Estimação

Julius Richard Dedekind (Braunschweig, Alemanha, 6 de outubro de 1831 — 12 de fevereiro de 1916)

Matemático alemão, Julius Wihelm Richard Dedekind nasceu a 6 de outubro de 1831, em Braunschweig, e faleceu a 12 de fevereiro de 1916, na mesma localidade. Os seus trabalhos sobre números irracionais conduziram a desenvolvimentos importantes em teoria dos números. Adepto das conceções de Cantor sobre infinito, contra as objeções de Kronecker, a sua análise sobre a natureza do número, sobre indução matemática e as definições de conjunto finito e infinito são contribuições de importância maior. Definiu um conjunto de axiomas que representam, formal e exatamente, o conceito lógico de número inteiro (estes axiomas seriam mais tarde atribuídos a Peano). Tendo publicado sobre outros temas - hidrodinâmica -, a sua edição sobre os trabalhos de Dirichlet são marcos inestimáveis para a Matemática.

FONTE: http://www.infopedia.pt

10 - ESTIMAÇÃO

10.1 - CONCEITOS

- Parâmetros: São valores característicos de uma população como, por exemplo, sua média (μ), seu desvio padrão (σ) ou qualquer outra medida dessa população. Os Parâmetros de uma população são valores fixos, uma vez que os elementos pertencentes a uma população também são fixos.
- Estatísticas: Qualquer valor calculado com base nos elementos de uma amostra é chamado *estatística*. Por exemplo, a média amostral, a mediana amostral, o mínimo amostral etc. As estatísticas variam de uma amostra para outra, sendo elas próprias variáveis aleatórias. Pode-se, assim, falar de em Distribuições de médias amostrais ou de desvios padrões amostrais.
- Estimação: É o processo que consiste em utilizar dados amostrais para estimar os valores de parâmetros populacionais desconhecidos. Qualquer característica de uma população pode ser estimada a partir de uma amostra aleatória, sendo que as características mais comuns são a média, o desvio padrão e a proporção populacional. A estimação de um parâmetro populacional comporta dois tipos. A Estimação Pontual procura fixar um valor numérico único que esteja satisfatoriamente próximo do verdadeiro valor do parâmetro. A Estimação Intervalar procura determinar intervalos com limites aleatórios, que abranjam o valor do parâmetro populacional, com uma margem de segurança prefixada.
- Estimador: Uma *estatística* destinada a estimar um parâmetro populacional é chamada de *Estimador* deste parâmetro.

10.2 - DISTRIBUIÇÕES AMOSTRAIS

Considerando todas as possíveis amostras de tamanho *"n"* de uma população, pode-se determinar para cada amostra a sua média e consequentemente obter com esses resultados a Distribuição por Amostragem da Média. De maneira semelhante pode-se obter as distribuições por amostragem da Variância, da Proporção ou de outras estatísticas.

TABELA 23.- EXEMPLO DE AMOSTRAS DE TAMANHO IGUAL A 10 DE UMA POPULAÇÃO

Amostra	Valores										$\bar{x}$	Md	Mínimos	σ
1	-5	-5	-4	-3	-2	-1	-1	1	2	3	-1,5	-1,5	-5	2,84
2	-5	-4	-4	-3	-3	0	0	4	4	5	-0,6	-1,5	-5	3,78
3	-5	-4	-4	-3	-2	0	0	1	1	4	-1,2	-1	-5	2,86
4	-5	-4	-3	-2	0	1	2	3	4	4	0	0,5	-5	3,33
5	-5	-3	-2	-2	-1	1	2	2	4	5	0,1	0	-5	3,21
6	-5	-5	-3	-3	-2	1	2	4	4	4	-0,3	-0,5	-5	3,71
7	-4	-4	-3	-3	0	0	0	2	3	4	-0,5	0	-4	2,92
8	-4	-3	-2	-1	-1	0	1	1	3	3	-0,3	-0,5	-4	2,36

10.2.1 - *AMOSTRAGEM COM E SEM REPOSIÇÃO*

Seja *"N"* o número de elementos de uma população e *"n"* o número de elementos da amostra. Então:

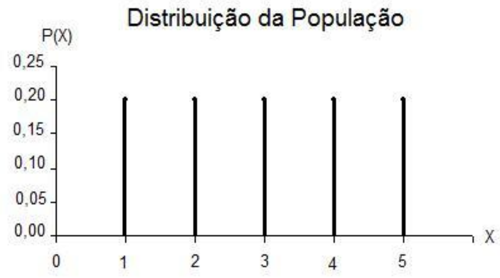
- Se o processo de retirada for com reposição, N^n dará o número de amostra de tamanho *"n"* que poderão ser extraídas com reposição da população de tamanho *"N"*.
- Se o processo de retirada for sem reposição o número combinatório C_N^n dará o total de amostras de tamanho *"n"* extraídas sem reposição da população de tamanho *"N"*.

10.2.2 - *DISTRIBUIÇÃO POR AMOSTRAGEM DA MÉDIA*

10.2.2.1 - Distribuição Amostral da Média com Amostragem com Reposição

Deve-se encontrar a média e a variância da Variável Aleatória $\overline{X}$ (Média Amostral).

Seja X uma população constituída dos seguintes elementos $\{1,2,3,4,5\}$. Essa população apresenta média $\mu = 3$ e variância $\sigma^2 = 2$. O gráfico representativo da Distribuição de Probabilidades dos valores desta população está ilustrado a seguir:



Observando o gráfico notamos que a população, apesar de Discreta, tem a forma de uma Distribuição Uniforme.

Extraindo-se, aleatoriamente, com reposição, amostras de 2 elementos pode-se obter $5^2 = 25$ amostras possíveis, que são:

$$\left\{\begin{matrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) \end{matrix}\right\}$$

Calculando-se para cada amostra a sua média tem-se a seguinte população de médias amostrais de tamanho 2:

$$\overline{x} = \left\{\begin{matrix} 1,0 & 1,5 & 2,0 & 2,5 & 3,0 \\ 1,5 & 2,0 & 2,5 & 3,0 & 3,5 \\ 2,0 & 2,5 & 3,0 & 3,5 & 4,0 \\ 2,5 & 3,0 & 3,5 & 4,0 & 4,5 \\ 3,0 & 3,5 & 4,0 & 4,5 & 5,0 \end{matrix}\right\}$$

A Tabela da Distribuição de Probabilidades da Variável Aleatória $\overline{x}$ é dada por:

$\overline{x}$	$P(\overline{x})$
1,0	$1/25$
1,5	$2/25$
2,0	$3/25$
2,5	$4/25$
3,0	$5/25$
3,5	$4/25$
4,0	$3/25$
4,5	$2/25$
5,0	$1/25$
	1

Calculando-se a média e a variância desta distribuição temos:

$$\mu_{\overline{X}} = E\left[\overline{X}\right] = \sum \overline{X}_i \times p\left(\overline{X}_i\right) \quad \rightarrow \quad \mu_{\overline{X}} = 3$$

$$\sigma^2\left(\overline{X}\right) = E\left[\left(\overline{X}_i - \mu_{\left(\overline{X}\right)}\right)^2\right] = E\left[\overline{X}_i^2\right] - \mu_{\left(\overline{X}\right)}^2$$

Mas, $E\left[\overline{X}_i^2\right] = \sum \overline{X}_i^2 \times p\left(\overline{X}_i\right) = 10$

Logo, $\sigma^2\left(\overline{X}\right) = 10 - 3^2 \rightarrow \sigma^2\left(\overline{X}\right) = 1$

Com a observação dos resultados das médias e das variâncias conclui-se que:

1º - A Média das Médias Amostrais $\left(\mu_{\overline{X}}\right)$ é igual à Média Populacional μ, ou $\mu_{\overline{x}} = \mu_x$

Demonstração:

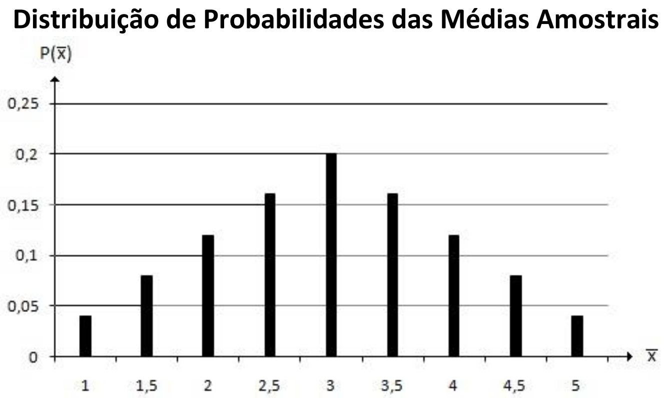
$$\mu_{\overline{X}} = E\left(\overline{X}\right) = E\left\{\frac{1}{n}\sum_{i=1}^{n} X_i\right\} = \frac{1}{n}E\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E\left(X_i\right) = \frac{1}{n}\sum_{i=1}^{n}\mu = \frac{n\mu}{n} = \mu$$

2º - A Variância da Média Amostral $\left(\sigma_{\overline{X}}^2\right)$ é igual à Variância Populacional dividida pelo tamanho da Amostra, ou. $\sigma_{\overline{X}}^2 = \dfrac{\sigma^2}{n}$ (também conhecido como Erro Padrão da Média).

Demonstração:

$$\sigma_{\overline{X}}^2 = \sigma^2\left\{\frac{1}{n}\sum_{i=1}^{n} X_i\right\} = \frac{1}{n^2}\sigma^2\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2\left(X_i\right) = \frac{1}{n^2}\times n\sigma^2 = \frac{\sigma^2}{n}$$

O Gráfico da Distribuição de Probabilidades das Médias Amostrais é:



Estes resultados ilustram o Teorema do Limite Central:

- Quando a população é Normal $X \overset{d}{=} N(\mu,\sigma)$ a média amostral $\overline{X}$, com amostras de tamanho *"n",* também tem distribuição Normal com $\overline{X} \overset{d}{=} N\left(\mu, \sigma/\sqrt{n}\right)$
- Quando a população tem distribuição Não-Normal com média μ e desvio padrão σ, a distribuição da média amostral $\overline{X}$ para amostras de tamanho *"n"* suficientemente grande será $\overline{X} \overset{d}{\approx} N\left(\mu, \sigma/\sqrt{n}\right)$.

A variável Normal Padronizada da variável $\bar{x}$ será $Z = \dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}$.

Graficamente, as Distribuições de X e $\overline{X}$ têm o seguinte formato:

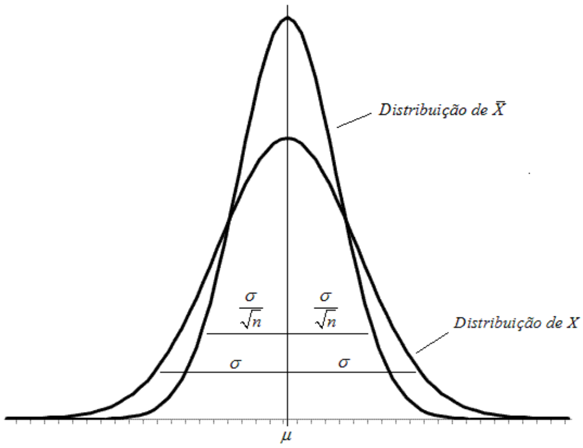


Figura 32.- Distribuição Normal de uma População e Distribuição Normal das Médias Amostrais

10.2.2.2 - Distribuição Amostral da Média com Amostragem Sem Reposição

A maior parte da amostragem se faz sem reposição, seja por motivos psicológicos, seja por conveniência e custo. Enquanto o tamanho da amostra for pequeno em relação ao da população, a amostragem sem reposição dará essencialmente a mesma variabilidade da amostragem com reposição. Entretanto, se o tamanho da amostra representar uma porcentagem apreciável da população (5% ou mais), os resultados dos dois tipos de amostragem começam a diferir. Isto porque, na amostragem sem reposição, a probabilidade de extração de itens varia de uma para a outra extração. Em tais condições, a distribuição adequada é a Distribuição Hipergeométrica.

