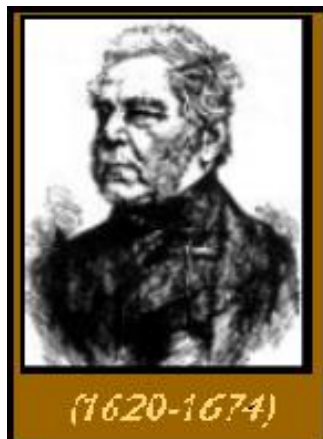


Teste de Hipóteses



John Graunt (Londres, Inglaterra, 24 de abril de 1620 — 18 de Abril de 1674)

John Graunt homem muito estudioso, um dos fundadores da Royal Society, viveu numa época marcada pelo nascimento da ciência moderna. Foi o primeiro estatístico a fazer o tratamento estatístico de dados demográficos e a tentar aplicar a teoria a problemas reais. Nasceu a 24 de Abril de 1620 em Londres, Inglaterra. Em 1662, Graunt publicou o famoso *Natural and Political Observations on the London Bills of Mortality* (Observações Naturais e Políticas da taxa de mortalidade londrina). Este foi o seu primeiro tratamento estatístico de dados demográficos. Estudou a mortalidade da cidade de Londres e as incidências das causas naturais, sociais e políticas nesse fenómeno. Através das Tábuas de Mortalidade realizadas na altura da peste na cidade de Londres, Graunt fez uma análise exaustiva do número de pessoas que morriam de várias doenças e estimou o número de nascimentos de homens e mulheres. Em 1664 foi eleito para o conselho da Royal Society, onde foi o primeiro participante regular das reuniões até Abril de 1666. Com o seu amigo William Petty (1623-1687) fundou a escola dos "Aritméticos Políticos" que se preocupava com o estudo numérico dos fenómenos sociais e políticos. Morreu a 18 de Abril de 1674 em Londres, Inglaterra.

FONTE: <http://www.alea.pt/html/nomesEdatas/swf/biografias.asp?art=10>

Introdução

Hipóteses Nula e Alternativa

Estatística de Teste e Valor Crítico

Tipos de Testes

Métodos para um Teste de Hipóteses

11 - TESTE DE HIPÓTESES

11.1 - INTRODUÇÃO

Em estatística, uma *Hipótese* é uma afirmativa sobre uma propriedade da população. Um *Teste de Hipótese* (ou *Teste de Significância*) é um procedimento padrão para testar uma afirmativa sobre uma propriedade da população.

Regra do Evento Raro para Inferência Estatística

Se, sob uma dada suposição, a probabilidade de um evento observado particular é excepcionalmente pequena, concluímos que a suposição provavelmente não é correta.

Testa-se uma afirmativa analisando dados amostrais na tentativa de distinguir entre resultados que podem *facilmente ocorrer por acaso* e resultados que são *altamente improváveis de ocorrer por acaso*. Pode-se explicar a ocorrência de resultados altamente improváveis dizendo que ou um evento raro realmente ocorreu, ou que a suposição subjacente não é verdadeira. Componente de um Teste de Hipóteses

11.2 - HIPÓTESES NULA E ALTERNATIVA

A **Hipótese Nula** (representado por H_0) é uma afirmativa de que o valor de um parâmetro populacional (com proporção, média ou desvio padrão) é *igual, maior ou igual*, ou ainda, *menor ou igual* a algum valor especificado. A Hipótese Nula é testada diretamente, no sentido de que se supõe que ela seja verdadeira e chega-se a uma conclusão para *rejeitar ou não rejeitar* H_0 .

A **Hipótese Alternativa** (representado por H_A ou H_1) é a afirmativa de que o parâmetro tem um valor que, de alguma forma, difere da hipótese nula. A forma simbólica da Hipótese Alternativa deve-se usar um dos símbolos: $>$ $<$ ou \neq .

Exemplo 11.1 - Exemplo: Um gerente de uma distribuidora de peças afirma que o volume médio de venda mensal é menor do que \$ 4000.

Solução:

A afirmação do problema é: *Afirm: $\mu < 4000$*

A Hipótese Nula será: $H_0: \mu \geq 4000$

A hipótese Alternativa será: $H_A: \mu < 4000$

11.3 - ESTATÍSTICA DE TESTE E VALOR CRÍTICO

A **Estatística de Teste** é um valor calculado a partir dos dados amostrais e é usada para se tomar uma decisão sobre a rejeição ou não da hipótese nula. A estatística de teste é encontrada pela conversão da estatística amostral (como a proporção amostral \hat{p} , a média amostral \bar{x} , ou o desvio padrão amostral s) em um escore (como Z_{teste} , t_{teste} ou χ^2_{teste}) com a suposição de que a hipótese nula seja verdadeira. A estatística de teste pode, portanto, ser usada para determinar se há evidência significativa contra a hipótese nula.

TABELA 25.- ESTATÍSTICAS DE TESTE

PARÂMETRO A SER TESTADO	ESTATÍSTICA DE TESTE
Proporção	$Z_{teste} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$
Média	$Z_{teste} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$ ou $t_{teste} = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}$
Desvio Padrão ou Variância	$\chi^2_{teste} = \frac{(n-1)S^2}{\sigma^2}$

A **Região Crítica** (ou **Região de Rejeição – RC**) é o conjunto de todos os valores da estatística de teste que nos fazem rejeitar a hipótese nula. A região crítica é delimitada pelo Valor Crítico (Z_c , t_c ou χ^2_c). A figura 35 a seguir ilustra estes conceitos para a Distribuição normal.

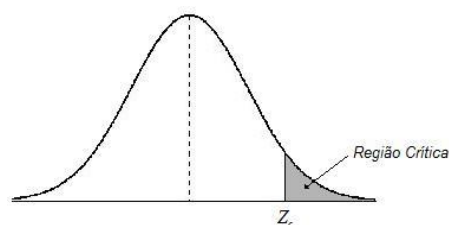


Figura 35.- Gráfico Normal com Região Crítica

11.4 - TIPOS DE TESTES

Dependendo das Hipóteses Nula e Alternativa um teste de hipóteses pode ser Bilateral, Unilateral Direito ou unilateral Esquerdo.

O Quadro a seguir ilustra os tipos de Testes. A letra θ representa qualquer estatística populacional (μ, σ, σ^2 ou p).

TABELA 26.- TIPOS DE TESTES DE HIPÓTESES E SUA HIPÓTESES NULA E ALTERNATIVA

Tipo de Teste	Afirmativa	Hipóteses	Gráficos
Bilateral	<i>Afirm:</i> $\theta = ?$	$H_0 : \theta = ?$ $H_A : \theta \neq ?$	
	<i>Afirm:</i> $\theta \neq ?$	$H_0 : \theta = ?$ $H_A : \theta \neq ?$	
Unilateral Esquerdo	<i>Afirm:</i> $\theta \geq ?$	$H_0 : \theta \geq ?$ $H_A : \theta < ?$	
	<i>Afirm:</i> $\theta < ?$	$H_0 : \theta \geq ?$ $H_A : \theta < ?$	
Unilateral Direito	<i>Afirm:</i> $\theta \leq ?$	$H_0 : \theta \leq ?$ $H_A : \theta > ?$	
	<i>Afirm:</i> $\theta > ?$	$H_0 : \theta \leq ?$ $H_A : \theta > ?$	

11.5 - MÉTODOS PARA UM TESTE DE HIPÓTESES

Existem três métodos para se fazer um Teste de Hipóteses. Qualquer um destes três métodos podem ser utilizados e apesar de seus procedimentos serem diferentes chega-se sempre às mesmas conclusões finais.

11.5.1 - TESTE DE HIPÓTESE PELO MÉTODO TRADICIONAL

O método Tradicional consiste em comparar graficamente a posição da estatística de teste em relação à região crítica. Se a estatística de teste pertencer à região crítica a hipótese nula deve ser rejeitada, caso contrário a hipótese nula Não será rejeitada.

