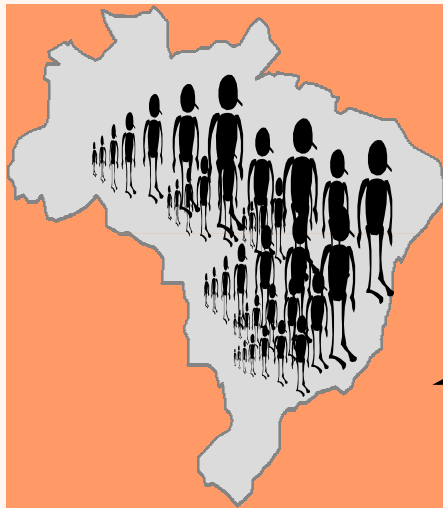


Parte IV

Inferência estatística

POPULAÇÃO:
eleitores brasileiros



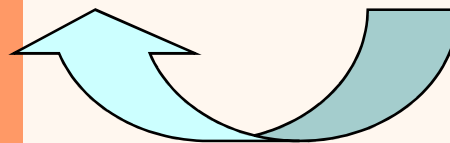
amostragem



AMOSTRA:
uma parte dos eleitores
brasileiros



inferência



- Como generalizar resultados de uma amostra para a população de onde ela foi extraída – *estimação de parâmetros* (Cap. 9).
- Como testar hipóteses com base em amostras – *testes de hipóteses* (Cap. 10)..

