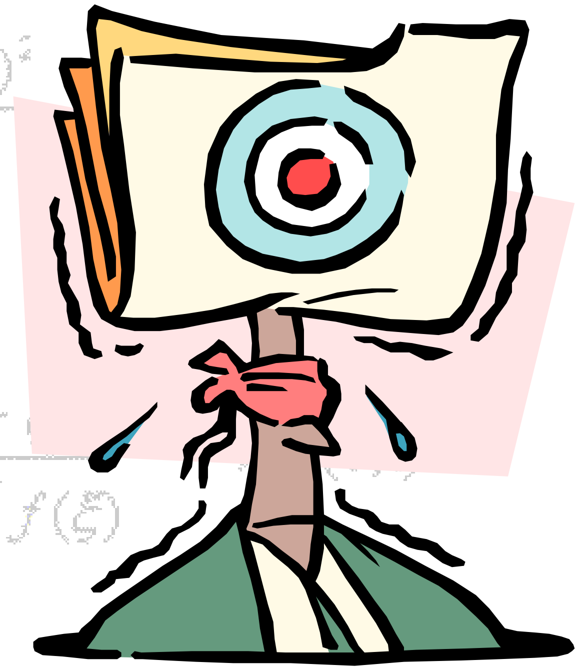
Estimação



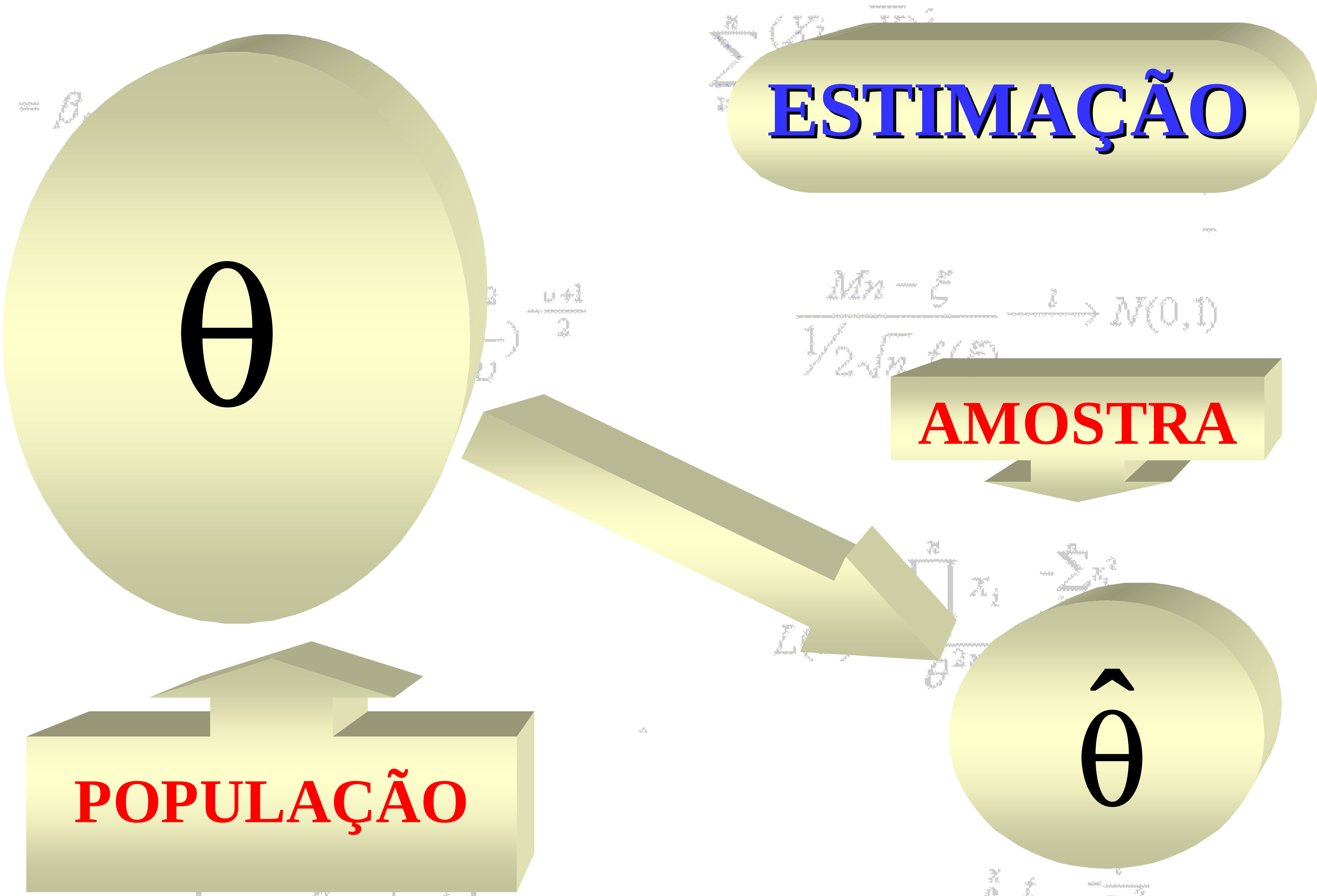
Prof. Lorí Viali, Dr.

http://www.pucrs.br/~viali/

viali@mat.pucrs.br

Estimação

Uma A estimação tem por objetivo fornecer informações sobre parâmetros populacionais, tendo como base uma amostra aleatória extraída da população de interesse.



ESTIMAÇÃO
θ
AMOSTRA
θ̂
POPULAÇÃO

Tipos de Estimação

Por Ponto

Por intervalo

ESTIMAÇÃO POR PONTO

A estimativa por ponto é feita através de um único valor.

ESTIMAÇÃO POR INTERVALO

A estimativa por intervalo, fornece um conjunto de valores.

Estimação por Ponto

As características básicas de um estimador são:

A média: $$\mu_{\hat{\theta}} = E(\hat{\theta})$$

A Variância: $$\sigma^2_{\hat{\theta}} = V(\hat{\theta}) =$$

$$= E[\hat{\theta} - E(\hat{\theta})]^2 =$$

$$= E(\hat{\theta}^2) - E(\hat{\theta})^2$$

Através da média, pode-se saber em torno de que valor o estimador está variando. O ideal é que ele varie em torno do parâmetro θ.

Através da raiz quadrada da variância tem-se uma idéia do erro cometido na estimação, isto é, o valor

$$\sigma_{\hat{\theta}} = \sqrt{V(\hat{\theta})}$$

é denominado de erro padrão de θ.