A fórmula do desvio padrão das médias amostrais deve ser modificada de modo a refletir a probabilidade, se o tamanho da amostra é superior a 5% da população. Assim o desvio padrão das amostras pode ser determinado por:

$$Se\ n \geq 0{,}05N \ \Rightarrow\ \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}}$$

Onde *"N"* é o tamanho da população e *"n"* o tamanho da amostra. Neste caso a variável Normal Padronizada Z será dada por:

$$Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}$$

A porção $\sqrt{\dfrac{N-n}{N-1}}$ é chamada de Coeficiente de Correção do desvio padrão.

10.2.3 - *Distribuição por Amostragem da Proporção (Frequência Relativa)*

Seja X uma população infinita e p a probabilidade de sucesso de certo evento. Logo, $1-p=q$ será a probabilidade fracasso deste evento.

Seja uma amostra aleatória de *n* elementos dessa população, e *x* o número de sucessos da amostra.

Então *x* (frequência absoluta) número de sucessos é uma variável aleatória Binomial de média *np* e variância *npq*.

Logo a distribuição da frequência relativa $\hat{p}=\frac{x}{n}$ será:

$$E(\hat{p})=E\left(\tfrac{x}{n}\right)=\frac{1}{n}E(x)=\frac{\mu_x}{n}=\frac{np}{n}=p$$

$$\sigma^2(\hat{p})=\sigma^2\left(\tfrac{x}{n}\right)=\frac{1}{n^2}\sigma^2(x)=\frac{npq}{n^2}=\frac{pq}{n}$$

Assim, tem-se que:

- A Média da Distribuição Amostral das Proporções é sempre igual à Proporção Populacional

- A Variância da Distribuição Amostral das Proporções é igual ao produto das probabilidades de sucesso e de fracasso da população dividido pelo tamanho da amostra.

Se *n* for grande pode-se aproximar a Binomial por uma Normal, onde:

$$\hat{p}\cong N\left(p;\tfrac{pq}{n}\right)$$

Sendo a variável Normal Padronizada da Frequência Relativa dada por:

$$Z=\frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}}$$

10.2.4 - *Distribuição por amostragem da Variância*

10.2.4.1 - Distribuição Qui-Quadrado: χ^2

Define-se a variável aleatória χ^2, com ϕ graus de liberdade[3], como sendo o somatório de ϕ normais padronizadas e independentes ao quadrado, ou seja:

$$\chi^2_\phi=\sum_{i=1}^{\phi}Z_i^2=\sum_{i=1}^{\phi}\left(\frac{X_i-\mu}{\sigma}\right)^2$$

Lembrando que $Z\overset{d}{=}N(0;1)$.

Conforme os graus de liberdade a função densidade dessa variável pode apresentar as seguintes formas:

[3] Graus de Liberdade: Número de valores que estão livres para variar depois de certas restrições terem sido impostas a todos os valores.

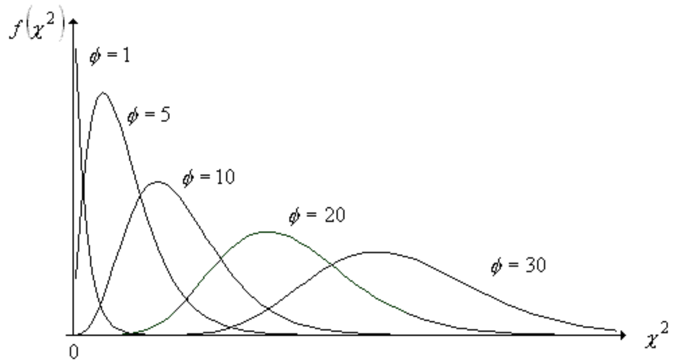


Figura 33.- Distribuição Qui-Quadrado com vários Graus de Liberdade

A média e a variância de uma Distribuição χ^2_ϕ são dadas por:

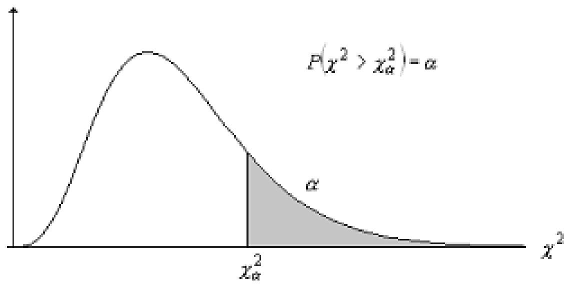
$$\mu\left(\chi^2_\phi\right) = \phi \qquad \sigma^2\left(\chi^2_\phi\right) = 2\phi$$

- **Propriedades da Distribuição χ^2**

1. Como a variável χ^2 é obtida pela soma de variáveis independentes, segue-se pelo Teorema do Limite Central que a família de distribuição do tipo χ^2 tende à distribuição Normal quando o número de graus de liberdade aumenta.
2. A soma algébrica de duas variáveis independentes com distribuição χ^2 com m e p graus de liberdade terá distribuição χ^2 com $m \pm p$ graus de liberdade. $\chi^2_r = \chi^2_m \pm \chi^2_p$ onde $r = m \pm p$

Os valores da variável χ^2 estão tabulados para as principais probabilidades, sendo que a tabela nos apresenta, para uma dada probabilidade α e um determinado número de graus de liberdade ϕ o valor $\chi^2_{\phi,\alpha}$, tal que

$$P\left(\chi^2 \geq \chi^2_{\varphi;\alpha}\right) = \alpha$$



10.2.4.2 - Distribuição por Amostragem da Variância Amostral

Seja $(X_1, X_2, X_3, \ldots, X_n)$ uma amostra extraída de uma população normal com média μ e variância σ^2.

Seja $S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$ a variância amostral, e $\bar{X} = \dfrac{\sum_{i=1}^{n} X_i}{n}$ a média amostral.

Tomando a soma $\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2$ e somando e subtraindo μ a cada termo, $\left(X_i-\bar{X}\right)^2$, tem-se:

$$\begin{aligned}\sum\left(X_i-\bar{X}\right)^2 &= \sum\left(X_i-\bar{X}+\mu-\mu\right)^2\\ &= \sum\left[\left(X_i-\mu\right)-\left(\bar{X}-\mu\right)\right]^2\\ &= \sum\left[\left(X_i-\mu\right)^2+\left(\bar{X}-\mu\right)^2-2\left(X_i-\mu\right)\left(\bar{X}-\mu\right)\right]\\ &= \sum\left(X_i-\mu\right)^2+\sum\left(\bar{X}-\mu\right)^2-2\sum\left[\left(X_i-\mu\right)\left(\bar{X}-\mu\right)\right]\end{aligned}$$

Como $\left(\bar{X}-\mu\right)$ é constante

$$\begin{aligned}\sum\left(X_i-\bar{X}\right)^2 &= \sum\left(X_i-\mu\right)^2+n\left(\bar{X}-\mu\right)^2-2\left(\bar{X}-\mu\right)\sum\left(X_i-\mu\right)\\ &= \sum\left(X_i-\mu\right)^2+n\left(\bar{X}-\mu\right)^2-2\left(\bar{X}-\mu\right)\left[\sum X_i-\sum\mu\right]\\ &= \sum\left(X_i-\mu\right)^2+n\left(\bar{X}-\mu\right)^2-2\left(\bar{X}-\mu\right)\left[n\bar{X}-n\mu\right]\\ &= \sum\left(X_i-\mu\right)^2+n\left(\bar{X}-\mu\right)^2-2n\left(\bar{X}-\mu\right)^2\end{aligned}$$

Assim

$$\sum\left(X_i-\bar{X}\right)^2=\sum\left(X_i-\mu\right)^2-n\left(\bar{X}-\mu\right)^2$$

e a variância amostral pode ser definida por

$$S^2=\frac{1}{n-1}\left[\sum\left(X_i-\mu\right)^2-n\left(\bar{X}-\mu\right)^2\right]$$

Multiplicando-se ambos os membros da expressão por $\frac{(n-1)}{\sigma^2}$, fica:

$$\frac{(n-1)S^2}{\sigma^2}=\frac{1}{\sigma^2}\left[\sum\left(X_i-\mu\right)^2-n\left(\bar{X}-\mu\right)^2\right]$$

Simplificando,

$$\frac{(n-1)S^2}{\sigma^2}=\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2-\left(\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right)^2$$

Os termos $\left(\frac{X_i-\mu}{\sigma}\right)$ são variáveis normais padronizadas com média 0 e variância 1. Da mesma maneira $\left(\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right)$ é também $N(0;1)$.

Da definição da Distribuição Qui-Quadrado, $\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2$ tem distribuição χ_n^2 e $\left(\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right)^2$ tem distribuição χ_1^2, logo

$$\frac{(n-1)S^2}{\sigma^2}=\chi_n^2-\chi_1^2$$

Portanto

$$\frac{(n-1)S^2}{\sigma^2}=\chi_{n-1}^2$$

ou seja

$$S^2=\frac{\sigma^2}{n-1}\chi_{n-1}^2$$

A variância amostral tem distribuição χ^2 com $(n-1)$ graus de liberdade.
Utilizando este resultado podemos obter a média e a variância de S^2.

$$E\left[S^2\right]=E\left[\frac{\sigma^2}{n-1}\chi^2_{n-1}\right]=\frac{\sigma^2}{n-1}E\left[\chi^2_{n-1}\right]=\frac{\sigma^2}{n-1}(n-1)=\sigma^2$$

$$VAR\left(S^2\right)=VAR\left[\frac{\sigma^2}{n-1}\chi^2_{n-1}\right]=\frac{\sigma^4}{(n-1)^2}VAR\left(\chi^2_{n-1}\right)=\frac{\sigma^4}{(n-1)^2}\left[2(n-1)\right]=\frac{2\sigma^4}{n-1}$$

10.3 - Estimação Pontual

A estimação pontual (por ponto) consistirá simplesmente em, à falta de melhor informação, adotar a estimativa disponível como sendo o valor do parâmetro. A ideia é, em sua essência, extremamente simples, porém a qualidade dos resultados irá depender fundamentalmente da conveniente escolha do estimador. Assim, dentre os vários estimadores razoáveis que poderemos imaginar para um determinado parâmetro, devemos ter a preocupação de escolher aquele que melhor satisfaça às propriedades de um bom estimador.

Suponhamos que uma amostra foi retirada de uma população e as estatísticas obtidas da amostra são: $\bar{x}=11{,}3$; $Md=15$; $Mo=15{,}7$; $s=1{,}7$; $s^2=2{,}89$. Uma estimativa pontual para a média populacional é $\mu=11{,}3$. Outra estimativa para média populacional seria $\mu=15$. Na primeira estimação foi utilizada a média amostra e na segunda a mediana amostral. É óbvio que a segunda estimativa possui um “erro” maior do que a primeira. Podemos concluir que:

> “A melhor estimativa pontual para um parâmetro populacional é a estatística de mesmo nome.”

10.4 - Intervalos de Confiança

A Estimação por pontos de um parâmetro não possui uma medida do possível erro cometido na estimação.

Uma maneira de expressar a precisão da estimação é estabelecer *limites,* que com certa probabilidade, incluam o verdadeiro valor do parâmetro da população.

Esses limites são chamados *“Limites de Confiança”*: Determinam um *Intervalo de Confiança (IC),* no qual deverá estar o verdadeiro valor do parâmetro.

Logo a Estimação por Intervalo consiste na fixação de dois valores tais que $(1-\alpha)$ seja a probabilidade de que o intervalo, por eles determinado, contenha o verdadeiro valor do parâmetro.

- α: Nível de Incerteza ou Grau de Desconfiança ou ainda Nível de significância;
- $(1-\alpha)$: Coeficiente de Confiança ou Nível de Confiabilidade.

Portanto, α fornece a medida da incerteza desta **inferência**.

Logo, a partir da informação da amostra, deve-se calcular os limites de um intervalo, valores críticos, que em $(1-\alpha)\%$ dos casos inclua o valor do parâmetro a estimar e em $\alpha\%$ dos casos não inclua o valor do parâmetro.

