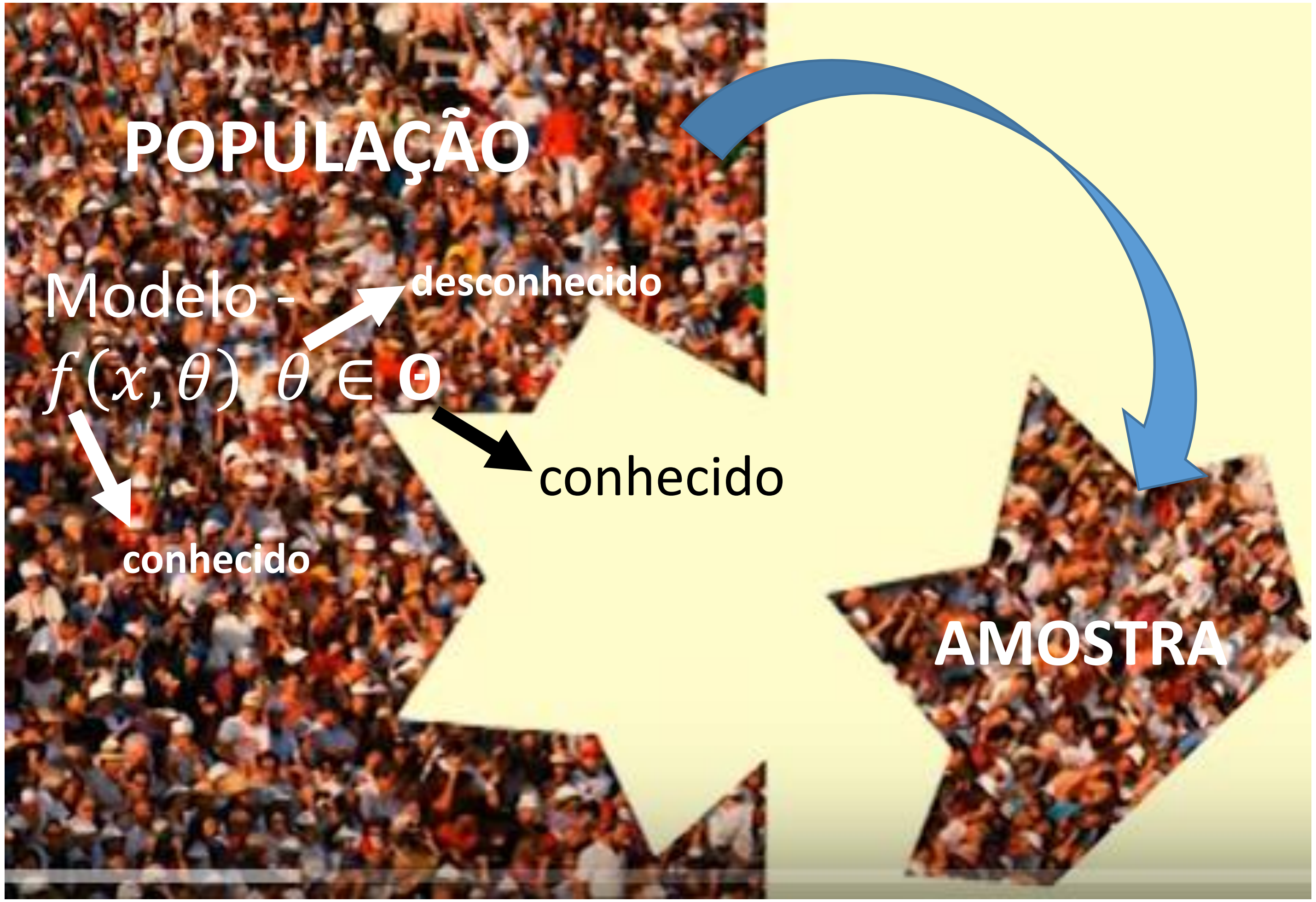
ESTIMAÇÃO



Estimar é utilizar a informação da amostra para "adivinhar" o valor de θ.

Dois aspectos a ter em, conta: a **precisão** e a **confiança.**

Ideia Importante $\rightarrow$ Fixada a dimensão da amostra, quanto mais precisa a resposta, menor a confiança que nela se deposita.

A estimação paramétrica desenvolve-se privilegiando: a precisão (**estimação por pontos**)
a confiança (**estimação por intervalos**).

ESTIMAÇÃO

Ter presente que o parâmetro de interesse θ pode ser:

- **Multidimensional**

Exemplo $\rightarrow$ Suponha que a valorização de um activo financeiro tem distribuição normal com média μ e variância σ^2.
Observada uma amostra casual pretende-se estimar μ (rendibilidade esperada) e σ^2(risco).

- **Função do(s) parâmetro(s) da distribuição**

. Exemplo $\rightarrow$ Suponha-se que o número de sinistros originados anualmente por uma apólice de seguro automóvel tem distribuição de Poisson de parâmetro $\boldsymbol{\lambda}$ (desconhecido).
Pode não nos interessar conhecer a média do fenómeno ($\boldsymbol{\lambda}$) mas sim a probabilidade de não se verificar nenhum sinistro, $P(X=0|\boldsymbol{\lambda})=e^{-\boldsymbol{\lambda}}$, uma função de $\boldsymbol{\lambda}$.

Nesse caso estima-se $\tau(\boldsymbol{\lambda})=e^{-\boldsymbol{\lambda}}$

ESTIMAÇÃO POR PONTOS

- **Conceitos Fundamentais:**

Estimador: é uma variável aleatória, função da amostra casual e representa-se por $T(X_1, X_2, \cdots, X_n) \Rightarrow$ é uma estatística.

Por exemplo: $T(X_1, X_2, \cdots, X_n) = \bar{X}$ média da amostra

Estimativa: é um número assumido pelo estimador para a particular amostra que se observou. Representa-se por $t(x_1, x_2, \cdots, x_n) \in \mathbb{R}$

Por exemplo: $t(x_1, x_2, \cdots, x_n) = \bar{x}$

- **Dois problemas em aberto**:
  - Como encontrar estimadores para determinado parâmetro?
  - Encontrado um ou mais estimadores, como avaliar a sua qualidade?

ESTIMAÇÃO POR PONTOS

Método dos momentos

$\Downarrow$

IDEIA

Utilizar estatísticas $\left(\bar{X}, \frac{\sum_{i=1}^{n} X_i^2}{n}\right)$ para estimar os correspondentes momentos da população $[E(X), E(X^2), Var(X)]$

- Existem outros métodos de estimação por pontos: Método da máxima verosimilhança. (Não vai ser estudado este semestre)

Método dos momentos

Momentos da população tem de existir

• Momentos da População	=	• Estatísticas
$E(X)$	=	$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$
$E(X^2)$	=	$\frac{\sum_{i=1}^{n} X_i^2}{n}$

Estatísticas existem sempre

Igualam-se os momentos da população às estatísticas correspondentes

ESTIMAÇÃO POR PONTOS

Método dos momentos

• Exemplo 1 Considere-se uma população de Bernoulli da qual se extraiu uma amostra casual de dimensão n com o objectivo de estimar θ.

Como se sabe:

- Média da população - $E(X) = \theta$
- Estatística para estimar média da população θ é $\bar{X}$.

$\rightarrow$ equação - $\bar{X} = \theta$

- solução: $\tilde{\theta}(X_1, X_2, \cdots, X_n) = \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$ Estimador pelo método dos momentos

$\tilde{\theta}(x_1, x_2, \cdots, x_n) = \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ Estimativa

Cuidado com a notação !

ESTIMAÇÃO POR PONTOS

Método dos momentos

- Exemplo 1 Considere-se uma população normal da qual se extraiu uma amostra casual de dimensão n com o objectivo de estimar μ e σ^2

Como se sabe:

- Momentos da população: $E(X) = \mu$; $E(X^2) = \sigma^2 + \mu^2$
- Estatísticas: $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$; $\frac{\sum_{i=1}^{n} X_i^2}{n}$
- sistema - $\begin{cases} \bar{X} = \mu \\ \frac{\sum_{i=1}^{n} X_i^2}{n} = \sigma^2 + \mu^2 \end{cases}$

Cuidado com a notação !

