

# Amostragem & Estimação



**Prof. Lorí Viali, Dr.**  
**viali@mat.ufrgs.br**

**<http://www.ufrgs.br/~viali/>**

# População



Uma coleção de todos os possíveis elementos, objetos ou medidas de interesse.





# Censo

Um levantamento efetuado sobre toda uma população é denominado de **levantamento censitário** ou simplesmente **censo**.





# Amostra

Um subconjunto finito  
de uma população de  
interesse.





# Amostragem

O processo de escolha de uma amostra da população é denominado de **amostragem**.



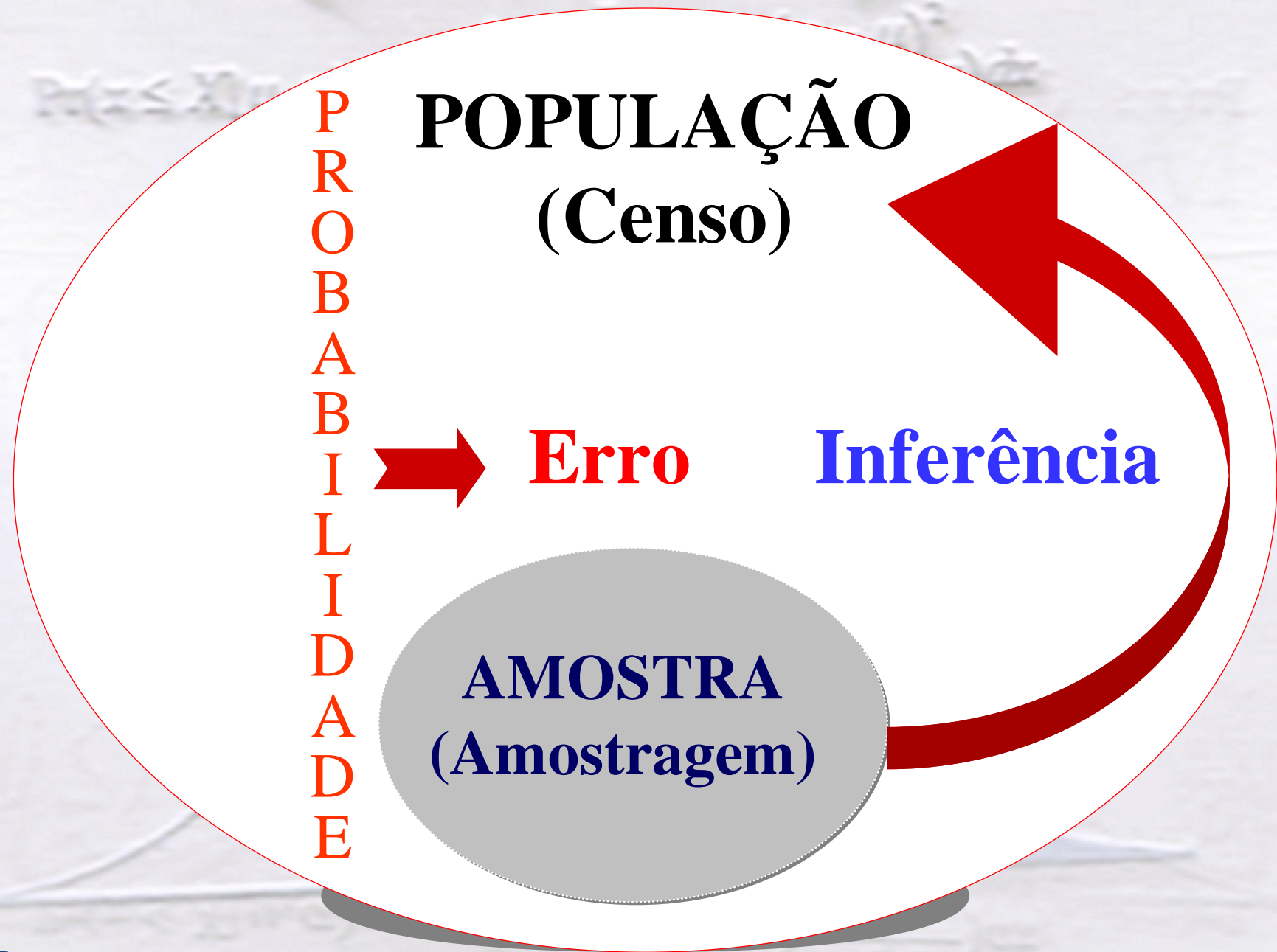
Método de se inferir sobre  
uma população a partir do  
conhecimento de pelo menos  
uma amostra dessa população.





Estudo das relações teóricas  
existentes entre uma população e  
as amostras dela extraídas.







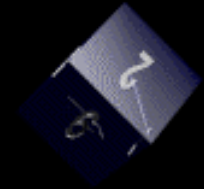
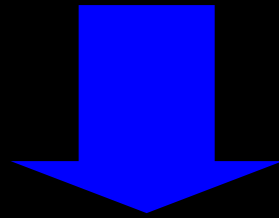
**A  
m  
o  
s  
t  
r  
a  
g  
e  
m**  
**T  
i  
p  
o  
s  
d  
e**

**Probabilística**

**Não Probabilística**



# Amostragem Probabilística



**Todos os elementos da população têm probabilidade conhecida (e diferente de zero) de fazer parte da amostra.**

# Métodos de Amostragem Probabilística

- Aleatória Simples
- Sistemática
- Estratificada
- Por Conglomerados



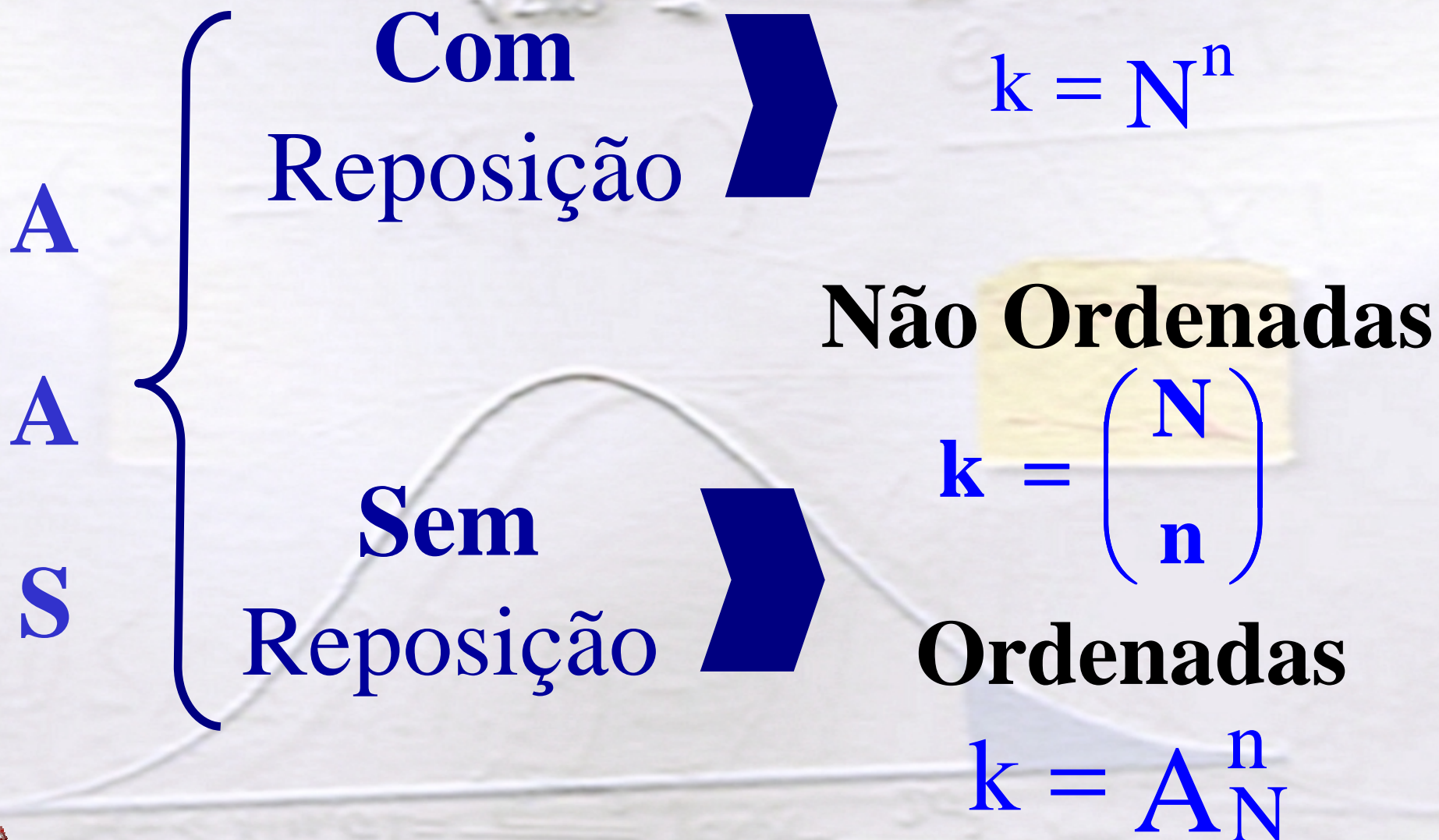


# Amostragem Ao Acaso (aa) ou Aleatória Simples (aas)

Uma amostra é dita “aleatória simples” ou “ao acaso” se todos os elementos da população tiverem a **mesma** probabilidade de pertencer a amostra



# Total de Amostras





# Estimador, Estimativa e Parâmetro

Uma característica da população é denominada de parâmetro.