A terceira informação necessária é a distribuição do estimador, isto é, qual o modelo teórico (probabilístico) do estimador.

OUTROS CONCEITOS IMPORTANTES

Erro amostral: $$\varepsilon = \theta - \hat{\theta}$$

Viés: $$B(\hat{\theta}) = E(\hat{\theta}) - \theta$$

EQM: $$EQM\ (\hat{\theta}) = E(\hat{\theta} - \theta)^2$$

Relação entre EQM e Variância

$$EQM(\hat{\theta}) = E(\hat{\theta} - \theta)^2 =$$

$$= E[\hat{\theta} - E(\hat{\theta}) + E(\hat{\theta}) - \theta]^2 =$$

$$= E\{[\hat{\theta} - E(\hat{\theta})] + [E(\hat{\theta}) - \theta]\}^2 =$$

$$= E[\hat{\theta} - E(\hat{\theta})]^2 + [E(\hat{\theta}) - \theta]^2 +$$

$$+ 2E[\hat{\theta} - E(\hat{\theta})][E(\hat{\theta}) - \theta]$$

Como:

$$E[\hat{\theta} - E(\hat{\theta})][E(\hat{\theta}) - \theta] = 0$$

Segue:

$$EQM(\hat{\theta}) = E(\hat{\theta} - \theta)^2 =$$

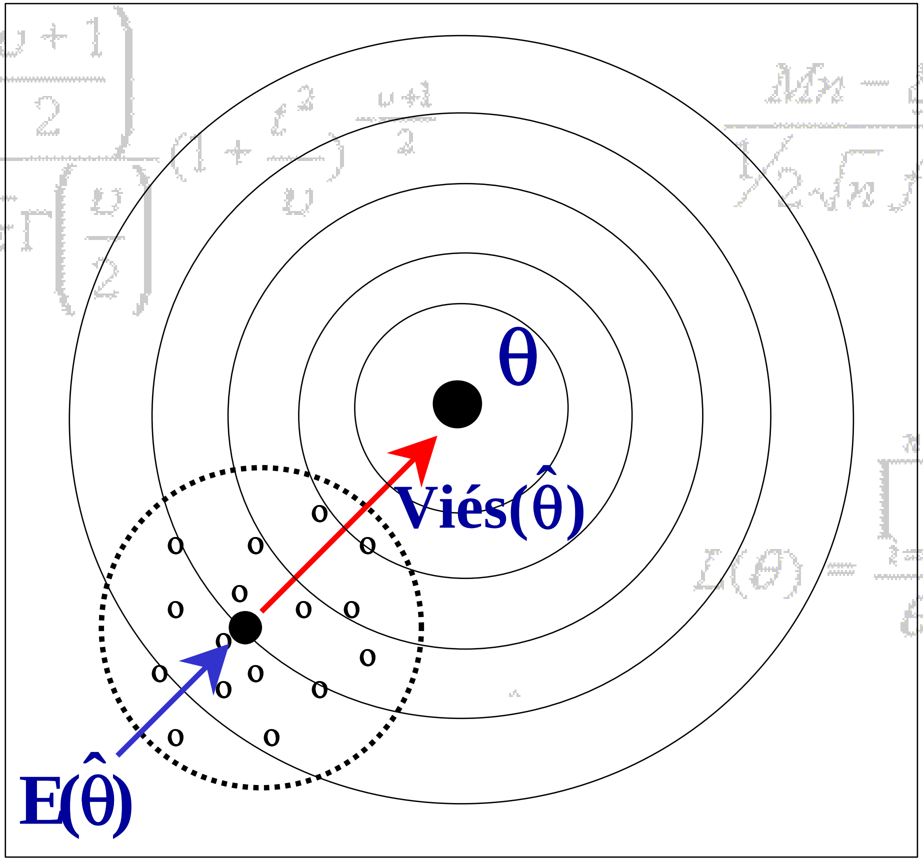
$$= E[\hat{\theta} - E(\hat{\theta})]^2 + E(\hat{\theta}) - \theta]^2 =$$

$$= V(\hat{\theta}) + B(\hat{\theta})^2$$

Isto é, o **Erro Quadrado Médio** de um estimador é a sua **Variância** **somada** com o **quadrado do Viés**.

$$\mathrm{EQM}(\hat{\theta}) = V(\hat{\theta}) + B(\hat{\theta})^2$$

Erro Quadrado Médio



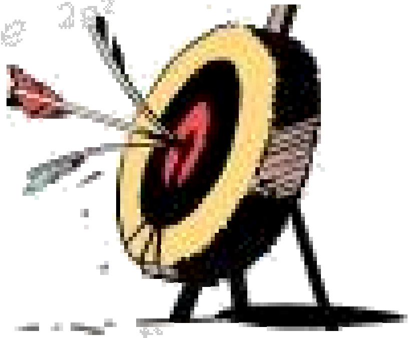
Propriedades dos Estimadores

Seja $(X_1, X_2, ..., X_n)$ uma amostra aleatória de uma variável (população) X, com um parâmetro de interesse θ. Seja $\hat{\theta}$ uma função da amostra (estimativa de θ).