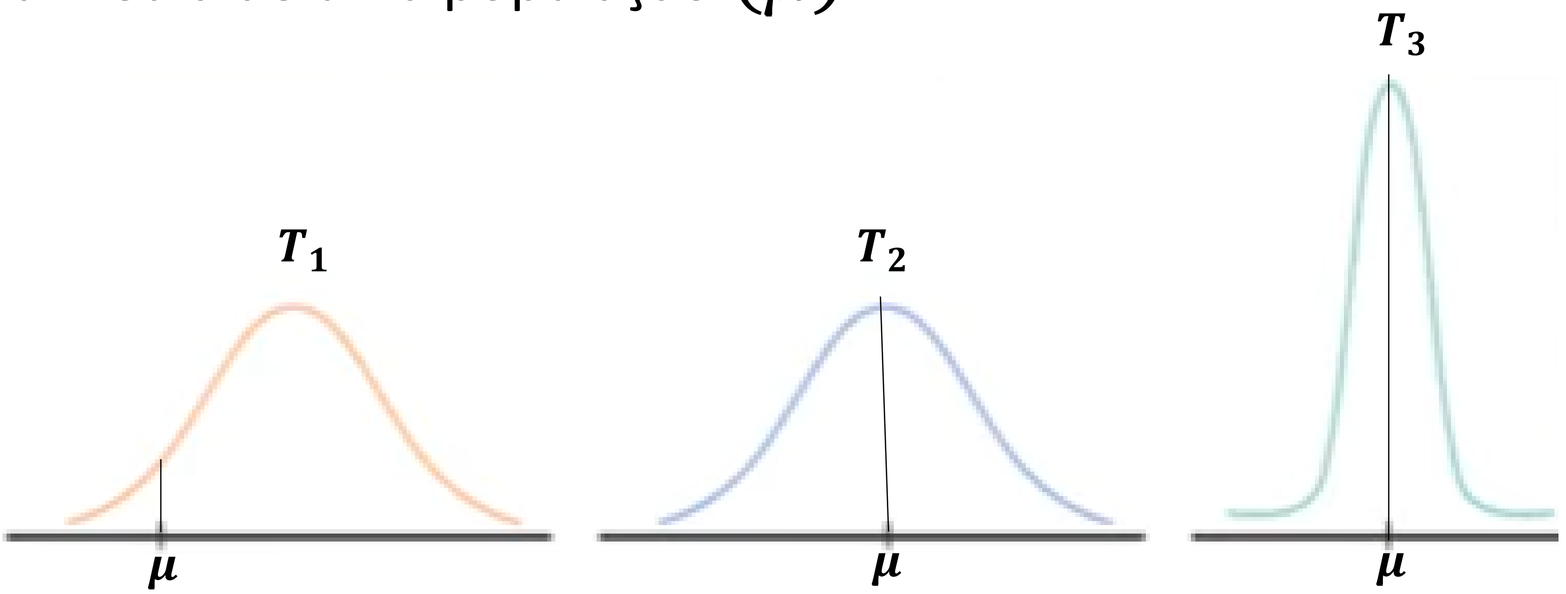
- solução: $\tilde{\mu} = \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$; $\widetilde{\sigma^2} = \frac{\sum_{i=1}^{n} X_i^2}{n} - \bar{X}^2 = S^2$ Estimadores

$\tilde{\mu} = \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$; $\widetilde{\sigma^2} = \frac{\sum_{i=1}^{n} x_i^2}{n} - \bar{x}^2 = s^2$ Estimativas

ESTIMAÇÃO POR PONTOS

Encontrado um ou mais estimadores, como avaliar a sua qualidade?

Distribuições por amostragem de 3 diferentes estatísticas para estimar a verdadeira média de uma população (μ)



Qual dos estimadores vos parece melhor entre T_1 e T_2? E entre T_2 e T_3

ESTIMAÇÃO POR PONTOS

Propriedades dos estimadores

- Não enviesamento

Definição: Um estimador $T(X_1, X_2, \cdots, X_n)$ para θ diz-se centrado ou não enviesado quando $E(T) = \theta \quad \forall\, \boldsymbol{\theta} \in \boldsymbol{\Theta}$

- Observação:

  O conceito de estimador centrado só se aplica se existir $E(T)$

Enviesamento: Se $E(T) \neq \theta$, então o estimador diz-se enviesado e a diferença $\mathrm{Env}(T) = E(T) - \theta$ mede o enviesamento.

ESTIMAÇÃO POR PONTOS

Propriedades dos estimadores

- Exemplo: Seja uma amostra casual $(X_1, X_2, \cdots, X_n)$ de uma população de Bernoulli $[X_i \sim B(1,\theta)]$. Será $\bar{X}$ um estimador centrado para θ?

$$E(\bar{X}) = E\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) = \frac{1}{n} E\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\left(\sum_{i=1}^{n} E(X_i)\right) = \frac{1}{n} * n\theta = \theta$$

Exemplo: Seja uma amostra casual $(X_1, X_2, \cdots, X_n)$ de uma população $N(\mu, \sigma^2)$. Será S^2 um estimador centrado para σ^2?

$$E(S^2) = \frac{n-1}{n}\sigma^2 \Rightarrow E(S^2) \neq \sigma^2$$

S^2 é um estimador enviesado para σ^2

$$Env(S^2) = E(S^2) - \sigma^2 = \frac{n-1}{n}\sigma^2 - \sigma^2 = -\sigma^2/n$$

ESTIMAÇÃO POR PONTOS

Propriedades dos estimadores

- Eficiência

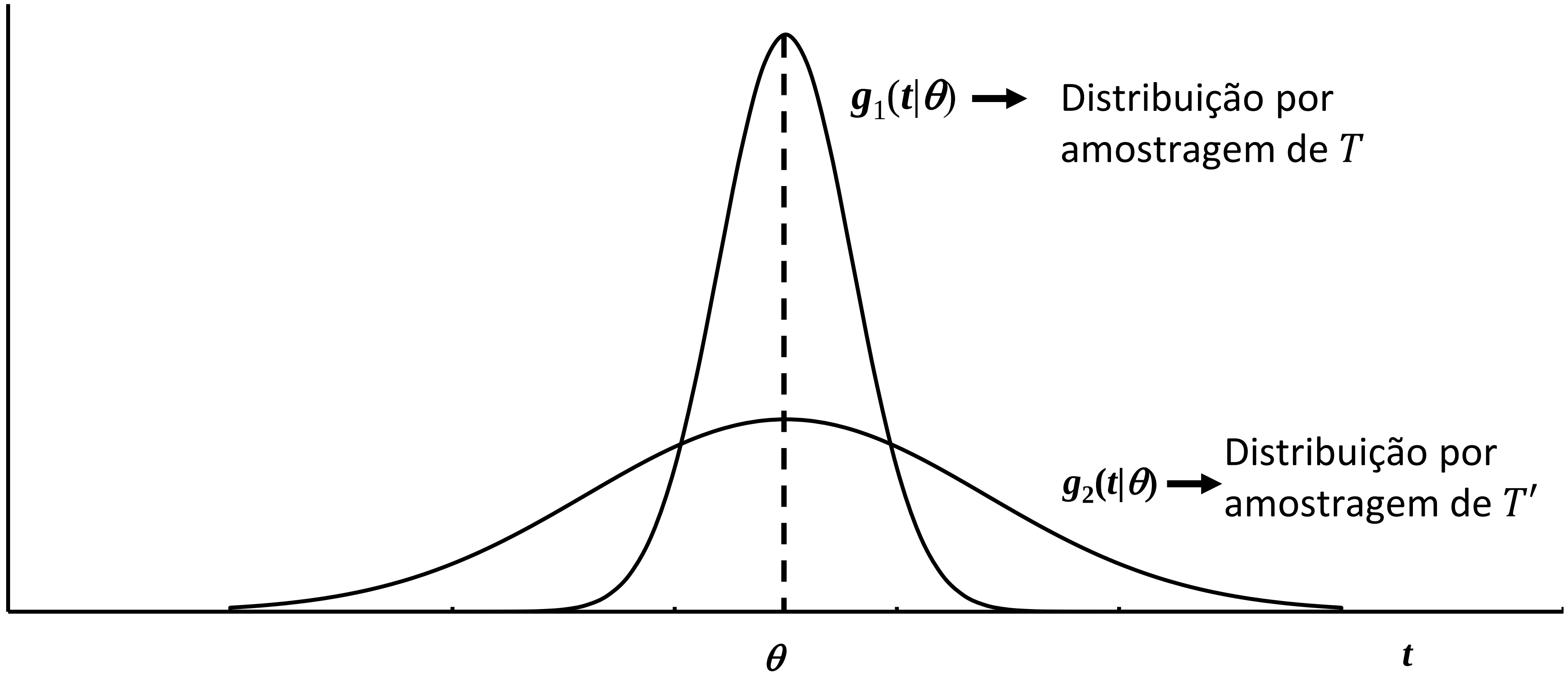
    - Compara dois estimadores centrados analisando a respectiva dispersão.