Um estimador é uma característica da amostra.

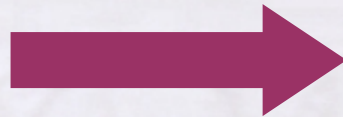
Uma estimativa é um valor particular de um estimador.





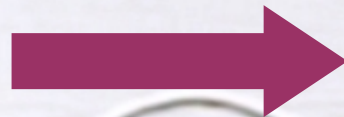
# Principais Parâmetros

$\mu$



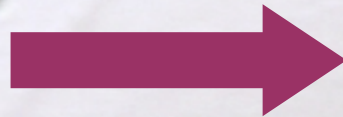
**A MÉDIA**

$\sigma^2$



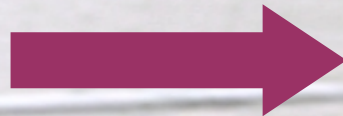
**A VARIÂNCIA**

$\sigma$



**O DESVIO  
PADRÃO**

$\pi$



**A PROPORÇÃO**



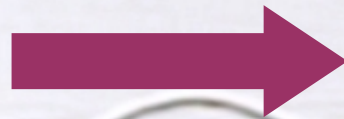
# Principais Estimadores

$\bar{X}$



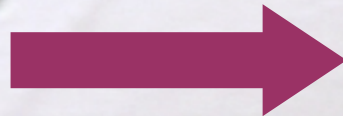
**A MÉDIA**

$S^2$



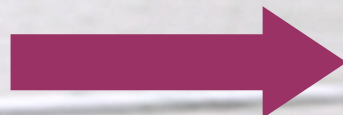
**A VARIÂNCIA**

$S$



**O DESVIO  
PADRÃO**

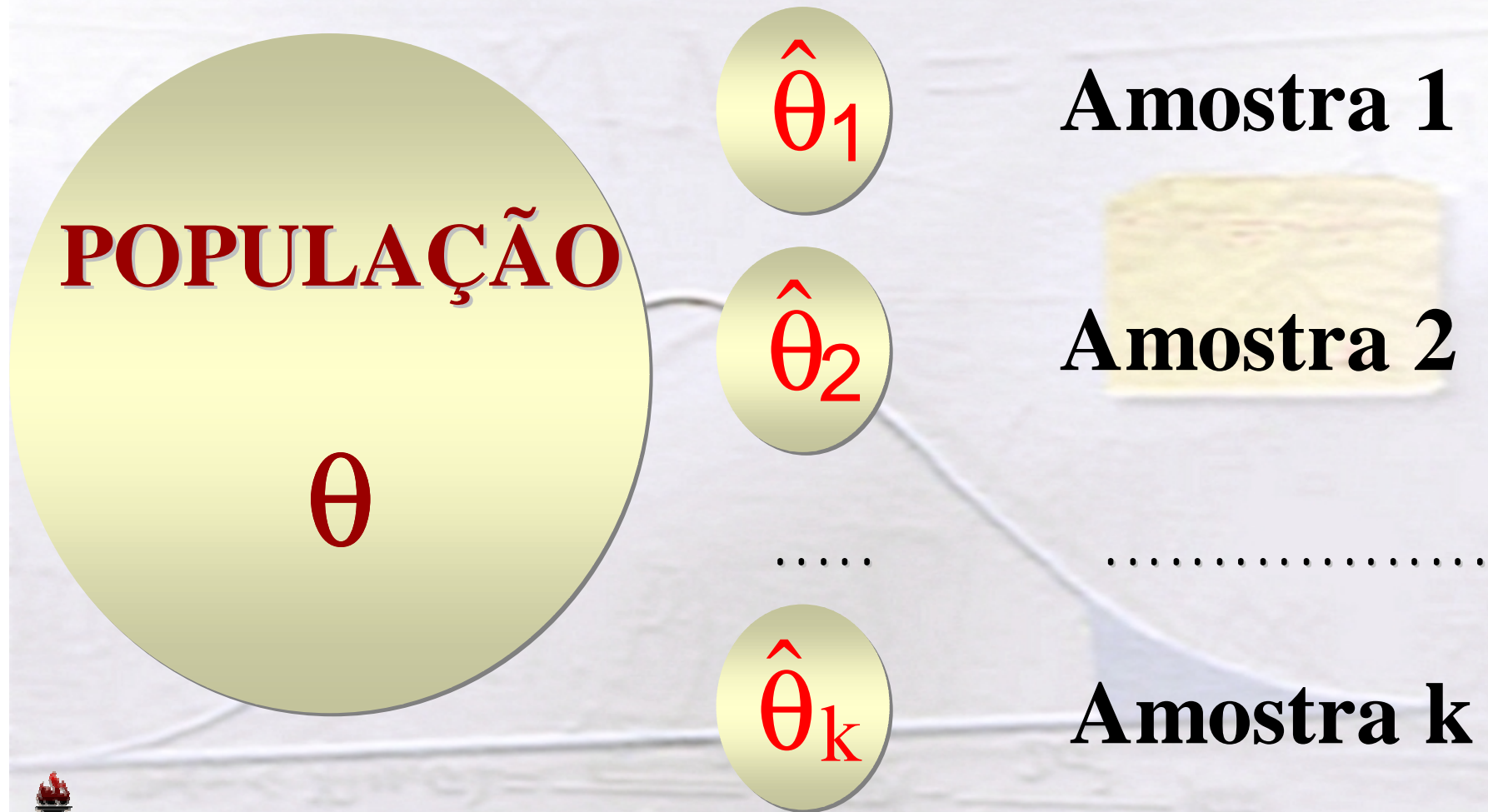
$P$



**A PROPORÇÃO**



# Distribuições Amostrais





# Distribuições Amostrais

A distribuição de probabilidade de um estimador (variável aleatória) é denominada de distribuição amostral desse estimador.