Estatística Aplicada às Ciências Sociais

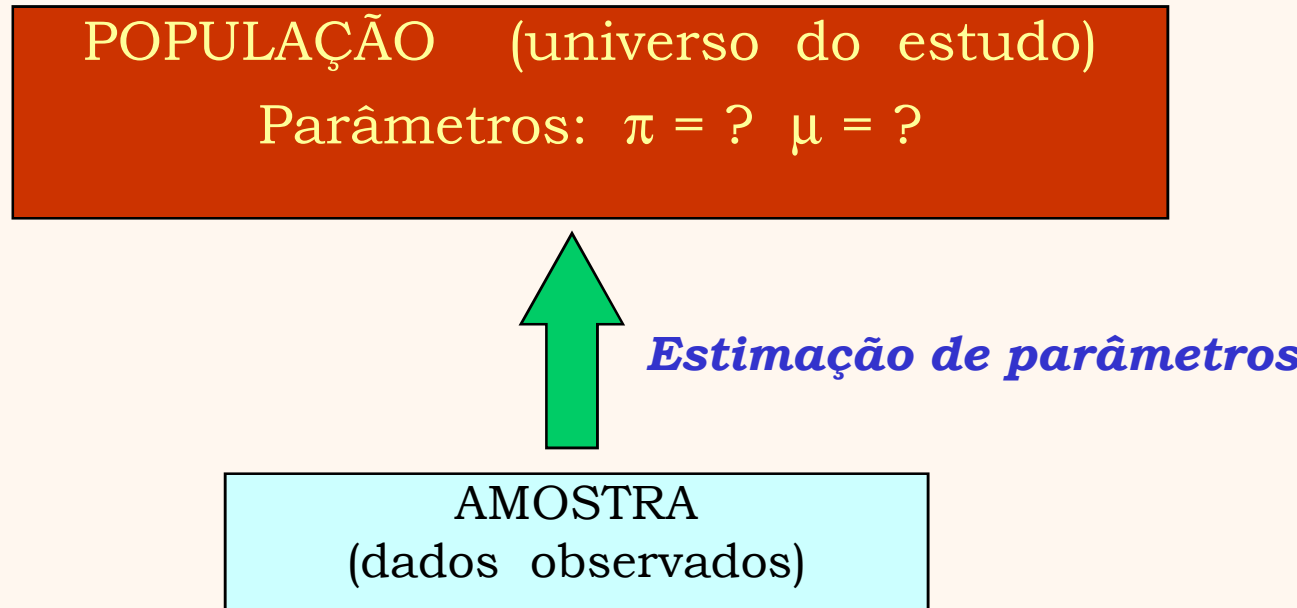
Sexta Edição

Pedro Alberto Barbetta

Florianópolis: Editora da UFSC, 2006

Cap. 9 – Estimação de parâmetros

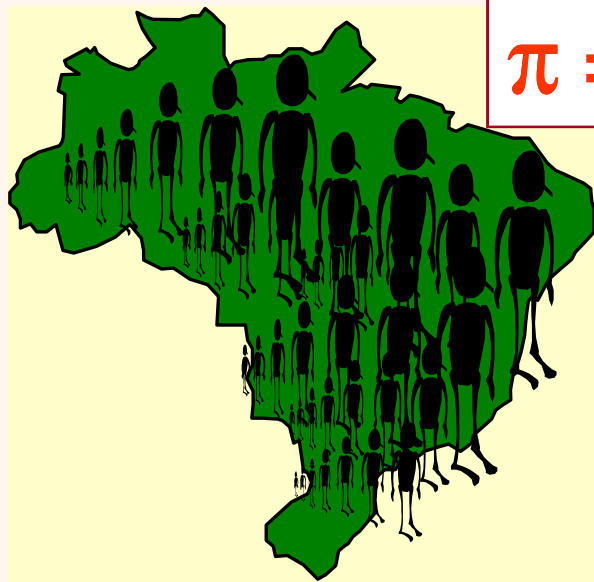
Estimação de Parâmetros



*O raciocínio **indutivo** da estimação de parâmetros*

Estimação de Parâmetros

POPULAÇÃO



$$\pi = ?$$

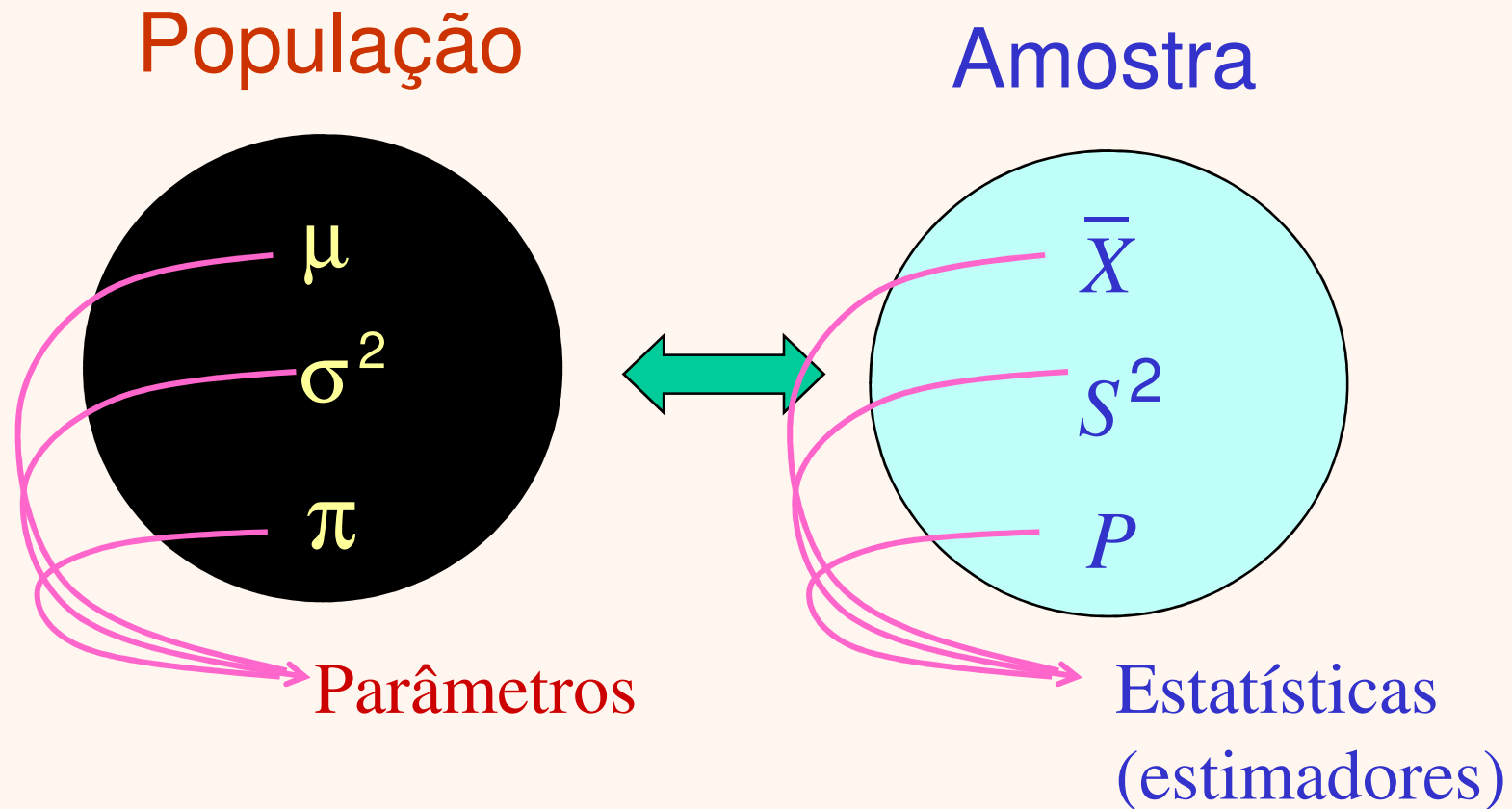
AMOSTRA



Observações: x_1 x_2 $x_3 \dots$ $\Rightarrow p$

$$\pi = p \pm \text{erro amostral}$$

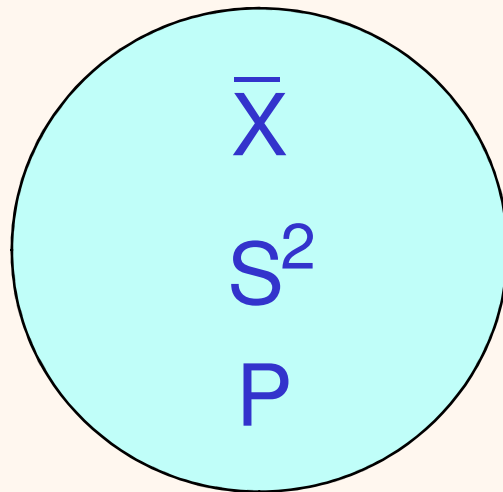
Estimação de Parâmetros



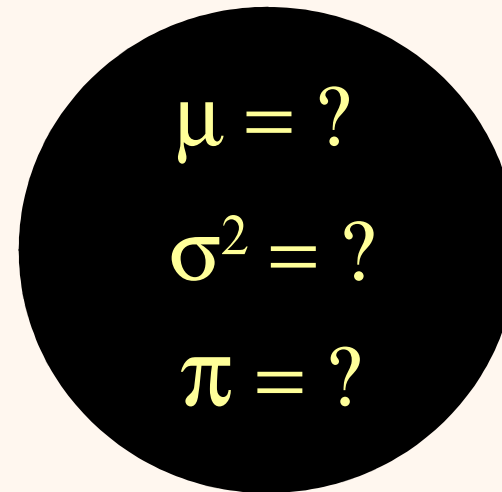
Estimação de Parâmetros. Objetivo:

- Com base em uma amostra, estimar os parâmetros populacionais.