Definição: Sejam T e T' dois estimadores centrados para θ. O estimador T é relativamente mais eficiente que T' quando $Var(T) \leq Var(T')\ \forall\ \boldsymbol{\theta} \in \boldsymbol{\Theta}$

- Observação: A eficiência exige a existência da variância dos estimadores.

ESTIMAÇÃO POR PONTOS

Propriedades dos estimadores



O estimador T é relativamente mais eficiente que T'

ESTIMAÇÃO POR PONTOS

Propriedades dos estimadores

- Consistência

  - Exige-se como condição mínima para um estimador ser um "bom" estimador que a precisão do estimador aumente quando aumenta a dimensão da amostra.

T 7.2: As condições $\lim_{n \to \infty} E(T) = \theta \;\; e \; \lim_{n \to \infty} Var(T) = 0$ são suficientes para que T seja estimador (simplesmente) consistente para θ

- Observações:
  - A consistência não é uma propriedade muito selectiva.
  - Um estimador que não seja consistente não deve ser utilizado.

ESTIMAÇÃO POR PONTOS

Propriedades dos estimadores

- Exemplo: Seja uma amostra casual $(X_1, X_2, \cdots, X_n)$ de uma população $Po(\boldsymbol{\lambda})$. O estimador centrado para$\mathrm{T}(X_1, X_2, \cdots, X_n) = \bar{X}$ será um estimador consistente para $\boldsymbol{\lambda}$?

$$E(\bar{X}) = E\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) = \frac{1}{n} E\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\left(\sum_{i=1}^{n} E(X_i)\right) = \frac{1}{n} * n\boldsymbol{\lambda} = \boldsymbol{\lambda}$$

$$\lim_{n\to\infty} E(\bar{X}) = \lim_{n\to\infty} \boldsymbol{\lambda} = \boldsymbol{\lambda}$$

$$Var(\bar{X}) = Var\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) = \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2} * n\boldsymbol{\lambda} = \frac{\boldsymbol{\lambda}}{n}$$

$$\lim_{n\to\infty} Var(\bar{X}) = \lim_{n\to\infty} \frac{\boldsymbol{\lambda}}{n} = 0$$