# Distribuição Amostral da Média

## Características

**Média** →

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

**Erro  
padrão** →

**COM  
Reposição**

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

**SEM  
Reposição**

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$





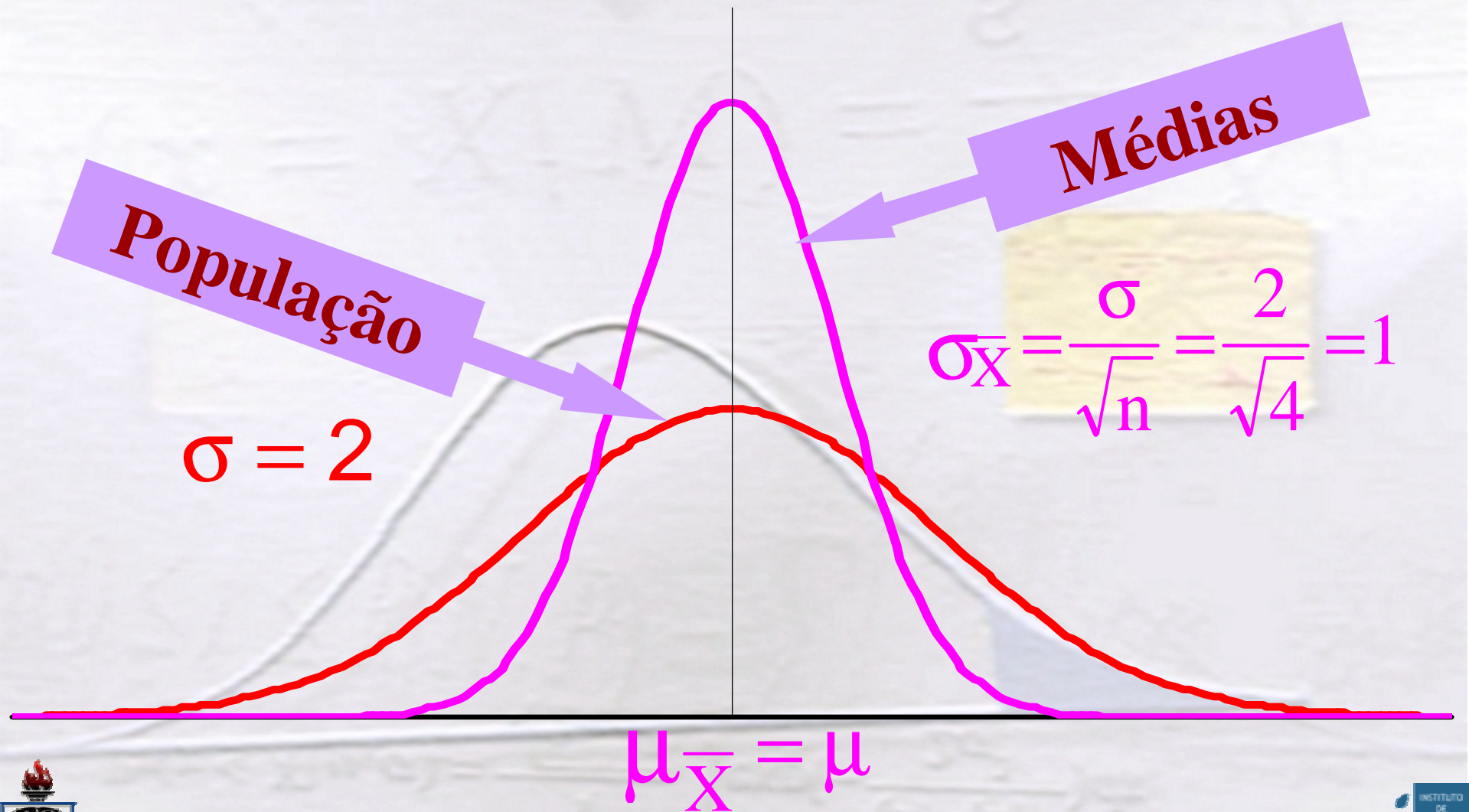
# Forma da Distribuição Amostral da Média

Se uma amostra aleatória de tamanho “n” for retirada de uma população  $X$  com uma distribuição  $N(\mu; \sigma)$ , então a distribuição de  $\bar{X}$ , média da amostra, tem uma distribuição  $N(\mu, \frac{\sigma}{\sqrt{n}})$





# Distribuição Amostral da Média

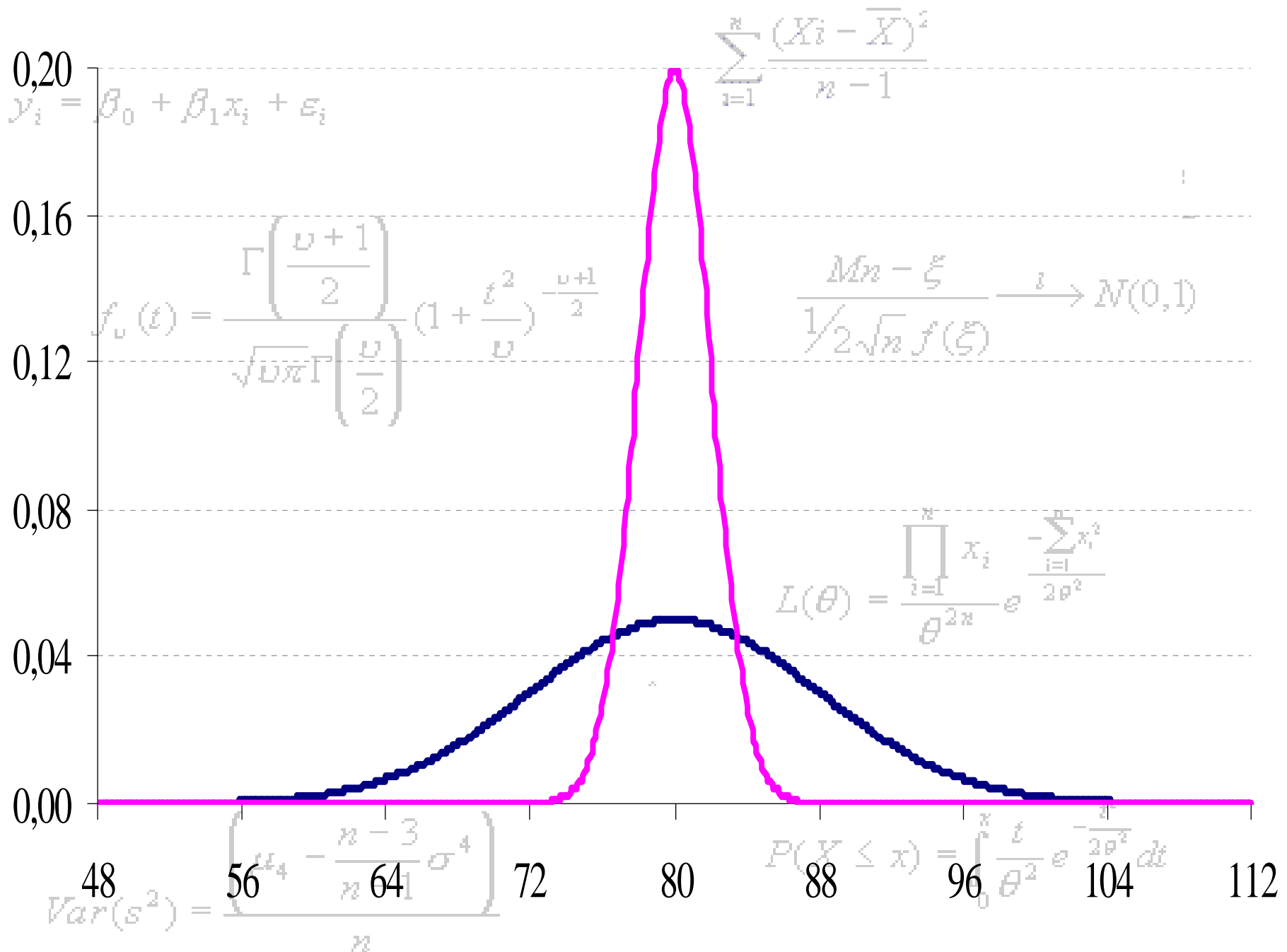


# Exemplo:

Uma amostra de  $n = 16$  elementos é retirada de uma população  $N(80; 8)$ . Determine:

$$P(76 < \bar{X} < 85)$$







# Solução:

Tem-se:  $\mu = 80$ ,  $\sigma = 8$

Sabe-se que:

$$\mu_{\bar{X}} = 80 \text{ e}$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{8}{\sqrt{16}} = 2$$