Amostra



População

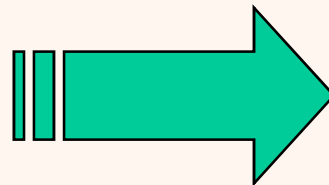
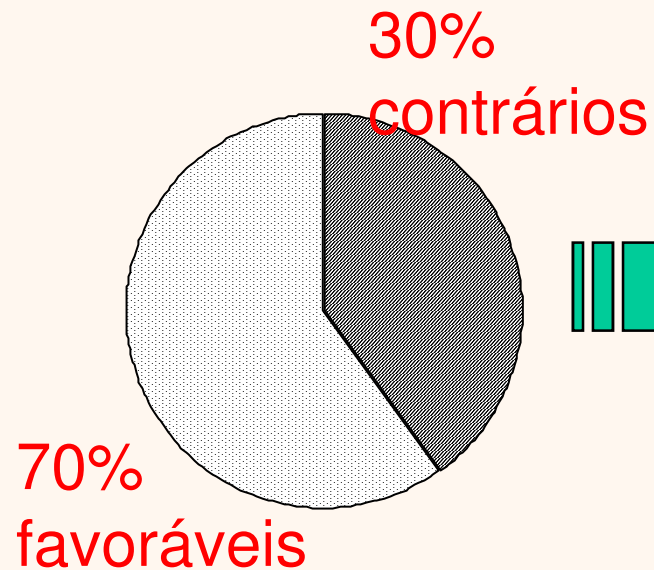


Estimação de uma proporção π

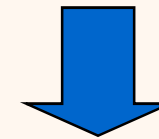
- Estimador: proporção amostral, P
- Relação entre o parâmetro π e a estatística P
→ base para calcular o erro amostral
- Considerar-se-á que a amostragem será (ou foi) aleatória simples.

Relação entre π e P. Uma ilustração

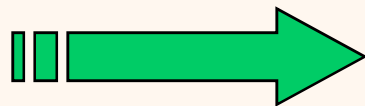
População



Amostra aleatória com
 $n = 400$ indivíduos

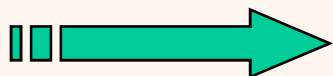
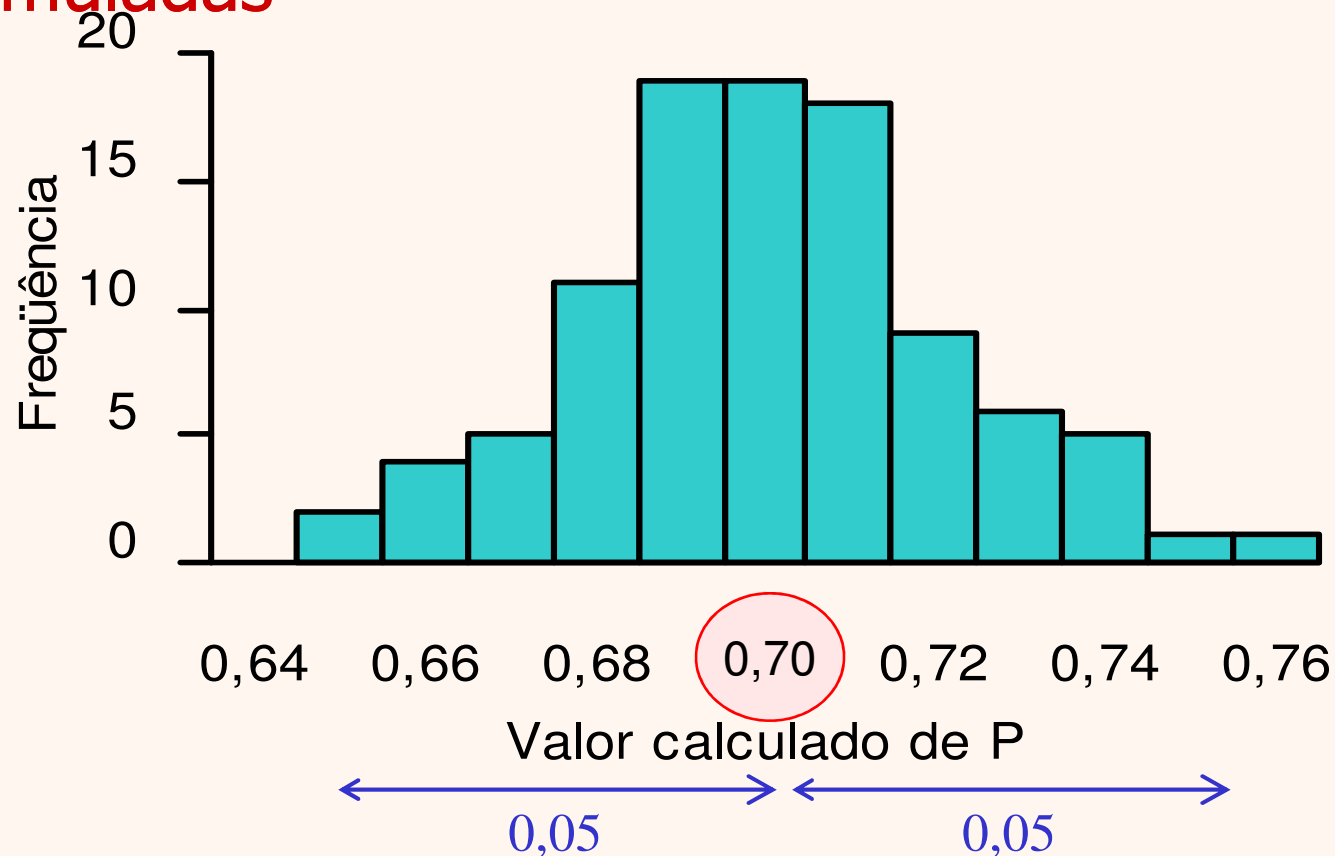