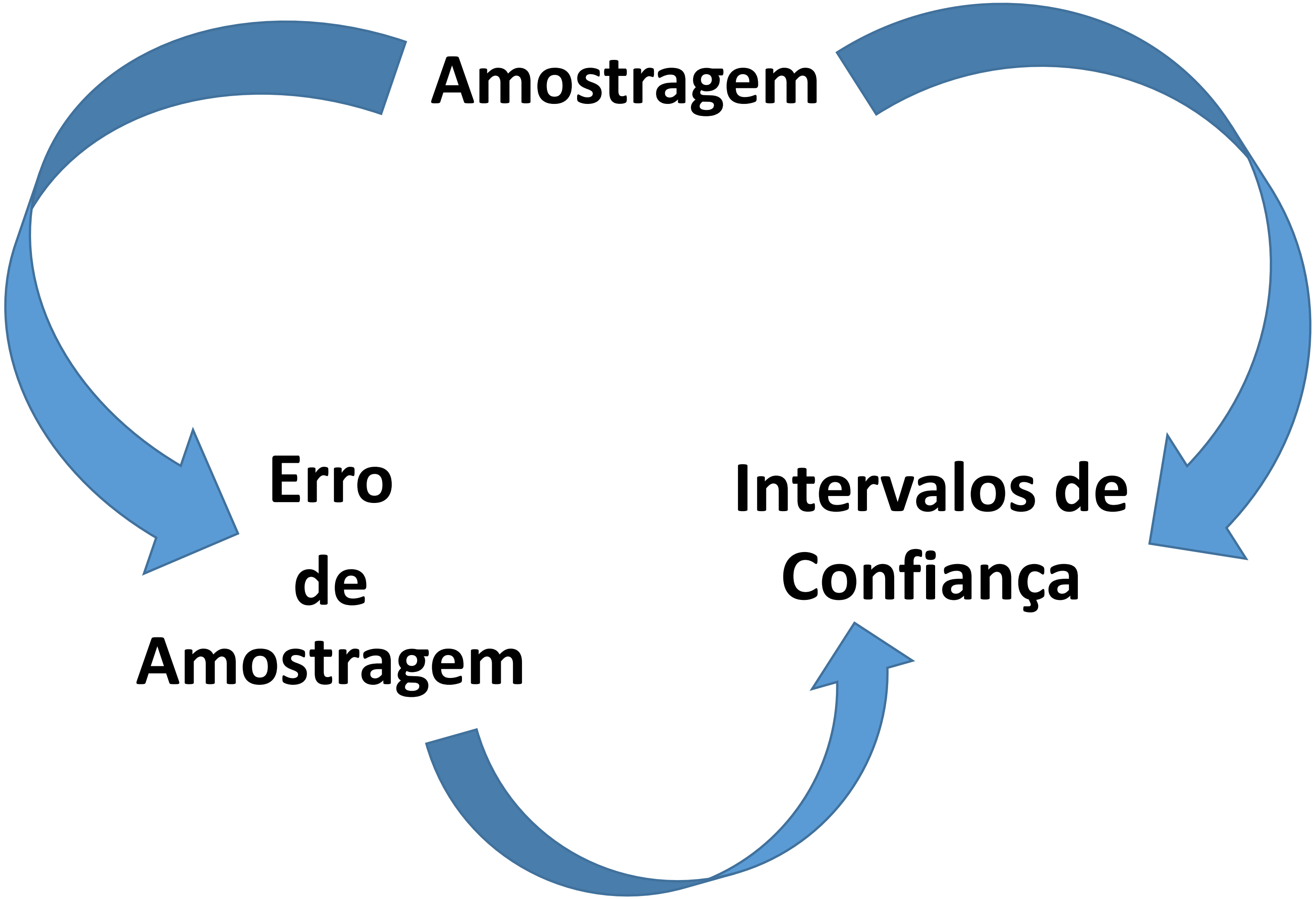
ESTIMAÇÃO POR PONTOS

Propriedades dos estimadores obtidos pelo Método dos Momentos

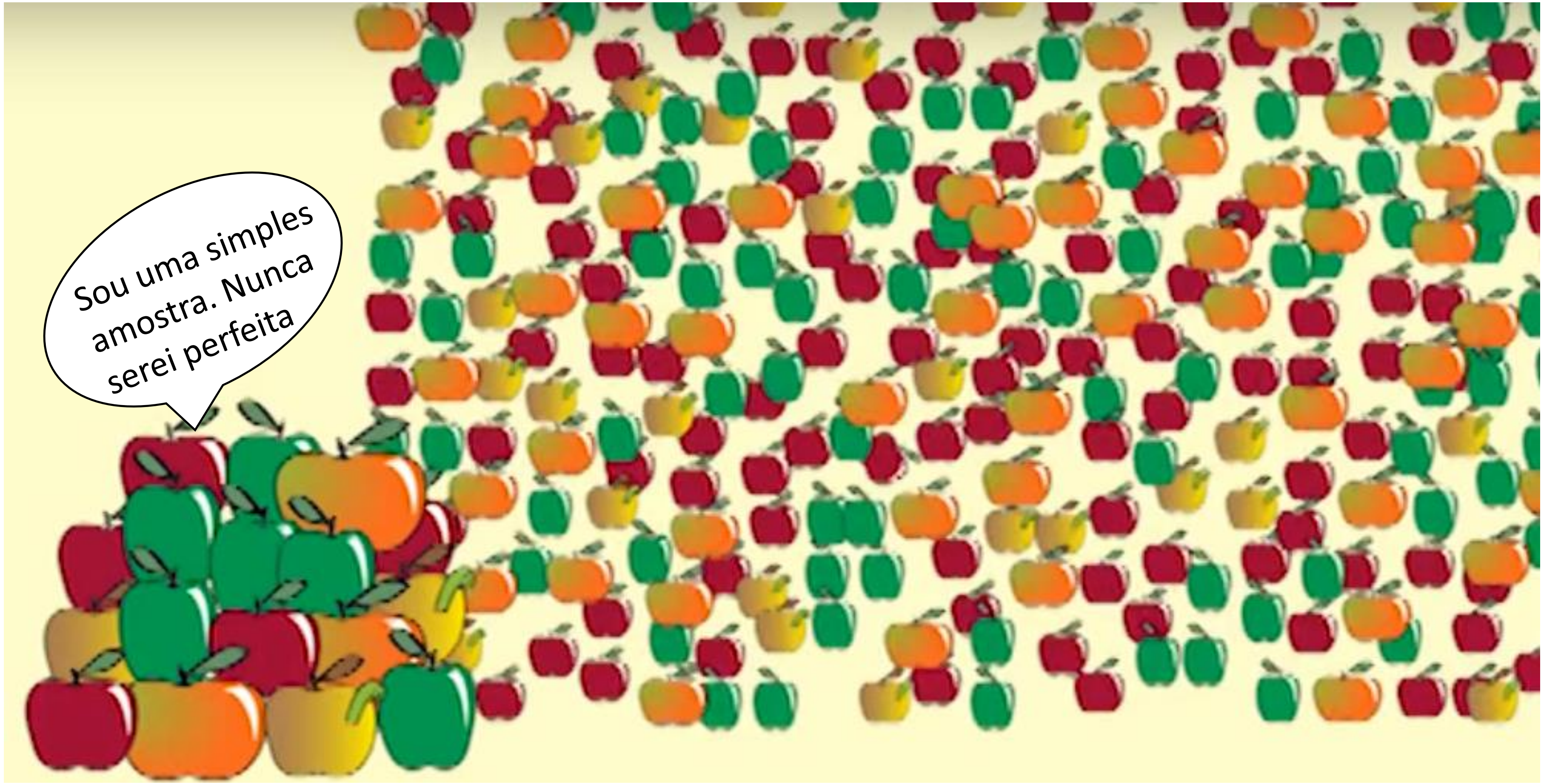
Em condições bastante gerais, são consistentes e possuem distribuição aproximadamente normal quando a dimensão da amostra é muito grande (distribuição assintótica).

ESTIMAÇÃO POR INTERVALOS

Estimação por intervalos privilegia a confiança



ESTIMAÇÃO POR INTERVALOS



ESTIMAÇÃO POR INTERVALOS

VARIABILIDADE DA AMOSTRA

Para diferentes amostras da mesma população as estimativas assumem diferentes valores

ERRO DE AMOSTRAGEM (ε)

Variação devido à amostragem

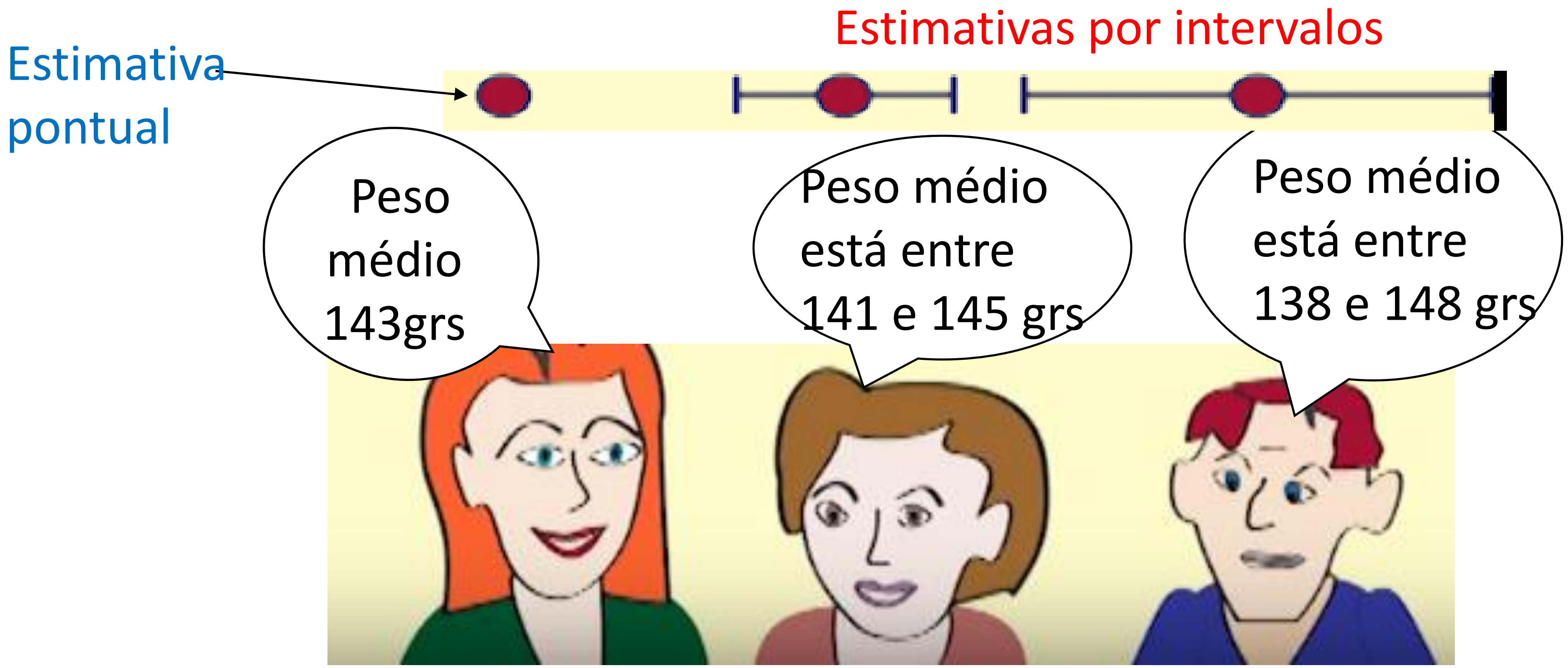
$\bar{x} = 149grs$

$\bar{x} = 130grs$

$\bar{x} = 153grs$

ESTIMAÇÃO POR INTERVALOS

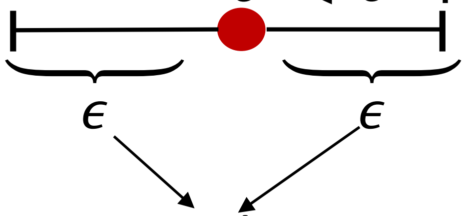
Quando estimamos um parâmetro de uma população é prática comum exprimir a estimativa sob a forma de um intervalo. O intervalo de confiança indica a **precisão** da estimativa associada a um **grau de confiança** fixado.



ESTIMAÇÃO POR INTERVALOS

O intervalo de confiança $(\hat{\theta} - \varepsilon, \hat{\theta} + \varepsilon)$ indica a **precisão** da estimativa associada a um **grau de confiança** fixado.

