Estatística Aplicada

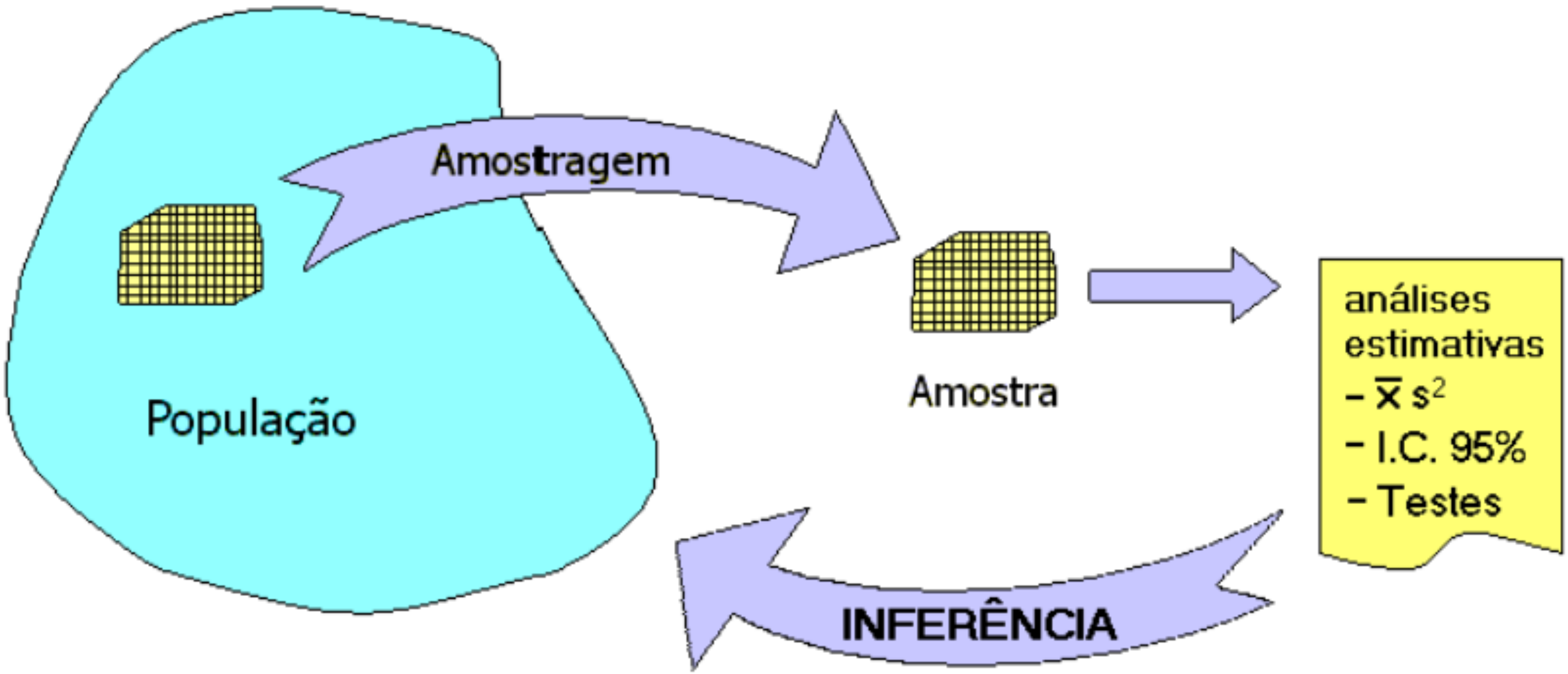
Inferência estatística com desvio padrão populacional conhecido

Amostragem

Capítulo 10

Itens 10.1 a 10.3

Relação entre amostragem e inferência



Inferência estatística

- **Inferência (ou estimação)** consiste em usar dados amostrais para estimar parâmetros populacionais desconhecidos
- Nem toda a amostra retirada de uma população reflete os parâmetros idênticos da população.
- O conjunto dos parâmetros obtidos através de amostras é chamado de **distribuição amostral**.
- Princípio básico:
  - A média da distribuição amostral é igual à média populacional