$$P(x \leq X | \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx$$

$$P(76 < \bar{X} < 85) =$$

$$= \text{DIST.NORM}(85; 80; 8; 1) -$$

$$- \text{DIST.NORM}(76; 80; 8; 1) =$$

$$= 73,40\% - 30,85 = 42,55\%$$





# Forma da Distribuição Amostral da Média

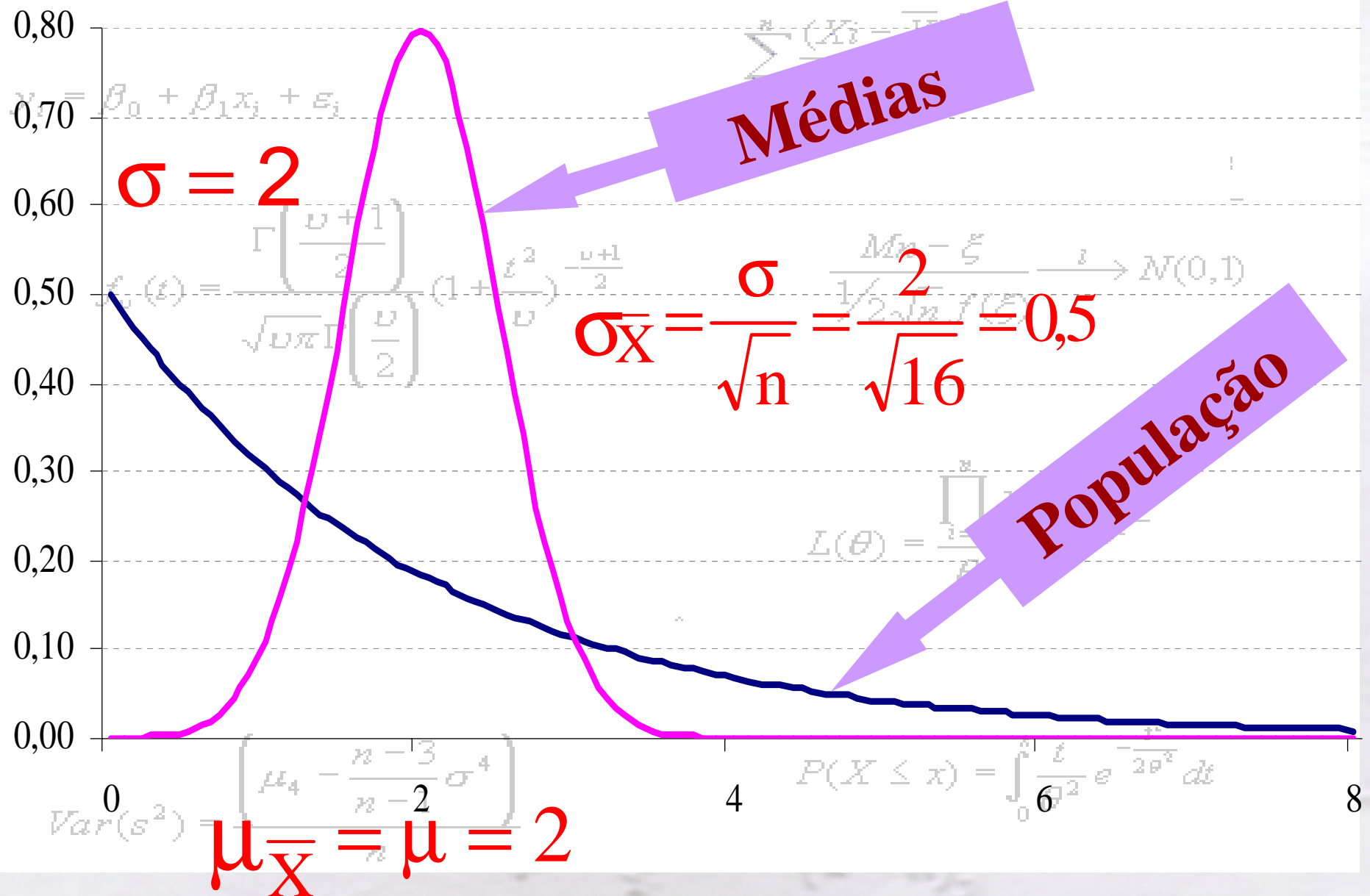
Se uma amostra aleatória de tamanho “ **$n > 30$** ” for retirada de uma população com **qualquer** distribuição de média  $\mu$  e desvio padrão  $\sigma$ , então a distribuição de  $\bar{X}$ , média da amostra, tem uma distribuição aproximadamente

$$N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$





# Distribuição Amostral da Média



# Exemplo:

Uma amostra de “n” elementos é retirada de uma população  $N(80; 4)$ .  
Determine “n” de forma que:

$$P(\bar{X} < 79) = 1,50 \%$$



# Solução:

Tem-se:  $\mu = 80$ ,  $\sigma = 4$

Sabe-se que:

$$\mu_{\bar{X}} = 80 \text{ e}$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{4}{\sqrt{n}}$$





# Então:

$$P(\bar{X} < 79) =$$

$$= P\left(\frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} < \frac{79 - 80}{\frac{4}{\sqrt{n}}}\right) =$$

$$= P\left(Z < -\frac{\sqrt{n}}{4}\right) = \Phi\left(-\frac{\sqrt{n}}{4}\right) = 1,50 \%$$



$$= \text{INV.NORMP}(1,5\%) = -2,17$$

Assim:

$$-\frac{\sqrt{n}}{4} = -2,17$$

$$\sqrt{n} = 2,17 \cdot 4 = 8,68$$

$$n \geq (8,68)^2 \cong 76$$

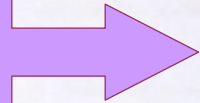




# Distribuição Amostral da Proporção

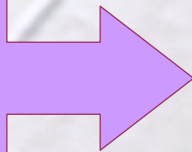
## Características

**Média**



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

**Erro  
padrão**



**COM  
Reposição**

$$\sigma_P = \sqrt{\frac{\pi(1-\pi)}{n}}$$

**SEM  
Reposição**

$$\sigma_P = \sqrt{\frac{\pi(1-\pi)}{n}} \sqrt{\frac{N-n}{N-1}}$$

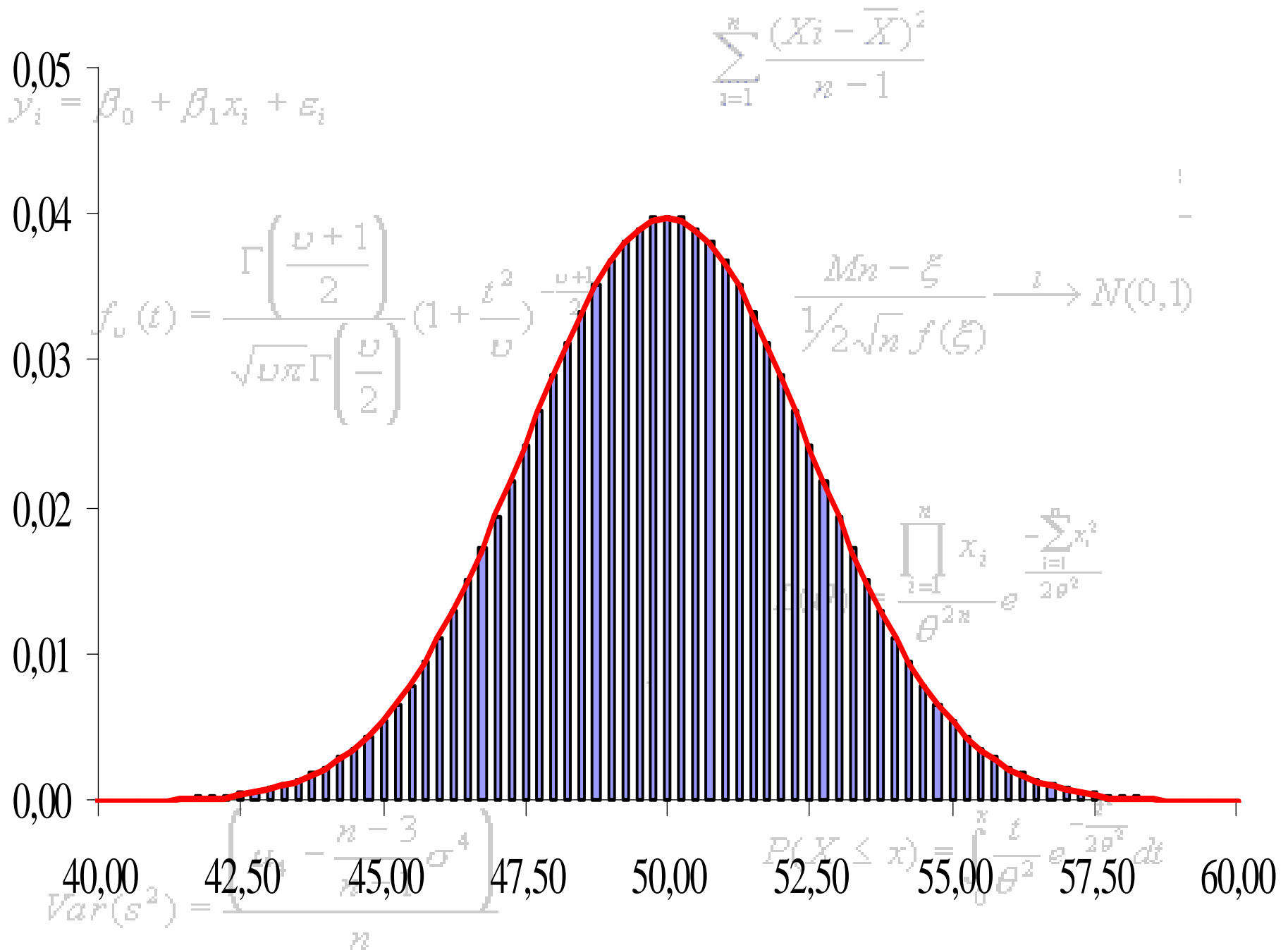




# Forma da Distribuição Amostral da Proporção

Se uma amostra aleatória de tamanho “ **$n > 100$** ” for retirada de uma população com proporção  **$\pi$** , então a distribuição de  **$P$** , **proporção na amostra**, tem uma distribuição aproximadamente  $N(\pi, \sqrt{\frac{\pi(1 - \pi)}{n}})$





# Exemplo:

Uma amostra de  $n = 400$  eleitores é retirada da população que prefere o candidato Zigoto com  $\pi = 50\%$   
Determine:

$$P(47\% < P < 54\%)$$





# Solução:

Tem-se:  $\pi = 50\%$

Sabe-se que:  $\mu_p = \pi = 50\%$

$$\begin{aligned}\sigma_p &= \sqrt{\frac{\pi(1-\pi)}{n}} = \\ &= \sqrt{\frac{0,45(1-0,45)}{400}} = \\ &= 0,025 = 2,50\%\end{aligned}$$



# Então:

$$P(47 < P < 54) =$$

$$= \text{DIST.NORM}(54; 50; 2,5; 1) -$$

$$- \text{DIST.NORM}(47; 50; 2,5; 1) =$$

$$= 94,52\% - 11,51 = 83,01\%$$





# Distribuição Amostral da Variância

## Características

Amostragem **com** reposição

**Média**

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

**Erro  
padrão**

$$\sigma_{S^2} = \sqrt{\frac{2\sigma^4}{n-1}} = \sigma^2 \sqrt{\frac{2}{n-1}}$$





# Forma da Distribuição Amostral da Variância

Se uma amostra aleatória de tamanho “**n**” (grande) for retirada de uma população com variância  $\sigma^2$ , então a distribuição de  **$S^2$ , variância da amostra**, tem uma distribuição aproximadamente  $\chi^2$  com “n-1” g.l., a menos de uma constante.