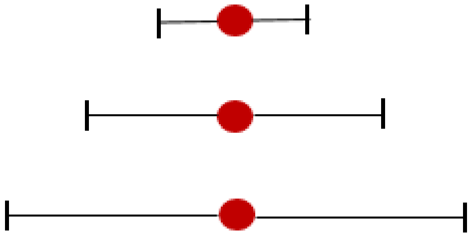
$$\hat{\theta} - \varepsilon \;<\; \hat{\theta} < \hat{\theta} + \varepsilon$$



Amplitude do Intervalo confiança:

$$\Delta IC_{\theta}^{1-\alpha} = (\hat{\theta} + \varepsilon) - (\hat{\theta} - \varepsilon)$$

Erro de amostragem\margem de erro – medida de precisão



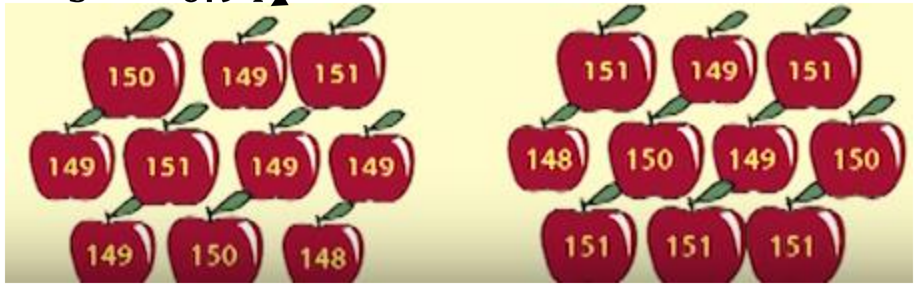
Que factores afectam a amplitude do Intervalo de confiança?

ESTIMAÇÃO POR INTERVALOS

1. Variação na população

População com
pequena variação

$s^2 = 0.94$ $\qquad$ $s^2 = 1.21$



Intervalos de confiança com
pequena amplitude

População com
grande variação

$s^2 = 152$ $\qquad$ $s^2 = 283$

Intervalos de confiança com
grande amplitude

ESTIMAÇÃO POR INTERVALOS

2. Variação na dimensão da amostra

$\bar{x} = 160$
$s^2 = 336.6$

$\bar{x} = 165.6$
$s^2 = 380$

Amostras pequenas contém pouca informação e variam mais de uma para outra

Intervalos com maior amplitude

$\bar{x} = 160.6$
$s^2 = 15.8$

$\bar{x} = 159.4$
$s^2 = 20.3$

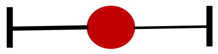
Amostras grandes contém mais informação e variam menos de uma para outra

Intervalos com menor amplitude

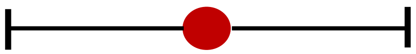
ESTIMAÇÃO POR INTERVALOS

3. Variação no grau de confiança

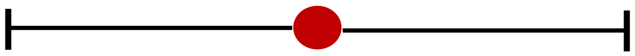
Intervalo de confiança a 90%



Intervalo de confiança a 95%



Intervalo de confiança a 99%



Quanto maior o grau de confiança maior a amplitude do intervalo

ESTIMAÇÃO POR INTERVALOS

A média da população varia entre …

A proporção na população varia entre …

A variância da população varia entre …

ESTIMAÇÃO POR INTERVALOS

Cálculo do intervalo de confiança a $(1-\alpha) * 100\%$ para o parâmetro θ

Seja $(X_1, X_2, \cdots, X_n)$ amostra casual de uma população $f(x, \theta), \theta \in \mathbf{\Theta}$

Intervalo aleatório para θ

Se $T_1 = T_1(X_1, X_2, \cdots, X_n), T_2 = T_2(X_1, X_2, \cdots, X_n)$, $T_1 < T_2$, com

$P(T_1 < \theta < T_2) = 1 - \alpha \quad \forall\theta \in \mathbf{\Theta},\ 0 < \alpha < 1$ (α não depende de θ),

Então (T_1, T_2) é um intervalo aleatório para θ de probabilidade $1-\alpha$.

Intervalo de confiança para θ

$(x_1, x_2, \cdots, x_n)$ amostra particular realização de $(X_1, X_2, \cdots, X_n)$

$t_1 = T_1(x_1, x_2, \cdots, x_n), t_2 = T_2(x_1, x_2, \cdots, x_n)$, valores assumidos por T_1 e T_2

A qualquer intervalo (t_1, t_2) concretização do intervalo aleatório (T_1, T_2) chama-se intervalo de confiança a $(1-\alpha)$*100% para θ

ESTIMAÇÃO POR INTERVALOS

Cálculo do intervalo de confiança a $(1-\alpha)*100\%$ para o parâmetro θ

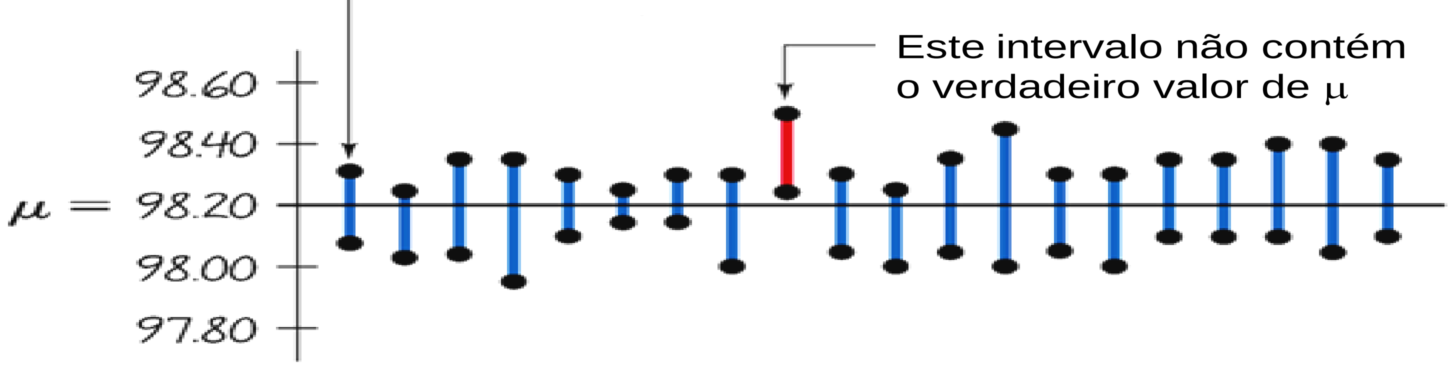
- Comentários:
  - As definições foram apresentadas para θ e não para $\tau(\theta)$ para simplificar a notação. A generalização é imediata.
  - **Só se atribui probabilidade ao intervalo aleatório**.
  - O conceito, tal como o vimos, é válido para $\mathbb{R}$. No caso de $\tau(\theta) \in \mathbb{R}^k, \;\; k > 1$ é necessário estendê-lo para **regiões de confiança**.

ESTIMAÇÃO POR INTERVALOS

Interpretação frequencista do intervalo de confiança

Seleccionadas várias amostras de idêntica dimensão, da população em estudo, e calculados os correspondentes intervalos de confiança, cerca de $100(1 - \alpha)\%$ dos intervalos calculados contêm o verdadeiro valor do parâmetro θ.