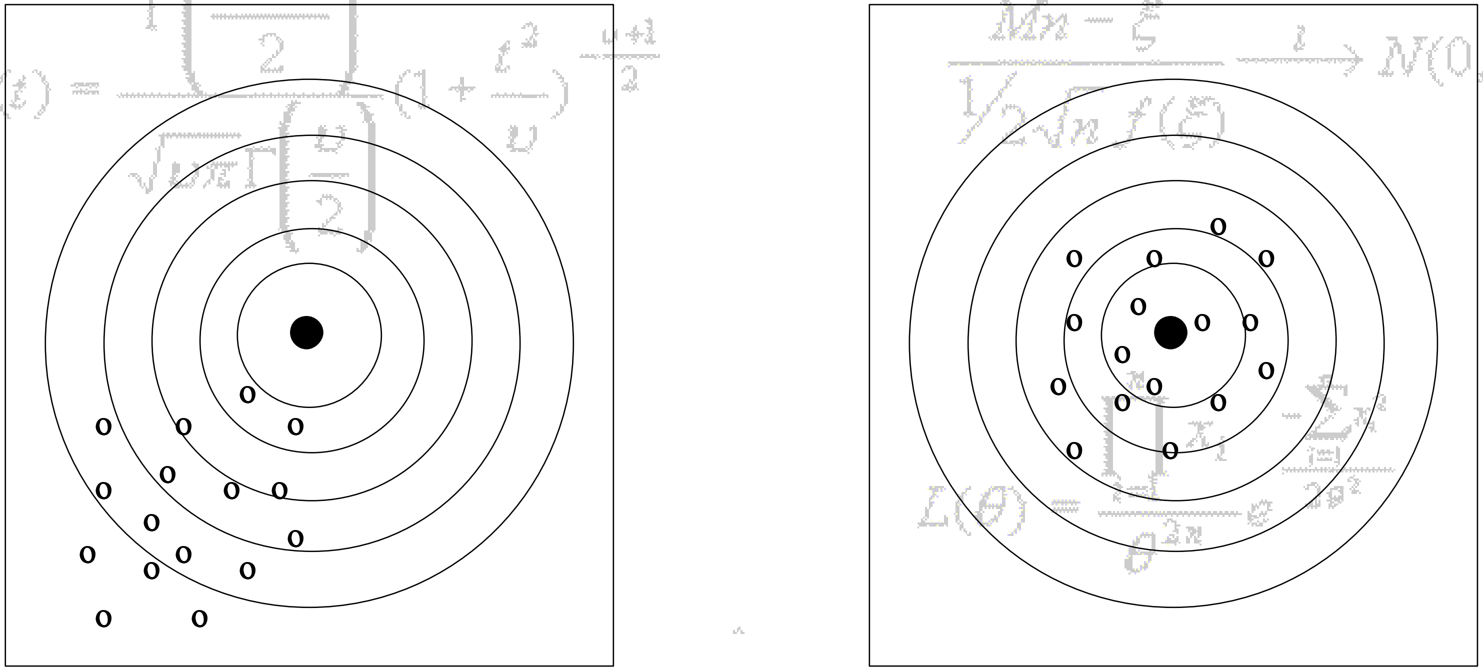
Não Tendenciosidade

Um estimador é dito **não-tendencioso**, **não-viciado**, **sem viés** ou **imparcial** se:

$$\mu_{\hat{\theta}} = E(\hat{\theta}) = \theta$$



Tendenciosidade



Tendencioso | **Não tendencioso**

Exemplos

A média da amostra $\overline{x}$ é um estimador não-viciado de μ, isto é:

$$\mu_{\overline{X}} = E(\overline{X}) = \mu$$

A proporção amostral "P" é um estimador não-viciado de π, isto é:

$$\mu_P = E(P) = \pi$$

A variância da amostra "S^2" é um estimador viciado de σ^2, isto é:

$$\mu_{S^2} = E(S^2) \neq \sigma^2$$

A variância da amostra "S^2", calculada com "**n-1**" **no denominador** é um estimador não viciado de σ^2, isto é:

$$\mu_{S^2} = \mathbf{E}(\mathbf{S}^2) = \sigma^2$$

Consistência

Um estimador **não viciado** é dito **consistente** se:

$$\lim_{n \to \infty} V(\hat{\theta}) = 0$$

Exemplos

A média da amostra $\bar{x}$ é um estimador consistente de μ, isto é:

$$\lim_{n \to \infty} V(\bar{X}) = \lim_{n \to \infty} \frac{\sigma^2}{n} = 0$$

A proporção amostral "P" é um estimador consistente de π, isto é:

$$\lim_{n \to \infty} V(P) = \lim_{n \to \infty} \sqrt{\frac{\pi(1-\pi)}{n}} = 0$$

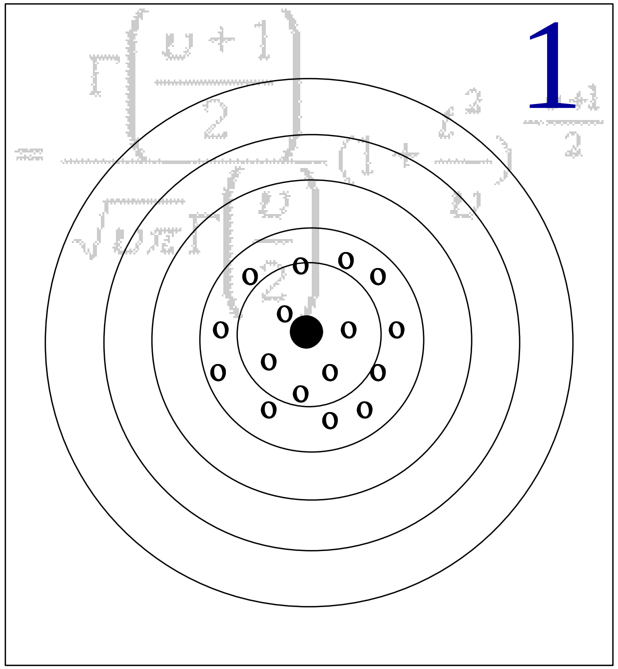
A variância da amostra "S^2" é um estimador consistente de σ^2, isto é:

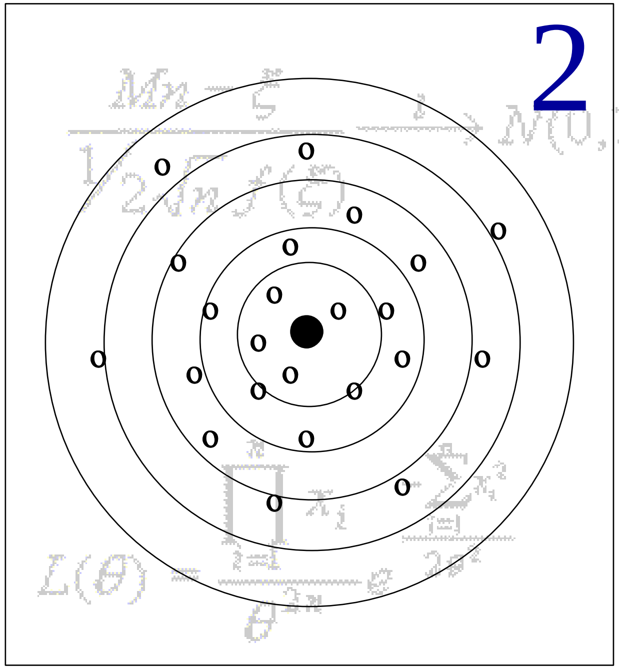
$$\lim_{n \to \infty} V(S^2) = \lim_{n \to \infty} \frac{2\sigma^4}{n-1} = 0$$

Eficiência

Dados dois estimadores **não-tendenciosos** de um mesmo parâmetro, o **mais eficiente** é o que apresenta menor variância.