Exemplo

- Seja uma população em que uma variável aleatória é composta pelos valores 2, 3, 6 e 8. Sua média será:

$$M = \frac{2+3+6+8}{4} = 4{,}75$$

- As amostras possíveis, tomadas de 3 em 3 são: a=2, 3 e 6; b=2, 3 e 8; c=3, 6 e 8; d=2, 6 e 8. Suas médias são:

$$M_a = \frac{2+3+6}{3} = 3{,}67$$

$$M_b = \frac{2+3+8}{3} = 4{,}33$$

$$M_c = \frac{3+6+8}{3} = 5{,}67$$

$$M_d = \frac{2+6+8}{3} = 5{,}33$$

A média das médias amostrais será:

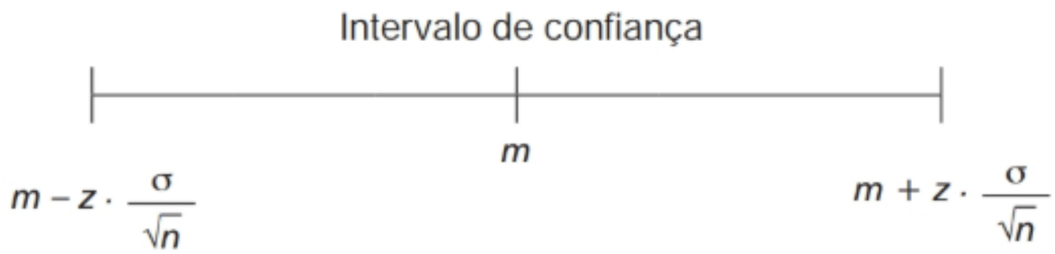
$$M_{amostral} = \frac{3{,}67 + 4{,}33 + 5{,}67 + 5{,}33}{4}$$

$$M_{amostral} = 4{,}75$$

Definição

- Assim, as médias de todas as amostras estarão dentro de um **intervalo de confiança**.
- A média populacional $\boldsymbol{M}$ estará dentro desse intervalo.
- Quando sabemos o desvio-padrão da população, o intervalo de confiança para uma **população infinita** será dado por:

$$I_C = m \pm z \cdot \frac{\sigma}{\sqrt{n}}$$



$m \rightarrow média\ da\ amostra$
$z \rightarrow coeficiente\ de\ confiança\ respectivo\ a\ um\ grau\ de\ confiança\ \alpha$
$\sigma \rightarrow desvio\ padrão\ da\ população$
$n \rightarrow número\ de\ elementos\ da\ amostra$

Significado do intervalo de confiança

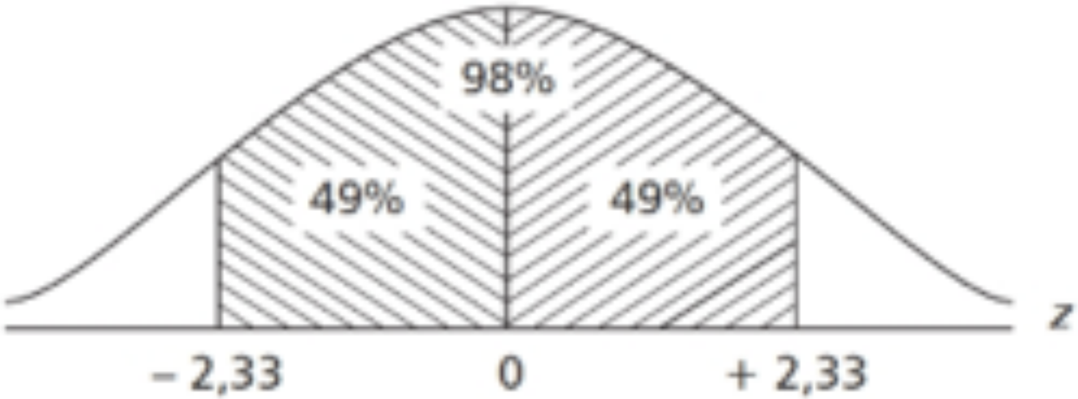
- O **grau de confiança** (α) é um valor que indica o **percentual** de médias amostrais que desejamos que caia dentro do intervalo.
- O **coeficente de confiança** é o valor de z associado ao grau de confiança, e é obtido na Tabela de Distribuição Normal.