➤ Procedimentos para um Teste de Hipótese pelo Método Tradicional:

- 1º - Extrair a afirmação do problema e representá-la em linguagem matemática;
- 2º - Construir as Hipóteses Nula e Alternativa;
- 3º - Determinar o tipo de teste: Bilateral, Unilateral Direito ou Unilateral Esquerdo;
- 4º - Determinar qual distribuição deve ser utilizada: Normal, T-Student ou Qui-Quadrado;
- 5º - Construir o gráfico adequado, de acordo com a Distribuição utilizada, indicando o nível de significância e o nível de confiabilidade;
- 6º - Determinar os valores críticos e plotar no gráfico;
- 7º - Calcular a estatística de teste apropriada e plotar no gráfico;
- 8º - Decidir se a Hipótese Nula deve ser Rejeitada ou não;
- 9º - Comparar a Hipótese Nula com a afirmação do problema e decidir se a afirmação deve ser rejeitada ou não.

Exemplo 11.2 - Um determinado fabricante produz implantes mamários utilizando gel de silicone. Uma amostra de 45 implantes apresentou uma tensão de ruptura média de $\bar{x} = 70,57$ kgf que foram obtidos por meio de testes físicos. Testes anterior indicam um desvio padrão da tensão de ruptura de $\sigma = 5,59$. Teste a afirmação de que os implantes possuem uma tensão de ruptura superior a 68 kgf, ao nível de 0,05 de significância.

Solução:

1º Passo) *Afirm : $\mu > 68$ kgf*

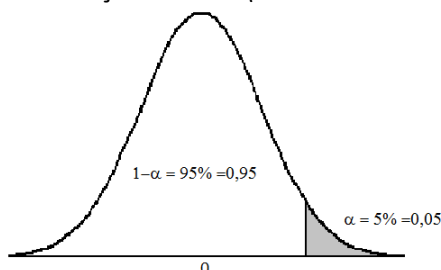
2º Passo) *$H_0 : \mu \leq 68$*

$H_A : \mu > 68$

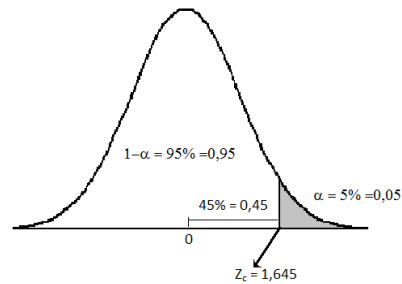
3º Passo) *Teste Unilateral Direito (A Hipótese Alternativa é ">")*

4º Passo) *Distribuição Normal (Além do $n > 30$ o desvio padrão populacional é CONHECIDO)*

5º Passo)

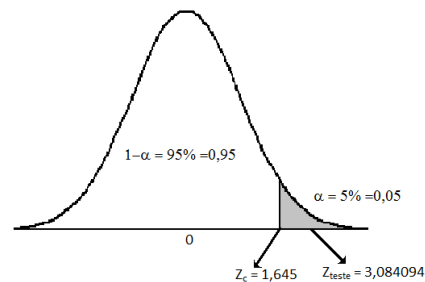


6º Passo) $Z_c = 1,645$ (da tabela da Distribuição Normal com uma probabilidade de 45%)



7º Passo)

$$Z_{Teste} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{70,57 - 68}{\frac{5,59}{\sqrt{45}}} = 3,084094$$



8º Passo) Rejeitar H_0 , pois $Z_{Teste} \in RC$

9º Passo) **Não rejeitar** a afirmação de que os implantes possuem uma tensão de ruptura superior a 68 kgf

11.5.2 - TESTE DE HIPÓTESE PELO MÉTODO DO INTERVALO DE CONFIANÇA

O Método do Intervalo de Confiança consiste em verificar se a Hipótese Nula pertence ao Intervalo de Confiança ou não. Se a Hipótese Nula pertencer ao Intervalo de Confiança ela não deve ser rejeitada, caso contrário a Hipótese nula será rejeitada.

➤ Procedimentos para um Teste de Hipótese pelo Método do Intervalo de Confiança:

- 1º - Extrair a afirmação do problema e representá-la em linguagem matemática;
- 2º - Construir as Hipóteses Nula e Alternativa;
- 3º - Determinar o tipo de teste: Bilateral, Unilateral Direito ou Unilateral Esquerdo;
- 4º - Determinar qual distribuição deve ser utilizada: Normal, T-Student ou Qui-Quadrado;
- 5º - Construir o Intervalo de Confiança;
- 6º - Verificar se a Hipótese Nula pertence ou não ao Intervalo de Confiança e decidir se a Hipótese Nula deve ser Rejeitada ou não;
- 7º - Comparar a Hipótese Nula com a afirmação do problema e decidir se a afirmação deve ser rejeitada ou não.

Observação: No caso de testes unilaterais deve-se construir o Intervalo de Confiança criando a região crítica que está faltando, sem, no entanto, alterar a significância do teste. Isto é, deve-se criar a outra região crítica com a mesma significância do teste.

Exemplo 11.3 - Um determinado fabricante produz implantes mamários utilizando gel de silicone. Uma amostra de 45 implantes apresentou uma tensão de ruptura média de $\bar{x} = 70,57$ kgf que foram obtidos por meio de testes físicos. Testes anterior indicam um desvio padrão da tensão de ruptura de $\sigma = 5,59$. Teste a afirmação de que os implantes possuem uma tensão de ruptura superior a 68 kgf, ao nível de 0,05 de significância.

Solução:

1º Passo) *Afirm* : $\mu > 68 \text{ kgf}$

2º Passo) $H_0 : \mu \leq 68$

$H_A : \mu > 68$

3º Passo) *Teste Unilateral Direito (A Hipótese Alternativa é ">")*

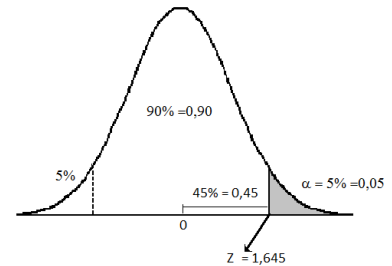
4º Passo) *Distribuição Normal (Além do $n > 30$ o desvio padrão populacional é CONHECIDO)*

5º Passo)
$$E = Z \frac{\sigma}{\sqrt{n}} = 1,645 \times \frac{5,59}{\sqrt{45}} = 1,370791659$$

$$\bar{x} - E = 70,57 - 1,370791659 = 69,19920834$$

$$\bar{x} + E = 70,57 + 1,370791659 = 71,94079166$$

$$IC(\mu; 90\%) = (69,199; 71,941)$$



6º Passo) *Rejeitar H_0 , pois $H_0 \notin IC$*

7º Passo) **Não rejeitar** a afirmação de que os implantes possuem uma tensão de ruptura superior a 68 kgf

11.5.3 - TESTE DE HIPÓTESES PELO MÉTODO DO VALOR P

O Método do Valor P consiste em calcular uma probabilidade e compará-la com o nível de significância. Se o Valor P calculado for maior ou igual à significância a Hipótese Nula não deve ser rejeitada, se for menor, a Hipótese Nula será rejeitada.

➤ Procedimentos para um Teste de Hipótese pelo Método Tradicional:

- 1º - Extrair a afirmação do problema e representá-la em linguagem matemática;
- 2º - Construir as Hipóteses Nula e Alternativa;
- 3º - Determinar o tipo de teste: Bilateral, Unilateral Direito ou Unilateral Esquerdo;
- 4º - Determinar qual distribuição deve ser utilizada: Normal, T-Student ou Qui-Quadrado;
- 5º - Calcular a estatística de teste apropriada;
- 6º - Calcular o Valor P. Se o teste for Unilateral Direito $P = P(\theta > \theta_{Teste})$, se o teste for Unilateral Esquerdo $P = P(\theta < \theta_{Teste})$, se o teste for Bilateral $P = 2 \times P(\theta < \theta_{Teste})$ se a estatística de teste estiver mais próxima da região crítica da esquerda e será $P = 2 \times P(\theta > \theta_{Teste})$ se a estatística de teste estiver mais próxima da região crítica da direita.
- 7º - Comparar o Valor P calculado com a significância. Se $P > \alpha$ a Hipótese Nula Não deve ser Rejeitada, se $P \leq \alpha$ a Hipótese Nula deve ser rejeitada;
- 8º - Comparar a Hipótese Nula com a afirmação do problema e decidir se a afirmação deve ser rejeitada ou não.