10.4.1 - *Intervalos de Confiança para a Média Populacional μ Quando a Variância Populacional σ^2 é Conhecida.*

Consideremos uma população Normal com média desconhecida, da qual se deseja estimar esta média, e com variância σ^2 conhecida, $X \overset{d}{=} N(\ ?\ ;\ \sigma\ )$.

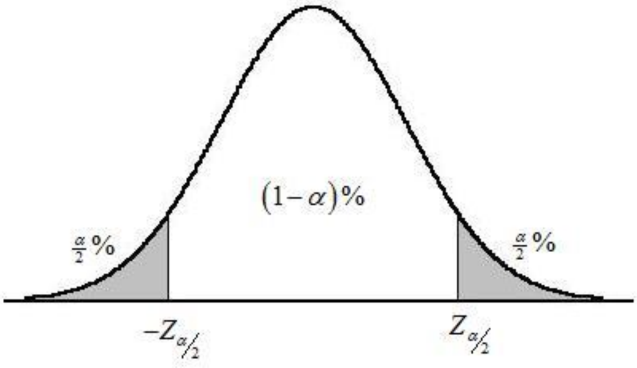
Tomando-se uma amostra aleatória de tamanho *n* tem-se que:

A média amostral é $\overline{X} = \mu$ e o desvio padrão amostral é $\sigma_{\bar{x}} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$. Assim, fixado o nível de confiança α, deve-se determinar Z_α tal que $P\left(|Z| > Z_{\alpha/2}\right) = \alpha$, ou seja $P\left(Z > Z_{\alpha/2}\right) = \alpha/2$ e $P\left(Z < Z_{\alpha/2}\right) = \alpha/2$. Ou ainda:

$$P\left(|Z| < Z_{\alpha/2}\right) = 1 - \alpha$$

ou, ainda,

$$P\left(-Z_{\alpha/2} < Z < Z_{\alpha/2}\right) = 1 - \alpha$$



Temos que, a variável Normal Padronizada é dada por $Z = \dfrac{X - \mu}{\sigma}$. Para a média amostral fica:

$$Z = \frac{\overline{X} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} \quad \text{ou} \quad Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Assim

$$P\left(-Z_{\alpha/2} < \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} < Z_{\alpha/2}\right) = 1 - \alpha$$

$$P\left(-Z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}} < \overline{X} - \mu < Z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

$$P\left(-\overline{X} - Z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}} < -\mu < -\overline{X} + Z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

Multiplicando a desigualdade por -1 temos:

$$\therefore \quad P\left(\overline{X} - Z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + Z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

Portanto os limites de confiança serão

$$\mu_1 = \overline{X} - Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$\mu_2 = \overline{X} + Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}.$$

A grandeza $Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ é conhecida com *Margem de Erro*[4] e será representada por *E*.

Usando uma notação simplificada, tem – se

$$IC\left(\mu,(1-\alpha)\%\right)=\left(\bar{X}-E;\bar{X}+E\right)$$

Onde $E=Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

A fórmula da margem de erro anterior foi obtida supondo-se que a população seja infinita ou que a amostra tenha um tamanho muito pequeno comparado ao tamanho de uma população finita. No caso onde a amostra tenha um tamanho comparável ao tamanho da população $(n\geq 0,05N)$ devemos usar o fator de correção para o desvio padrão e, portanto, a fórmula da margem de erro fica:

$$E=Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

10.4.2 - Distribuição de t - Student e Intervalos de Confiança para a Média Populacional μ Quando a Variância Populacional σ^2 é Desconhecida

A variável $Z=\frac{\bar{x}-\mu}{\sigma_{\bar{x}}}$ tem distribuição normal. Quando a variância σ^2 é desconhecida, deve-se usar s^2, estimador de σ^2.

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2 \quad \text{e} \quad S_{\bar{x}}\cong\sqrt{\frac{S^2}{n}}\cong\frac{S}{\sqrt{n}}$$

Neste caso existem duas situações. A primeira ocorre para amostras grandes, isto é, $n>30$.

Quando a amostra for grande simplesmente substituímos σ por S, assim as fórmulas do Intervalo de confiança ficam:

$$IC\left(\mu,(1-\alpha)\%\right)=\left(\bar{X}-E;\bar{X}+E\right)$$

Onde $E=Z_{\alpha/2}\frac{s}{\sqrt{n}}$

A outra situação ocorre para amostras pequenas, isto é, $n\leq 30$. Neste caso a Distribuição Normal não oferece uma estimação muito precisa e devemos utilizar a Distribuição t – Student.

A variável definida como $t_\varphi=\frac{\bar{x}-\mu}{s_{\bar{x}}}$ é denominada uma variável com distribuição de "t de Student" com φ graus de liberdade.

O número de informações independentes da amostra dá o número de graus de liberdade φ da distribuição de *t* e este número será igual ao número de informações independentes da amostra (n) subtraído pelo número de parâmetros da população a serem estimados (K) além do parâmetro inerente ao estudo.

$$\varphi=n-K$$

No caso de estimar a média com a variância desconhecida, além de $\bar{X}$, estimador inerente ao estudo, estima-se σ^2, um parâmetro a mais. Isto significa que deve-se usar a distribuição de *t* com $n-1$ graus de liberdade.

Para cada valor de φ tem-se uma curva diferente de *t*, e quando $n\to\infty \;\Rightarrow\; t\to N(0;1)$.

[4] Margem de Erro: Diferença máxima provável (com probabilidade $1-\alpha$) entre a estatística amostral observada e o verdadeiro valor do parâmetro populacional.

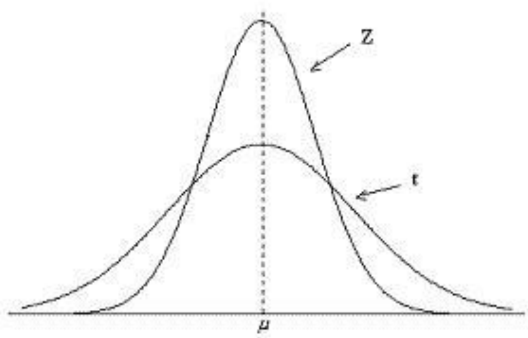
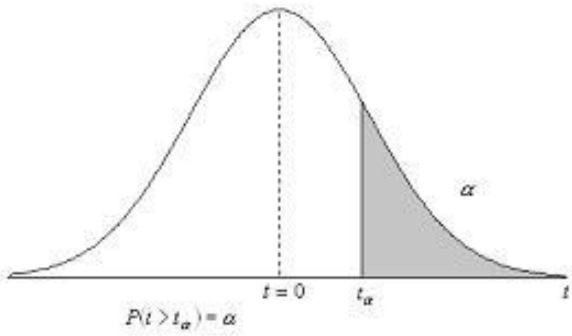


Figura 34.- Distribuição Normal Padronizada e Distribuição t-Student

A tabela da Distribuição de t - Student fornece o valor de t_α onde $P(t > t_\alpha) = \alpha$.



Os Intervalos de Confiança para a média populacional μ, quando a variância populacional σ^2 é desconhecida, são obtidos de maneira similar à anterior, apenas substituindo-se $Z_{\alpha/2}$ pelo valor de $t_{\varphi;\alpha/2}$. O Intervalo de Confiança será dado por:

$$IC\left(\mu,(1-\alpha)\%\right) = \left(\bar{X} - E; \bar{X} + E\right)$$

Onde $E = t_{\phi;\alpha/2} \dfrac{S}{\sqrt{n}}$

10.4.3 - *Intervalos de Confiança para Proporções*

Quando a proporção populacional *p* é conhecida o estimador $\hat{p} = \dfrac{x}{n}$ tem distribuição

$$\hat{p} \cong N\left(p, \sqrt{\frac{pq}{n}}\right)$$

Para obter-se o Intervalo de Confiança para *p* desconhecida, determina-se $\hat{p}$ na amostra e considerando

$$\sigma_{\hat{p}} \cong \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}}.$$

Logo, ao nível α de significância,

$$P\left(|Z| \le Z_{\alpha/2}\right) = 1 - \alpha \quad \therefore \text{ Sendo } Z = \frac{\hat{p} - p}{\sqrt{\frac{\hat{p}\hat{q}}{n}}}$$

Substituindo-se *Z* e desenvolvendo tem-se:

$$IC\left(p,(1-\alpha)\%\right)=\left[\hat{p}-E\ ;\hat{p}+E\ \right]$$

Onde $E=Z_{\alpha/2}\sqrt{\frac{\hat{p}\cdot\hat{q}}{n}}$

Se o tamanho da amostra for superior a 5% do tamanho da população o Fator de Correção deve ser aplicado e a fórmula da margem de erro fica:

$$E=Z_{\alpha/2}\sqrt{\frac{\hat{p}\cdot\hat{q}}{n}}\sqrt{\frac{N-n}{N-1}}$$

10.4.4 - *Determinação do Tamanho da Amostra*

Se o objetivo é estimar a média, ou uma proporção, pode-se usar os Intervalos de Confiança estabelecidos para determinar o tamanho da amostra *"n"*. Para isto, deve-se fixar o maior erro aceitável.

As fórmulas para determinação do tamanho de uma amostra são obtidas através das margens de erro:

TABELA 24.- FÓRMULAS PARA DETERMINAÇÃO DO TAMANHO DA AMOSTRA

PARÂMETRO A SER ESTIMADO		TAMANHO DA POPULAÇÃO (N)	
		Desconhecido	Conhecido
Média Populacional μ		$n=\left(\frac{Z_{\alpha/2}\sigma}{E}\right)^2$	$n=\frac{N\sigma^2Z^2}{(N-1)E^2+\sigma^2Z^2}$
Proporção Populacional p	Proporção Amostral $\hat{p}$ conhecida	$n=\frac{Z^2_{\alpha/2}\hat{p}\hat{q}}{E^2}$	$n=\frac{N\hat{p}\hat{q}Z^2}{\hat{p}\hat{q}Z^2+(N-1)E^2}$
	Proporção Amostral $\hat{p}$ Desconhecida	$n=\frac{0,25\times Z^2_{\alpha/2}}{E^2}$	$n=\frac{0,25NZ^2}{0,25Z^2+(N-1)E^2}$

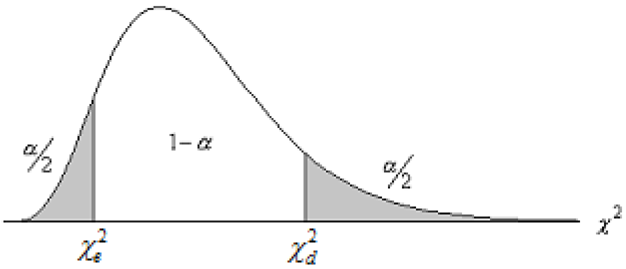
10.4.5 - *Intervalo de Confiança para a Variância Populacional e para o Desvio Padrão Populacional*

Seja uma amostra extraída de uma população normal com média μ e variância σ^2. Seja $\bar{X}=\frac{\sum_{i=1}^{n}X_i}{n}$ a média amostral e seja $S^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2$ a variância amostral.

Da distribuição χ^2 temos que:

$$\chi^2_{n-1}=\frac{(n-1)S^2}{\sigma^2}$$

Então, o Intervalo de Confiança será:



$$P\left(\chi_e^2 \le \chi^2 \le \chi_d^2\right) = 1-\alpha$$
$$P\left(\chi_e^2 \le \frac{(n-1)s^2}{\sigma^2} \le \chi_d^2\right) = 1-\alpha$$

Resolvendo as desigualdades separadamente tem-se:

$$\chi_e^2 \le \frac{(n-1)s^2}{\sigma^2} \quad \to \quad \chi_e^2 \cdot \sigma^2 \le (n-1)s^2 \quad \to \quad \sigma^2 \le \frac{(n-1)s^2}{\chi_e^2}$$
$$\frac{(n-1)s^2}{\sigma^2} \le \chi_d^2 \quad \to \quad \frac{(n-1)s^2}{\chi_e^2} \le \sigma^2$$

Portanto

$$P\left(\frac{(n-1)s^2}{\chi_d^2} \le \sigma^2 \le \frac{(n-1)s^2}{\chi_e^2}\right) = 1-\alpha$$

que é o Intervalo de Confiança para a Variância Populacional.