20 amostras $\Rightarrow$ 20 I. C. a 95% para μ



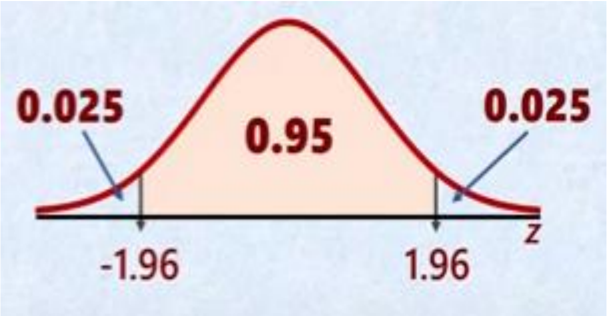
ESTIMAÇÃO POR INTERVALOS

Intervalo de confiança para a média (variância conhecida) – Populações normais

Grau\nível de confiança$= 1-\alpha = 0.95 \Rightarrow \alpha = 0.05$

Variável Fulcral $T = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$

$z_{\alpha/2}: P\left(-z_{\alpha/2} < T < z_{\alpha/2}\right) = 1-\alpha$



$-z_{\alpha/2}$ $\quad z_{\alpha/2}$

$$-z_{\alpha/2} < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < z_{\alpha/2}$$

$$\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Intervalo aleatório para $\mu = \left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$

Margem de erro (ϵ)

Intervalo confiança para μ a $(1-\alpha) * 100\% = \left(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$

ESTIMAÇÃO POR INTERVALOS

Intervalo de confiança para a média (variância conhecida)

Exemplo: População $X \sim N(\mu, \sigma^2 = 605)$

Seja amostra casual (807.7, 790.7, 818.8, 853.4, 858.6)$\Rightarrow n = 5,\ \bar{x} = 825{,}84$

Grau\nível de confiança$= 95\% \Leftrightarrow 1 - \alpha = 0.95 \Leftrightarrow \alpha = 0.05$

Variável fulcral $- T = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0{,}1)$

$\Rightarrow z_{0.025} = \text{invnorm}(0.025, 0, 1) = 1.96$

Margem de erro - $\varepsilon = 1.96 \dfrac{24{,}597}{\sqrt{5}} = 21{,}56 \ \Rightarrow\ \bar{X} - 21{,}56 < \mu < \bar{X} + 21{,}56$

Intervalo aleatório para $\mu = (\bar{X} - 21{,}56 < \mu < \bar{X} + 21{,}56)$

Intervalo confiança para $\mu = (825.84 - 21.56,\ 825.84 + 21.56) = (804.28,\ \ 847.4)$

ESTIMAÇÃO POR INTERVALOS

Efeito da dimensão da amostra na amplitude do Intervalo de confiança para a média

Exemplo (continuação): População $X \sim N(\mu, \sigma^2 = 605)$

Grau\nível de confiança$= 0.95 \Leftrightarrow 1 - \alpha = 0.95 \Leftrightarrow \alpha = 0.05$

Amostra anterior: $n = 5$, $\bar{x} = 825{,}84$

$I.C.$ para $\mu = (804.28, \quad 847.4)$

Amplitude do $I.C. = 43.12$

Amostra:
$$\begin{Bmatrix} 807.7, 790.7, 818.8, 853.4, 858.6, \\ 798.6, 812.2, 813.1, 839.4, 812.8 \end{Bmatrix}$$

$n = 10$, $\bar{x} = 820.53$

$I.C.$ para $\mu = (805.28, \quad 835.78)$

Amplitude do $I.C. = 30.49$

Amplitude do intervalo de confiança reduz-se quando aumenta a dimensão da amostra

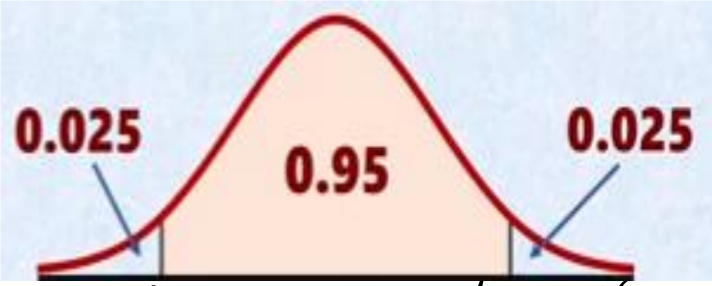
ESTIMAÇÃO POR INTERVALOS

Intervalo de confiança para a média (variância desconhecida) – Populações normais

Grau\nível de confiança$= 1-\alpha = 0.95 \Rightarrow \alpha = 0.05$

Variável Fulcral $T = \dfrac{\bar{X} - \mu}{S'/\sqrt{n}} \sim t(n-1)$

$t_{\alpha/2}: P\left(-t_{\alpha/2} < T < t_{\alpha/2}\right) = 1-\alpha$



$$-t_{\alpha/2} < \frac{\bar{X} - \mu}{S'/\sqrt{n}} < t_{\alpha/2}$$

$$\bar{X} - t_{\alpha/2}\frac{S'}{\sqrt{n}} < \mu < \bar{X} + t_{\alpha/2}\frac{S'}{\sqrt{n}}$$

Intervalo aleatório para $\mu = \left(\bar{X} - t_{\alpha/2}\frac{S'}{\sqrt{n}}, \bar{X} + t_{\alpha/2}\frac{S'}{\sqrt{n}}\right)$

Intervalo confiança para $\mu = \left(\bar{x} - t_{\alpha/2}\frac{s'}{\sqrt{n}}, \bar{x} + t_{\alpha/2}\frac{s'}{\sqrt{n}}\right)$

ESTIMAÇÃO POR INTERVALOS

Intervalo de confiança para a média (variância desconhecida)

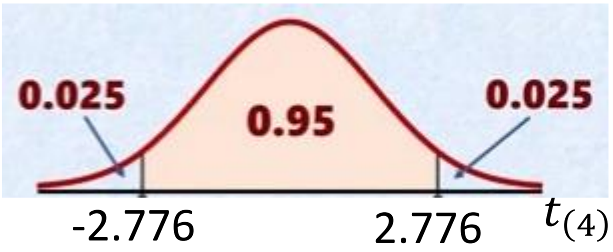
Exemplo: População $X \sim N(\mu, \sigma^2)$, μ, σ^2 desconhecidos

Grau\nível de confiança$= 0.95 \Leftrightarrow 1-\alpha = 0.95 \Leftrightarrow \alpha = 0.05$

Seja amostra casual (807.7, 790.7, 818.8, 853.4, 858.6)$\Rightarrow n = 5,\ \bar{x} = 825{,}84, s' = 29.352$

Variável fulcral $-T = \dfrac{\bar{X} - \mu}{s'/\sqrt{n}} \sim t\left(\underbrace{5-1}_{4}\right)$

$\Rightarrow t_{(4)}^{0.025} = \mathrm{invt}(0.025, 4) = 2.776$



Margem de erro - $\varepsilon = 2.776 \dfrac{29.352}{\sqrt{5}} = 36.44 \Rightarrow \bar{X} - 36.44 < \mu < \bar{X} + 36.44$

Intervalo aleatório para $\mu = (\bar{X} - 36.44 < \mu < \bar{X} + 36.44)$

Intervalo confiança para $\mu = (825{,}84 - 36.44,\ 825.84 - 36.44) = (789.40,\ \ 862.28)$

ESTIMAÇÃO POR INTERVALOS

Efeito da variação do grau de confiança na amplitude do Intervalo de confiança para a média (variância desconhecida)

Exemplo: População $X \sim N(\mu, \sigma^2)$, μ, σ^2 desconhecidos

Amostra anterior: $n = 5$, $\bar{x} = 825{,}84$, $s' = 29.352$

Grau de confiança $= 0.95$

$\Leftrightarrow 1 - \alpha = 0.95 \Leftrightarrow \alpha = 0.05$

$t_{(4)}^{0.025} = 2.776$

$I.C.$ para $\mu = (789.40,\ 862.28)$

Amplitude do $I.C. = 72.88$

Grau de confiança$= 0.99 \Leftrightarrow 1 - \alpha = 0.99$

$\Leftrightarrow \alpha = 0.01$

$t_{(4)}^{0.005} = 4.604$

$I.C.$ para $\mu = (765.40,\ \ 886.28)$

Amplitude do $I.C. = 120.87$

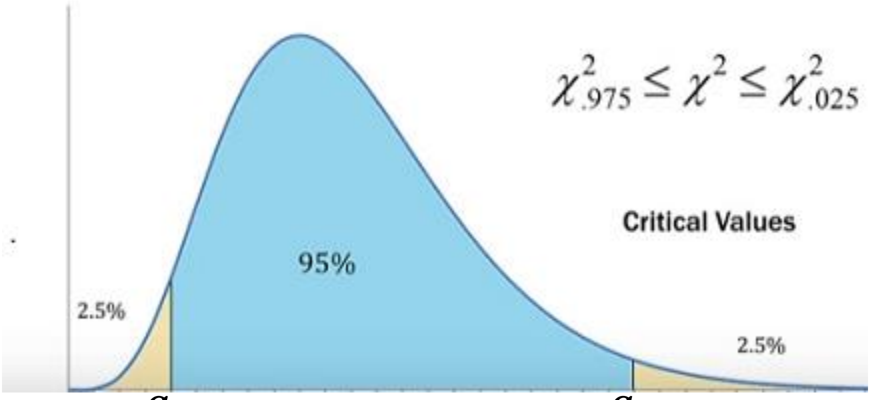
A amplitude do intervalo de confiança aumenta quando se aumenta o grau de confiança