Exemplo 11.4 - Um determinado fabricante produz implantes mamários utilizando gel de silicone. Uma amostra de 45 implantes apresentou uma tensão de ruptura média de $\bar{x} = 70,57 \text{ kgf}$ que foram obtidos por meio de testes físicos. Testes anterior indicam um desvio padrão da tensão de ruptura de $\sigma = 5,59$. Teste a afirmação de que os implantes possuem uma tensão de ruptura superior a 68 kgf, ao nível de 0,05 de significância.

Solução:

1º Passo) *Afirm : $\mu > 68$ kgf*

2º Passo) *$H_0 : \mu \leq 68$*

$H_A : \mu > 68$

3º Passo) *Teste Unilateral Direito (A Hipótese Alternativa é ">")*

4º Passo) *Distribuição Normal (Além do $n > 30$ o desvio padrão populacional é CONHECIDO)*

$$5^\circ \text{ Passo) } Z_{\text{Teste}} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{70,57 - 68}{\frac{5,59}{\sqrt{45}}} = 3,084094$$

$$P = P(Z > Z_{\text{Teste}}) = P(Z > 3,08409375) = P(Z > 3,085)$$

$$P = 0,5 - P(0 < Z < 3,085) = 0,5 - 0,498962 = 0,001038$$

6º Passo) *Rejeitar H_0 , pois $P \leq \alpha$*

9º Passo) **Não rejeitar** a afirmação de que os implantes possuem uma tensão de ruptura superior a 68 kgf

11.6 - EXERCÍCIOS PROPOSTOS

11.6.1) Um teste de hipótese foi aplicado e, ao nível de significância de 5%, rejeitou-se H_0 . O que acontecerá, se forem adotados os níveis de significância de 1% e de 10%, respectivamente:

- Aceitar-se-á a 1% e rejeitar-se-á a 10%.
- Nada se pode afirmar em ambos os casos.
- Nada se pode afirmar quanto ao de 1% e rejeitar-se-á H_0 a 10%.
- Rejeitar-se-á H_0 a 1% e nada se pode afirmar quanto ao de 10%.
- Rejeitar-se-á H_0 em ambos os casos.

➤ **Teste de Hipóteses para a Média Populacional μ**

11.6.2) Estudantes de estatística do autor selecionados aleatoriamente participaram de um experimento para testar a habilidade de determinarem quando 1 minuto (ou 60 segundos) havia se passado. Quarenta estudantes resultaram em uma média amostral de 58,3 segundos. Supondo que $\sigma = 9,5$ segundos, use o nível significância de 0,05 para testar a hipótese afirmativa de que a média populacional seja igual a 60 segundos. Com base no resultado, a percepção geral de um minuto parece razoavelmente precisa?

11.6.3) A análise dos últimos dígitos de dados amostrais revela, algumas vezes, se esses dados foram criteriosamente medidos e relatados. Como se seleciona aleatoriamente, com reposição, dígitos de 0 a 9, a média deve ser 4,50 e o desvio padrão deve ser 2,87. Dados relatados (tais como pesos ou alturas) são, em geral, arredondados, de modo que os últimos dígitos incluem, de modo desproporcional, mais dígitos 0 e 5. Use os últimos dígitos de alturas relatadas (em pés) dos 73 *home runs* batidos por Barry Bonds em 2001, para testar afirmativa de que eles sejam provenientes de uma população com média de 4,50. Use o nível significância de 0,05. As distâncias dos 73 *home runs* (listadas no Conjunto de Dados 17 no Apêndice B) têm últimos dígitos com média de 1,753. Com base nos resultados, parece que as distâncias têm sido acuradamente medidas?

- 11.6.4) Quando 40 pessoas usaram a dieta Atkins por um ano, a mudança média de seus pesos foi de -2,1 lb (com base em dados de "*Comparison of the Atkins, Ornish, Weigh Watchers, and Zone Diets for Weight Loss and Heart Disease Reduction*", de Dansinger et al., *Journal of the American Medical Association*, Vol. 293, No. 1). Suponha que o desvio padrão de todas as mudanças de peso seja $\sigma = 4,8 \text{ lb}$ e use o nível de significância de 0,05 para testar afirmativa de que a mudança média de peso seja menor do que 0. Com base nos resultados, a dieta parece ser eficaz? A mudança média de peso parece ser substancial o suficiente para justificar a dieta especial?
- 11.6.5) A carga axial de uma lata de alumínio é o peso máximo que os lados podem suportar antes de cederem. A carga axial é uma medida importante, porque as beiradas superiores são pressionadas sobre os lados com pressões que variam entre 158 lb e 165 lb. A Pepsi fez a experiência com latas de alumínio mais fino, e umas amostras aleatórias de 175 dessas latas têm uma carga axial média de 267,11 lb. As latas padrões têm uma carga axial média de 281,81 lb e um desvio padrão de 27,77 lb. Use o nível de significância de 0,01 para testar a afirmativa de que as latas mais finas têm uma carga axial média menor do que 281,81 lb. Suponha $\sigma = 27,77 \text{ lb}$. As latas mais finas parecem ter uma carga axial média menor do que dos 281,81 lb? As latas mais finas parecem ser fortes o bastante para não cederem quando as beiradas superiores são pressionadas sobre os lados?
- 11.6.6) Quando 14 estudantes do segundo ano de medicina do Hospital de Bellevue mediram a pressão sanguínea sistólica da mesma pessoa, obtiveram os resultados listados abaixo (em mmHg). Supondo que o desvio padrão da população seja conhecido e igual a 10 mmHg, use o nível de significância de 0,05 para testar afirmativa de que o nível da pressão sanguínea seja de 140 mmHg. A hipertensão é definida como um nível de pressão sanguínea muito alta, maior ou igual a 140 mmHg. Com base nos resultados do teste de hipótese, podemos concluir com segurança que a pessoa não seja hipertensa?

138 130 135 140 120 125 120 130 130 144 143 140 130 150

- 11.6.7) O menor mamífero do mundo é o morcego mamangaba, também conhecido como morcego de focinho de porco (ou *Craseonucteris thonglongyai*). Tais morcegos são, no máximo, do tamanho de um mamangaba (ou abelhão). Abaixo, estão listados pesos (em gramas) de uma amostra desses morcegos. Supondo que tais pesos têm um desvio padrão de 0,30 g, use o nível significância de 0,05 para testar a afirmativa de que esses morcegos são de uma mesma população com um peso médio de 1,8 g. Os morcegos parecem provir de uma mesma população?

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

- 11.6.8) Use os pesos de moedas de um quarto de dólar cunhados depois de 1964, listados no Conjunto de Dados 14 do Apêndice B. Supondo que as moedas sejam cunhadas para terem pesos com desvio padrão populacional de 0,068 g, use os pesos amostrais com nível de significância de 0,01 para testar a afirmativa de que as moedas de um quarto de dólar sejam provenientes de uma população com uma média de 5,670 g. Os quartos de dólar do conjunto parecem ter sido fabricados de acordo com a especificação americana de cunhagem de que a média seja igual a 5,670 g?