Eficiência





O estimador "**1**" é mais eficiente que o "**2**"

Exemplo

A média (simples) da amostra $\bar{x}$ é um estimador mais eficiente de μ, do que **qualquer** média ponderada.

Métodos de Estimação

- Momentos
- Mínimos Quadrados
- Máxima Verossimilhança
- MELNT (Melhor Estimativa Linear Não Tendenciosa)
- Bayes

Métodos dos Momentos

É o mais antigo dos métodos para determinar estimadores (Pearson, 1894). Baseia-se no princípio de que se deve estimar o **momento de uma distribuição populacional** pelo **momento correspondente da amostra**.

Desta forma a **média populacional** deve ser estimada pela **média amostral**, a **variância populacional** pela **variância amostral** e assim por diante.

Este método produz estimadores que são consistentes e assintoticamente normais.

Exercício um

Considere o seguinte conjunto de valores:

-3 -1,2 -0,5 0,9 1,1 2,2 2,8 4,5

Determine estimativas da:

(a) Média

(b) Variabilidade

(c) Da proporção de positivos

A média

A melhor estimativa da média é dada pela média da amostra. Assim:

$$\bar{x} = \frac{\sum x_i}{n} =$$

$$= \frac{-3 - 1{,}2 - 0{,}5 + 0{,}9 + 1{,}1 + 2{,}2 + 2{,}8 + 4{,}5}{8} =$$

$$= \frac{6{,}8}{8} = 0{,}85$$

A variância

A melhor estimativa da variância (σ^2) é dada pela variância amostral (s^2). Assim:

$$s^2 = \frac{\sum x_i^2 - n\,\overline{x}^2}{n-1} == \frac{45{,}64 - 8.(0{,}85)^2}{8-1} =$$

$$= \frac{45{,}64 - 5{,}78}{8-1} = \frac{39{,}86}{7} \cong 5{,}69$$

O desvio padrão

Extraindo a raiz quadrada da variância, tem-se uma estimativa do desvio padrão:

$$s = \sqrt{\frac{\sum x_i^2 - n\,\overline{x}^2}{n-1}} = \sqrt{\frac{45,64 - 8.(0,85)^2}{8-1}} =$$

$$= \sqrt{\frac{45,64 - 5,78}{8-1}} = \sqrt{\frac{39,86}{7}} = \sqrt{5,6943} \cong 2,39$$

A proporção

A melhor estimativa de $\boldsymbol{\pi}$ é dada pela proporção amostral (**p**):

$$p = \frac{f}{n} = \frac{5}{8} = 0{,}625 = 62{,}50\,\%$$

Exercício dois

Com base na distribuição da velocidades de uma amostra de **120** carros andando na estrada POA/Osório, determine estimativas da:

(a) velocidade média

(b) variabilidade da velocidade

(c) da proporção de carros acima dos 100 km/h

Velocidades	Freqüência
80 ├── 85	8
85 ├── 90	13
90 ├── 95	24
95 ├── 100	33
100 ├── 105	29
105 ├── 110	13
Total	120

Velocidades	Freqüência	x_i	$f_i x_i$
80 ├─ 85	8	82,5	660,0
85 ├─ 90	13	87,5	1137,5
90 ├─ 95	24	92,5	2220,0
95 ├─ 100	33	97,5	3217,5
100 ├─ 105	29	102,5	2972,5
105 ├─ 110	13	107,5	13,97,5
Total	120	—	**11605**

A média

A melhor estimativa da média é dada pela média da amostra. Assim:

$$\bar{x} = \frac{\sum f_i x_i}{n} = \frac{11605}{120} = 96{,}71 \ \ \text{km/h}$$

Velocidades	Freqüência	x_i	$f_i x_i^2$
80 ├── 85	8	82,5	54450,00
85 ├── 90	13	87,5	99531,25
90 ├── 95	24	92,5	205350,00
95 ├── 100	33	97,5	313706,25
100 ├── 105	29	102,5	304681,25
105 ├── 110	13	107,5	150231,25
Total	120	—	**1127950**

O desvio padrão

$$s = \sqrt{\frac{\sum f_i x_i^2 - n\,\overline{x}^2}{n-1}} =$$

$$\sqrt{\frac{1127950 - 120.(96{,}7083)^2}{120-1}} =$$

$$= \sqrt{\frac{5649{,}7917}{119}} = \sqrt{47{,}4772} \cong 6{,}89 \text{ km/h}$$

A proporção

A melhor estimativa de π é dada pela proporção amostral (p):