ESTIMAÇÃO POR INTERVALOS

Intervalo de confiança para a variância – Populações normais

Grau\nível de confiança= $1-\alpha = 0.95 \Rightarrow \alpha = 0.05$

Variável Fulcral $T = \dfrac{(n-1)S'^2}{\sigma^2} \sim \chi^2_{(n-1)}$



q_1 $\quad$ q_2

$\chi^2_{(0.975)}$ $\quad$ $\chi^2_{(0.025)}$

$$q_1 < \frac{(n-1)S'^2}{\sigma^2} < q_2$$

$$\frac{(n-1)S'^2}{q_1} < \sigma^2 < \frac{(n-1)S'^2}{q_2}$$

Intervalo aleatório para $\sigma^2 = \left(\dfrac{(n-1)S'^2}{q_1}, \dfrac{(n-1)S'^2}{q_2}\right)$

Intervalo confiança para $\sigma^2 = \left(\dfrac{(n-1)s'^2}{q_1}, \dfrac{(n-1)s'^2}{q_2}\right)$

ESTIMAÇÃO POR INTERVALOS

Intervalo de confiança para a variância – Populações normais

Exemplo: Ao investir em acções, geralmente há um trade-off: risco vs rendimento.

Em termos financeiros, "risco" é sinónimo de variância no valor das acções.

Algumas acções são estáveis (baixo risco), mas oferecem baixos retornos potenciais (GE), outras variam descontroladamente (maior risco), mas oferecem retornos potenciais mais elevados (Apple).

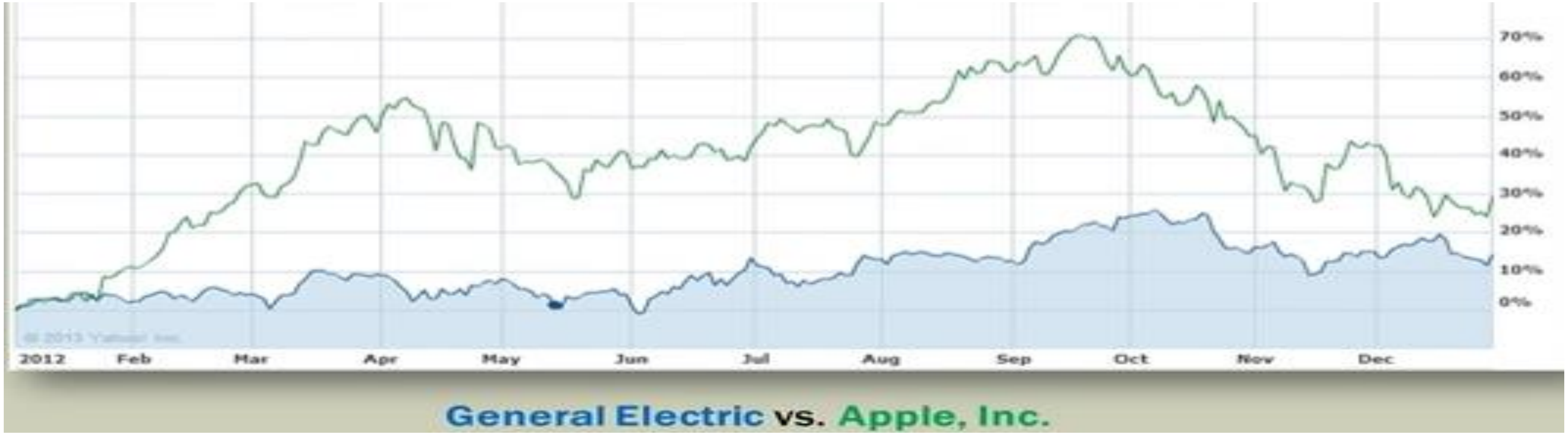
Suponham que compramos acções da GE e da Apple e as mantemos por um ano.

ESTIMAÇÃO POR INTERVALOS

Intervalo de confiança para a variância – Populações normais

Exemplo (continuação):

Rendimentos das acções em percentagem



Pense em cada tipo de acção como um vôo. Qual o vôo que vos deixará mais enjoados devido à turbulência?

ESTIMAÇÃO POR INTERVALOS

Intervalo de confiança para a variância – Populações normais

Exemplo (continuação):

Date	GE	APPL	GE%	AAPL%
January 2012	0.043869	0.119655	4.39%	11.97%
February 2012	0.026947	0.172533	2.69%	17.25%
March 2012	0.051809	0.100103	5.18%	10.01%
April 2012	-0.02451	-0.026306	-2.45%	-2.63%
May 2012	-0.02567	-0.010762	-2.57%	-1.08%
June 2012	0.096491	0.010796	9.65%	1.08%
July 2012	-0.00444	0.044801	-0.44%	4.48%
August 2012	-0.00198	0.089724	-0.20%	8.97%
September 2012	0.099801	0.002791	9.98%	0.28%
October 2012	-0.07535	-0.113841	-7.53%	-11.38%
November 2012	0.003376	-0.012450	0.34%	-1.25%
December 2012	0.002404	-0.095123	0.24%	-9.51%
Mean	0.016063	0.023493	1.61%	2.35%
Variance	0.002590	0.007330	25.89	73.30
Standard Dev.	0.050890	0.085618	5.09%	8.56%

Mean Monthly Return

GE = 1.61%   AAPL = 2.35%

Monthly Return Variance, s^2

GE = 25.89   AAPL = 73.30

ESTIMAÇÃO POR INTERVALOS

Intervalo de confiança para a variância – Populações normais

Exemplo (continuação):

$$\frac{(n-1)s^2}{\chi^2_{.025}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{.975}}$$