- 11.6.9) Consulte o Conjunto de Dados 8 do Apêndice B e subtraia cada temperatura máxima real da temperatura máxima prevista um dia antes. O resultado é uma lista de erros. Supondo que todos esses erros tenham um desvio padrão de $2,5^\circ$, use o que nível de significância de 0,05 para testar a afirmativa de que tais erros tenham uma média igual a 0. O que o resultado sugere sobre a precisão das temperaturas previstas?
- 11.6.10) O Conjunto de Dados 2 no Apêndice B inclui 107 temperaturas corporais (Temperatura Dia 2 – 12 H) com uma média de $98,20^\circ\text{F}$ e um desvio padrão de $0,68^\circ\text{F}$. Use o nível significância de 0,05 para testar a afirmativa de que a temperatura média do corpo é menor do que $98,6^\circ\text{F}$. Com base nesses resultados, parece errada a média de $98,6^\circ\text{F}$ comumente usada?
- 11.6.11) Em testes anteriores, bolas de beisebol foram jogadas de uma altura de 24 pés sobre uma superfície de concreto e quicaram a uma altura média de 92,84 in. Em um teste de uma amostra de 40 novas bolas, as alturas do quiques tiveram uma média de 92,67 in e um desvio padrão de 1,79 in (com base em dados do Laboratório Nacional de Brookhaven e do *USA Today*). Use o nível de significância de 0,05 para determinar se há evidência suficiente para apoiar a afirmativa de que as novas bolas têm uma altura de quiques com uma média diferente 92,84 in. Será que as novas bolas são diferentes?
- 11.6.12) Em um estudo relativo aos efeitos sobre bebês do uso de cocaína no período pré-natal, obtiveram-se os seguintes dados amostrais para os pesos ao nascer: $n=190$; $\bar{x}=2700\text{g}$, $s=645\text{g}$ (com base em dados de “*Cognitive Outcomes of Preschool Children with Prenatal Cocaine Exposure*”, de Singer ET al., *Journal of the American Medical Association*, Vol. 291, No. 20). Use o nível de significância de 0,01 para testar a afirmativa de que os pesos de bebês que nascem de mães usuárias de cocaína têm uma média menor do que a média de 3.103 g para bebês nascidos de mães que não usam cocaína. Com base nos resultados, parece que os pesos ao nascer são afetados pelo uso de cocaína?
- 11.6.13) Quando consumidores solicitam crédito, seus níveis de crédito são classificados usando-se os escores FICO (*Fair, Isaac and Company*). As classificações de crédito para uma amostra de pretendentes a empréstimos para compra de carro são apresentadas abaixo. Obtêm-se uma amostra aleatória de classificação de crédito e os escores FICO são resumidos com essas estatísticas: $n=18$; $\bar{x}=660,3$, $s=95,9$. Use o nível de significância de 0,05 para testar a afirmativa de que as classificações de crédito são provenientes de uma população com média igual a 700. Se o Banco de Newport pede uma classificação de 700 ou mais para um empréstimo para compra de carro, os resultados indicam que todas as pessoas tenham a chance de conseguir tal empréstimo? Por que sim ou por que não?
- 11.6.14) Pacientes com a síndrome da fadiga crônica foram testados e depois retestados após serem tratados com fludrocortisona. Uma escala padrão de -7 a +7 é usada para medir a fadiga antes e depois do tratamento. As mudanças são resumidas por estas estatísticas: $n=21$; $\bar{x}=4,00$, $s=2,17$ (com base em dados de “*The Relationship Between Neurally Mediated Hypotension and the Chronic Fatigue Syndrome*”, de Bou-Holaigah, Rowe, Kan e Calkins, *Journal of the American Medical Association*, Vol. 274, No. 12). As mudanças foram calculadas de maneira tal que os valores positivos representem melhoras. Use o nível de significância de 0,01 para testar a afirmativa de que a mudança média seja positiva. O tratamento parece eficaz?

- 11.6.15) Quarenta sujeitos seguiram a dieta dos Vigilantes do Peso por um ano. Suas mudanças de peso estão resumidas por estas estatísticas: $\bar{x} = -6,6 \text{ lb}$, $s = 10,8 \text{ lb}$ (com base em dados de “*Comparison of the Atkins, Ornish, Weight Watchers, and Zone Diets for Weight Loss and Heart Disease Risk Reduction*”, de Dansinger et al., *Journal of the American Medical Association*, Vol. 293, No. 1). Use o nível e significância de 0,01 para testar a afirmativa de que a dieta não tenha qualquer efeito. Com base nos resultados, a dieta parece eficaz?
- 11.6.16) Medem-se as alturas das supermodelos Niki Taylor, Nadia Avermann, Claudia Schiffer, Elle MacPherson, Christy Turlington, Bridget Hall, Kate Moss, Valeria Mazza e Kristy Hume. Elas têm uma média de 70,2 in (ou 178,308 cm) e um desvio padrão de 1,5 in. Use o nível significância de 0,01 para testar a afirmativa de que as supermodelos têm altura com média superior à média de 63,6 in (ou 161,544 cm) para mulheres da população em geral. Uma vez que há apenas nove alturas representadas, podemos realmente comente concluir que é as supermodelos sejam mais altas do que as mulheres típicas?
- 11.6.17) Um artigo no *New York Times* observou que a vida média de 35 regentes homens de orquestras sinfônicas era de 73,4 anos, em contraste com a vida média de 69,5 anos para os homens da população geral. Supondo que os 35 homens tenham tempo médio de vida com um desvio padrão de 8,7 anos, use o nível de significância de 0,05 para testar a afirmativa de que regentes homens de orquestras sinfônicas têm um tempo médio de vida maior do que 69,5 com anos. Parece que os regentes homens de orquestras sinfônicas vivem mais do que os homens em geral? Por que a experiência de ser um regente de orquestra não faz com que os homens vivam mais? (Sugestão: Os homens que são regentes de orquestra nascem regentes ou se tornam regentes depois de muito tempo?)
- 11.6.18) A seguir, estão listadas as quantidades medidas de chumbo (em microgramas por metro cúbico ou $\mu\text{g} / \text{m}^3$) no ar. A Agência de Proteção Ambiental estabeleceu um padrão de qualidade do ar para o chumbo de $1,5 \mu\text{g} / \text{m}^3$. As medidas mostradas a seguir foram registradas no Edifício 5 do local do World Trade Center, em dias diferentes imediatamente após a destruição causada pelos ataques terroristas de 11 de setembro de 2001. Após o colapso dos dois edifícios do World Trade Center, houve muita preocupação sobre a qualidade do ar. Use o nível de significância 0,05 para testar a afirmativa de que a amostra provém de uma população com uma média maior do que o padrão da APA de $1,5 \mu\text{g} / \text{m}^3$. Há alguma coisa sobre esse conjunto de dados que sugira que a suposição de uma população normalmente distribuída possa não ser válida?

5,40 1,10 0,42 0,73 0,48 1,10

- 11.6.19) Selecionou-se cereais aleatoriamente e mede-se o conteúdo de açúcar (gramas de açúcar por grama de cereal) para cada cereal, com os resultados dados abaixo para Cheerios, Harmony, Smart Start, Cocoa Puffs, Lucky Charms, Corn Flakes, Fruit Loops, Wheaties, Cap'n Crunch, Frosted Flakes, Apple Jacks, Bran Flakes, Special K, Rice Krispies, Corn Pops e Trix. Use o nível de significância de 0,05 para testar a afirmativa de um lobista de que a média para todos cereais é menor do que 0,3 g.