- Por exemplo, quando dizemos que desejamos um intervalo com um grau de confiança de 95%, significa que desejamos os limites inferior e superior, dentro dos quais cairiam 95% das amostras.
- Aplicação: em um lote de peças produzidas, o intervalo de confiança de 95 % para uma certa dimensão é $32 \pm 3\ mm$. Isso significa que, se colhermos 100 peças, esperamos que 95 apresentem valores entre 29 e 35 mm.

Como determinar o valor de z

- Exemplo: desejamos um intervalo de confiança de 98%.
  - Os $98\ \% \Rightarrow \alpha = 0{,}9800$, (grau de confiança) estão distribuídos bilateralmente, ou seja, $\frac{\alpha}{2} = 0{,}4900$ para cada lado.
  - Assim, procuramos na tabela o valor 0,4900 ou o mais próximo **acima**.
  - Encontramos que os limites são $z = \pm 2{,}33$ (coeficientes de confiança)

z	0,00	0,01	0,02	0,03	0,04
0,0				↑	
				↑	
				↑	
2,2				↑	
2,2				↑	
2,3	←	←	←	0,4901	
2,4					



Exercício

Encontre os valores críticos do coeficiente ***z*** para construir um intervalo de $\propto$% de confiança para estimar a média populacional, quando:

1) $\propto$ **= 90%**	**2)** $\propto$ **= 99%**	**3)** $\propto$ **= 95%**	**4)** $\propto$ **= 80%**

Respostas:

1) $z = \pm 1{,}65$; **2)** $z = \pm 2{,}58$; **3)** $z = \pm 1{,}96$; **4)** $z = \pm 1{,}29$.

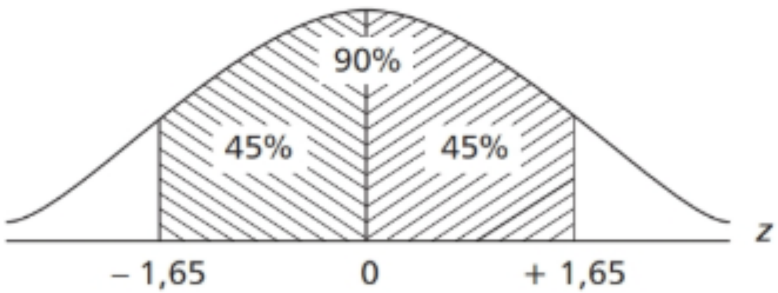
Aplicação

- Uma **amostra** aleatória de 60 unidades de certo produto químico apresentou preço médio, por kg, de R$ 45,21. Determinar um intervalo de 90% de confiança para estimar a média populacional desse produto, sabendo que o desvio padrão **populacional** é R$ 6,32.

1° Cálculo dos limites de z

$$90\ \% \Rightarrow \ \alpha = 0{,}900 \Rightarrow \frac{\alpha}{2} = 0{,}4500$$

$$Tabela \Rightarrow z_{\frac{\alpha}{2}=0{,}4500} = 1{,}65$$

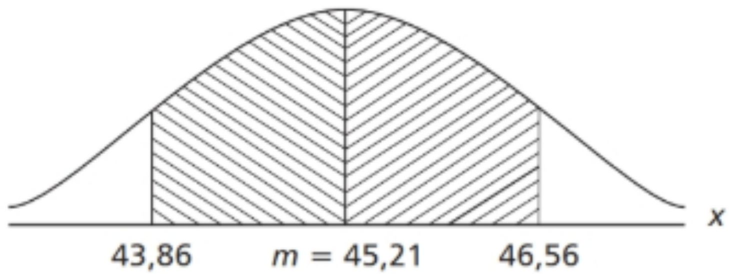


2° Cálculo do I_C

$$I_C = m \pm z \cdot \frac{\sigma}{\sqrt{n}}$$

$$I_C = 45{,}21 \pm 1{,}65 \cdot \frac{6{,}32}{\sqrt{60}}$$

$$I_C = 45{,}21 \pm 1{,}34$$



Resposta: 90 % das médias amostrais deverão estar entre R$ 43,86 e R$ 46,56

Aplicação

- No exercício anterior, calcule os intervalos de confiança de 95% e de 99%. Compare os resultados e analise.