# Distribuição Amostral da Variância

**Isto é:**

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

**Este resultado é conhecido  
como Teorema de Fisher**





# Exemplo:

Uma amostra de  $n = 81$  elementos é retirada de uma população com variância  $\sigma^2 = 10$ . Determine a probabilidade de que  $P(S^2 > 15)$ .





# Solução:

Tem-se:

$$n = 81$$

$$\sigma^2 = 10$$

Sabe-se que:

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



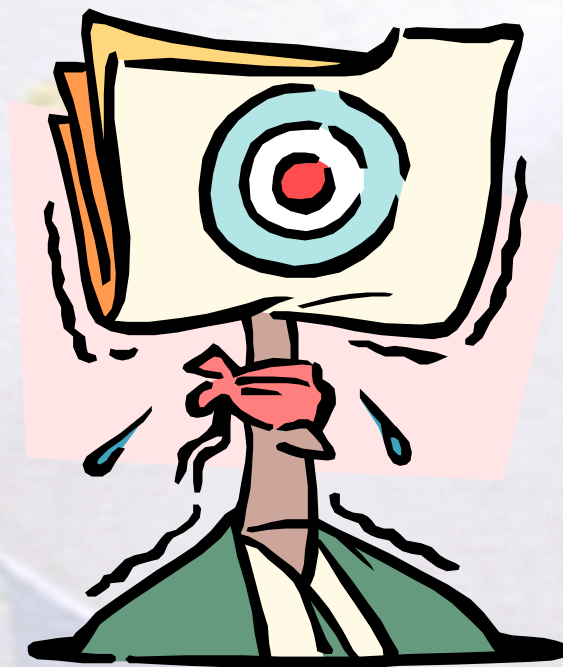
# Então:

$$\begin{aligned}P(S^2 > 15) &= P\left[\frac{\sigma^2}{(n-1)}\chi_{n-1}^2 > 15\right] = \\&= P\left[\chi_{n-1}^2 > \frac{15 \cdot (n-1)}{\sigma^2}\right] = \\&= P\left(\chi_{80}^2 > \frac{15 \cdot 80}{10}\right) = P\left(\chi_{80}^2 > \frac{15 \cdot 80}{10}\right) = \\P(\chi_{80}^2 > 120) &= \text{DIST.QUI}(120; 80) = 0,25\%\end{aligned}$$





# Estiminação





# Estimação

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



**ESTIMAÇÃO**

$\theta$

**AMOSTRA**

**POPULAÇÃO**

$\hat{\theta}$



# Tipos de Estimacao

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.



$$P(x \leq X | \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx$$

# Estimação por Ponto





As características básicas de um estimador são:

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

A Variância:  $\sigma_{\hat{\theta}}^2 = V(\hat{\theta}) =$   
 $= E[\hat{\theta} - E(\hat{\theta})]^2 =$   
 $= E(\hat{\theta}^2) - E(\hat{\theta})^2$





$$P(x \leq X | \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx$$

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  $\theta$ .



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  $\theta$ .





$$P(x \leq X | \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx$$

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





# OUTROS CONCEITOS IMPORTANTES

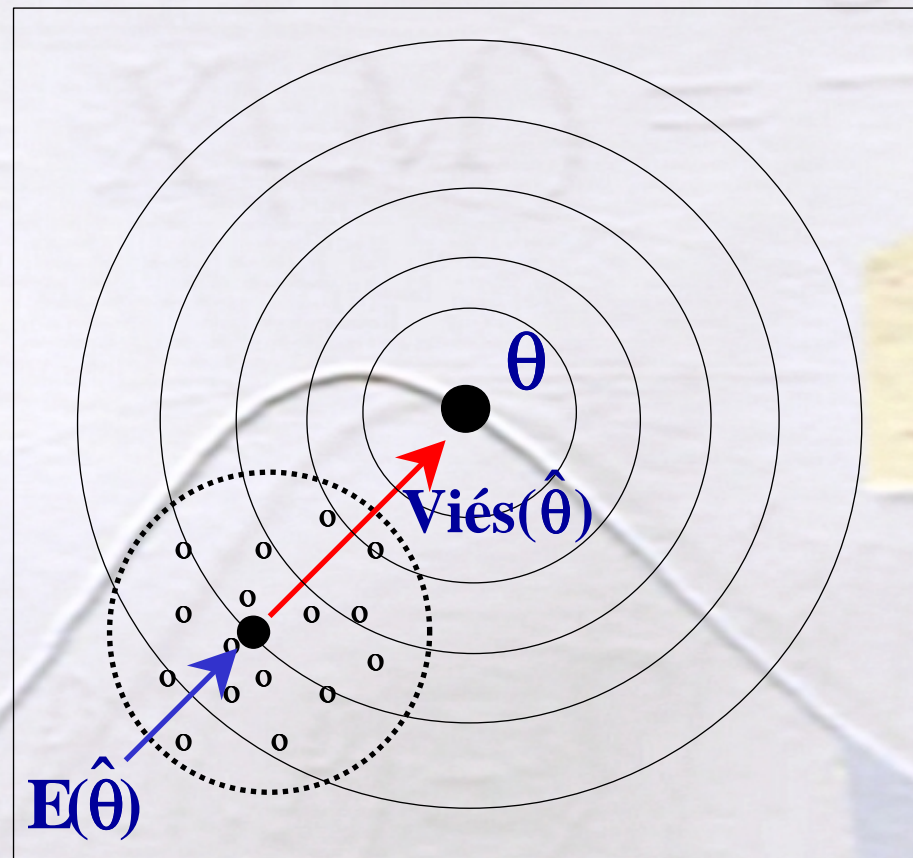
Viés:  $B(\hat{\theta}) = E(\hat{\theta}) - \theta$

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

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



# Erro Quadrado Médio





# Propriedades dos Estimadores





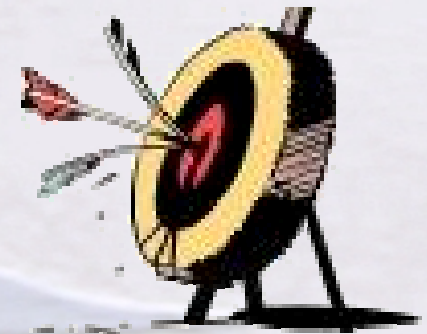
Seja  $(X_1, X_2, \dots, X_n)$  uma amostra aleatória de uma variável (população)  $X$ , com um parâmetro de interesse  $\theta$ . Seja  $\hat{\theta}$  uma função da amostra (estimativa de  $\theta$ ).