$$p = \frac{f}{n} = \frac{(29 + 13)}{120} =$$

$$= \frac{42}{120} = 0{,}35 = 35\%$$

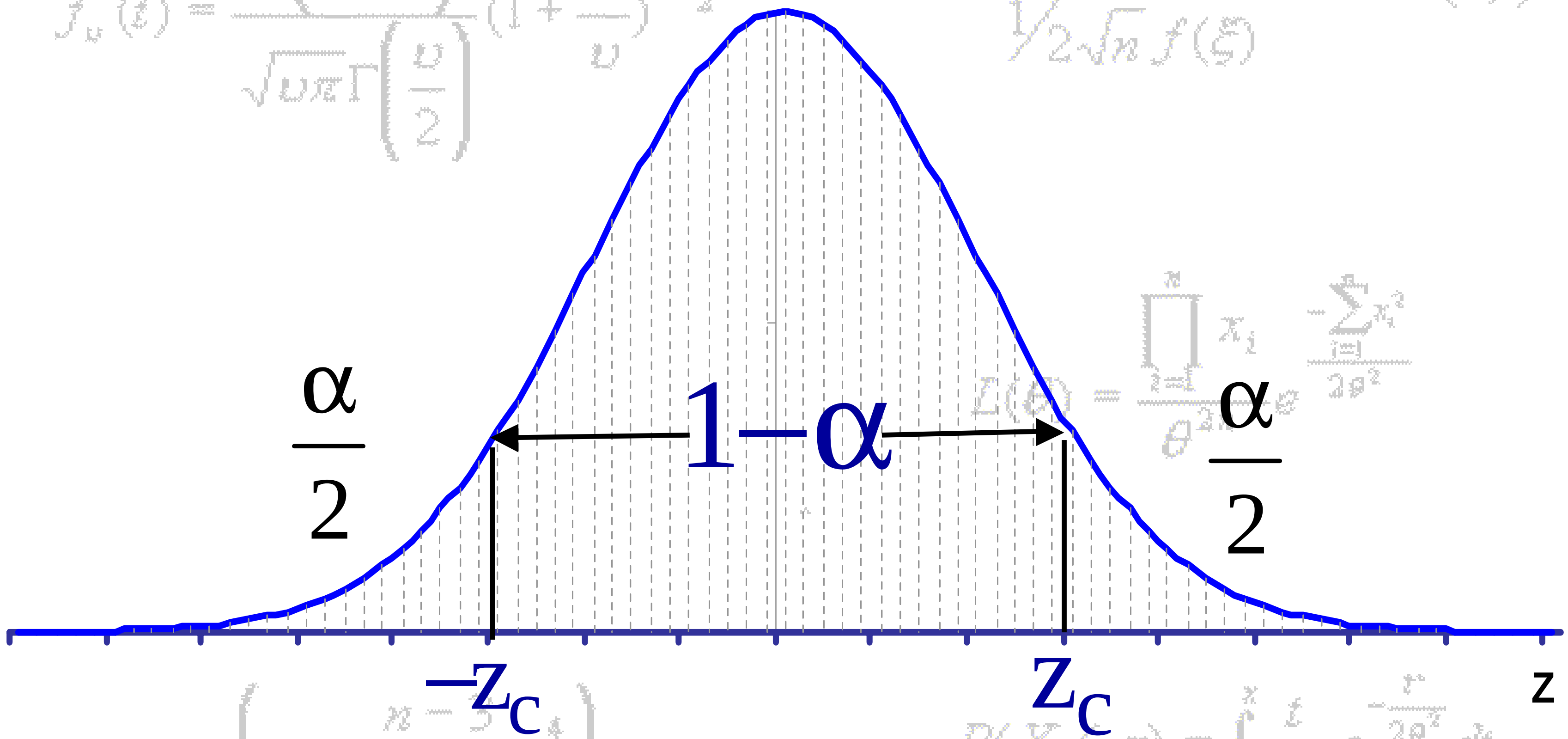
Estimação por Intervalo

(A)

Da Média

Supondo σ conhecido

$$P(-z_c < Z < z_c) = 1 - \alpha$$



De $P(-z_c < Z < z_c) = 1 - \alpha$

Tem-se:

$$P\left(-z_c < \frac{\overline{X} - \mu_{\overline{X}}}{\sigma_{\overline{X}}} < z_c\right) = 1 - \alpha$$

$$P(-z_c \cdot \sigma_{\overline{X}} < \overline{X} - \mu < z_c \cdot \sigma_{\overline{X}}) = 1 - \alpha$$

$$P(-\overline{X} - z_c \cdot \sigma_{\overline{X}} < -\mu < -\overline{X} + z_c \cdot \sigma_{\overline{X}}) = 1 - \alpha$$

Assim:

$$P(-\overline{X} - z_c \cdot \sigma_{\overline{X}} < -\mu < -\overline{X} + z_c \cdot \sigma_{\overline{X}}) = 1 - \alpha$$

$$P(\overline{X} - z_c \cdot \sigma_{\overline{X}} < \mu < \overline{X} + z_c \cdot \sigma_{\overline{X}}) = 1 - \alpha$$

Então, o IC de "1 – α" para μ é calculado por:

$$\overline{X} \pm \varepsilon_{\overline{X}} \qquad \varepsilon_{\overline{X}} = z_c \sigma_{\overline{x}} \qquad \sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

Exemplo

Com base na distribuição das velocidades de uma amostra de **120** carros andando na estrada POA/Osório, e supondo que o desvio padrão populacional é igual a sete km/h determine uma estimativa para a velocidade média, com uma confiabilidade de 95%.

Tem-se:

$$\bar{X} \pm \varepsilon_{\bar{X}} \qquad \varepsilon_{\bar{X}} = z_c \sigma_{\bar{X}} \qquad \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

Mas: $z_c = 1{,}96$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{7}{\sqrt{120}} = 0{,}6390$$

$$\varepsilon_{\bar{X}} = 1{,}96.0{,}6390 = 1{,}25$$

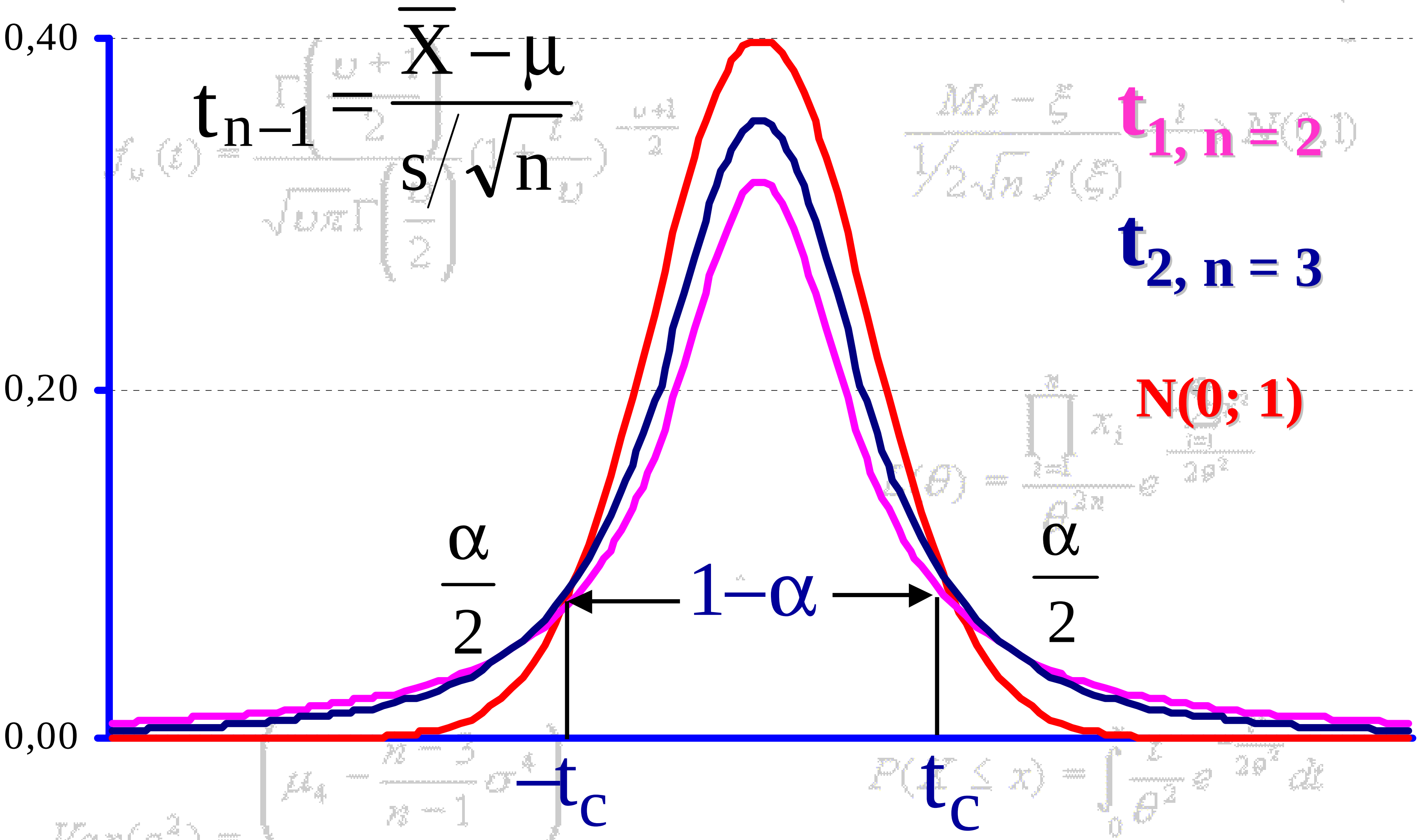
O IC de "1 – α" para μ é calculado por:

$$\left[\overline{X} - \varepsilon_{\overline{X}} ;\ \overline{X} + \varepsilon_{\overline{X}}\right]$$

$$\left[96{,}71 - 1{,}25 ;\ 96{,}71 + 1{,}25\right]$$

$$\left[95{,}46 ;\ 97{,}96\right]$$

σ desconhecido



De $P(-t_c < t < t_c) = 1 - \alpha$

Tem-se:

$$P\left(-t_c < \frac{\overline{X} - \mu_{\overline{X}}}{\hat{\sigma}_{\overline{X}}} < t_c\right) = 1 - \alpha$$

$$P(-t_c \cdot \hat{\sigma}_{\overline{X}} < \overline{X} - \mu < t_c \cdot \hat{\sigma}_{\overline{X}}) = 1 - \alpha$$

$$P(-\overline{X} - t_c \cdot \hat{\sigma}_{\overline{X}} < -\mu < -\overline{X} + z_t \cdot \hat{\sigma}_{\overline{X}}) = 1 - \alpha$$

Assim:

$$P(-\overline{X} - t_c \cdot \hat{\sigma}_{\overline{X}} < -\mu < -\overline{X} + t_c \cdot \hat{\sigma}_{\overline{X}}) = 1 - \alpha$$

$$P(\overline{X} - t_c \cdot \hat{\sigma}_{\overline{X}} < \mu < \overline{X} + t_c \cdot \hat{\sigma}_{\overline{X}}) = 1 - \alpha$$

Então, o IC de "1 – α" para μ, se σ for desconhecido é calculado por:

$$\overline{X} \pm \hat{\varepsilon}_{\overline{X}} \qquad \hat{\varepsilon}_{\overline{X}} = t_c \hat{\sigma}_{\overline{X}} \qquad \hat{\sigma}_{\overline{X}} = \frac{s}{\sqrt{n}}$$

Exemplo

Com base na distribuição das velocidades de uma amostra de **120** carros andando na estrada POA/Osório, determine uma estimativa para a **velocidade média**, com uma confiabilidade de 95%.

Tem-se:

$$\overline{X} \pm \hat{\varepsilon}_{\overline{X}} \qquad \hat{\varepsilon}_{\overline{X}} = t_c \hat{\sigma}_{\overline{X}} \qquad \hat{\sigma}_{\overline{X}} = \frac{s}{\sqrt{n}}$$

Mas:

$$t_c = 1{,}98$$

$$\hat{\sigma}_{\overline{X}} = \frac{s}{\sqrt{n}} = \sqrt{\frac{47{,}4772}{120}} = 0{,}6290$$

$$\hat{\varepsilon}_{\overline{X}} = 1{,}98.0{,}6290 = 1{,}25$$

O IC de "1 – α" para μ é calculado por:

$$\left[\overline{X} - \hat{\varepsilon}_{\overline{X}}\,;\ \ \overline{X} + \hat{\varepsilon}_{\overline{X}}\right]$$