De forma mais simplificada, o Intervalo de Confiança será dado por:

$$IC\left(\sigma^2;(1-\alpha)\%\right) = \left(\frac{(n-1)s^2}{\chi_d^2};\frac{(n-1)s^2}{\chi_e^2}\right)$$

Para obter o Intervalo de Confiança para o desvio Padrão populacional basta extrairmos a raiz quadrada.

$$IC\left(\sigma;(1-\alpha)\%\right) = \left(\sqrt{\frac{(n-1)s^2}{\chi_d^2}};\sqrt{\frac{(n-1)s^2}{\chi_e^2}}\right)$$

10.5 - Exercícios Propostos

- **Probabilidades para Média Amostral**

10.5.1) Seja $X \overset{d}{=} N\left(80;\sqrt{26}\right)$. Dessa população retiramos uma amostra de $n=25$. Calcular:

a) $P(\bar{x} > 83)$

b) $P(\bar{x} \le 82)$

c) $P(78 \le \bar{x} \le 81)$

10.5.2) Seja $X \overset{d}{=} N\left(100;\sqrt{85}\right)$. Retiramos uma amostra de tamanho $n=20$. Determinar:

a) $P(96 < \bar{x} < 105)$

b) $P(\bar{x} < 103)$

10.5.3) Seja $X : N\left(900;\sqrt{642}\right)$. Retiramos uma amostra de tamanho igual a 30. Determinar:

a) $P(\bar{x} \leq 894)$
b) $P(896 \leq \bar{x} \leq 903)$
c) $P(\bar{x} \geq 890)$

10.5.4) Seja $X : N\left(1200;\sqrt{1444}\right)$. Retiramos uma amostra de tamanho igual a 15. Determinar:

a) $P(1194 < \bar{x} < 1206)$
b) $P(\bar{x} \geq 1150)$

10.5.5) Uma fábrica de peças especifica em suas embalagens que a proporção de defeitos é de 4%. Um cliente dessa fábrica inspeciona uma amostra de 200 peças e constata que 12 são defeituosas. Baseado nesses dados em quantas amostras o cliente encontraria uma proporção de defeitos maior que o especificado pelo fabricante?

➢ **Intervalos de Confiança para Média Populacional** μ

10.5.6) De uma população normal com parâmetros desconhecidos, tiramos uma amostra de tamanho igual a 100, obtendo-se $\bar{x} = 112$ e $s = 11$. Fazer um Intervalo de Confiança para μ ao nível de 10%.

10.5.7) Uma amostra aleatória de 200 possuidores de cartão mostra que o débito médio anual nesses cartões, para contas individuais, é $ 1592, com desvio – padrão de $ 997 (com base em dados do *USA Today*). Com essas estatísticas, construa um intervalo de 94% de confiança para o débito médio anual em cartões de crédito para a população de todas as contas.

10.5.8) Uma papelaria gostaria de calcular o valor médio do preço de cartões existentes em seu estoque. Uma amostra de 45 cartões indica um valor médio de R$ 1,67 com desvio – padrão de R$ 0,32. Construa um intervalo de confiança, para a verdadeira média do preço de um cartão, ao nível de 99%.

10.5.9) Em um esforço para estimar a quantia média gasta por cliente para jantar em um grande restaurante de Atlanta, foram coletados os dados de uma amostra de 49 clientes em um período de três semanas. Considere um desvio-padrão populacional de US$2,50.

a) Qual é o erro-padrão da média?
b) Qual é a margem de erro para um intervalo de confiança de 95%?
c) Se a média da amostra é US$22,60, qual é o intervalo de confiança de 95% para a média da população?

10.5.10) Para uma amostra de 32 consumidores foi levantada a seguinte questão: “Um vendedor seria responsável por um produto defeituoso ainda que envidasse os melhores esforços e cuidados na venda e produção?”. As respostas eram graduadas de 1 (discordam inteiramente) a 5 (concordam plenamente), e a média foi de 3,81 e o desvio padrão de 1,34.

a) Determine um intervalo de confiança de 90% para a média de pontos;
b) Sem precisar recorrer a cálculos, um intervalo de 80% para a média de respostas seria maior, menor ou igual àquela encontrada em (a)?

10.5.11) Um teste específico foi aplicado sobre uma amostra de 352 contadores. A média amostral foi de 60,41 e o desvio padrão de 11,28. Encontre um intervalo de confiança para a média populacional dos pontos dos contadores com 99% de certeza.

10.5.13) Para uma amostra de 457 fabricantes japoneses foi levantada a seguinte questão: "Reduzir níveis de defeitos dos produtos significa maiores custos?" As respostas eram graduadas de 1 (concordam inteiramente) a 5 (discordam plenamente), e a média foi de 3,59 e o desvio padrão de 1,045. Um intervalo de confiança entre 3,49 e 3,69 foi calculado para a média populacional. Encontre o nível de confiança associado a esse intervalo.

10.5.14) A uma amostra de 174 profissionais foi pedido que respondessem, numa escala de 1 (concordo fortemente) até 7 (discordo fortemente), à colocação: "Algumas vezes eu uso técnicas de pesquisa que garantem a obtenção de resultados que meus clientes gostariam de alcançar." A média de resposta foi de 6,06 e o desvio padrão de 1,43. Além disso, o intervalo de confiança para a média variou de 5,86 a 6,26. Determine o nível de confiança associado a esse intervalo.

10.5.15) Em um esforço para estimar a quantia média que cada cliente gasta por jantar em um grande restaurante de Atlanta foram coletados dados de uma amostra de 49 clientes. Suponha um desvio padrão de US$ 5,00 para população.

a) Para um grau de confiança de 95%, qual é a margem de erro?
b) Se a média amostral é US$ 24,80, qual é o intervalo de confiança de 95% para a média populacional?

10.5.16) A *Nielsen Media Research* relatou que o tempo médio que as famílias passam assistindo à televisão, no período das 8 h às 11 h da noite, é de 8,5 horas por semana (*The World Almanac 2003*). Dado um tamanho de amostra de 300 famílias e um desvio padrão σ da população igual a 3,5 horas, qual é a estimação por intervalo de confiança de 95% da média de tempo que as pessoas assistem à televisão durante o período das 8 h às 11 h da noite?

10.5.17) A equipe de vendas da *Skillings Distributors* apresenta semanalmente relatórios que relacionam os contatos feitos com clientes durante a semana. Uma amostra de 65 relatórios semanais exibiu uma média amostral de 19,5 contatos com clientes por semana. O desvio padrão da amostra foi de 5,2. Forneça os intervalos de confiança de 90% e 95% correspondentes ao número médio da população de contatos semanais com clientes feitos pela equipe de vendas.

10.5.18) Em um estudo sobre o tempo que os estudantes gastam para obter o grau de bacharel, 80 estudantes foram selecionados aleatoriamente e verificou-se que tinham uma média de 4,8 anos (com base em dados do *National Center for Educations Statistics*). Supondo $\sigma = 2{,}2$ anos, construa um intervalo de confiança de 97% para a média populacional. O intervalo de confiança contradiz o fato de que 39% dos estudantes obtêm seu grau de bacharel em quatro anos?

10.5.19) Um estudo das idades de motociclistas mortos em acidentes envolve a seleção aleatória de 150 motociclistas com idade média de 37,1 anos (com base em dados do *Institute of Highway Safety*). Supondo $\sigma = 12{,}0$ anos, construa um intervalo de confiança de 99% de confiança para a idade média de todos aos motociclistas mortos em acidentes. Se os limites do intervalo de confiança não incluem idades abaixo de 20 anos, isso significa que os motociclistas nessa faixa de idade raramente morrem em acidentes?

10.5.20) Estudantes de estatística do autor, selecionados aleatoriamente, participaram de um experimento para testar sua habilidade em determinar quando, a partir de certo instante, se passou um minuto (ou 60 segundos). Quarenta estudantes resultaram em uma média de 58,3.s. Supondo $\sigma = 9{,}5$ s, construa um intervalo de confiança de 95% de confiança para média populacional de todos os estudantes de estatística. Com base no resultado, é provável que essas estimativas tenham uma média que seja razoavelmente próxima de 60 segundos?

10.5.21) Quando as pessoas fumam, a nicotina que absorvem é convertida em cotinina, que pode ser medida. Uma amostra de 40 fumantes tem um nível médio de cotinina de 172,5. Supondo que σ seja conhecido como 119,5, ache um intervalo de confiança de 96% de confiabilidade para o nível médio de cotinina de todos os fumantes. Que aspecto desse problema não é realista?

10.5.22) Use os pesos de quartos de dólar pós 1964 do Conjunto de Dados 14 do Apêndice B. Supondo que os quartos de dólar sejam cunhados para produzir pesos com um desvio padrão populacional de 0,068 g, use a amostra de pesos para construir um intervalo de confiança de 99% de confiança para o peso médio. As especificações de cunhagem nos Estados Unidos exigem que os quartos de dólar tenham pesos entre 5,443 g e 5,987 g. O que o intervalo de confiança sugere sobre o processo de produção?

10.5.23) 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 tais erros tenham um desvio padrão de 2,5°, construa um intervalo de confiança de 95% de confiança para média de todos tais erros. O que o resultado sugere sobre a precisão das temperaturas previstas?

10.5.24) Uma amostra aleatória de peso ao nascer de 186 bebês têm uma média de 3103 g e um desvio padrão de 696 g (com base em dados de *"Cognitive Outcomes of Preschool Childre with Prenatal Cocaine Expoure", de Singer et al., Journal of the American Medical Association, Vol. 291, No.20*). Esses bebês nasceram de mães que não ousaram cocaína na gravidez. Construa um intervalo de confiança de 95% de confiabilidade para o peso médio ao nascer para todos tais bebês. Se o intervalo de confiança para o peso de bebês que nasceram de mães que usaram cocaína durante a gravidez foi determinado como $IC(\mu;95\%)=(2608;2792)g$, o uso de cocaína parece afetar o peso ao nascer de um bebê?

10.5.25) O Conjunto de Dados 2 no Apêndice B inclui 107 temperaturas corporais para as quais $\bar{x}=98,20^{o}F$ e $s=0,62^{o}F$. Usando as estatísticas amostrais, construa um intervalo de confiança de 99% de confiança para a temperatura média corporal para todas as pessoas sadias. Os limites do intervalo de confiança contêm 98,6 °F? O que a amostra sugere sobre o uso de 98,6°F como temperatura média do corpo?

10.5.26) O Conjunto de Dados 8 do Apêndice B inclui uma lista de temperaturas máximas reais e a lista correspondente das previsões com três dias de temperaturas máximas. Se a diferença para cada dia é encontrada subtraindo-se as temperaturas máximas previstas com três dias das temperaturas máximas reais, o resultado é uma lista de 35 valores como a média de -1,3 ° e desvio padrão de 4,7 °.

a) Construa um intervalo de confiança de 99% confiança para a diferença média entre as temperaturas máximas reais e as temperaturas máximas previstas com três dias.
b) O intervalo de confiança inclui o 0°? Se o meteorologista afirma que a previsão de três dias das temperaturas máximas tende a ser muito alta porque a diferença média da amostra é –1,3°, essa afirmativa parece ser válida? Por que sim ou por que não?

10.5.27) Uma médica deseja desenvolver um critério para determinar se a taxa de pulsação de um paciente é atípica e deseja determinar também se há diferenças significativas entre homens e mulheres. Use a amostra de taxas de pulsação do Conjunto de Dados 1 no Apêndice B.

a) Construa um intervalo de Confiança de 95% de confiança para a taxa de pulsação média dos homens;
b) Construa um intervalo de Confiança de 95% de confiança para a taxa de pulsação média das mulheres;
c) Compare os resultados precedentes. Podemos concluir que as médias populacionais para homens e mulheres sejam diferentes? Por que sim ou por que não?