# Não Tendenciosidade

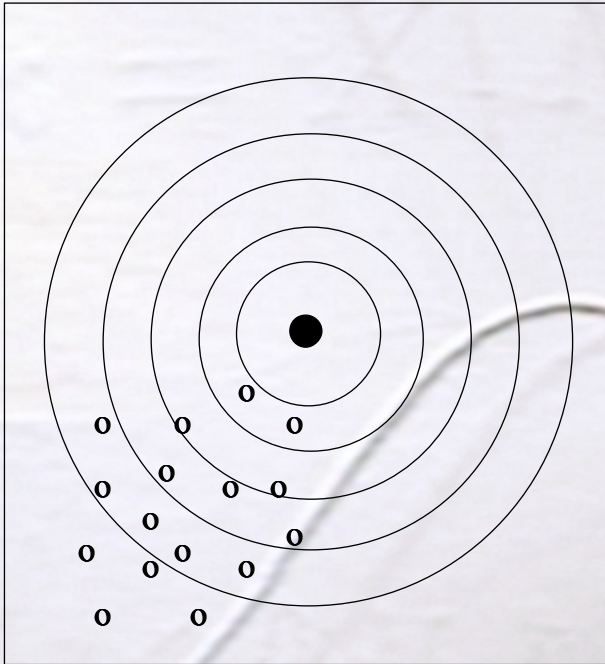
Um estimador é dito **não-tendencioso**, **não-viciado**, **sem viés** ou **imparcial** se:

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

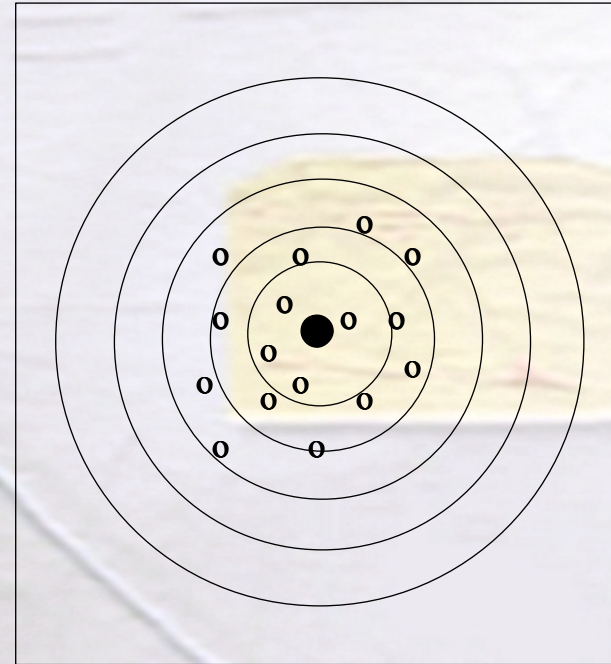




# Tendenciosidade



**Tendencioso**



**Não tendencioso**





$$Pr(x \leq X | \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx$$

# 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.



# Exemplos





A média da amostra  $\bar{X}$  é um estimador não-viciado de  $\mu$ , isto é:

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





A proporção amostral  
“P” é um estimador não-  
viciado de  $\pi$ , isto é:

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



A variância da amostra  
“ $S^2$ ” é um estimador viciado  
de  $\sigma^2$ , 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  $\sigma^2$ ,  
isto é:

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





# Consistência

Um estimador não  
viciado é dito consistente  
se:

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



# Exemplos





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

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





A proporção amostral  
“P” é um estimador  
consistente de  $\pi$ , isto é:

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



A variância da amostra  
“ $S^2$ ” é um estimador  
consistente de  $\sigma^2$ , isto é:

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





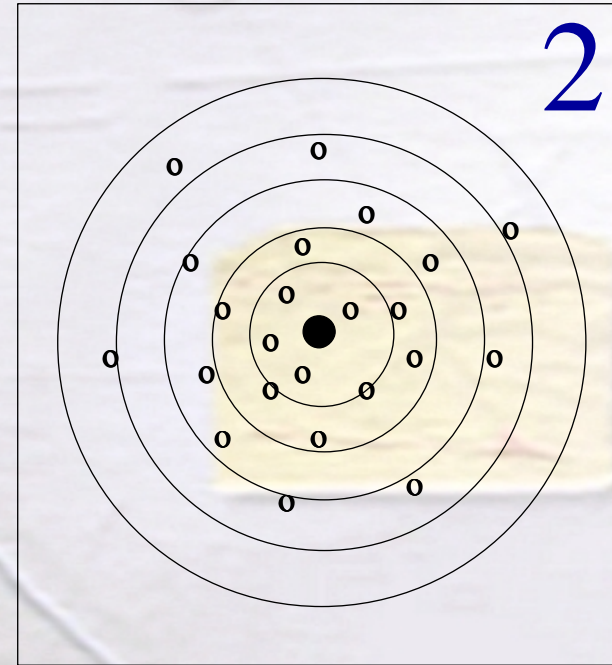
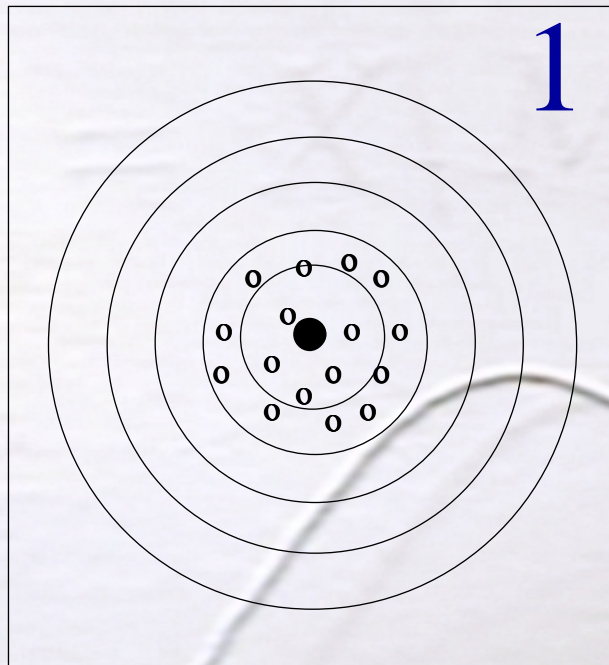
# Eficiência

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  $\mu$ , do que **qualquer** média ponderada.





# Exercício um



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**





**(d) Estimativas dos erros padrão da:**

**(i) Média**

**(ii) Proporção**

**(iii) Variância**





Velocidades			Frequência
80	————	85	8
85	————	90	13
90	————	95	24
95	————	100	33
100	————	105	29
105	————	110	13
Total			120



Velocidades	Frequência	$x_i$	$f_i x_i$
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	1397,5
<b>Total</b>	<b>120</b>	<b>—</b>	<b>11605</b>



# 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	Frequência	$x_i$	$f_i x_i^2$
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
<b>Total</b>	<b>120</b>	<b>—</b>	<b>1127950</b>

# O desvio padrão

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

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

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





# A proporção

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

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

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





O erro padrão da média é dado por:

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

Então uma estimativa do erro é

dada por:

$$\hat{\sigma}_{\bar{X}} = \frac{\hat{\sigma}}{\sqrt{n}} = \frac{s}{\sqrt{n}} =$$

$$\sqrt{\frac{47,4772}{120}} = 0,63 \text{ km/h}$$



O erro padrão da proporção é dado por:

$$\sigma_P = \sqrt{\frac{\pi.(1 - \pi)}{n}}$$

Então uma estimativa do erro é dada por:

$$\hat{\sigma}_P = \sqrt{\frac{\hat{\pi}.(1 - \hat{\pi})}{n}} = \sqrt{\frac{p.(1 - p)}{n}} =$$

$$\sqrt{\frac{0,35.0,65}{120}} = 0,0435 = 4,35 \%$$





O erro padrão da variância é dado por:

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

Então uma estimativa do erro é dada por:

$$\begin{aligned}\hat{\sigma}_{S^2} &= \hat{\sigma}^2 \sqrt{\frac{2}{n-1}} = s^2 \sqrt{\frac{2}{n-1}} = \\ &= 47,4772 \sqrt{\frac{2}{120-1}} = 6,15 \text{ (km/h)}^2\end{aligned}$$





$$Pr(x \leq X | \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx$$

# Estimação por Intervalo



$$P(x \leq X | \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx$$

(A)