$$95\ \% \Rightarrow \ \alpha = 0{,}950 \Rightarrow \frac{\alpha}{2} = 0{,}4750$$

$$Tabela \Rightarrow z_{\frac{\alpha}{2}=0{,}4750} = 1{,}96$$

$$I_C = 45{,}21 \pm 1{,}96 \cdot \frac{6{,}32}{\sqrt{60}}$$

$$I_C = 45{,}21 \pm 1{,}60$$

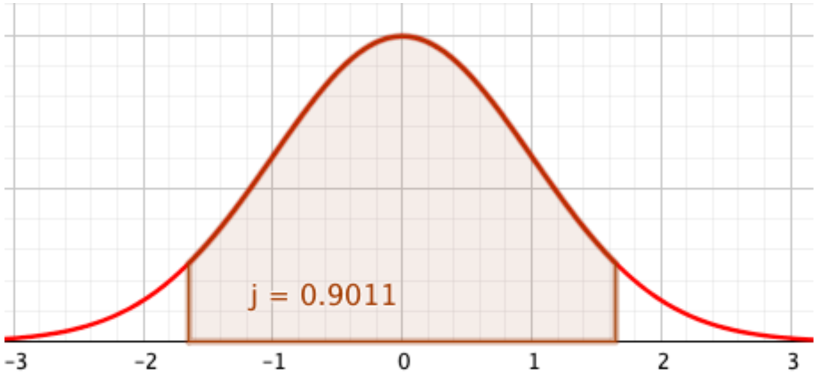
$$99\ \% \Rightarrow \ \alpha = 0{,}990 \Rightarrow \frac{\alpha}{2} = 0{,}4950$$

$$Tabela \Rightarrow z_{\frac{\alpha}{2}=0{,}4950} = 2{,}58$$

$$I_C = 45{,}21 \pm 2{,}58 \cdot \frac{6{,}32}{\sqrt{60}}$$

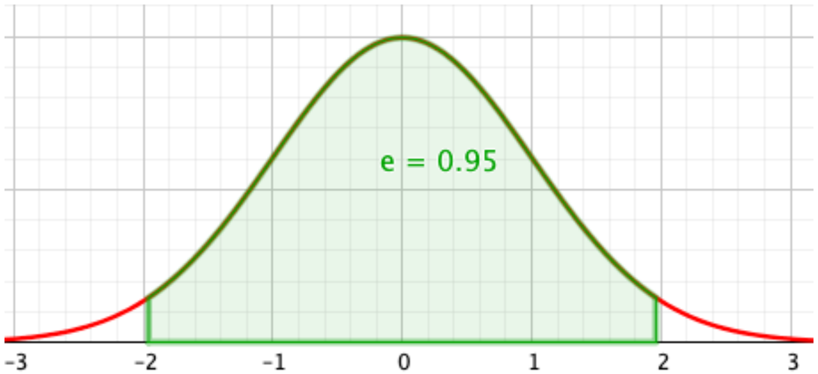
$$I_C = 45{,}21 \pm 2{,}11$$

Aplicação



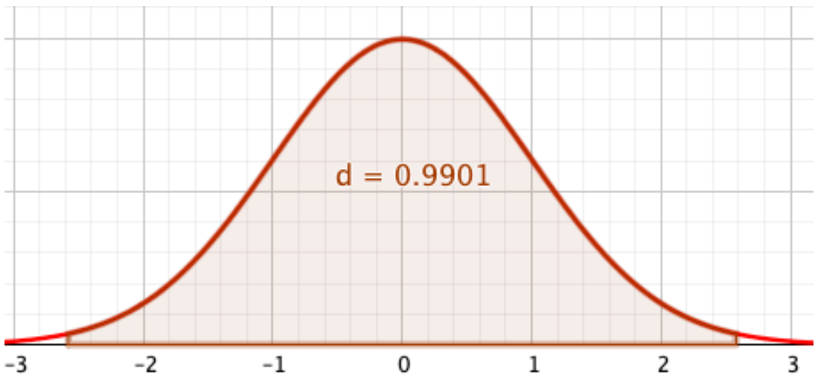
Intervalo de confiança de 90%:

R$ 43,86 e R$ 46,56



Intervalo de confiança de 95%:

R$ 43,61 e R$ 46,81



Intervalo de confiança de 99%:

R$ 43,10 e R$ 47,32

Exercícios

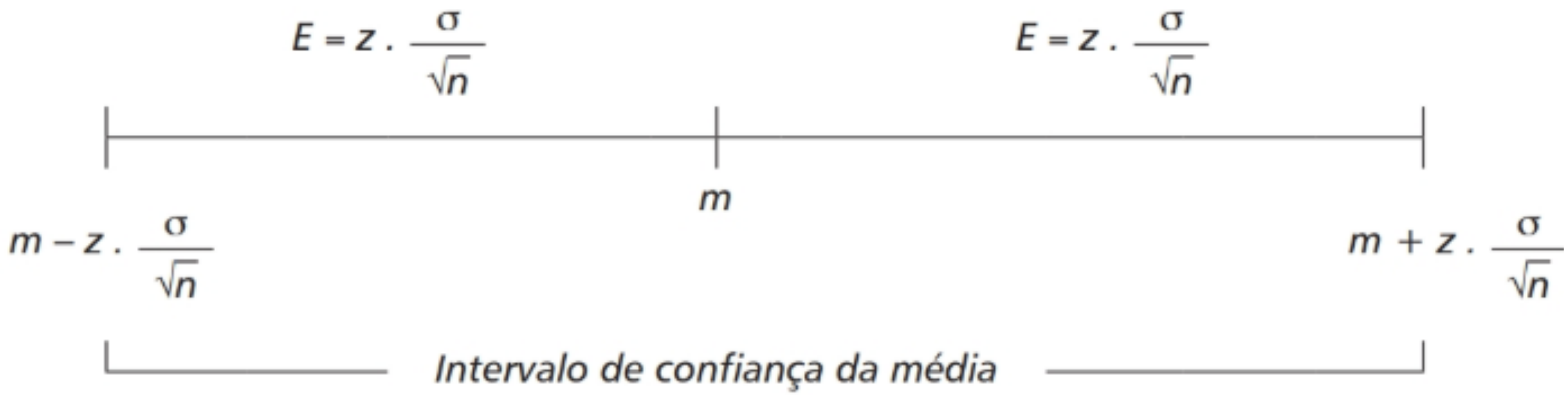
- Pg. 163
- 1 a 6

Considerações adicionais

- Na fórmula

$$I_C = m \pm z \cdot \frac{\sigma}{\sqrt{n}}$$

Chamamos o valor $E = z \cdot \frac{\sigma}{\sqrt{n}}$ de **erro de estimativa da média** ou **margem de erro**



E chamamos o valor $e = \frac{\sigma}{\sqrt{n}}$ é o **desvio-padrão da média amostral** ou **erro-padrão da média**

Exemplo (pg. 167)

- Um engenheiro químico pretende estimar o tempo médio de secagem de uma nova marca de tinta látex. Um amostra de 25 pinturas feitas em condições semelhantes revelou um tempo médio de secagem de 4,7 horas. O desvio padrão populacional é estimado em 1,1 hora. Determinar

a) O tempo médio pontual de secagem

b) O intervalo de confiança de 98% para o verdadeiro tempo de secagem dessa tinta

c) O erro máximo da estimativa da média populacional (ou margem de erro) para uma confiança de 90%

d) O erro-padrão da média do tempo de secagem

Exercícios

- Exercício pg. 168

Correção para populações finitas

- Quando a população é finita e a amostra é superior a 5% da população, todas as fórmulas serão multiplicadas pelo fator de correção $\sqrt{\frac{N-n}{N-1}}$, onde N é a população e n a amostra.