10.5.28) Consulte o Conjunto de Dados 12 do Apêndice B e use os dados amostrais.

a) Construa um intervalo de confiança de 95% de confiança para o peso médio de cola em latas de Pepsi normal.
b) Construa um intervalo de confiança de 95% de confiança para o peso médio de cola em latas de Pepsi dietética.
c) Compare os resultados das partes (a) e (b) e os interprete. Parece haver alguma diferença? Caso positivo, identifique uma razão para a diferença.

10.5.29) Uma amostra aleatória de 80 notas de matemática, de uma população com distribuição normal de 500 notas, apresenta média de 5,5 e desvio padrão de 1,25.

a) Quais os limites de confiança de 95% para a média das 500 notas?
b) Com que grau de confiança pode-se dizer que a média das notas é maior que 5,0 e menor que 6,0?

10.5.30) De uma população normal com $\sigma^2 = 16$, levantou-se uma amostra, obtendo-se as observações: 10, 5, 10, 15. Determinar ao nível de 13% um Intervalo de Confiança para a média da população.

10.5.31) Dada uma população normal com $\sigma^2 = 3$, levantou-se uma amostra de 4 elementos, tal que $\sum_{i=1}^{4} x_i = 0{,}8$. Construir um Intervalo de Confiança para a verdadeira média populacional μ ao nível de 1%.

10.5.32) A experiência com trabalhadores de certa indústria indica que o tempo necessário para que um trabalhador, aleatoriamente selecionada, realize uma tarefa é distribuído de maneira aproximadamente normal, com desvio padrão de 12 minutos. Uma amostra de 25 trabalhadores forneceu $\overline{x} = 140\,\text{min}$. Determinar os limites de confiança de 95% para a μ da população de todos os trabalhadores que fazem aquele determinado serviço.

10.5.33) A secretaria de um curso de graduação determinou que historicamente os candidatos apresentam notas com desvio padrão 0,45. Uma amostra de 25 fichas de inscrição foi levantada, mostrando uma média geral de 2,9.

a) Encontre um intervalo de 95% de certeza para a média populacional.
b) Uma pessoa determinou o intervalo de confiança para a média da população como sendo de 2,81 a 2,99. Qual o nível de confiança associado a este IC (intervalo de confiança).

10.5.34) Um gerente de RH determinou que historicamente os escores dos testes de aptidão para os candidatos a emprego apresentavam desvio padrão de 32,4 pontos. Uma amostra de 9 elementos apresentou média de 187,9.

a) Determine um intervalo de confiança com 80% de certeza para a média de pontos.
b) Um analista de RH determinou um intervalo de confiança com valores entre 165,8 e 210,0 pontos. Encontre o nível de confiança associado a este intervalo.

10.5.35) Uma escola de administração está estudando os níveis salariais de seus ex-alunos. Sabe-se que o desvio padrão dos salários é de $4.780. Uma amostra de 25 graduados foi levantada e anotou-se uma média de $42.740. Determine um IC a 90% de certeza para a média de salários dos graduados.

10.5.36) Em um estudo sobre aplicação do tempo, constatou-se que 20 administradores selecionados aleatoriamente gastam uma média de 2,40 horas por dia com serviço burocrático. Estudos anteriores indicam que o desvio padrão seja de 1,30 h. Considerando que os dados apresentam uma distribuição aproximadamente normal, determine um intervalo de confiança de 95% para o tempo médio gasto em trabalho burocrático por todos os administradores.

10.5.37) Uma Amostra aleatória de 19 mulheres acusou altura média de 63,85 in. Outro estudo revelou que as alturas das mulheres têm distribuição normal com desvio padrão de 2,5 in. Construa um intervalo de confiança de 98% para altura média de todas as mulheres.

10.5.38) Quando 14 estudantes de segundo ano do curso de medicina, no Hospital Bellevue, mediram a pressão sanguínea da mesma pessoa, obtiveram os resultados listados abaixo. Supondo que o desvio padrão da população seja conhecido e igual a 10 mmHg, construa um intervalo de confiança, com uma confiabilidade de 93%, para a média populacional. Idealmente, qual deveria ser o intervalo de confiança nessa situação?

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

10.5.39) O menor mamífero do mundo é o morcego mamangaba, também conhecido como morcego com nariz de porco (ou *Craseonycteris thonglonyai*). Tais morcegos são basicamente do tamanho de uma mamangaba (abelha). Abaixo, listam-se os pesos (em gramas) de uma amostra desses morcegos. Supondo que os pesos desses morcegos tenham um desvio padrão de 0,30 g, construa um intervalo de confiança de 95% de confiança para seu peso médio. Use o intervalo de confiança para determinar se essa amostra de morcegos é proveniente da mesma população com uma média conhecida de 1,8 g.

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

10.5.40) Seja X uma variável aleatória normal com parâmetros desconhecidos. Dessa população foi retirada uma amostra $x_i: 10,\ 12,\ 14,\ 15,\ 9,\ 12,\ 16,\ 11,\ 8,\ 13.$ Construir um Intervalo de Confiança para μ ao nível de 5%.

10.5.41) Dado que uma amostra apresenta $\bar{x} = 20,\ s = 24 \text{ e } n = 16,$ com X normalmente distribuída, determinar os limites de confiança de 95% para a média.

10.5.42) Construir um Intervalo de Confiança de 90% para a média de uma população normal com variância desconhecida, sabendo-se que uma amostra de 26 observações fornece $\bar{x} = 15{,}6$ e $s = 2{,}58.$

10.5.43) Um gerente de uma agencia deseja calcular a quantia média mantida no banco pelos depositantes em contas de cadernetas de poupança. Uma amostra aleatória de 25 depositantes é selecionada, e os resultados indicam uma Média de R$ 4.700,00 e um Desvio – Padrão de R$ 1.200,00. Determine um Intervalo de Confiança para a verdadeira média da população ao nível de 95%.

10.5.44) A Associação Americana de Agências de Propaganda registra dados sobre minutos sem programação nos programas de meia hora no horário nobre de televisão. Os dados representativos em minutos para uma amostra de 20 programas de horário nobre nas maiores redes às 20:30 h são apresentados a seguir:

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

Forneça um intervalo de confiança de 93% para o número médio de minutos sem programação nos programas de meia hora no horário nobre de televisão às 20:30h.

10.5.45) Foi levantada uma amostra de 25 pessoas, às quais se fez a pergunta: “A maior parte das campanhas de marketing agridem a inteligência média das pessoas”. As respostas eram graduadas de 1 (discordam inteiramente) a 7 (concordam plenamente), e a média foi de 3,92 e o desvio padrão de 1,57. Determine um intervalo de confiança com 95% de certeza para a média de respostas da população.

10.5.46) Trinta restaurantes de *fast-food*, incluindo o Wendy’s, o McDonald’s e o Burger King, foram frequentados durante o verão de 2000 (*The Cincinnatti Enquirer*, 9 de julho de 2000). Durante cada visita, o cliente ia ao *drive-through* e pedi uma refeição básica, por exemplo, uma refeição “combo” ou um sanduíche, batatas fritas e um *Milk-shake*. Foi registrado o tempo decorrido entre escolher a opção do cardápio e receber o pedido. Os tempos, em minutos, para as 30 visitas foram os seguintes:

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

a) Apresente uma estimação por ponto da média populacional de tempo gasto nos *drive-througths* dos restaurantes de *fast-food*;
b) Qual é a estimação por intervalo de confiança de 97% para a média populacional?

10.5.47) Como as mortes por problemas cardíacos parecem aumentar após fortes nevascas, planejou se um experimento para comparar as necessidades cardíacas de pessoas que usam o limpador de neve manual com as pessoas que limpam a neve usando um limpador elétrico. Dez sujeitos limparam áreas com neve usando os dois métodos, suas taxas cardíacas máximas (batimentos por minuto) foram registradas durante ambas as atividades. Obtiveram-se os seguintes resultados (com base em dados que *"Cardiac Demands of Heavy Snow Shoveling", de Franklin et al.*, Vol. 273, No. 11);

Taxas máximas de batimentos cardíacos para limpeza manual	$n=10$	$\bar{x}=175$	$s=15$
Taxas máximas de batimentos cardíacos para limpeza com limpador elétrico	$n=10$	$\bar{x}=124$	$s=18$

a) Ache um intervalo de confiança de 95% de confiança para a média populacional das pessoas que limpam neve manualmente;
b) Ache um intervalo de confiança de 95% de confiança para média populacional das pessoas que e limpam neve com limpador elétrico;
c) Se você fosse um médico preocupado com as mortes por causas cardíacas devidas à limpeza de neve por métodos manuais, que valor um intervalo de confiança da parte (a) seria de maior interesse?
d) Compare os intervalos de confiança das partes (a) e (b) e interprete o resultado.

10.5.48) A seguir estão listadas as quantidades medidas de chumbo no ar (em microgramas por metro cúbico ou $\mu g/m^3$). A agência de proteção ambiental estabeleceu um padrão de qualidade do ar para o chumbo de 1,5 $\mu g/m^3$. As medidas mostradas a seguir foram feitas no Edifício 5 do local do *Word Trade Center*, em dias diferentes, após o atentado terrorista de 11 de setembro de 2001. Depois do colapso das duas torres do *Word Trade Center*, houve muita preocupação sobre a qualidade do ar. Use os valores dados para construir um intervalo de confiança de 95% de confiança para a quantidade média de chumbo no ar. Há alguma coisa nesses dados que sugira que o intervalo de confiança possa não ser muito bom? Explique.

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

10.5.49) O ramo-e-folhas abaixo lista as idades de candidatos a promoções que foram bem-sucedidos (com base em *"Debating the Use of Statistical Evidence in Allegation of Age Discrimination", de Barry e Boland, American Statistician, Vol. 58, No. 2*). Suponha que a amostra seja amostra aleatória simples e construa num intervalo de confiança de 95% de confiança para a idade média de todas tais pessoas bem-sucedidas.

3	3 6 7 8 8 9
4	2 2 3 3 4 4 4 5 5 5 5 6 6 7 7 8 8 9 9
5	1 1 2 4

10.5.50) 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. Use os dados amostrais para construir um intervalo de 99% de confiança para o escore FICO médio de todos pretendentes a empréstimos. Se um banco exige uma classificação de crédito de, pelo menos, 620 para a compra de um carro, parece que quase todos pretendentes terão classificação de crédito adequadas?

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

10.5.51) Em uma amostra de sete carros, cada um foi testado em relação à emissão de óxido de nitrogênio em gramas por milha) e obtiveram se os seguintes resultados: 0,06; 0,11; 0,16; 0,15; 0,14; 0,08; 0,15 (com base em dados da Agência de Proteção Ambiental dos Estados Unidos). Supondo que essa amostra seja representativa dos carros em uso, construa um intervalo de confiança de 98% de confiança para a quantidade média das emissões de dióxido de o nitrogênio para todos os carros. Se a Agência de Proteção Ambiental exige que as emissões de óxido de nitrogênio sejam menores do que 0,165 g/ milha, pode-se concluir o com segurança que essa exigência esteja sendo satisfeita?

10.5.52) A seguir estão listadas as larguras máxima de amostras de crânios de homens egípcios de 4000 a.C. e 750 d.C. (com base em dados de *Ancient Races of the Thebaid, de Thonson e Randall-Macivery*):

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

Mudanças nos tamanhos nas cabeças ao longo do tempo sugerem o cruzamento com pessoas de outras regiões. Use o intervalo de confiança de 97% para determinar se os tamanhos das cabeças parecem ter mudado de 4000 a.C. até 150 d.C. Explique seu resultado.