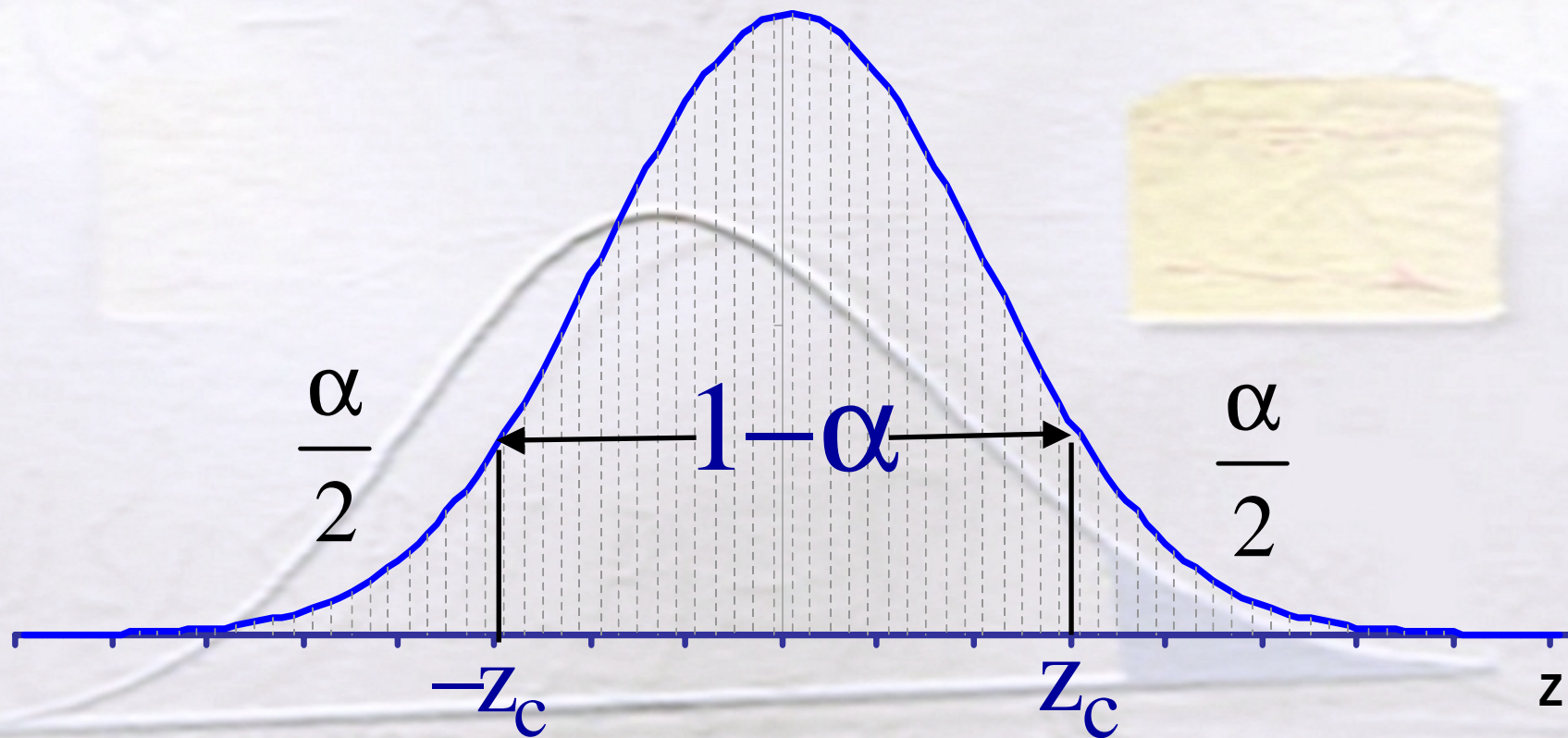
Da Média





# Supondo $\sigma$ conhecido

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





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

**Tem-se:**

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

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

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



**Assim:**

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

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

Então, o IC de “ $1 - \alpha$ ” para  $\mu$  é calculado por:

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





# 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 seis km/h determine uma estimativa para a velocidade média, com uma confiabilidade de 95%.



**Tem-se:**

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

**Mas:**  $z_c = 1,96$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{6}{\sqrt{120}} = 0,5477$$

$$\varepsilon_{\bar{X}} = 1,96 \cdot 0,5477 = 1,07$$





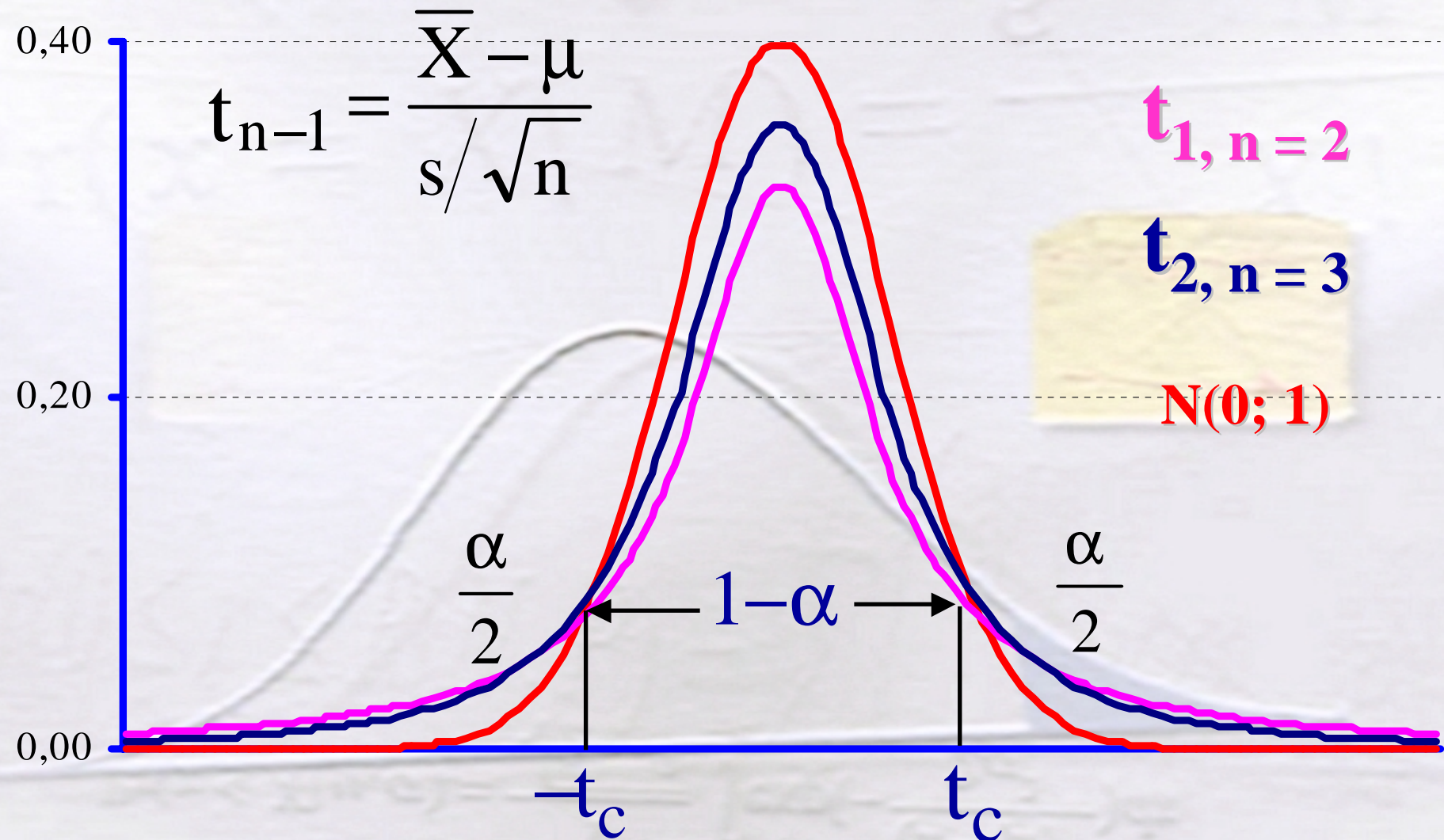
O IC de “ $1 - \alpha$ ” para  $\mu$  é calculado por:

$$\begin{aligned} & [\bar{X} - \varepsilon_{\bar{X}}; \bar{X} + \varepsilon_{\bar{X}}] \\ & [96,71 - 1,07; 96,71 + 1,07] \\ & [95,64; 97,78] \end{aligned}$$





# $\sigma$ desconhecido



**De**  $P(-t_c < t < t_c) = 1 - \alpha$   
**Tem-se:**

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

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

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





**Assim:**

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

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

Então, o IC de “ $1 - \alpha$ ” para  $\mu$ , se  $\sigma$  for desconhecido é calculado por:

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





# 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:**

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

**Mas:**  $t_c = 1,98$

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

$$\hat{\varepsilon}_{\bar{X}} = 1,98 \cdot 0,6290 = 1,25$$



O IC de “ $1 - \alpha$ ” para  $\mu$  é calculado por:

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

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

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





(B)