Calcula-se P

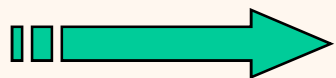


Simulou-se 100 amostras desta forma

Histograma dos resultados das 100 amostras simuladas



Em geral, erro amostral $< 0,05$



Em geral, o intervalo $P \pm 0,05$ contém π

Estimação de parâmetros

- Na prática, examinamos apenas *uma* amostra, resultando em um único valor para a estatística – uma **estimativa**. Porém, o conhecimento da *distribuição amostral* da estatística permite avaliarmos um limite superior para o erro amostral (*margem de erro*), com certo *nível de confiança*.
- A **distribuição amostral** de uma estatística é a distribuição dos possíveis valores dessa estatística, *se examinássemos todas* as possíveis amostras de tamanho n , extraídas aleatoriamente de uma população.

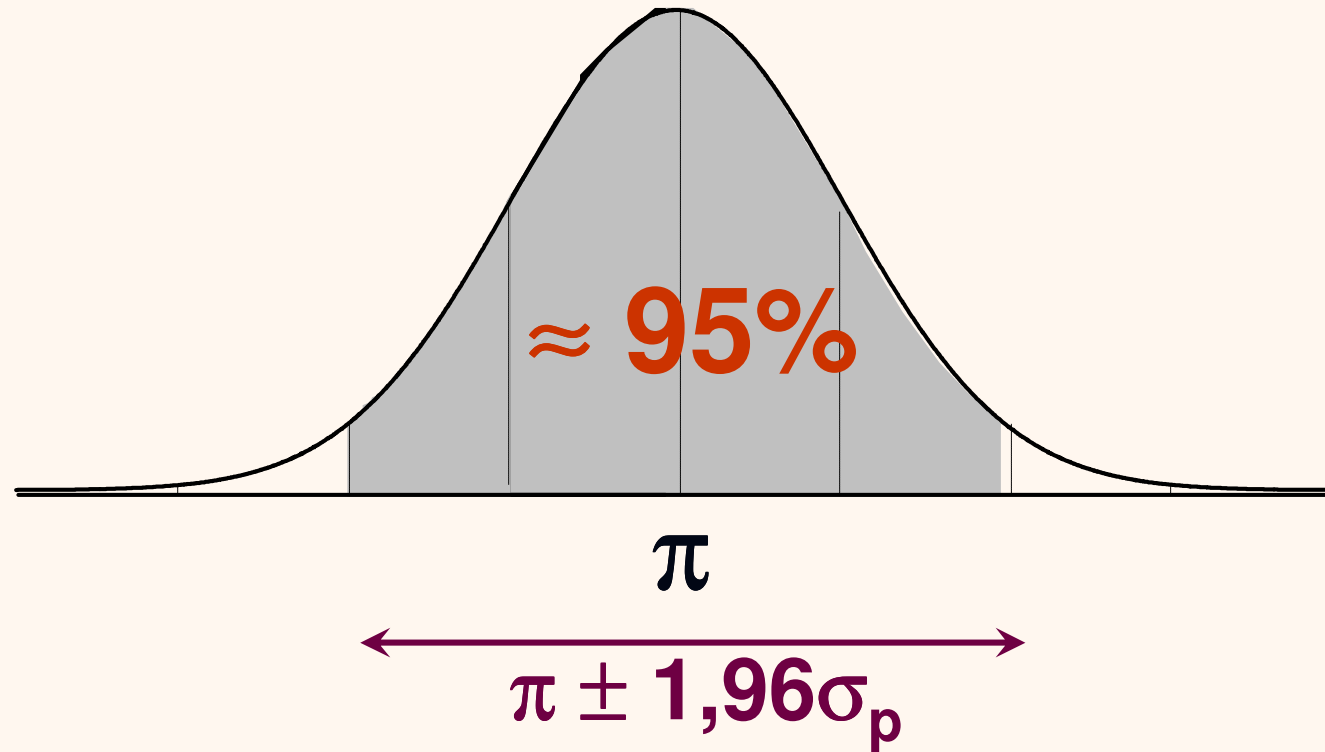
Distribuição amostral de uma proporção P

- Desde que a amostra seja aleatória e razoavelmente grande ($n > 30$, na maioria dos casos), tem-se:
 - Os possíveis valores de **P** seguem uma **distribuição** (aproximada) **normal** com média e desvio padrão dados por:

$$\mu_P = \pi$$

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

Distribuição de P



- O desvio padrão da distribuição amostral de uma estatística é comumente chamado de **erro padrão** da estatística.

Erro padrão da proporção

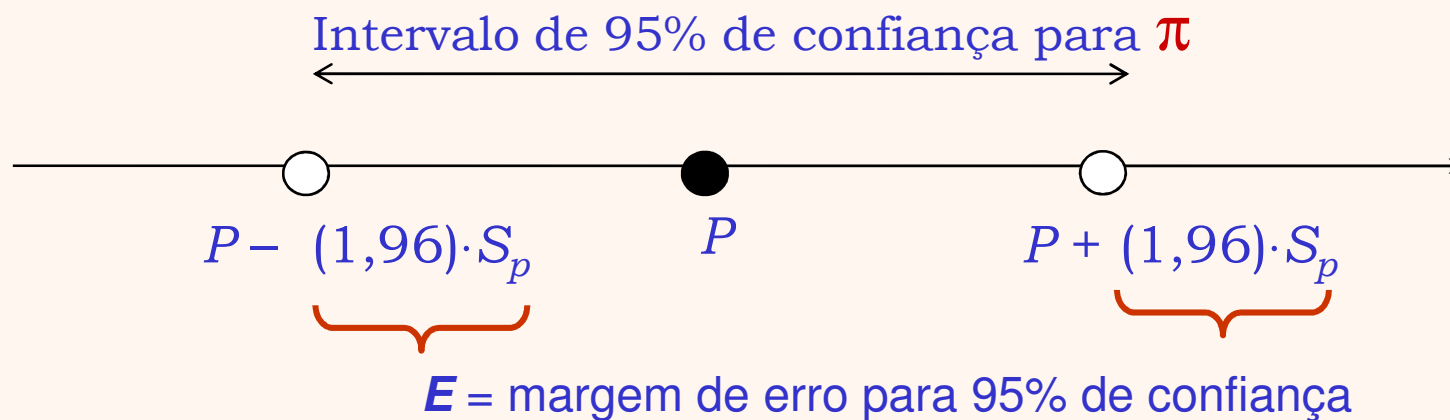
Na prática, estima-se o erro padrão da proporção por:

$$S_P = \sqrt{\frac{p \cdot (1-p)}{n}}$$

Se o tamanho da população, **N**, for conhecido e não muito grande ($N < 20n$), então estima-se o erro padrão da proporção por:

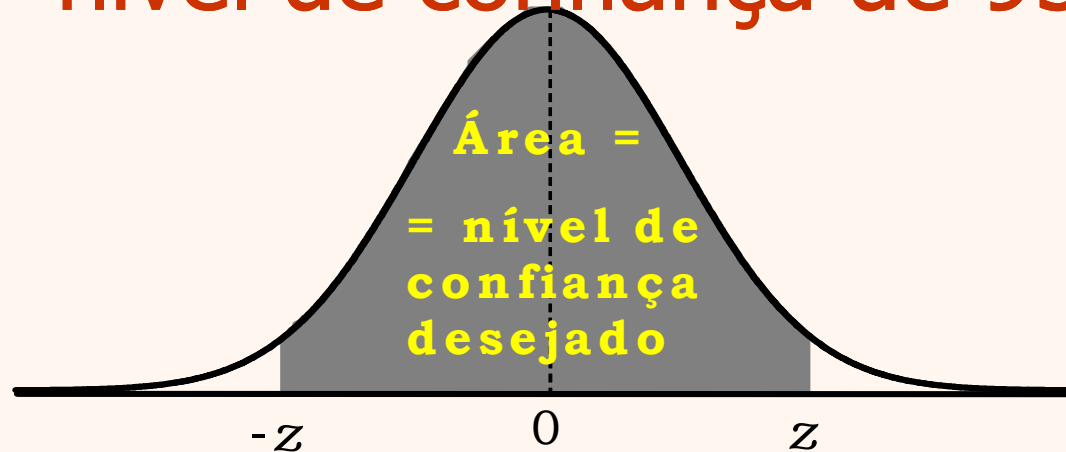
$$S_p = \sqrt{\frac{P \cdot (1-P)}{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

Intervalo de confiança para π , nível de confiança de 95%:



Ver Exemplo 9.1

Intervalo de confiança para π , nível de confiança de 95%:



Área	0,800	0,900	0,950	0,980	0,990	0,995	0,998
z	1,282	1,645	1,960	2,326	2,576	2,807	3,090

Margem de erro:

$$E = z \cdot S_P$$

Intervalo de confiança para π :

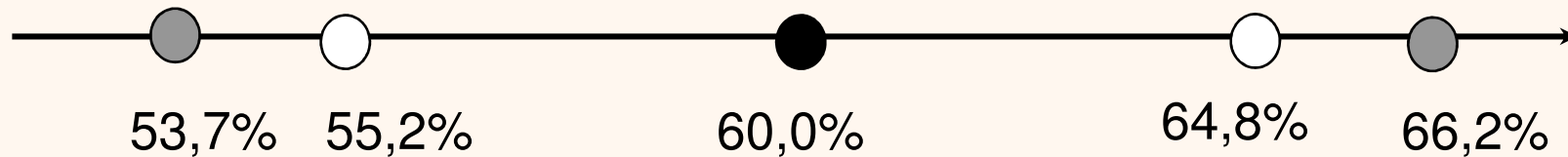
$$P \pm E$$

Resultados do Exemplo 9.1:

Intervalo de 99% de confiança para π
(60,0 \pm 6,3%)

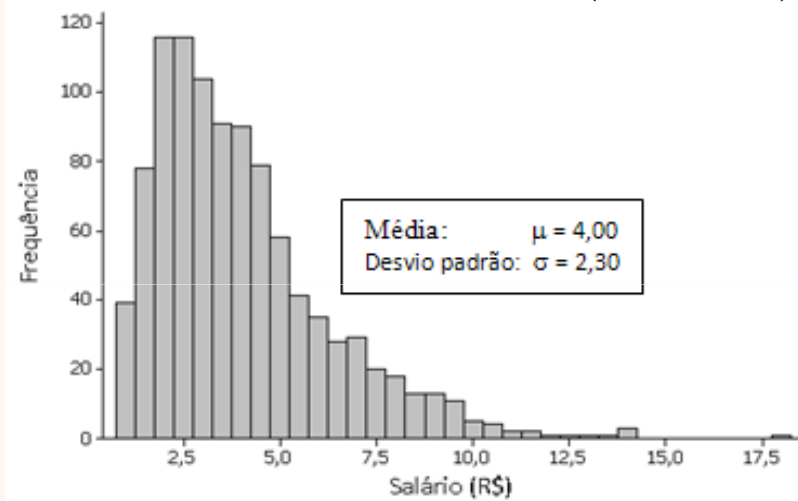


Intervalo de 95% de confiança para π
(60,0 \pm 4,8%)