- **Intervalos de Confiança para a Proporção Populacional**

10.5.53) Para se estimar a porcentagem de alunos de um curso favoráveis à modificação do currículo escolar, tomou-se uma amostra de 100 alunos, dos quais 80 foram favoráveis.

a) Fazer um Intervalo de Confiança para a proporção de todos os alunos do curso favoráveis à modificação ao nível de 4%.
b) Qual o valor do erro de estimação cometido na letra a?

10.5.54) Em uma linha de produção de certa peça mecânica, colheu-se uma amostra de 100 itens, constatando-se que 4 peças eram defeituosas. Construir um Intervalo de Confiança para a proporção "p" das peças defeituosas ao nível de 10%.

10.5.55) Em uma pesquisa de opinião, entre 600 pessoas pesquisadas, 240 responderam "sim" a determinada pergunta. Estimar a porcentagem de pessoas com essa mesma opinião na população, dando um intervalo de 95% de confiabilidade.

10.5.56) Uma votação realizada entre 400 eleitores, escolhidos ao acaso dentre todos os eleitores de um determinado distrito, indicou que 55% deles são a favor do candidato "A". Determinar os limites de confiança de 99% para a proporção de eleitores do distrito favoráveis ao candidato "A".

10.5.57) Um levantamento da Time/CNN solicitou a 814 adultos que respondessem a uma série de questões sobre suas perspectivas em relação à situação financeira dos Estados unidos. Um total de 562 adultos respondeu "Sim" à questão:

a) Você acha que as coisas estão indo bem nos Estados Unidos de Hoje?
b) Qual é o intervalo de confiança de 90% para a proporção de adultos que sentem que as coisas estão indo bem nos Estados Unidos?

10.5.58) De uma amostra de 134 auditores empregados numa grande empresa, 82 disseram que quando recebe um novo cliente, eles sempre perguntam ao auditor anterior a razão da mudança de empresa de auditoria. Determine um IC a 95% de certeza para a proporção de auditores que adotam tal procedimento.

10.5.59) De uma amostra de 323 membros de sindicatos, 47,9% concordavam com a declaração: "Trabalhadores sindicalizados deveriam se recusar a trabalhar caso trabalhadores não sindicalizados sejam empregados pela empresa". Baseado nesta informação, uma pessoa determinou um IC para a proporção de empregados sindicalizados com esta visão como sendo de 45,8% a 50,0%. Determine o nível de confiança associado a este intervalo.

10.5.60) De uma amostra de 95 empresas, 29 indicaram melhorias de qualidade como a mais importante ação tomada para revitalizar produtos ou melhorar a desempenho competitivo. Determine um IC a 99% para a proporção populacional dessas empresas. Além disso, e sem proceder a cálculos, estabeleça se um IC a 90% para a proporção populacional seria maior ou menor do que a encontrada para 99% de certeza.

10.5.61) Foi levantada uma amostra de 96 empresas que usavam distribuidores americanos para seus produtos. Dos membros presentes na amostra, 32 disseram que os distribuidores não eram capazes de exercer a função de assistência técnica. Determine um IC a 80% de certeza para a proporção populacional dessas empresas.

10.5.62) De uma amostra de 198 estudantes, 98 apontaram como ponto negativo colocar "floreios" em currículos profissionais. Um pesquisador levantou um intervalo de confiança para a proporção populacional destes estudantes como sendo de 0,445 a 0,545. Qual o nível de confiança associado a esse intervalo.

10.5.63) Numa amostra de 1158 executivos promovidos recentemente, 47,9% afirmaram que cursos de pós-graduação eram importantes, como parte da preparação para uma carreira de negócios.

a) Determine um IC a 99% de certeza para a proporção populacional de executivos com este ponto de vista.
b) Baseado na informação amostral, um gerente desenvolveu um IC para a proporção populacional como sendo de 0,458 a 0,500. Qual o nível de confiança associado a esse intervalo?

10.5.64) De uma amostra aleatória de 113 consumidores, 57 disseram possuir um aparelho de TV nacional. Determine um IC a 95% de certeza para a proporção populacional de todos os consumidores que diriam tal coisa.

10.5.65) Em uma amostra aleatória de 200 alunos de uma grande universidade, 144 são contrários ao aumento do número de disciplinas curriculares na área de ciências sociais, enquanto que 56 são favoráveis a tal aumento. Construa um intervalo de 95% de confiança para a proporção da população que se opõe e a que é a favor ao aumento da carga curricular.

10.5.66) O *Genetics and IVF Institute* realizou um experimento clínico do método XSORT, projetado para aumentar a probabilidade de se conceber uma menina. Nascem 325 bebês de pais que usam o método XSORT, e 295 deles são meninas. Use os dados amostrais para construir um intervalo de confiança de 99% de confiança para a porcentagem de meninas que nascem de pais que usam o método XSORT. Com base nesse resultado, o método XSORT parece ser eficaz? Por que sim ou por que não?

10.5.67) O *Genetics and IVF Institute* realizou um experimento clínico do método YSORT, projetado para aumentar a probabilidade de se conceber um menino. Enquanto este livro estava sendo escrito, 51 bebês nasceram de pais que usavam o método YSORT, e 39 desses bebês eram meninos. Use os dados amostrais para construir um intervalo de confiança de 99% de confiança para a porcentagem de meninos nascidos de pais que usam o método YSORT. Com base nesse resultado, o método YSORT parece ser eficaz? Por que sim ou por que não?

10.5.68) Uma hipótese interessante e popular é a de que as pessoas podem adiar temporariamente a morte para sobreviverem a um feriado maior ou a um evento importante, como um aniversário. Em um estudo desse fenômeno, encontrou-se que na semana anterior e na semana posterior ao feriado de Ação de Graças houve um total de 12.000 mortes, das quais 6.062 ocorreram na semana anterior ao feriado (com base em dados de "*Holidays, Birthdays and Postponement of Cancer Death*", de *Young e Hade, Journal of the American medical Association*, Vol. 292, No.24.) Construa um intervalo de confiança de 95% de confiança para a proporção de mortes na semana anterior ao feriado de Ação de Graças em relação ao total de mortes nas semanas antes e depois do feriado. Com base no resultado, parece haver alguma indicação de que as pessoas adiem sua morte para sobreviverem ao feriado de Ação de Graças? Por que sim ou por que não?

10.5.69) Um importante problema que os americanos enfrentam é o grande número de processos relativos à imperícia médica e às despesas que eles geram. Em um estudo de 1228 processos de imperícia médica selecionados aleatoriamente, encontrou-se que 856 deles foram, mais tarde, retirados ou rejeitados (com base em dados da *Physician Insurers Association of America*). Construa um intervalo de confiança de 99% de confiança para a proporção de processos sobre imperícia médica que são retirados ou rejeitados. Parece que a maioria de tais processos é retirada ou rejeitada?

10.5.70) Quando Mendel realizou seus famosos experimentos em genética com ervilhas, uma amostra das descendentes consistia de 428 ervilhas verdes e 152 amarelas.

a) Ache um intervalo de confiança de 95% de confiança para a porcentagem de ervilhas amarelas;
b) Com base na teoria genética, Mendel esperava que 25% das ervilhas descendentes fossem amarelas. Dado que a porcentagem das ervilhas descendentes não é 25%, os resultados contradizem a teoria de Mendel? Por que sim ou por que não?

10.5.71) Em uma pesquisa com 1002 pessoas, 701 disseram que votaram em uma recente eleição presidencial (com base em dados do *ICR Research Group*). Os registros da votação mostram que 61% dos eleitores habilitados realmente votaram.

a) Ache um intervalo de confiança de 99% de confiança para a proporção de pessoas que dizem ter votado;
b) Os resultados da pesquisa estão de acordo com o resultado real de votantes de 61%? Por que sim ou por que não?

10.5.72) Um estudo com 420.095 dinamarqueses usuários de telefones celulares descobriu que 135 deles tinham desenvolvido câncer no cérebro ou no sistema nervoso. Anteriormente a esse estudo do uso de telefone celular, a taxa desse tipo de câncer era de 0,0340% para aqueles que não usavam o telefone celular. Os dados são do *Journal of the National Cancer Institute*.

a) Use os dados amostrais para construir um intervalo de confiança de 95% de confiança para a porcentagem de usuários de telefone celular que desenvolveram câncer de cérebro ou do sistema nervoso;
b) Os usuários de telefones celulares parecem ter uma taxa de câncer do cérebro ou do sistema nervoso que seja diferente da taxa de tais cânceres entre os que não usam telefones celulares? Por que sim ou por que não?

10.5.73) Uma recente pesquisa do Gallup perguntou a 1012 adultos selecionados aleatoriamente se "a clonagem humana deve ou não deve ser permitida". Os resultados mostram que 901 dos entrevistados indicam que a clonagem não deveria ser permitida. Um repórter deseja determinar se esses resultados de pesquisa constituem forte evidência de que a maioria (mais de 50%) das pessoas seja contra tal clonagem. Construa um intervalo de confiança de 99% de confiança para a proporção de adultos que acreditam que a clonagem de humanos não deva ser permitida. Com base nesse resultado, há evidência forte que apoie a tese de que a maioria se oponha a tal clonagem?

10.5.74) No caso *Castaneda vs. Partida*, descobriu-se que, no Condado de Hidalgo, Texas, durante um período de 11 anos, 870 pessoas haviam sido escolhidas para o júri e 39% delas eram mexicano-americanos. Use os dados amostrais para construir um intervalo de confiança de 99% de confiança para a proporção de membros do júri que eram mexicano-americanos. Dado que, entre as pessoas elegíveis para o júri, 79,1% eram mexicano-americanos, parece que o processo para seleção para o júri era, de alguma maneira, viesado contra os mexicanos-americanos? Por que sim ou pro que não?

10.5.75) Quando trabalhava para a promotoria do distrito do Brooklyn, o investigador Robert Burton analisou os dígitos iniciais das quantias em cheques de companhias que estavam sob suspeita de fraude. Entre 784 cheques, 61% tinham quantias com o dígito inicial igual a 5. Construa um intervalo de confiança de 90% de confiabilidade para a proporção de cheques que têm quantias com o dígito inicial igual a 5. Quando cheques são emitidos no curso normal de transações honestas, espera-se que 7,9% desses cheques tenham quantias cujos dígitos iniciais sejam 5. O que o intervalo de confiança sugere?

10.5.76) Em 1920, apenas 35% das residências nos Estados Unidos tinham telefone, mas essa taxa é, agora, muito maior. Uma pesquisa recente com 4276 residências selecionadas aleatoriamente mostrou que 94% delas tinham telefone (com base em dados do *U.S. Census Bureau*). Usando esses resultados da pesquisa, construa um intervalo de confiança de 99% de confiança para a proporção de residências com telefones. Uma vez que a pesquisa envolva apenas 4276 de um total de 115 milhões de residências, temos realmente evidência suficiente para dizer que a porcentagem de residências com telefone é agora maior que a taxa de 35% de 1920?

10.5.77) Em uma pesquisa do Gallup, 1025 adultos selecionados aleatoriamente foram entrevistados e 29% deles disseram usar a internet para compras pelo menos algumas vezes por ano.

a) Ache uma estimativa pontual da porcentagem de adultos que usam a internet para compras;
b) Ache um intervalo de confiança de 99% de confiança para a porcentagem de adultos que usam a internet para compras;
c) Se uma loja de varejo tradicional deseja estimar a porcentagem de adultos que compram pela internet para determinar o impacto máximo desses compradores em suas vendas, qual porcentagem de compradores pela internet deve ser usada?

10.5.78) Consulte o Conjunto de Dados 13 no Apêndice B e ache a proporção amostral de balas M&M que são azuis. Use este resultado para construir um intervalo de confiança de 95% de confiança para a porcentagem populacional das balas M&M que são azuis. Esse resultado é condizente com a taxa de 24% apresentada pelo fabricante de balas Mars?