0,03 0,24 0,30 0,47 0,43 0,07 0,47 0,13
0,44 0,39 0,48 0,17 0,13 0,09 0,45 0,43

- 11.6.20) A abaixo estão listados tempos (em segundos) de homens de vencedores do 100 metros rasos nos jogos olímpicos de verão consecutivos, listados em ordem por linha. Supondo que esses resultados sejam dados amostrais selecionados aleatoriamente de uma população de todos passados e futuros jogos Olímpicos, teste a afirmativa de que o tempo médio é menor do que a 10,5 s. O que você observa sobre a precisão dos números? A Qual característica extremamente importante do conjunto de dados não está sendo considerada no teste de hipótese? Os resultados do teste de hipótese sugerem que os futuros tempos dos campeões devam ficar em torno de 10,5 s? Essa conclusão é válida? (Considere uma significância de 5%)

12,0 11,0 11,0 11,2 10,8 10,8 10,8 10,6 10,8 10,3 10,3 10,3
10,4 10,5 10,2 10,0 9,95 10,14 10,06 10,25 9,99 9,92 9,96

- 11.6.21) Use os pesos de moedas de quartos de dólar cunhados depois de 1964 do Conjunto de Dados 14 do Apêndice B. Teste a afirmativa de que esses quartos de dólar sejam fabricados de acordo com a especificação americana de cunhagem de peso médio e igual a 5,670 g. (Considere uma significância de 0,01)
- 11.6.22) Consulte o Conjunto de Dados 8 do Apêndice B e subtraia a temperatura máxima real da temperatura máxima prevista um dia antes. O resultado é uma lista de erros. Teste a afirmativa de que todos esses erros tenham uma média igual a 0. O que o resultado sugere sobre a precisão das temperaturas previstas? (Considere uma significância de 0,05)
- 11.6.23) O autor, no auge de um programa de exercícios, afirmou que sua taxa de pulsação era menor do que a taxa média de pulsação de um homem típico. A taxa de pulsação do autor foi medida em 60 batimentos por minuto. Use as taxas de pulsação de homens listadas no Conjunto de Dados 1 do Apêndice B e teste a afirmativa de que tais taxas de pulsação sejam provenientes de uma população com uma média superior a 60. (Considere uma significância de 0,05)
- 11.6.24) Consulte o Conjunto de Dados 5 do Apêndice B e use apenas os filmes que mostram o uso do fumo. Teste a afirmativa de um crítico de cinema de que "entre tais filmes que mostram o uso do fumo, a exposição média é de 2 minutos". Em face dos dados amostrais, a afirmativa é enganosa? (Considere uma significância de 0,05)
- 11.6.25) O gerente de uma empresa de transporte suspeita da afirmação de um vendedor de pneus de que o seu produto tem uma vida média de 28.000 milhas. Para verificar a afirmação, a firma instala 40 desses pneus em seus caminhões, obtendo uma vida média de 27.563 milhas, com desvio padrão de 1.348 milhas. Com 99% de confiança, qual a conclusão do gerente?
- 11.6.26) Um estudante de direito deseja verificar a alegação de seu professor de que os condenados por crime de falsificação passam, em média, 8,8 meses na cadeia. Para isto, obtém uma amostra de 70 desses casos dos arquivos do tribunal e encontra média de 8,2 meses e desvio padrão de 2,2 meses. Pode-se concluir com 95% de confiança que o tempo encontrado é diferente do informado pelo professor?
- 11.6.27) Você é responsável por um museu na Grã-Bretanha que mostra como as pessoas viviam na América. Cada sala apresenta uma cena diferente. Você sempre preferiu as salas relativas aos índios americanos e pediu, então, ao diretor para aumentar esse espaço, restringindo a parte destinada à história da Virgínia. O diretor não gostou da idéia e insiste em que os visitantes despendem em média 3,5 minutos nessa sala, o que é bastante, em termos dos padrões do museu. Você monitora uma amostra de 25 visitantes e acha uma média de 2,9 minutos, com um desvio padrão de 1 minuto. Os resultados de sua amostra estão de acordo com a afirmativa de seu diretor? Teste com nível de confiança de 95%.

- 11.6.28) Os salários dos funcionários de uma empresa apresentam uma média histórica de 4 salários mínimos e desvio padrão de 1,2 salários mínimo. Uma política de aumento dos salários mais baixos combinada com a contenção dos salários mais altos visa diminuir as diferenças salariais, implementando uma distribuição salarial melhor. Após algum tempo de execução desta política, uma amostra de 25 salários selecionados forneceu média de 5 salários mínimos. Teste ao nível de 99% de confiança, se a política de aumento dos salários foi efetiva.
- 11.6.29) Um industrial alega que com o uso de um aditivo seria possível rodar mais 3 km, em média, com um litro de combustível. Uma amostra de 100 automóveis foi usada para avaliar este produto e a média de aumento ficou em 2,4 e o desvio padrão em 1,8 km por litro. Teste a hipótese de que a média da população é de 3 km por litro, com nível de confiança de 95%.
- 11.6.30) Perguntou-se a uma amostra de 1562 estudantes de graduação que respondessem numa escala de um (discordo frontalmente) a sete (concordo inteiramente) à proposição: "Propaganda pode elevar nosso padrão de vida". A resposta média da amostra foi de 4,27 e o desvio padrão foi de 1,32. Teste ao nível de 1% de significância, num teste bicaudal, a hipótese de que a média da população é de 4.
- 11.6.31) A contabilidade de uma empresa revela que na média, suas contas a receber são de \$125,32. Um auditor checkou uma amostra de 16 dessas contas e a média encontrada foi de \$131,78 e o desvio padrão de \$25,41. Teste ao nível de significância de 5%, a hipótese de que a média da população é de \$125,32 (teste bicaudal).
- 11.6.32) Um distribuidor de cerveja alega que um novo rótulo aumentará as vendas nos supermercados numa média de 50 caixas por semana com desvio padrão de 12,2 caixas. De uma amostra de 20 supermercados, constatou-se um aumento médio de 41,3 caixas. Teste ao nível de 5% de significância a hipótese de que a média de crescimento da população é menor do que 50 caixas.
- 11.6.33) A gerência de uma empresa afirma que os bônus pagos aos funcionários são, em média, de R\$ 650. Você coleta uma amostra de 100 funcionários que trabalham nesta empresa e verifica que eles receberam bônus de, em média, R\$ 600, com um desvio padrão de R\$ 100. Pergunta-se:
- Faça um teste ao nível de 5% de significância e comente a afirmativa da gerência.
 - Estabeleça um intervalo de confiança para a média populacional dos bônus pagos aos funcionários com nível de confiança de 99%.
 - Um gerente estabeleceu um intervalo de confiança para a média populacional como sendo de 582 a 618. Calcule o nível de confiança utilizado.
 - Calcule o tamanho da amostra para atender o intervalo 590 a 610 com nível de confiança de 90%.
- 11.6.34) Uma empresa recebeu de seu fornecedor um lote de 800 produtos para a venda. Interessada em analisar o lote, a empresa amostrou 20% do lote e encontrou uma média de 50 com desvio padrão igual a 20.
- Calcule o intervalo de confiança para a média populacional com 90% de confiança.
 - Um gerente encontrou um intervalo de confiança como sendo de 49 a 51. Calcule o nível de confiança utilizado.
 - Calcule o tamanho da amostra para atender ao intervalo 47,8 a 52,2, a 99% de certeza.
 - A gerência afirma que espera para o lote uma média de 53. Teste com 95% de confiança se a informação do gerente está correta com base na amostra utilizada.