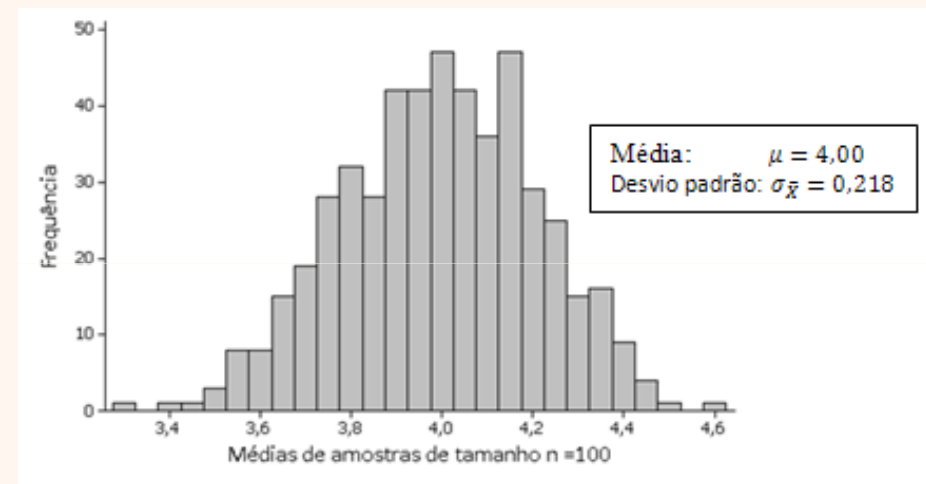


Distribuição da média amostral

População dos salários dos empregados de certo setor da economia (N = 1.000)



Distribuição de frequências de médias de amostras de tamanho n = 100



$$E(X) = \mu$$

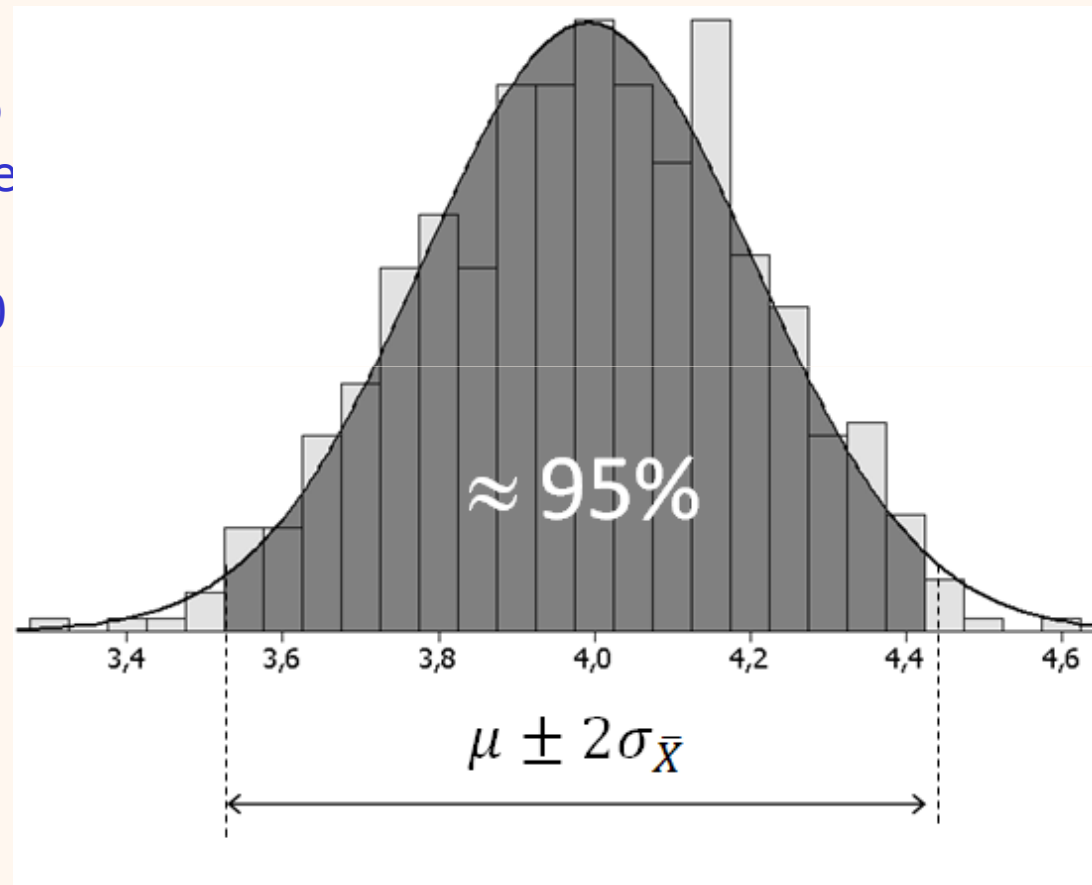
$$DP(X) = \sigma$$

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

$$DP(\bar{X}) = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

Distribuição da média amostral

Intervalo que abrange, aproximadamente, 95% das médias amostrais de tamanho $n = 100$ da população de $N = 1.000$ salários