10.5.79) Consulte o Conjunto de Dados 5 no Apêndice B.

a) Construa um intervalo de confiança de 95% de confiança para a porcentagem de filmes de desenhos animados infantis que mostram algum uso de fumo;
b) Construa um intervalo de confiança de 95% de confiança para a porcentagem de filmes de desenhos animados infantis que mostram algum uso de álcool;
c) Compare os resultados precedentes. O fumo ou o álcool aparecem em porcentagens, uma maior do que a outra, em filmes de desenhos animados infantis?
d) No uso dos resultados (a) e (b) como medidas para a descrição de hábitos não saudáveis, qual característica importante dos dados não está incluída?

10.5.80) Consulte o Conjunto de dados 10 no Apêndice B e considere os dias com valores de precipitação diferentes de 0 como dia com precipitação. Construa um intervalo de confiança de 95% de confiança para a proporção de quartas-feiras com precipitação, e construa também um intervalo de confiança para a proporção de domingos com precipitação. Compare os resultados. A precipitação parece ocorrer mais em algum dos dias?

10.5.81) Consulte o Conjunto de Dados 8 no Apêndice B. Construa um intervalo de confiança de 95% para a proporção de dias em que a temperatura máxima real difere por mais de 2° da temperatura máxima prevista um dia antes. Construa então um intervalo de confiança de 95% de confiança para a proporção de dias com a temperatura máxima real diferente por mais de 2° da temperatura máxima prevista cinco dias antes. Compare os resultados.

- **Intervalo de Confiança Para o Desvio Padrão e para a Variância Populacional**

10.5.82) De uma população normal levantou – se uma amostra casual de 27 elementos, obtendo – se:

6,599	8,756	9,184	10,510
10,738	12,881	14,215	14,646
16,412	6,989	8,922	9,568
10,601	12,029	13,331	14,253
16,176	16,437	8,685	8,987
10,396	10,643	12,809	14,210
14,602	16,411	16,649	

Construa um intervalo de confiança para a média populacional e um intervalo de confiança para o desvio – padrão populacional, ambos com 95% de nível de confiança.

10.5.83) De uma população normal com média igual a 20, levantou-se uma amostra de 24 elementos, obtendo-se $\sum_{i=1}^{24}(x_i-\mu)^2=423,42$. Construa um intervalo de confiança de 90% para a variância populacional.

10.5.84) Sabe-se que o tempo de vida de certo tipo de válvula tem distribuição aproximadamente normal. Uma amostra de 25 válvulas forneceu $\overline{x}=500h$ e $s=50h$. Construir um Intervalo de Confiança para σ^2, ao nível de 2%.

10.5.85) De uma população normal levantou – se uma amostra de 10 observações, obtendo – se os seguintes valores:

10, 8, 15, 11, 13, 19, 21, 13, 15 e 14.

Sabendo – se que a população tem média $\mu=14$, construir um Intervalo de Confiança para a σ^2 populacional ao nível de 5%.

10.5.86) O Desvio Padrão de 10 lâmpadas produzidas por uma fábrica é de 120 horas. Construir um Intervalo de Confiança para a variância de todas as lâmpadas fabricadas ao nível de 10%.

10.5.87) De uma população normal com média 4, levantou – se uma amostra casual de 21 elementos, obtendo – se:

1, 2, 2, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 6, 6, 7.

Ao nível de 10%, construir um Intervalo de Confiança para a variância populacional.

10.5.88) De uma população normal com média desconhecida, levantou – se uma amostra de tamanho igual a 20, obtendo – se $\sum_{i=1}^{20}x_i=114$ e $\sum_{i=1}^{20}x_i^2=846$. Ao nível de 10%, construir um Intervalo de Confiança para a variância da população.

10.5.89) Construir um Intervalo de Confiança para a variância populacional, ao nível de 5%, para as seguintes amostras.

a) $X = \{2,2,3,4,4,5,5,5,6,6,7,7,7,7,8,8,9,9,9,10\}$

b) $X = \{1,2;1,3;1,5;1,5;1,5;1,5;1,7;1,7;1,8;1,9;2,1;2,2;2,2;2,5;2,8\}$

c) $X = \{23,4;23,6;24,0;25,7;26,2;26,3;26,3;27,9;28,2;28,5;28,6;29,0\}$ `

d) $X = \{3;3;3;3;3;5;5;5;6;6;6;8;8;8;8;8;8;9;9;11;11;11;11;12;12;12;15;15;15;16;16;16;17\}$

e) $X = \left\{ \begin{matrix} 8;11;14;15;16;17;19;20;22;25;29;31;32;35;36;37;38;41;43;44;45;46;47; \\ 49;52;54;57;58;59;60;63;67;68;69;71;72;74;75;77;78;79;82;83 \end{matrix} \right\}$

10.5.90) Em um estudo dos efeitos sobre bebês do uso de cocaína pelas mães no período pré-natal, 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 os dados amostrais para construir um intervalo de confiança de 95% para o desvio padrão de todos os pesos ao nascer de bebês nascidos de mulheres que usaram cocaína durante a gravidez. Com base no resultado, o desvio padrão parece ser diferente do desvio padrão de 696 g para pesos ao nascer de bebês nascidos de mães que não usaram cocaína na gravidez?

10.5.91) Quartos de dólar são cunhados atualmente com pesos que têm uma média de 5,670 g e um desvio padrão de 0,062 g. Um novo equipamento está sendo testado numa tentativa de melhorar a qualidade pela redução da variação. Uma amostra aleatória simples de 24 quartos de dólar é obtida daqueles fabricados pelo novo equipamento, e essa amostra tem um desvio padrão de 0,049 g. Use os resultados amostrais para construir um intervalo de confiança de 95% para o desvio padrão dos pesos de quartos de dólar cunhados pelo novo equipamento. Com base no intervalo de confiança, parece que o novo equipamento produz um desvio padrão que seja claramente mais baixo do que o desvio padrão de 0,062 g do equipamento antigo? Com base nos resultados, o novo equipamento parece ser eficaz na redução da variação dos pesos?

10.5.92) O Conjunto de Dados 2 do Apêndice B inclui 107 temperaturas do corpo para as quais $\bar{x} = 98,20^o F$ e $s = 0,62^o F$. Usando as estatísticas amostrais, construa um intervalo de confiança de 95% de confiabilidade para o desvio padrão da temperatura do corpo de todas as pessoas saudáveis. Com base no resultado, podemos concluir com segurança que o desvio padrão da população seja menor do que 2,10 °F? (Se o desvio padrão da população for de 2,10 ou mais, a variação será grande o bastante de modo que a média amostral de 98,20 °F não difira de 98,6 °F por uma quantidade significativa).

10.5.93) Como as mortes por problemas cardíacos parecem aumentar depois de fortes nevascas, planejou-se um experimento para comparar as demandas cardíacas para limpar neve manualmente com as demandas dos que usam um limpador elétrico. Dez sujeitos limparam áreas de neve usando ambos os métodos, e seus batimentos cardíacos máximos (batidas por minuto) foram registrados durante as atividades. Obtiveram-se os seguintes resultados (com base em dados que *"Cardiac Demands of Heavy Snow Shoveling", de Franklin et al., Vol. 273, No. 11*);

Taxas máximas de batimentos cardíacos para limpeza manual	$n = 10 \quad \bar{x} = 175 \quad s = 15$
Taxas máximas de batimentos cardíacos para limpeza com limpador elétrico	$n = 10 \quad \bar{x} = 124 \quad s = 18$

a) Construa um intervalo de confiança de 95% de confiança para o desvio padrão populacional σ para os que limpam neve manualmente;

b) Construa um intervalo de confiança de 95% de confiança para o desvio padrão populacional σ para os que e limpam neve com limpador elétrico;

c) Compare os resultados. A variação parece ser diferente para os dois grupos?

10.5.94) Quando consumidores pedem crédito, seu crédito é classificado usando-se os escores FICO (Fair, Isaac, and Company). Abaixo estão as classificações de crédito para uma amostra de candidatos a empréstimos para a compra de carro. Use os dados amostrais para construir um intervalo de confiança para o desvio padrão dos escores FICO para todos os pretendentes a crédito. (***** Suponha uma confiabilidade de 99% *****)

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

10.5.95) O menor mamífero do mundo é o morcego mamangaba, também conhecido como morcego de focinho de porco (ou *Craseonycteris thonglonguai*). Tais morcegos são basicamente do tamanho de uma mamangaba. Abaixo, a lista os seus pesos (em gramas) de uma amostra desses morcegos. Construa um intervalo de confiança de 95% de confiança para o desvio padrão dos pesos de todos os morcegos deste tipo.

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

10.5.96) A seguir estão de listadas as quantidades medidas de chumbo no ar (em microgramas por metro cúbico ou $\mu g / m^3$). a agência de proteção ambiental estabeleceu um padrão de qualidade do ar para o chumbo fundo de 1,5 $\mu g / m^3$. As medidas mostradas a seguir foram feitas no Edifício 5 do local do *Word Trade Center*, em dias diferentes, após o atentado terrorista de 11 de setembro de 2001. Após o colapso das duas torres do *Word Trade Center*, houve muita preocupação sobre a qualidade do ar. Use os valores dados para construir um intervalo de confiança de 95% de confiança para o desvio padrão das quantidades de chumbo no ar. Há alguma coisa nesses dados que sugira que o intervalo de confiança possa não ser muito bom? Explique.

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

10.5.97) a) Os valores listados são tempos de espera (em minutos) de clientes do banco Jeferson Valley, onde os clientes fazem uma única fila que leva a três caixas. Construa um intervalo de confiança de 95% de confiança para o desvio padrão populacional σ.

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

b) Os valores listados são tempos de espera (em minutos) de clientes do Banco Providence, onde os clientes fazem três filas diferentes para cada um dos três caixas. Construa um intervalo de confiança de 95% de confiança para o desvio padrão populacional σ.

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

c) Interprete os resultados encontrados nas partes (a) e (b). Os Intervalos de confiança sugerem uma diferença na variação entre os tempos de espera? Que arranjo parece ser melhor: o de fila única ou o de múltiplas filas?

10.5.98) Consulte o Conjunto de Dados 1 do Apêndice B e use os dados amostrais.

a) Construa um intervalo de confiança de 99% de confiança para o desvio padrão dos IMCs para os homens;
b) Construa um intervalo de confiança de 99% de confiança para o desvio padrão dos IMCs para as mulheres;
c) Compare e interprete os resultados.

10.5.99) Consulte o Conjunto de Dados 14 do Apêndice B e use os dados amostrais.

a) Construa um intervalo de confiança de 99% para o desvio padrão dos pesos de quartos de dólar cunhados depois de 1964;
b) Construa um intervalo de confiança de 99% para o desvio padrão dos pesos de quartos de dólar de prata cunhados antes de 1964;
c) Compare e interprete os resultados.

- **Determinação do Tamanho da Amostra para Estimar a Média Populacional**

10.5.100) A Nielsen Media Research deseja estimar o tempo médio (em horas) que os estudantes universitários de tempo integral passam vendo televisão em cada dia da semana. Determine o tamanho da amostra necessário para estimar essa média com uma margem de erro de 0,25 (ou 15 minutos). Suponha que se exija um grau de 96% de confiança. Suponha também que um estudo piloto tenha indicado que o desvio – padrão é estimado em 1,87 horas.

10.5.101) O teste Weschler é planejado de modo que a média seja 100 e o desvio padrão seja 15 para a população de adultos normais. Ache o tamanho de amostra necessário para estimar o escore de QI médio de estudantes de estatística. Desejamos ter 95% de confiança em que nossa média amostral esteja a menos de 2 pontos de QI do verdadeiro valor da média. A média para essa população é claramente maior do que 100. O desvio padrão para essa população é, provavelmente, menor do que 15, porque é um grupo com menos variação do que um grupo selecionado aleatoriamente da população geral; assim, se usarmos $\sigma = 15$, estaremos sendo conservadores, pois estaremos usando um valor que torna o tamanho amostral no mínimo tão grande quanto necessário. Suponha $\sigma = 15$ e determine o tamanho amostral requerido.