- 11.6.35) Um processo, para estar operando sob controle, deve apresentar uma média de 1000. Um analista amostrou 25 produtos e encontrou média igual a 1120 e desvio padrão de 300. Pedese:
- Com 10% de significância, teste a partir da amostra obtida se o processo está sob controle.
 - Encontre um intervalo de confiança para a média populacional. Faça nível de confiança de 95%.
- 11.6.36) De um lote com 6000 peças, retira-se uma amostra de 500 peças. Após análise verificou-se que a durabilidade média das peças era de 2000 horas com desvio padrão de 800 horas. A gerência espera uma durabilidade média para o lote de 1900 horas conforme previsto juntamente com o fornecedor.
- Estabeleça um intervalo de confiança para a durabilidade média do lote com 90% de confiança.
 - Visando aumentar o tamanho da amostra, a diretoria sugeriu que se trabalhasse com uma margem de erro de 15 horas ao invés do valor encontrado na letra a desse exercício. Isto é possível a 90% de confiança? Justifique.
 - Teste com 99% de confiança se a informação da gerência é diferente da durabilidade média encontrada na amostra.
- 11.6.37) Um processo industrial produz peças que têm uma média de 40 cm quando o processo está sob controle. O processo tem um desvio padrão conhecido de 2,2 cm, que permanece constante, quer o processo esteja sob controle quer não. Uma amostra de 18 peças apresentou uma média de 42 cm. A partir dessas informações, pede-se:
- Estabeleça um intervalo de confiança para a média populacional das peças com nível de confiança de 99%.
 - Se a margem de erro for alterada para 0,70 cm, qual deverá ser o tamanho da amostra com nível de confiança de 99%?
 - A diretoria argumenta que a média encontrada para a amostra de 18 peças não é verdadeira. Teste a afirmação da diretoria com nível de confiança de 95%.
- 11.6.38) Após o último reajuste no preço da gasolina, a Secretaria Especial de Acompanhamento de Preços relatou que o novo preço médio deste produto é de R\$ 2,37. Por meio da informação de que os postos de bandeira branca e de pequenos distribuidores são os que oferecem gasolina mais barata em BH, uma pesquisa foi desenvolvida em uma amostra de 24 postos e encontrou-se um preço médio de R\$ 2,31 com desvio padrão de R\$ 0,38.
- Teste com nível de significância de 10% se os preços dos postos sob a bandeira branca e de pequenos distribuidores são menores do que os demais.
 - Estabeleça um intervalo de confiança para o preço médio da gasolina em BH, com nível de confiança de 80%.

➤ Teste de Hipótese para Proporção Populacional

- 11.6.39) Uma empresa de entregas afirma que 95% das encomendas “especiais” são entregues no mesmo dia, se remetidas antes das 10:00 h da manhã. Você envia uma amostra de 50 encomendas e observa que a proporção de encomendas entregue no mesmo dia foi de 92%. Pode-se concluir com 99% de confiança que este serviço é mesmo “especial” como a empresa afirma?
- 11.6.40) Uma empresa de consultoria recebe trimestralmente 100 projetos para avaliação de viabilidade econômica. Um indicador de que a economia está em expansão é o aumento do número de projetos viáveis enviados para a avaliação, o que mostra a intenção efetiva de investimento

das empresas. Uma amostra de 31 projetos foi selecionada e avaliada. Obteve-se uma proporção de 60% de projetos viáveis. Pode-se concluir com 90% de confiança de que a proporção é ainda de 50% de projetos economicamente viáveis, conforme dados históricos?

- 11.6.41) Um banco de investimento calcula que pelo menos 20% dos empréstimos são concedidos a pequenas empresas. Com a finalidade de aumentar esta porcentagem, o governo resolveu subsidiar os juros para este tipo de empresa. Após algum tempo de vigência desta política, o banco amostrou 40 projetos de financiamento e verificou que 12 deles se destinavam a pequenas empresas. Teste ao nível de confiança de 80% a hipótese de que a política tenha sido bem sucedida.
- 11.6.42) De uma amostra de 361 proprietários de empresas de varejo que faliram, 105 disseram não ter tido qualquer assessoria prévia antes de abrir um mal fadado negócio. Teste a hipótese de que 25% de todos os membros desta população não tiveram assistência profissional antes de abrir o negócio com nível de confiança de 90%.
- 11.6.43) Numa amostra de 1000 adultos, 17,3% indicaram desacordo com a declaração: “O capitalismo é mais do que um sistema econômico – ele é um complexo de instituições, atitudes e cultura”. Teste ao nível de 5% de significância a hipótese de que 25% da população não concordam com a declaração.
- 11.6.44) Numa amostra de 160 executivos, 72 indicaram alguma medida de desacordo com a declaração: “Uma reputação de conduta ética é menos importante para que um gerente consiga promoções do que uma reputação de ganhador de dinheiro para a empresa”. Teste a hipótese de que metade de todos os executivos concordaria com a declaração num teste bicaudal, com nível de confiança de 99%.
- 11.6.45) Numa amostra de 200 auditores, 105 indicaram alguma medida de desacordo com a declaração: “Fluxo de caixa das operações é uma medida válida de rentabilidade”. Teste ao nível de 10% de significância (teste bicaudal) a hipótese de que metade dos membros desta população concorda com a declaração.
- 11.6.46) Um estudo está sendo feito visando a satisfação dos consumidores frente ao lançamento de um novo produto. Para isso, em um supermercado, uma equipe de vendas selecionou uma amostra de 1200 pessoas e, após conhecer o produto, 660 pessoas mencionaram que ficaram satisfeitas com o novo produto.
 - a) Estabeleça um intervalo de confiança para a proporção de todas as pessoas que estão satisfeitas com o novo produto, com nível de confiança de 99%.
 - b) O gerente de vendas apresentou um intervalo para a diretoria dos clientes que estão insatisfeitos com o produto como sendo de 0,43 a 0,47. Calcule o nível de confiança adotado pelo gerente.
 - c) A diretoria menciona que a satisfação dos clientes, conforme dados históricos, é de 61%. Teste com 90% de confiança se a informação do diretor está correta com base no estudo realizado.
- 11.6.47) Uma empresa produz anualmente 32.000 produtos. Após estudos com uma amostra de 2500 produtos verificou-se que 1000 deles estavam fora da especificação. Pergunta-se:
 - a) Qual o intervalo de confiança de 99% para a proporção dos produtos produzidos anualmente dentro da especificação?
 - b) Quando se interpreta o intervalo encontrado na letra “a” desse exercício, encontramos qual erro?

- c) Com que nível de confiança dir-se-ia que a proporção dos produtos fora da especificação é de 0,38 a 0,42?
 - d) Que tamanho deve ter a amostra para que seja de 95% a confiança na estimativa da proporção dos produtos fora da especificação como sendo de $40 \pm 1,40$?
 - e) A empresa tem divulgado que sua produção tem apresentado nos últimos anos 35% de rejeição, ou seja, de produtos fora da especificação. Teste se a afirmação da empresa é diferente do encontrado no estudo amostral, com nível de confiança de 90%.
- 11.6.48) O *Genetics and IVF Institute* realizou um experimento clínico do método XSORT, projetado para aumentar a probabilidade de concepção de bebês do sexo feminino. Quando este livro estava sendo escrito, 325 bebês nasceram de pais que usavam o método XSORT, e 295 deles eram meninas. Use os dados amostrais com nível de significância de 0,01 para testar a afirmativa de que, com esse método, a probabilidade de um bebê ser uma menina é maior do que 0,5. O método parece funcionar?
- 11.6.49) O *Genetics and IVF Institute* realizou um experimento clínico do método YSORT, projetado para aumentar a probabilidade de concepção de bebês do sexo masculino. Quando este livro estava sendo escrito, 51 bebês nasceram de pais que usavam o método YSORT, e 39 deles eram meninos. Use os dados amostrais com um nível de significância de 0,01 para testar a afirmativa de que, com esse método, a probabilidade de um bebê ser um menino é maior do que 0,5. O método parece funcionar?
- 11.6.50) Em um estudo de 11.000 acidentes de carro, encontrou-se que 5720 deles ocorreram a menos de 5 milhas de casa (com base em dados do *Progressive Insurance*). Use um nível de significância de 0,01 para testar a afirmativa de que mais de 50% dos acidentes de carro ocorrem a menos de 5 milhas de casa. Os resultados são questionáveis por se basearem em uma pesquisa patrocinada por uma companhia de seguros?
- 11.6.51) Entre 734 usuários de Internet selecionados aleatoriamente, verificou-se que 360 deles usam a Internet para fazerem planos de viagem (com base em dados de uma pesquisa do Gallup). Use um nível de significância de 0,01 para testar a afirmativa de que, entre usuários da Internet, menos de 50% usam-na para fazer planos de viagem. Os resultados são importantes para os agentes de viagem?
- 11.6.52) A tecnologia está mudando drasticamente a forma como nos comunicamos. Em 1997, uma pesquisa em 880 domicílios americanos mostrou que 149 deles usavam e-mail (conforme dados do *The World Almanac and Book of Facts*). Use esses resultados amostrais para testar a afirmativa de que, em mais de 15% dos domicílios americanos, usa-se o e-mail. Use o nível de significância de 0,05. A conclusão é válida ainda hoje? Por que sim ou por que não?
- 11.6.53) Em 1990, 5,8% dos candidatos a emprego que foram submetidos ao teste de drogas não passaram no teste. Ao nível de significância de 0,01, teste a afirmativa de que a taxa de falha agora é menor, se uma amostra aleatória simples de 1520 candidatos atuais a emprego resultou em 58 falhas (conforme dados da Associação de Administração Americana). O resultado parece sugerir que menos candidatos a emprego agora usam drogas?
- 11.6.54) Em um estudo com 420.095 usuários de telefone celular na Dinamarca, 135 pessoas desenvolveram câncer de cérebro ou do sistema nervoso (conforme dados do *Journal of the National Cancer Institute*, como relatado no *USA Today*). Teste a afirmativa de uma crença popular que diz que tais cânceres são afetados pelo uso do telefone celular. Isto é, teste a afirmativa de que os usuários de telefone celular desenvolvem câncer de cérebro ou do sistema nervoso a uma taxa que é diferente da taxa de 0,0340% entre os não-usuários de telefone celular.