Estimação de uma média μ

- Estimador: média amostral,

$$\bar{X} = \frac{\sum X}{n}$$

- Estimativa do erro padrão:

$$S_{\bar{X}} = \frac{S}{\sqrt{n}}$$

- Se o tamanho da população, N , for conhecido e não muito grande ($N < 20n$), então estima-se o erro padrão da média por:

$$S_{\bar{X}} = \frac{S}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

Margem de erro na estimação de uma média

- Se a amostra for grande ($n > 30$):

$$E = z \cdot S_{\bar{X}}$$

➤ onde z vem da distribuição normal padrão

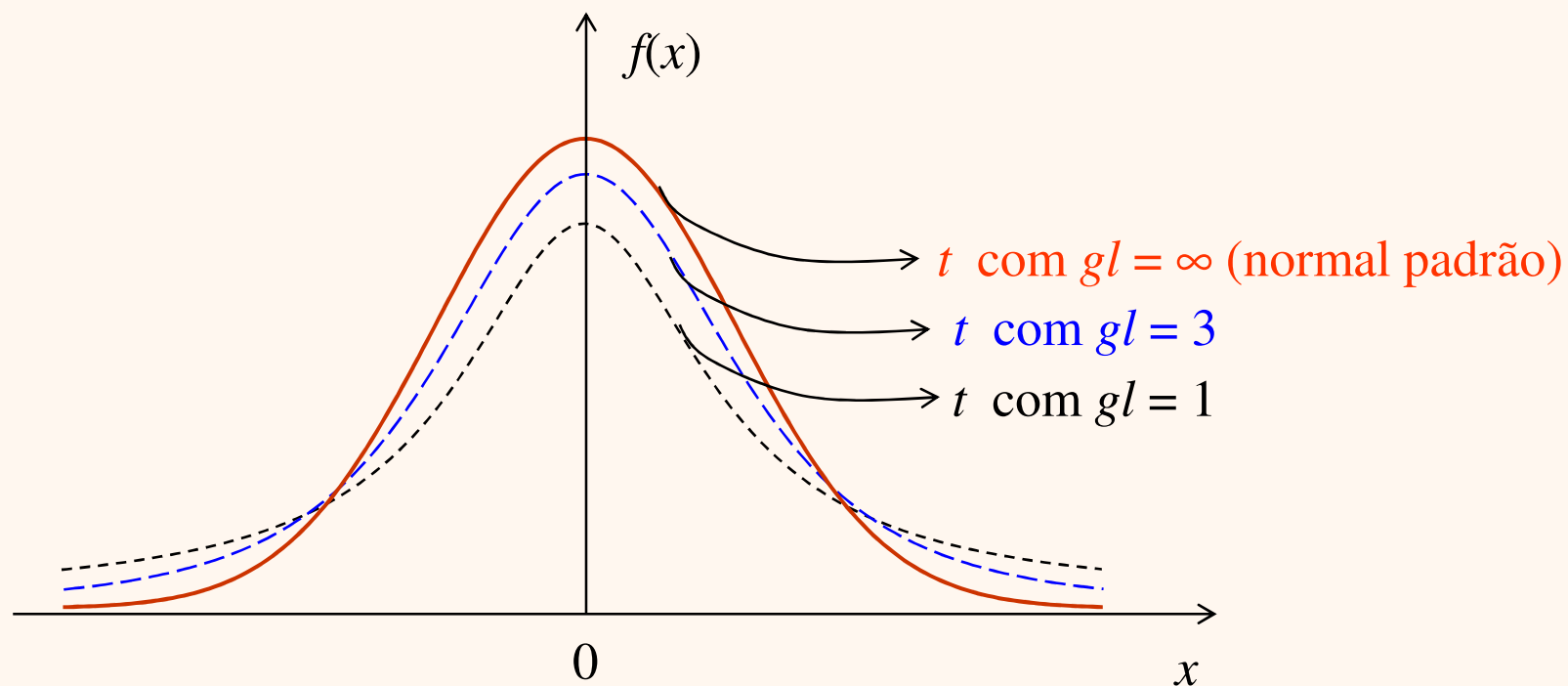
- Se a amostra for pequena ($n < 30$):

$$E = t \cdot S_{\bar{X}}$$

➤ onde t vem da distribuição t com $gl = n - 1$

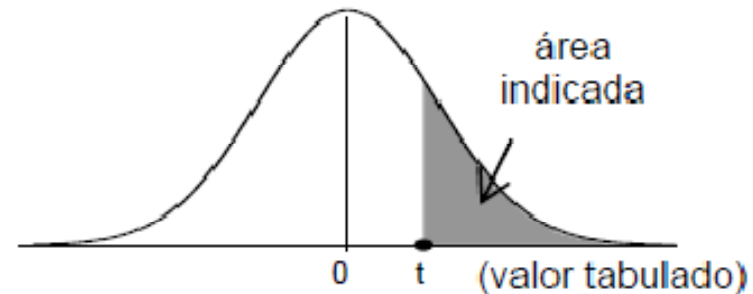
NOTA: Para $n > 30$, $t \cong z$, assim, pode-se sempre usar t .