10.5.102) A Corporação Tyco de Videogame acha que está perdendo renda por causa de fichas usadas em seus videogames. As máquinas devem ser ajustadas para aceitar moedas apenas se estiverem dentro de certos limites. Para a determinação desses limites, o peso médio de quartos de dólar em circulação deve ser estimado. Uma amostra de quartos de Dolores será pesada para se determinar a média. Quantos quartos de dólar devem ser selecionados aleatoriamente e pesados, se desejamos 99% de confiança em que a média amostral estará a menos de 0,025 g da verdadeira média populacional de todos os quartos de dólar? Com base nos resultados de uma amostra de quartos de dólar, podemos estimar o desvio padrão população como 0,068 g.

10.5.103) Desejamos estimar a perda média de peso de pessoas um ano após usarem a dieta de Atkins. Quantas pessoas que fizeram a dieta devem ser entrevistadas se desejamos estar a 95% confiantes em que a perda de peso média amostral estará a 0,25 lb da verdadeira média populacional? Suponha que o desvio padrão populacional seja conhecido e igual a 10,6 lb (com base em dados de *"Comparison of the Atkins, Ornish, Weigth Watchers, and Zone Diets for Weigth Loss and Heart Disease Risk Redution", de Dansinger et al., Journal of the American Medical Association, Vol. 293, No. 1*).

10.5.104) Você acaba de ser contratado pela divisão de marketing da General Motors para estimar a quantidade média de dinheiro gasto, agora, na compra de carros novos, nos Estados Unidos. Primeiro, use a regra empírica para fazer uma estimativa grosseira do desvio padrão da quantidade gasta. É razoável supor que as quantidades típicas variem de 12.000 dólares a cerca de 70.000 dólares. Use, então, esse desvio padrão estimado para determinar o tamanho amostral correspondente a 95% de confiança e a uma margem de erro de US$100. Esse tamanho amostral é prático? Se não, o que deve ser mudado para se obter um tamanho prático de amostra?

10.5.105) Você deseja estimar a taxa média de pulsação de homens adultos. Consulte o Conjunto de Dados 1 do Apêndice B, ache as taxas mínima e máxima de pulsação para homens e use esses valores com a regra empírica da amplitude para estimar σ. Quantos homens adultos devem ser selecionados aleatoriamente e testados, se você deseja 95% de confiança em que a taxa média amostral de pulsação esteja a menos de 2 batidas (por minuto) da verdadeira média populacional μ? Se, em vez de se usar a regra empírica e amplitude, for usado o desvio padrão das taxas de pulsação dos homens do Conjunto de Dados 1 com estimativa de σ, o tamanho amostral requerido será muito diferente? Qual o tamanho amostral estará, provavelmente, mais próximo do valor correto do tamanho amostral?

- **Determinação do Tamanho da Amostra para Estimar a Proporção Populacional**

10.5.106) Suponha que estejamos interessados em estimar a porcentagem de consumidores de certo produto. Se uma amostra de tamanho igual a 300 forneceu 100 indivíduos que consomem o dado produto, determine o IC da proporção populacional, a 95% de certeza. Qual deverá ser o tamanho da amostra para que a margem de erro seja de 0,02?

10.5.107) Um fabricante de flashes deseja estimar a probabilidade de um flash funcionar. Como se trata de um teste destrutivo, ele deseja manter o tamanho da amostra o menor possível. Determine o número de observações que devem ser feitas para estimar a probabilidade, ao nível de confiança de 90%, sabendo-se que a proporção de defeituosos é de 6% e a margem de erro varia de 2% a 7%.

10.5.108) Muitos estados estão considerando cuidadosamente medidas que os ajudariam a recolher taxas de vendas de itens comprados pela internet. Quantas transações de vendas selecionadas aleatoriamente devem ser pesquisadas para se determinar a porcentagem que ocorre pela internet? Suponha que desejamos estar 99% confiantes em que a porcentagem amostral esteja a dois pontos percentuais da verdadeira porcentagem populacional de todas as transações de vendas.

10.5.109) A indústria de música deve se ajustar à prática crescente dos consumidores de baixar músicas em vez de comprar CDs. Assim, torna-se importante estimar a proporção de músicas que são atualmente baixadas. Quantas aquisições de músicas selecionadas aleatoriamente devem ser pesquisadas para se determinar a porcentagem das que foram obtidas por *download*? Suponha que desejamos estar 95% confiantes em que a porcentagem amostral esteja a um ponto percentual da verdadeira porcentagem populacional.

10.5.110) Recentemente foi projetada uma campanha para convencer os donos de carros de que eles devem encher os pneus com nitrogênio em vez de ar. A um custo de 5 dólares por pneu, o nitrogênio supostamente tem a vantagem de vazar a uma taxa mais lenta do que o ar, de modo que a pressão ideal do pneu se mantém por mias tempo. Antes de se gastarem grandes somas de dinheiro para a propaganda do nitrogênio, seria aconselhável realizar uma pesquisa para determinar a porcentagem de donos de carros que pagariam pelo nitrogênio. Quantos proprietários de carros, selecionados aleatoriamente, devem ser entrevistados? Suponha que desejamos estar 98% confiantes em que a porcentagem amostral esteja a três pontos percentuais da verdadeira porcentagem de todos os proprietários de carros que pagariam pelo nitrogênio.

10.5.111) A Toyota disponibiliza uma opção de teto solar e *air bags* laterais para seu modelo Corolla. Esse pacote custa 1400 dólares (1159 faturados). Suponha que, antes de oferecer esse pacote opcional, a Toyota deseje determinar a porcentagem de compradores de Corolla que pagariam 1400 dólares extras pelo teto e *air bags* laterais. Quantos compradores de Corolla devem ser entrevistados se desejamos estar 95% confiantes em que a porcentagem amostral esteja a quatro pontos percentuais para todos os compradores de Corolla?

10.5.112) No exercício 11.7.63 levantou-se um IC para a proporção de consumidores que dizem possuir um aparelho de TV nacional. Quantas observações seriam necessárias para estar certo que um IC a 95% de certeza para a proporção populacional não se estenderia além de 0,05 de cada lado da proporção amostral?

- **Exercícios de Revisão**

10.5.113) Para um determinado intervalo de confiança, marque a alternativa correta:

a) Aumentando n, aumenta-se o intervalo.
b) Quanto maior o intervalo, melhor a estimação da média populacional.
c) Aumentando n, diminui-se o intervalo.
d) O tamanho da amostra não está relacionado com o intervalo de confiança.
e) O tamanho da amostra é diretamente proporcional a margem de erro.

10.5.114) Um estudo desenvolvido por um instituto de pesquisa com uma amostra de 38 escolas mostrou que a média de notas dos alunos que estão formando é de 78 pontos com desvio padrão de 8 pontos.

a) Estabeleça um intervalo de confiança para a média de notas, com 99% de certeza, dos alunos destas escolas que estão formando.
b) A diretora de uma destas escolas levantou um intervalo de confiança para a média das notas como sendo de 76,5 a 79,5 pontos. Pergunta-se: qual o nível de confiança adotado pela diretora?
c) Determine o tamanho da amostra necessário para alcançar uma margem de erro de 2 pontos com nível de confiança de 95%.

10.5.115) Antes de uma eleição, um determinado partido está interessado em estimar a proporção de eleitores favoráveis a seu partido. Uma amostra piloto de tamanho 100 revelou que 60% dos eleitores eram favoráveis ao candidato em questão.

a) Determine o tamanho da amostra necessário para que a margem de erro seja de 0,01 com nível de confiança de 80%.
b) Se na amostra final, com tamanho igual ao obtido na letra a, observou-se que 55% dos eleitores eram favoráveis ao candidato em questão, construa um IC para a proporção dos eleitores simpáticos ao candidato (a 95% de certeza).

10.5.116) Está sendo planejado um levantamento sobre as atitudes dos consumidores em relação a um novo supermercado. Existe interesse particular na resposta para a questão de se precisar ou não ser providenciado um posto de gasolina. Acredita-se que cerca de 20% dos compradores farão uso do posto. A empresa quer margem de erro que não ultrapasse mais ou menos 3% das respostas à questão, e está sendo preparada para aceitar um nível de 95% de confiança para os resultados.

a) Que tamanho de amostra será necessário?
b) Um gerente argumenta que sempre foi usado um tamanho de amostra de 500. Que margem de erro essa amostra produz?

10.5.117) A prefeitura de uma cidade quer estimar a proporção p dos moradores favoráveis à mudança do horário comercial. Essa proporção deverá ser estimada com uma margem de erro de 5%, a um nível de confiança de 90%.

a) Que tamanho deverá ter a amostra se a proporção p esperada deve estar entre 20% e 50%?
b) Numa amostra de 400 moradores, 160 foram favoráveis à mudança. Qual será o intervalo de confiança para p, com nível de confiança de 95%?

10.5.118) De 5.000 produtos fabricados por uma companhia retira-se uma amostra de 400 produtos, e obtém-se a vida média de 800 horas e desvio padrão de 100 horas.

a) Qual intervalo de confiança de 99% para a vida média dos 5.000 produtos?
b) Com que nível de confiança dir-se-ia que a vida média é $800 \pm 9{,}8$?
c) Que tamanho deve ter a amostra para que seja de 95% a confiança na estimativa $800 \pm 7{,}84$?

10.5.119) Quarenta dentre uma amostra de 52 pessoas que responderam a um questionário indicaram dificuldade em conseguir que um novo revendedor corrigisse deficiências em seus carros novos.

a) Construa um intervalo de 98% de confiança para a proporção da população, da qual se pode esperar resposta semelhante a tal pergunta.
b) Com que nível de confiança dir-se-ia que a proporção é $0{,}77 \pm 0{,}07$?
c) Qual deverá ser o tamanho da amostra para atender ao intervalo 0,74 a 0,80 com nível de confiança de 95%?

10.5.120) Um levantamento com 369 pais que trabalham, revelou que 200 deles disseram passar pouco tempo com os filhos por causa dos compromissos de trabalho.

a) Qual é a estimativa pontual da proporção da população de pais que trabalham que sentem que passam pouco tempo com os filhos por causa dos compromissos de trabalho?
b) Qual é a margem de erro para um nível de confiança de 95%?
c) Qual é a estimativa por intervalo de confiança da proporção (nível de confiança de 99%) da população de pais que trabalham que sentem que passam pouco tempo com os filhos por causa dos compromissos de trabalho?
d) De acordo com a resposta da letra c, qual o erro dessa pesquisa?

10.5.121) Um analista de marketing preocupado com o número de defeitos apresentados por um produto lançado recentemente no mercado de manufaturas, resolveu levantar a proporção de defeituosos num lote de 2000 produtos que acaba de receber. Para isto, amostra aleatoriamente 50 produtos, testando cada um deles, obtendo 5 defeituosos. Pergunta-se:

a) Estabeleça um intervalo de confiança para a proporção de produtos defeituosos no lote, com 90% de certeza.
b) Se cada produto vendido dá um lucro de R$ 5,00 e cada defeituoso provoca um prejuízo de R$ 3,00, é possível esperar um lucro maior que R$ 7.000 na venda do lote de produtos, ao nível de confiança adotado?

10.5.122) 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.

a) Determine o intervalo de confiança para a média populacional com 90% de confiança.
b) Qual deverá ser o tamanho da amostra para atender ao intervalo 47,8 a 52,2, com nível de confiança de 99%?
c) Um gerente determinou um intervalo de confiança para a média populacional como sendo de 48,1 a 51,9. Qual o nível de confiança adotado?

10.5.123) A diretoria de uma escola tem como objetivo conhecer a nota média dos alunos que estão no último período. Sabe-se que o desvio padrão das notas dos alunos é de 6 pontos. Para atender ao objetivo, o responsável pelo estudo amostrou 18 alunos e encontrou nota média de 75 pontos. Pergunta-se:

a) Encontre um intervalo de confiança para a nota média de todos os alunos que estão no último período. Faça nível de confiança de 95%.
b) Através do resultado encontrado na letra a, a diretoria sugeriu que se trabalhasse com 50% da margem de erro obtida. Qual deverá ser o tamanho da amostra para atender a sugestão da diretoria? Use NC = 95%.