, a equipe decidiu encontrar meios de aumentar a duração das chamadas telefônicas.

Inicialmente, a empresa investigou o impacto que o horário da chamada telefônica teria sobre a duração da chamada. Nos padrões atuais, as chamadas eram realizadas no horário noturno, entre 17h e 19h, de segunda a sexta-feira. A equipe queria comparar a duração das chamadas feitas no começo da noite (antes das 19h) com aquelas realizadas mais tarde (depois das 19h), para determinar se um desses períodos de tempo seria mais propício para uma chamada mais longa e, consequentemente, para um aumento na venda de assinaturas. A equipe selecionou uma amostra de 30 atendentes do sexo feminino, que compõem o banco de atendimento nas noites de quarta-feira, e designou 15 delas aleatoriamente para o grupo das chamadas feitas "mais cedo" e 15 para o grupo das chamadas feitas "mais tarde". As atendentes sabiam que a equipe estava observando seu trabalho naquela noite, mas não sabiam quais as chamadas que eram monitoradas. As atendentes foram treinadas para realizar suas apresentações por telefone de forma estruturada. Elas deveriam ler um roteiro, e seu cumprimento seria pessoal, porém informal ("Oi, aqui é Leigh Richardson, do *Springville Herald*. Posso falar com Stuart Knoll?")

Foram realizadas medições para a duração da chamada (definidas como a diferença, em segundos, entre o momento em que a pessoa atendeu ao telefone e o momento em que ela o desligou). Os resultados (armazenados em **SH10**) são apresentados na Tabela SH10.1.

TABELA SH10.1

Duração das Chamadas, em Segundos, com Base no Horário da Chamada – Cedo *Versus* Tarde à Noite

Tempo da Chamada		Tempo da Chamada	
Cedo	**Tarde**	**Cedo**	**Tarde**
41,3	37,1	40,6	40,7
37,5	38,9	33,3	38,0
39,3	42,2	39,6	43,6
37,4	45,7	35,7	43,8
33,6	42,4	31,3	34,9
38,5	39,0	36,8	35,7
32,6	40,9	36,3	47,4
37,3	40,5		

EXERCÍCIOS

SH10.1 Analise os dados na Tabela SH10.1 e redija um relatório para a equipe do departamento de marketing que indique as suas descobertas. Anexe um apêndice no qual você discute as razões pelas quais selecionou um determinado teste estatístico para comparar os dois grupos independentes de atendentes.

SH10.2 Suponha que, em vez do modelo de pesquisa aqui descrito, existissem somente 15 atendentes selecionadas para fins de amostra e cada uma das atendentes tivesse que ser monitorada duas vezes à noite — uma vez no período mais cedo e uma vez no período mais tarde. Por conseguinte, na Tabela SH10.1, cada par de valores representa as duas medições de uma determinada atendente. Analise novamente esses dados e redija um relatório para a equipe indicando suas descobertas.

SH10.3 Quais outras variáveis deveriam ser investigadas em seguida? Por quê?

CASO DE INTERNET

Aplique os seus conhecimentos sobre testes de hipóteses neste Caso de Internet, que dá continuidade ao Caso de Internet dos Capítulos 7 e 9, que tratam da contenda sobre a quantidade abastecida de cereais.

Mesmo depois do recente experimento de domínio público sobre os pesos das caixas de cereais, a CCACC — Consumers Concerned About Cereal Cheaters (Consumidores Preocupados com Fraudadores de Cereais) continua convencida de que a Oxford Cereals enganou o público. O grupo criou e enviou um documento no qual afirma que as caixas de cereais produzidas na Unidade de Produção 2, em Springville, pesam menos do que a média aritmética de 368 gramas declarada. Utilizando o seu navegador na web, acesse o site da LTC Editora e abra a página da Internet para o Caso de Internet do Capítulo 10, ou abra diretamente o arquivo **MoreCheating.htm**, caso já tenha baixado para o seu computador os arquivos da pasta Caso de Internet, para visitar a página virtual Mais Fraudes da CCACC. Depois disso, responda às seguintes questões:

1. Os resultados da CCACC comprovam que existe uma diferença estatística entre as médias aritméticas dos pesos das caixas de cereais produzidas nas Unidades de Produção 1 e 2?
2. Realize a análise apropriada para testar a hipótese da CCACC. A que conclusões você consegue chegar com base nos dados?

REFERÊNCIAS

1. Conover, W. J., *Practical Nonparametric Statistics*, 3rd ed. (New York: Wiley, 2000).
2. Daniel, W., *Applied Nonparametric Statistics*, 2nd ed. (Boston: Houghton Mifflin, 1990).
3. *Microsoft Excel 2007* (Redmond, WA: Microsoft Corp., 2007).
4. Satterthwaite, F. E., "An Approximate Distribution of Estimates of Variance Components," *Biometrics Bulletin*, 2(1946): 110-114.
5. Snedecor, G. W., and W. G. Cochran, *Statistical Methods*, 8th ed. (Ames, IA: Iowa State University Press, 1989).
6. Winer, B. J., D. R. Brown, and K. M. Michels, *Statistical Principles in Experimental Design*, 3rd ed. (New York: McGraw-Hill, 1989).