A distribuição t de Student



A distribuição *t* de *Student*

TABELA Distribuição *t* de *Student*

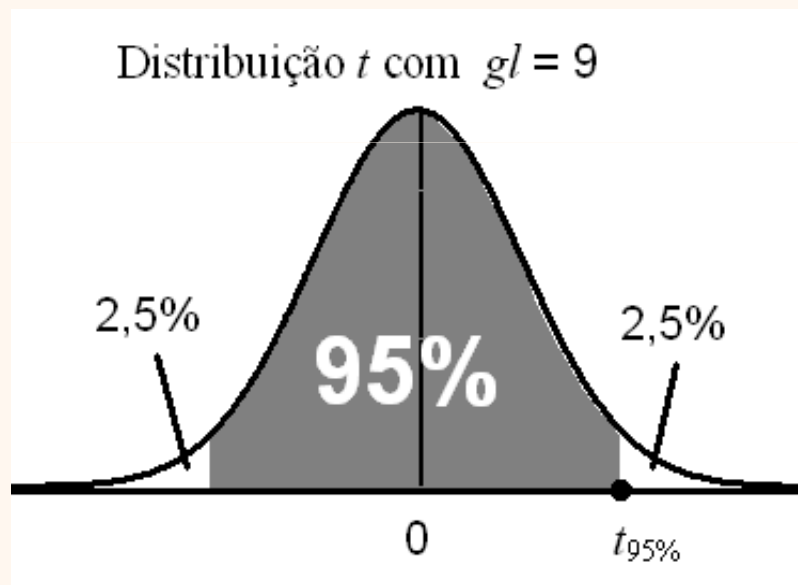


<i>gl</i>	Área na cauda superior								
	0,25	0,10	0,05	0,025	0,01	0,005	0,0025	0,001	0,0005
1	1,000	3,078	6,314	12,71	31,82	63,66	127,3	318,3	636,6
2	0,816	1,886	2,920	4,303	6,965	9,925	14,09	22,33	31,60
3	0,765	1,638	2,353	3,182	4,541	5,841	7,453	10,21	12,92
4	0,741	1,533	2,132	2,776	3,747	4,604	5,598	7,173	8,610
5	0,727	1,476	2,015	2,571	3,365	4,032	4,773	5,894	6,869
6	0,718	1,440	1,943	2,447	3,143	3,707	4,317	5,208	5,959
7	0,711	1,415	1,895	2,365	2,998	3,499	4,029	4,785	5,408
8	0,706	1,397	1,860	2,306	2,896	3,355	3,833	4,501	5,041
9	0,703	1,383	1,833	2,262	2,821	3,250	3,690	4,297	4,781
45	0,680	1,301	1,679	2,014	2,412	2,690	2,952	3,281	3,520
50	0,679	1,299	1,676	2,009	2,403	2,678	2,937	3,261	3,496
<i>z</i>	0,674	1,282	1,645	1,960	2,326	2,576	2,807	3,090	3,291

Como usar a Tabela t

(Tabela 5 do Apêndice)

- Ilustração com $gl = 9$ e nível de confiança de 95%.



gl	Área na cauda superior
	... 0,025 ...
...	
9	2,262
...	

→ $t_{95\%} = 2,262$

Exemplo

- Encontrar o valor de t para:
 - a) nível de confiança de 99%, com 19 graus de liberdade (amostra com 20 elementos)

Resp. $t = 2,861$

Ver exemplos 9.2 e 9.3

Exemplo 9.3

- Para verificar a eficácia de um programa de prevenção de acidentes de trabalho, foi realizado um estudo experimental, implementando esse programa em dez empresas da construção civil, escolhidas ao acaso, numa certa região. Os dados abaixo se referem aos *percentuais de redução de acidentes de trabalho*, nas dez empresas observadas.
- Estimar o parâmetro: μ = média da redução percentual de acidentes de trabalho, em todas as empresas da construção civil da região, que venham a serem submetidas ao programa preventivo.

Amostra	Estatísticas
20 15 23 11 29	Média: $\bar{X} = 18$
5 20 22 18 17	Desvio padrão: $S = 6,65$

Tamanho de Amostras

- No planejamento de uma pesquisa, uma pergunta natural é:
 - Qual é o tamanho da amostra necessário? ($n = ?$)
- No que segue, considerar-se-á que a amostragem será aleatória simples.

Tamanho de amostra

- No caso de estimação de μ , podemos exigir erro amostral menor que um dado E_0 , isto é:

$$|\bar{X} - \mu| \leq E_0$$

ou: $z \frac{\sigma}{\sqrt{n}} \leq E_0$

ou:

$$n \geq \frac{z^2 \sigma^2}{E_0^2}$$

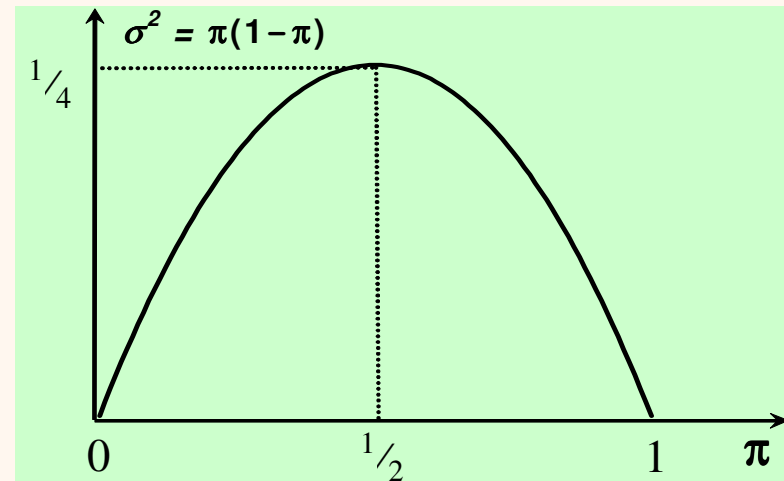
Tamanho de amostra

- No caso de estimação de π :

$$\sigma^2 = \pi.(1-\pi) \leq \frac{1}{4}$$

Assim:

$$n \geq \frac{z^2 \pi(1-\pi)}{E_0^2}$$



Então, tomamos:

$$n = \frac{z^2}{4E_0^2}$$

Ver discussão no livro.

RESUMO: tamanho mínimo de uma amostra aleatória simples

Parâmetro de interesse	Valor inicial do tamanho da amostra
uma média (μ):	$n_0 = \frac{z^2 \sigma^2}{E_0^2}$
uma proporção (π):	$n_0 = \frac{z^2 \pi(1-\pi)}{E_0^2}$
várias proporções (π_1, π_2, \dots):	$n_0 = \frac{z^2}{4E_0^2}$
Tamanho da amostra	
População muito grande ($N > 20n$):	$n = n_0$
População de tamanho N :	$n = \frac{N \cdot n_0}{N + n_0}$