- Teremos então:

$$I_C = m \pm z \cdot \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

$$E = z \cdot \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

$$e = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

Exemplo

1. De um lote de rolamentos, um engenheiro seleciona 150 peças, cujo diâmetro médio é de 0,725 polegadas. O desvio padrão do lote é tido como 0,035 polegadas. Determinar os limites de um intervalo de confiança de 90% para o verdadeiro diâmetro do lote, considerando que ele tenha:
   a) N=2000 peças
   b) N=5000 peças

2. Determine o erro máximo das estimativa em cada caso

3. Determine o erro padrão em cada caso

4. Compare e analise os resultados

Exercícios

- Pg. 172, ex. 1 a 5

Tamanho da amostra

- A partir das fórmulas da margem de erro, podemos estimar o tamanho mínimo da amostra.
- Para populações infinitas, a partir de $E = z \cdot \frac{\sigma}{\sqrt{n}}$, chegamos a:

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

- Para populações finitas (n>5% de N), a partir de $E = z \cdot \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$, chegamos a:

$$n = \frac{N \cdot z^2 \cdot \sigma^2}{(N-1) \cdot E^2 + z^2 \cdot \sigma^2}$$

Exemplo

- O engenheiro de produção responsável pela fabricação de um componente precisa determinar o **tempo médio** para a usinagem de 10 orifícios. Considerando um desvio padrão de 40 s, um erro admissível de 14 s e uma confiança de 95% calcule qual a amostra mínima (ou seja, em quantas placas ele deve medir o tempo de perfuração).

$$95\ \% \Rightarrow\ \alpha = 0{,}950 \Rightarrow \frac{\alpha}{2} = 0{,}475$$

$$Tabela \Rightarrow z_{\frac{\alpha}{2}=0{,}4750} = 1{,}96$$

$$n = \left(\frac{z \cdot \sigma}{E}\right)^2 \Rightarrow n = \left(\frac{1{,}96 \cdot 40}{14}\right)^2 \Rightarrow n = 32\ placas$$

Exemplo

- Repita o cálculo anterior, com uma confiança de 90% e de 99%, e analise os resultados

$$90\ \% \Rightarrow\ \alpha = 0{,}900 \Rightarrow \frac{\alpha}{2} = 0{,}4500$$

$$Tabela \Rightarrow z_{\frac{\alpha}{2}=0{,}4500} = 1{,}65$$

$$n = \left(\frac{z \cdot \sigma}{E}\right)^2 \Rightarrow n = \left(\frac{1{,}65 \cdot 40}{14}\right)^2$$

$$n = 23\ placas$$

$$99\ \% \Rightarrow\ \alpha = 0{,}990 \Rightarrow \frac{\alpha}{2} = 0{,}4950$$

$$Tabela \Rightarrow z_{\frac{\alpha}{2}=0{,}4950} = 2{,}58$$

$$n = \left(\frac{z \cdot \sigma}{E}\right)^2 \Rightarrow n = \left(\frac{2{,}58 \cdot 40}{14}\right)^2$$

$$n = 54\ placas$$

Conclusão: o tamanho da amostra é maior, caso queiramos uma maior confiança.

Resumo: Estimação com DP populacional conhecido

	População Infinita	**População Finita**
Intervalo de confiança	$I_C = m \pm z \cdot \frac{\sigma}{\sqrt{n}}$	$I_C = m \pm z \cdot \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$
Margem de erro	$E = z \cdot \frac{\sigma}{\sqrt{n}}$	$E = z \cdot \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$
Desvio padrão da média amostral	$e = \frac{\sigma}{\sqrt{n}}$	$e = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$
Amostra mínima	$n = \left(\frac{z \cdot \sigma}{E}\right)^2$	$n = \frac{N \cdot z^2 \cdot \sigma^2}{(N-1) \cdot E^2 + z^2 \cdot \sigma^2}$

Exercícios

- Intervalo de confiança, população infinita:
  - Pg. 163, ex. 1 a 6
- Erro médio e erro-padrão:
  - Pg. 168
- População finita:
  - Pg. 172, ex. 1 a 5
- Amostragem:
  - Pg. 175