Como essa questão é de grande importância, use um nível de significância de 0,005. Os usuários de telefone celular devem se preocupar sobre o câncer de cérebro ou do sistema nervoso?

- 11.6.55) Em um estudo com fumantes que tentaram parar de fumar com a terapia dos adesivos de nicotina, 39 estavam fumando após um ano de tratamento e 32 não estavam fumando um ano após o tratamento (conforme dados de *"High-Dose nicotine Patch Therapy"*, de Dale ET al., *Journal of the American medical Association*, Vol. 274, No. 17). Use o nível de significância de 0,10 para testar a afirmativa de que, entre fumantes que tentam deixar o vício com a terapia do adesivo, a maioria volta a fumar um ano após o tratamento. Esses resultados sugerem que a terapia dos adesivos de nicotina seja eficaz?
- 11.6.56) Em um estudo dos leitores de códigos de barras nas lojas, 1234 itens foram verificados, constatando-se que vinte deles foram cobrados excesso e 1214 itens verificados não o foram (com base em dados de *"UPC Scanner Pricing Systems: Are they Accurate?"* de Goodstein, *Journal of Marketing*, Vol. 58). Use o nível significância 0,05 para testar afirmativa de que, com leitores, 1% das vendas é cobrado com excesso. (Antes dos leitores, a taxa de cobrança excessiva era estimada em cerca de 1%.) Com base nesses resultados, os leitores parecem auxiliar o consumidor a evitar cobrança excessiva?
- 11.6.57) Uma hipótese popular interessante é de que as pessoas podem, temporariamente, adiar suas mortes para sobreviverem ao feriado ou a um evento importante, como um aniversário. Em um estudo desse fenômeno, verificou-se ter havido 6062 mortes na semana anterior ao dia de Ação de Graças, e 5938 mortes na semana seguinte ao dia de Ação de Graças (com base em dados de *"Holidays, Birthdays, and Postponement of Cancer Death"*, de Young e Hade, *Journal of the American Association*, Vol. 292, No. 24). Se as pessoas puderem adiar suas mortes até depois do feriado de Ação de Graças, então a proporção de mortes na semana anterior ao feriado deverá ser menor do que 0,5. Use o nível significância de 0,05 para testar afirmativa de que a proporção de mortes na semana anterior ao feriado é menor do que 0,5. Com base no resultado, parece haver alguma indicação de que as pessoas possam adiar temporariamente suas mortes para sobreviverem ao feriado de Ação de Graças?
- 11.6.58) Uma recente pesquisa do Gallup com 936 adultos selecionados aleatoriamente mostrou que 312 deles nunca haviam bebido. Use esses resultados de pesquisa para testar afirmativa de que menos de 1/3 de todos adultos nunca bebeu. Use o nível significância de 0,05. Examinem, também, a seguinte redação da questão real e determine se é provável que faça surgir respostas honestas: "Com que frequência, se alguma, você toma bebidas alcoólicas tais como licor, vinho ou cerveja - todo dia, algumas vezes por semana, cerca de uma vez por semana, apenas em ocasiões especiais como Ano Novo e feriados, o nunca?"
- 11.6.59) Em uma pesquisa do Gallup com 1018 adultos, verificou-se que 22% haviam fumado cigarros na semana anterior. Use o nível significância de 0,05 para testar afirmativa de que menos de 25% dos adultos têm fumado na semana anterior. A conclusão mudaria si, em vez de uma pesquisa do Gallup, os resultados tivessem sido obtidos através de uma pesquisa pela Internet na qual se pedisse aos usuários dessa mídia que respondessem?
- 11.6.60) Em uma pesquisa do Gallup com 1125 adultos, verificou-se que 47% nunca ou raramente haviam voado. Use o nível de significância de 0,05 para testar afirmativa de que a porcentagem de adultos que nunca ou raramente voaram seja igual a 50%. Uma vez que os sujeitos à pesquisa responderam voluntariamente à questão, é possível que os resultados amostrais estejam em erro por uma quantidade razoável?

- 11.6.61) No caso *Castanella vs. Partida*, verificou-se que, durante um período de onze anos no condado de Hidalgo, Texas, 70 pessoas haviam sido selecionadas para a função do grande júri, e 39% delas eram mexicano-americanas. Entre as pessoas elegíveis para o Grande júri, 79,1% eram mexicano-americanas. Use o nível de significância ser de 0,01 para testar afirmativa de que o processo de seleção é viesado contra os mexicanos-americanos.
- 11.6.62) Clarinex é uma droga usada para tratamento da asma. Em testes clínicos dessa droga, 1655 pacientes foram tratados com dose de 5 mg de Clarinex e 2,1% deles tiveram fadiga aparentes com base em dados da *Schering Corporation*). Use o nível de significância de 0,01 para testar afirmativa de que a porcentagem de usuários de Clarinex que demonstram fadiga seja maior do que a taxa de 1,2% daqueles que não usam Clarinex. A fadiga parece ser uma reação adversa do Clarinex?
- 11.6.63) Uma pesquisa mostrou que entre 785 sujeitos com quatro anos de faculdade completos, aleatoriamente selecionados, 18,3% fumavam e 81,7% não fumavam (conforme dados da Associação Médica Americana). Use o nível de significância de 0,01 para testar afirmativa de que a taxa de fumantes entre as pessoas com quatro anos completos de faculdade é menor que a taxa de 27% da população geral. Por que os alunos de faculdades fumariam a uma taxa menor que os outros?
- 11.6.64) Consulte o Conjunto de Dados 13 do Apêndice B e ache a proporção amostral de balas M&M azuis. Use o resultado para testar afirmativa da Mars de que 24% das suas balas M&M são azuis. (Use uma significância de 5%)
- 11.6.65) Consulte o Conjunto de Dados 10 no Apêndice B e note que dias com alguma precipitação tem valor diferente de 0. Use o nível significância de 0,05 para testar a afirmativa de que em Boston, aos domingos, há precipitação em mais de 25% dos dias.
- 11.6.66) Consulte o Conjunto de Dados 8 no Apêndice B. Ache a proporção dos dias em que a temperatura máxima real seja diferente, em mais de 2°, da temperatura máxima prevista um dia antes. Seja p essa proporção. Use o nível de significância de 0,05 para testar afirmativa de que $p < 0,5$. O que o resultado indica sobre a precisão das previsões?
- 11.6.67) Usando os resultados listados no Conjunto de Dados 5 do Apêndice B, teste a afirmativa de que a maioria dos desenhos animados para crianças mostram o uso de álcool ou fumo (ou ambos). Use o nível significância 0,05.