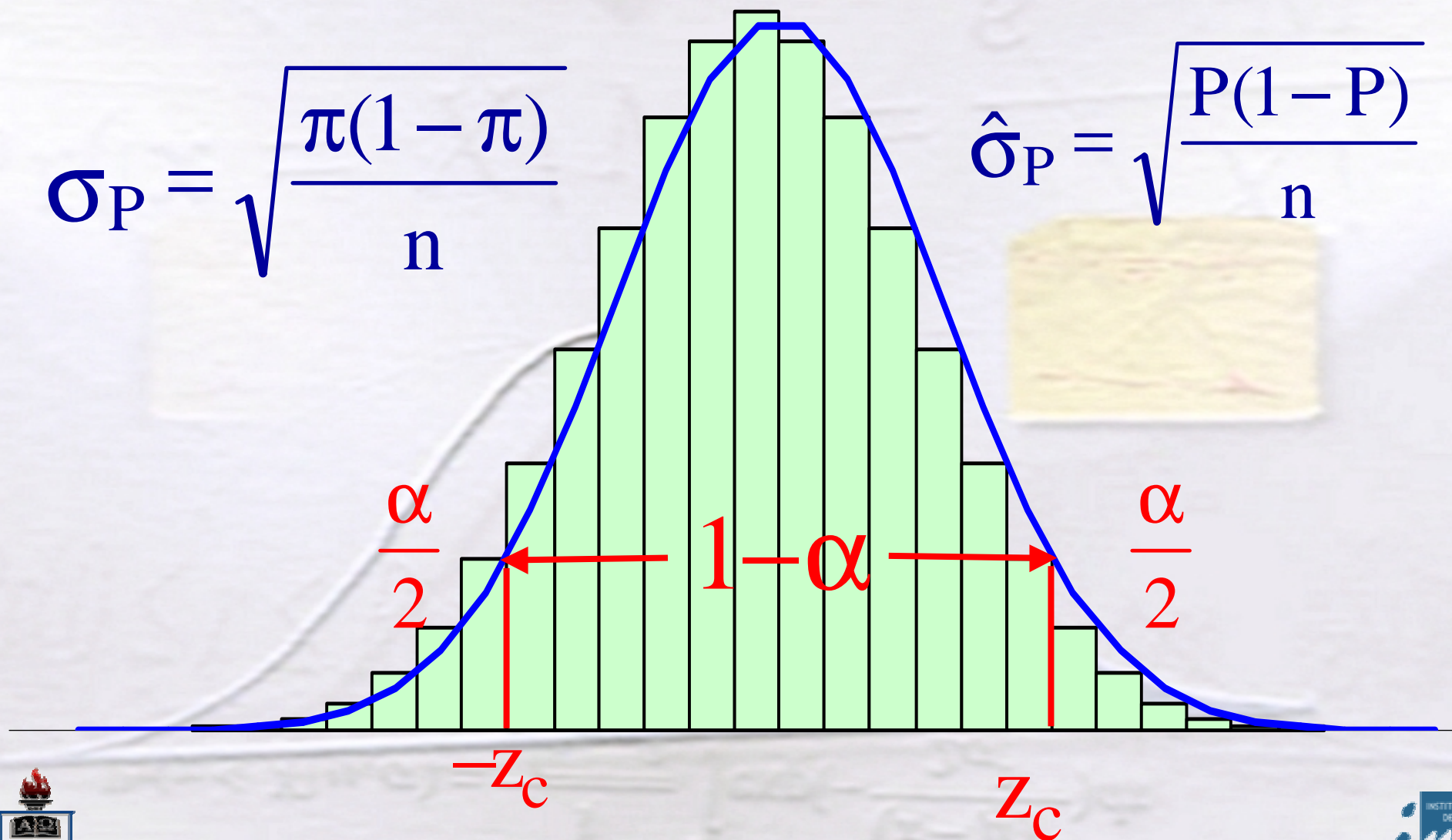
# Da Proporção



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

$$\sigma_P = \sqrt{\frac{\pi(1-\pi)}{n}}$$

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





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

**Tem-se:**

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

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

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



**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 “ $1 - \alpha$ ” para  $\pi$  é calculado por:

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





# 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:**

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

**Mas:**  $z_c = 1,96$

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

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



O IC de “ $1 - \alpha$ ” para  $\pi$  é calculado por:

$$[P - \hat{\epsilon}_P; P + \hat{\epsilon}_P]$$

$$[35\% - 8,53\%; 35\% + 8,53\%]$$

$$[26,47\%; 43,53\%]$$



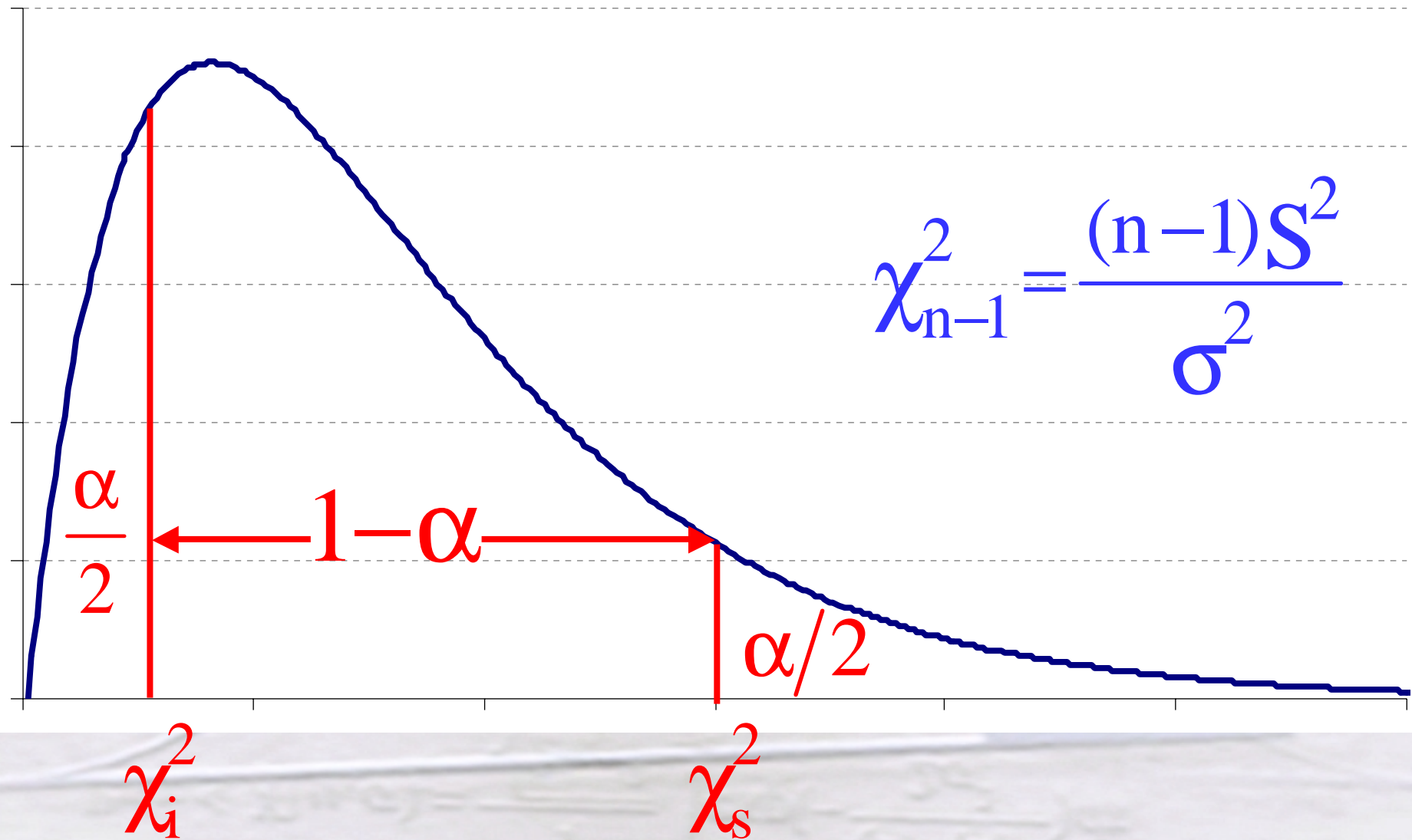


(C)

# Da Variância (Desvio Padrão)



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





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

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

$$P(\frac{1}{\chi_s^2} < \frac{\sigma^2}{(n-1)S^2} < \frac{1}{\chi_1^2}) = 1 - \alpha$$

$$P(\frac{(n-1)S^2}{\chi_s^2} < \sigma^2 < \frac{(n-1)S^2}{\chi_1^2}) = 1 - \alpha$$



Então o IC de “ $1 - \alpha$ ” para  $\sigma^2$  é calculado por:

$$\left[ \frac{(n - 1) S^2}{\chi_s^2}; \frac{(n - 1) S^2}{\chi_1^2} \right]$$





Então o IC de “ $1 - \alpha$ ” para  $\sigma$  é calculado por:

$$\left[ \sqrt{\frac{(n-1)S^2}{\chi_s^2}}; \sqrt{\frac{(n-1)S^2}{\chi_1^2}} \right]$$



# 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:**

$$\left[ \sqrt{\frac{(n-1)S^2}{\chi_s^2}}; \sqrt{\frac{(n-1)S^2}{\chi_1^2}} \right]$$

**Mas:**

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

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





O IC de “ $1 - \alpha$ ” para  $\sigma$  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  $(1 - \alpha)$  e pequena amplitude  $(\varepsilon)$  . 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**  $(\varepsilon)$  especificados.



Para estimar a média de uma população, supondo  $\sigma$  conhecido

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

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

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





Para estimar a média de uma população, com  $\sigma$  conhecido

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

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

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

$t_c$  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)

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

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

$$n \geq 1068$$



# Solução

(b)

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

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

$$n \geq 385$$