n = sample size

s^2 = sample variance

$$\chi^2_{.975} \le \chi^2 \le \chi^2_{.025}$$

$$3.82 \le \chi^2 \le 21.92$$

Critical Values

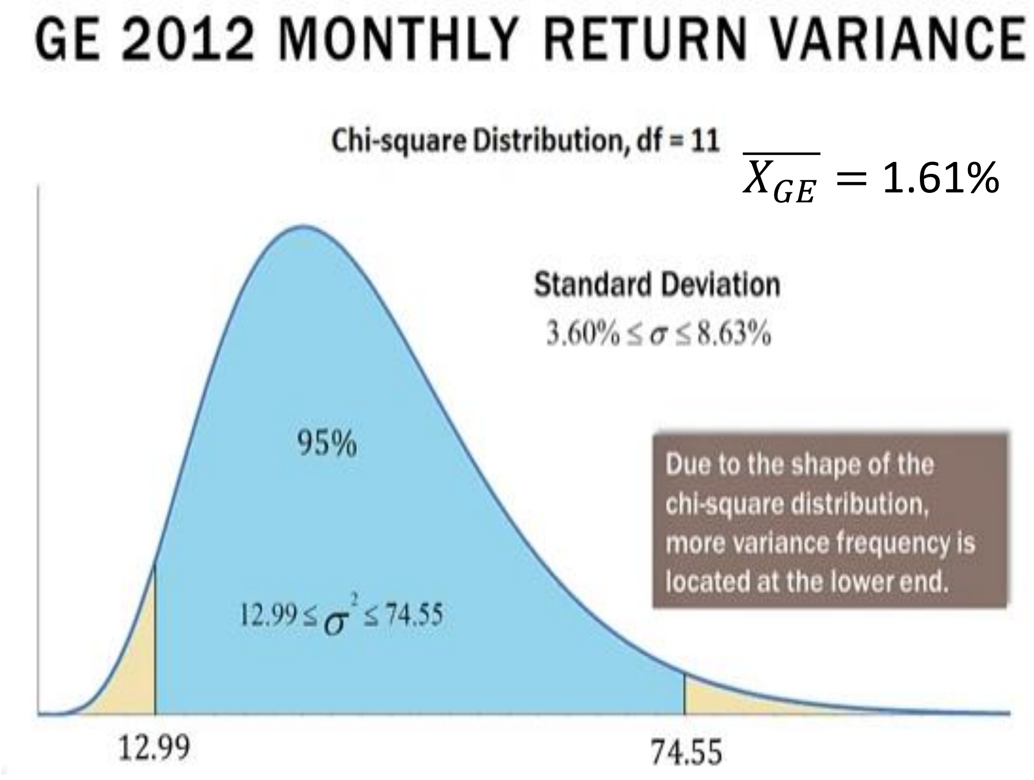
Monthly Return Variance, s^2

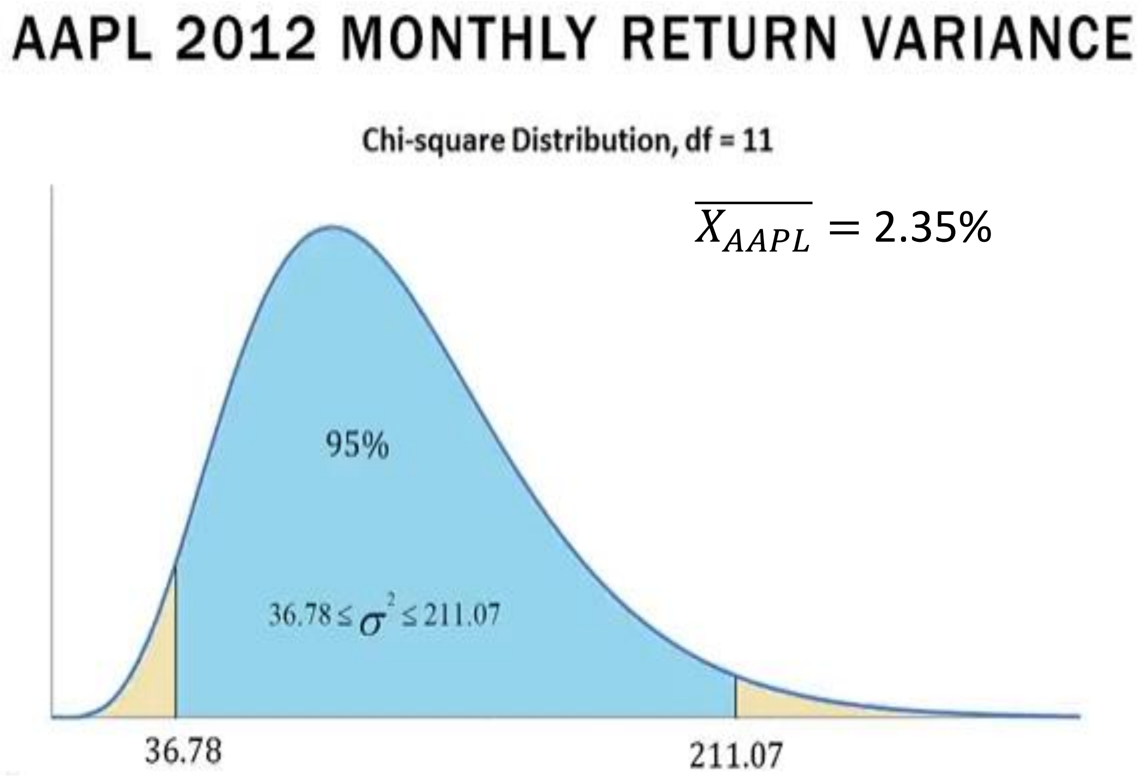
GE = 25.89	AAPL = 73.30
$\frac{(12-1)25.89}{21.92} \le \sigma^2 \le \frac{(12-1)25.89}{3.82}$	$\frac{(12-1)73.30}{21.92} \le \sigma^2 \le \frac{(12-1)73.30}{3.82}$
$12.99 \le \sigma^2 \le 74.55$	$36.78 \le \sigma^2 \le 211.07$
$3.60\% \le \sigma \le 8.63\%$	$6.06\% \le \sigma \le 14.53\%$

ESTIMAÇÃO POR INTERVALOS

Intervalo de confiança para a variância – Populações normais

Exemplo (continuação):





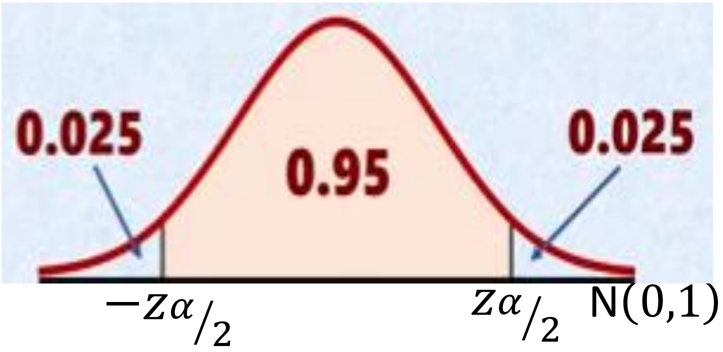
Amplitude do Intervalo = 61.56

Amplitude do Intervalo = 174.29

ESTIMAÇÃO POR INTERVALOS

Intervalo de confiança (grandes amostras) – qualquer População

Grau\nível de confiança$= 1-\alpha = 0.95 \Rightarrow \alpha = 0.05$



Variável Fulcral $Z = \dfrac{\bar{X} - \mu}{S'/\sqrt{n}} \dot{\sim} N(0,1)$

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

$$-z_{\alpha/2} < \frac{\bar{X} - \mu}{S'/\sqrt{n}} < z_{\alpha/2}$$

Intervalo aleatório para $\theta = \left(\bar{X} - z_{\alpha/2}\, S'/\sqrt{n}, \bar{X} + z_{\alpha/2}\, S'/\sqrt{n}\right)$

Intervalo confiança para $\theta = \left(\bar{x} - z_{\alpha/2}\, s'/\sqrt{n}, \bar{x} + z_{\alpha/2}\, s'/\sqrt{n}\right)$

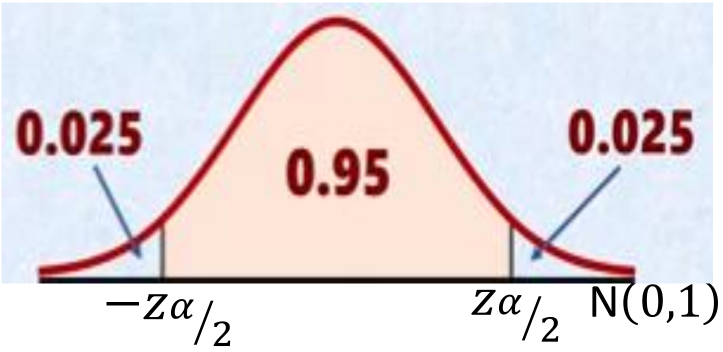
ESTIMAÇÃO POR INTERVALOS

Intervalo de confiança para a proporção – Populações Bernoulli

Grau\nível de confiança$= 1-\alpha = 0.95 \Rightarrow \alpha = 0.05$

Variável Fulcral $Z = \dfrac{\bar{X} - \theta}{\sqrt{\dfrac{\bar{X}\,(1-\bar{X}\,)}{n}}} \dot{\sim} N(0,1)$

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



$$-z_{\alpha/2} < \frac{\bar{X} - \theta}{\sqrt{\dfrac{\bar{X}\,(1-\bar{X}\,)}{n}}} < z_{\alpha/2}$$

Intervalo aleatório para $\theta = \left(\bar{X} - z_{\alpha/2}\sqrt{\dfrac{\bar{X}\,(1-\bar{X}\,)}{n}}, \bar{X} + z_{\alpha/2}\sqrt{\dfrac{\bar{X}\,(1-\bar{X}\,)}{n}}\right)$

Intervalo confiança para $\theta = \left(\bar{x} - z_{\alpha/2}\sqrt{\dfrac{\bar{x}\,(1-\bar{x}\,)}{n}}, \bar{x} + z_{\alpha/2}\sqrt{\dfrac{\bar{x}\,(1-\bar{x}\,)}{n}}\right)$