$$\left[96{,}71 - 1{,}25\,;\ \ 96{,}71 + 1{,}25\right]$$

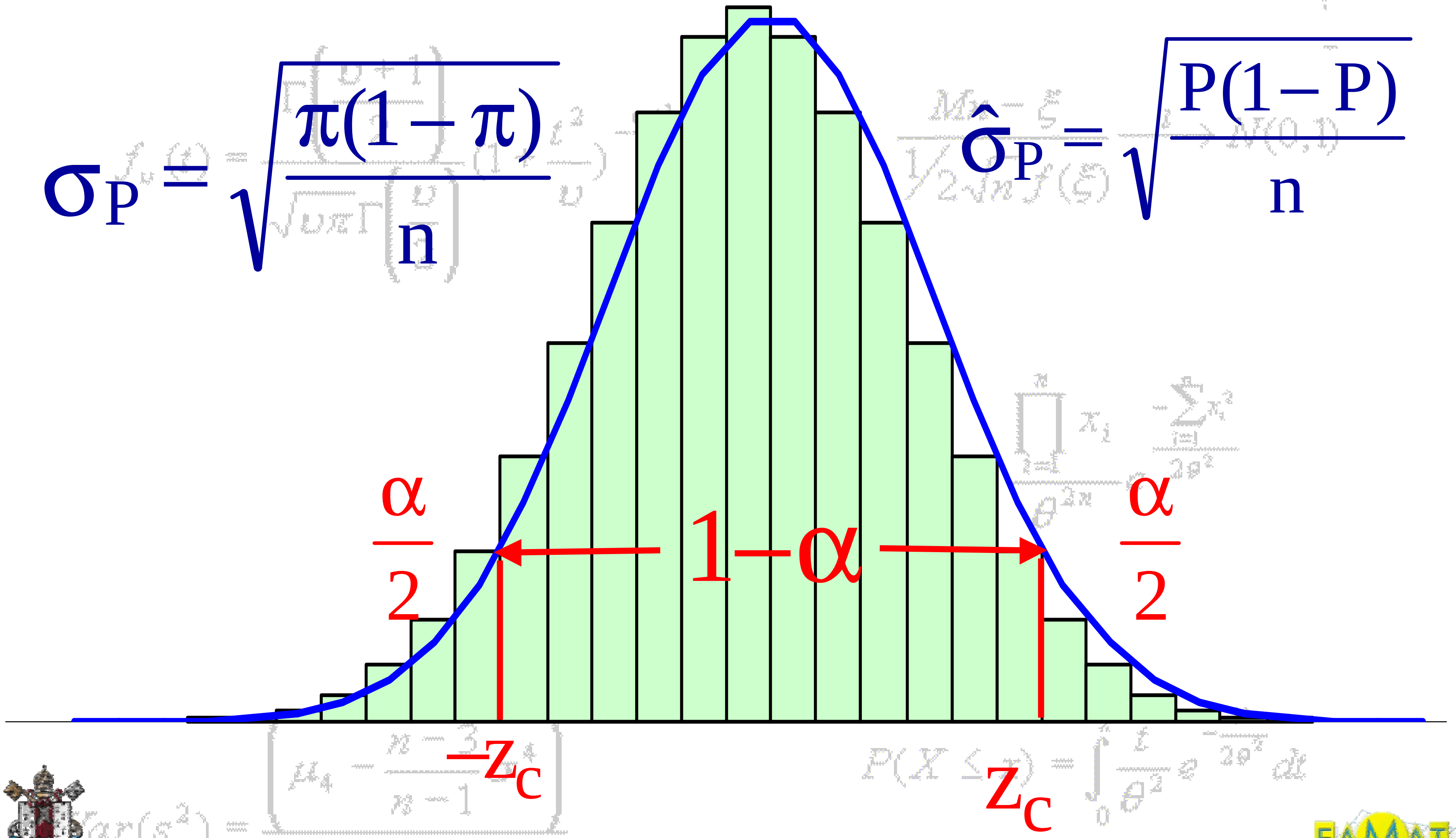
$$\left[95{,}46\,;\ 97{,}96\right]$$

(B)

Da Proporção

$$P(-z_c < Z < z_c) = 1 - \alpha$$

$$\sigma_P = \sqrt{\frac{\pi(1-\pi)}{n}} \qquad \hat{\sigma}_P = \sqrt{\frac{P(1-P)}{n}}$$



De $$P(-z_c < Z < z_c) = 1-\alpha$$

Tem-se:

$$P\left(-z_c < \frac{\overline{X} - \mu_P}{\sigma_P} < z_c\right) = 1-\alpha$$

$$P(-z_c \cdot \sigma_P < P - \mu < z_c \cdot \sigma_P) = 1-\alpha$$

$$P(-P - z_c \cdot \sigma_P < -\mu < -P + z_c \cdot \sigma_P) = 1-\alpha$$

Assim:

$$P(-P - z_c \cdot \sigma_P < -\mu < -P + z_c \cdot \sigma_P) = 1 - \alpha$$

$$P(P - z_c \cdot \sigma_P < \mu < P + z_c \cdot \sigma_P) = 1 - \alpha$$

Então, o IC de “1 – α” para π é calculado por:

$$P \pm \hat{\varepsilon}_P \qquad \hat{\varepsilon}_P = z_c \hat{\sigma}_P \qquad \hat{\sigma}_P = \sqrt{\frac{P(1-P)}{n}}$$

Exemplo

Com base na distribuição das velocidades de uma amostra de **120** carros andando na estrada POA/Osório, determine uma estimativa para a **proporção de carros com velocidade acima de 100 km/h**, com uma confiabilidade de 95%.

Tem-se:

$$P \pm \hat{\varepsilon}_P \qquad \hat{\varepsilon}_P = z_c \hat{\sigma}_P \qquad \hat{\sigma}_P = \sqrt{\frac{P(1-P)}{n}}$$

Mas:

$$z_c = 1{,}96$$

$$\hat{\sigma}_P = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0{,}35.(1-0{,}35)}{120}} = 4{,}3541\%$$

$$\hat{\varepsilon}_{\overline{X}} = 1{,}96.4{,}3541 = 8{,}53\%$$

O IC de "1 – α" para π é calculado por:

$$\left[P-\hat{\varepsilon}_P;\ P+\hat{\varepsilon}_P\right]$$

$$\left[35\%-8{,}53\%;\quad 35\%+8{,}53\%\right]$$

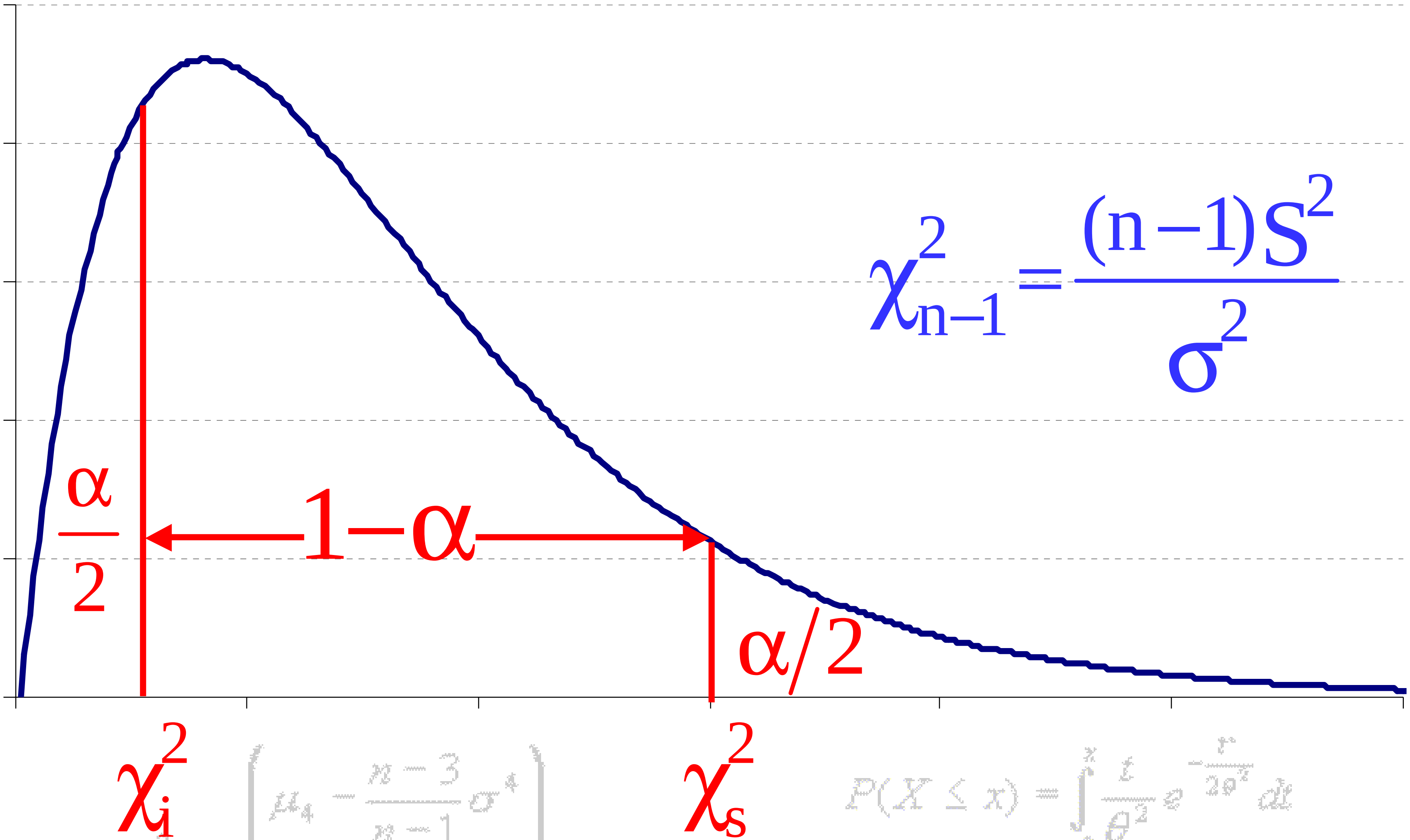
$$\left[26{,}47\%;\ 43{,}53\%\right]$$

(C)

Da Variância (Desvio Padrão)

$$P(\chi_i^2 < \chi_{n-1}^2 < \chi_s^2) = 1-\alpha$$

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



De $$P(\chi_i^2 < \chi_{n-1}^2 < \chi_s^2) = 1 - \alpha$$

Tem-se:

$$P(\chi_i^2 < \frac{(n-1)S^2}{\sigma^2} < \chi_s^2) = 1 - \alpha$$

$$P(\frac{1}{\chi_s^2} < \frac{\sigma^2}{(n-1)S^2} < \frac{1}{\chi_i^2}) = 1 - \alpha$$

$$P(\frac{(n-1)S^2}{\chi_s^2} < \sigma^2 < \frac{(n-1)S^2}{\chi_i^2}) = 1 - \alpha$$

Então o IC de "1 – α" para σ^2 é calculado por:

$$\left[\frac{(n-1)S^2}{\chi^2_s};\frac{(n-1)S^2}{\chi^2_i}\right]$$

Então o IC de “1 – α” para σ é calculado por:

$$\left[\sqrt{\frac{(n-1)S^2}{\chi_s^2}};\ \sqrt{\frac{(n-1)S^2}{\chi_i^2}}\right]$$

Exemplo

Com base na distribuição das velocidades de uma amostra de **120** carros andando na estrada POA/Osório, determine uma estimativa para a **variabilidade da velocidade**, com uma confiabilidade de 95%.

Tem-se:

$$\left[\sqrt{\frac{(n-1)S^2}{\chi_s^2}};\ \sqrt{\frac{(n-1)S^2}{\chi_i^2}}\right]$$

Mas:

$$\chi_i^2 = 94{,}81$$

$$\chi_s^2 = 145{,}46$$

O IC de “1 – α” para σ é calculado por:

$$\left[\sqrt{\frac{119.47,4772}{145,46}};\ \sqrt{\frac{119.47,4772}{94,81}}\right]$$

$$[6,23;\ 7,72]$$

Dimensionamento da Amostra

É desejável um IC com alta confiabilidade (1 - α) e pequena amplitude (ε) . Isto requer uma amostra suficientemente grande, pois, para "n" fixo, confiança e precisão varia inversamente.

A seguir os tamanhos mínimos necessários de amostras para estimar os principais parâmetros dentro de uma **confiabilidade** $(1 - \alpha)$ e uma **precisão** (ε) especificados.

Para estimar a média de uma população, supondo σ conhecido

$$\varepsilon = z_c \sigma_{\bar{x}} = z_c \frac{\sigma}{\sqrt{n}}$$

$$\sqrt{n} = \frac{\sigma . z_c}{\varepsilon}$$

$$n \geq \left(\frac{\sigma . z_c}{\varepsilon} \right)^2$$

Para estimar a média de uma população, com σ conhecido

$$\varepsilon = t_c s_{\bar{x}} = t_c \frac{s}{\sqrt{n}}$$

$$\sqrt{n} = \frac{s . t_c}{\varepsilon}$$

$$n \geq \left(\frac{s . t_c}{\varepsilon}\right)^2$$

t_c será obtido através de uma amostra piloto n’

Para estimar a proporção populacional.

$$\varepsilon = z_c \sigma_{\bar{x}} = z_c \sqrt{\frac{P(1-P)}{n}}$$

$$\sqrt{n} = \frac{z_c}{\varepsilon}\sqrt{P(1-P)}$$

$$n \geq \left(\frac{z_c}{\varepsilon}\right)^2 P(1-P)$$

"p" será estimado através de uma amostra piloto n'

Exemplo

Qual o tamanho mínimo de uma amostra para estimarmos a proporção de defeituosos de uma máquina com uma precisão de 3% e uma confiabilidade de 95%. Se **(a)** nada se sabe sobre esta proporção **(b)** ela não é superior a 10%.

Solução

(a)

$$n \geq \left(\frac{z_c}{\varepsilon}\right)^2 P(1-P)$$

$$n \geq \left(\frac{1,96}{0,03}\right)^2 0,50 . 0,5$$

$$n \geq 1068$$

Solução

(b)

$$n \geq \left(\frac{z_c}{\varepsilon}\right)^2 P(1-P)$$

$$n \geq \left(\frac{1,96}{0,03}\right)^2 0,1 . 0,9$$

$$n \geq 385$$

Até a próxima!