➤ **Teste de Hipóteses para o Desvio Padrão Populacional σ e para a Variância Populacional σ^2**

- 11.6.68) Realizou-se um estudo com bebês nascidos de mães que usaram cocaína durante a gravidez, e obtiveram-se os seguintes dados amostrais para pesos ao nascer: $n = 190$; $\bar{x} = 2700g$, $s = 645g$ (com base em dados de "*Cognitive Outcomes of Preschool Children with Prenatal Cocaine Exposure*", de Singer ET al., *Journal of the American Medical Association*, Vol. 291, No. 20). Use o nível de significância de 0,05 para testar a afirmativa de que o desvio padrão dos pesos ao nascer de bebês nascidos de mães que usavam cocaína é diferente do desvio padrão de 696 g para os pesos ao nascer de bebês de mães que não usaram cocaína durante a gravidez. Com base no resultado, o uso de cocaína pelas mães parece afetar a variação dos pesos de seus bebês?
- 11.6.69) Os quartos de dólar são, atualmente, cunhados com pesos que têm uma média de 5,670 g e desvio padrão de 0,062 g. Novos equipamentos estão sendo testados para tentar melhorar a qualidade pela redução da variação. Obtém-se uma amostra aleatória simples de 24 quartos

dentre aqueles fabricados com o novo equipamento, e essa amostra tem um desvio padrão de 0,049 g. Use o nível de significância de 0,05 para testar a afirmativa de que os quartos de dólar fabricados pelo novo equipamento tenham pesos com um desvio padrão menor do que 0,062 g. O novo equipamento parece eficaz na redução da variação dos pesos? Qual seria a consequência adversa em se ter quartos com os pesos que variassem muito?

- 11.6.70) Use o nível de significância 0,01 para testar a afirmativa de que os amendoins M&M têm pesos que variam mais do que os pesos das balas M&M. O desvio padrão para os pesos das balas M&M é de 0,056 g. Uma amostra de 41 amendoins M&M tem pesos com um desvio padrão de 0,31 g. Por que os amendoins M&M devem ter pesos que variam mais do que os pesos das balas M&M?
- 11.6.71) A *Stewart Aviation of Products Company* usa um novo método de produção para fabricar altímetros de aeronaves. Uma amostra aleatória simples de 81 altímetros é testada em uma câmara de pressão, e os erros na altitude são registrados como valores positivos (para leituras que são muito altas) ou negativos (para leituras que são muito baixas). A amostra tem um desvio padrão de $s = 52,3$ pés. Ao nível de significância de 0,05, teste a afirmativa de que a nova linha de produção tem erros com um desvio padrão diferente de 43,7 pés, que era o desvio padrão para o método antigo de produção. Se parece que o desvio padrão mudou, o novo método de produção parece ser melhor ou pior do que o antigo?
- 11.6.72) Quando consumidores se candidatam a um financiamento, seus créditos são classificados através dos escores FICO (*Fair, Isaac and Company*). Lista-se, abaixo, as classificações de crédito para uma amostra de candidatos a financiamentos de carro de uma nova agência do *Bank of Newport*. Use os dados amostrais para testar a afirmativa de que essas classificações de crédito sejam provenientes de uma população com um desvio padrão diferente de 83, que é o desvio padrão dos candidatos clientes da matriz. Use o nível de significância de 0,05. Com base nos resultados, parece que os candidatos a financiamentos da agência tenham classificações que variem mais do que as dos clientes da matriz?

661 595 548 730 791 678 672 491 492 583 762 624 769 729 734 706

- 11.6.73) O menor mamífero do mundo é o morcego mamangaba, também conhecido como morcego de focinho de porco (ou *Craseonucteris thonglongyai*). Tais morcegos são, no máximo, do tamanho de um mamangaba (ou abelhão). Abaixo, estão listados os pesos (em gramas) de uma amostra desses morcegos. O usando o nível de significância de 0,05, teste a afirmativa de que esses pesos sejam provenientes de uma população com um desvio padrão igual a 0,30 g, que é o desvio padrão dos pesos dos morcegos mamangaba de uma região da Tailândia. Esses morcegos da mostra parecem ter pesos com a mesma variação que os morcegos daquela região da Tailândia?

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

- 11.6.74) Use o nível de significância 0,01 para testar a afirmativa de que os pesos de modelos femininas variam menos do que os pesos das mulheres em geral. O desvio padrão dos pesos da população de mulheres é de 29 lb. Abaixo estão listados os pesos (em libras) de nove modelos selecionadas aleatoriamente.

125 (Taylor)	119 (Auermann)	128 (Schiffer)	128 (MacPherson)
119 (Turlington)	127 (Hall)	105 (Moss)	123 (Mazza)
115 (Hume)			

- 11.6.75) Use o nível de significância 0,05 para testar a afirmativa de que as alturas de modelos femininas variam menos do que as alturas das mulheres em geral. O desvio padrão das alturas da população de mulheres é de 2,5 in. Abaixo estão listadas as alturas (em polegadas) de modelos selecionadas aleatoriamente (Taylor, Harlow, Muldor, Goff, Evangelista, Auermann, Schiffer, MacPherson, Turlington, Hall, Crawford, Campbell, Herzigova, Saymour, Banks, Moss, Mazza e Hume).

71	71	70	69	69,5	70,5	71	72	70
70	69	69,5	69	70	70	66,5	70	71

- 11.6.76) Abaixo estão listadas as medidas das quantidades de chumbo (em microgramas por metro cúbico ou $\mu\text{g} / \text{m}^3$) no ar. A Agência de Proteção Ambiental estabeleceu um padrão de qualidade do ar para o chumbo de $1,5 \mu\text{g} / \text{m}^3$. As medidas mostradas a seguir foram registradas no Edifício 5 do local do *World Trade Center*, em dias diferentes imediatamente após a destruição causada pelos ataques terroristas de 11 de setembro de 2001. Teste a afirmativa de que essas quantidades sejam provenientes de uma população com um desvio padrão maior do que $0,4 \mu\text{g} / \text{m}^3$. Use o nível de significância 0,05. Há alguma coisa sobre os valores amostrais que possa violar um requisito para uso do teste Qui-Quadrado?

5,40	1,10	0,42	0,73	0,48	1,10
------	------	------	------	------	------

- 11.6.77) Os valores listados abaixo são tempos de espera (em minutos) de clientes do *Jefferson Valley Bank*, onde aguardam atendimento em fila única que conduz a três caixas. Teste a afirmativa de que o desvio padrão dos tempos de espera seja menor do que 1,9 min., que é o desvio padrão dos tempos de espera no mesmo banco quando se espera em filas individuais para cada caixa. Use o nível de significância de 0,05. O uso da fila única parece diminuir a variação entre os tempos de espera? Qual a vantagem em se reduzir essa variação? .

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

- 11.6.78) Consulte o Conjunto de Dados 14 do Apêndice B e use os dados amostrais que consistem em pesos de moedas de quartos de dólar feitas depois de 1964. Teste a afirmativa de que esses quartos de dólar sejam provenientes de uma população com desvio padrão igual ao valor especificado de 0,068 g. Use o nível de significância de 0,05. A variação dos pesos parece ser a desejada?
- 11.6.79) Consulte o Conjunto de Dados a 8 do Apêndice B e use os dados amostrais que consistem em quantidades de precipitação. Teste a afirmativa de que essas quantidades sejam provenientes de uma população com um desvio padrão menor do que 1,00 in. Use o nível de significância de 0,05. O requisito de uma distribuição normal parece ser satisfeita e como ele afeta os resultados?
- 11.6.80) A *Kansas Farm Products Company* usa uma máquina que enche sacos de 50 lb de milho. No passado, a máquina apresentava um desvio padrão de 0,75 lb. Com o intuito de obter pesos mais compatíveis, os mecânicos substituíram algumas peças gastas da máquina. Uma amostra de 61 sacos, extraída da máquina já reparada, acusou média amostral de 50,13 lb e desvio padrão amostral de 0,48 lb. Ao nível de 0,05 de significância, teste a afirmação de que os pesos são mais compatíveis com a máquina reparada em comparação com a situação anterior. Se os pesos são mais compatíveis, qual é o efeito sobre o desvio padrão?