



ESALQ

Escola Superior de Agricultura Luiz de Queiroz
Universidade de São Paulo



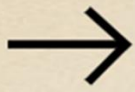
LCE2112 – Estatística Aplicada às Ciências Sociais e Ambientais

Profa. Taciana Villela Savian

tvsvian@usp.br

Elemento

Variáveis Aleatórias



- Altura [m]
- Massa [kg]
- Concentração ferro no sangue [mg/L]
- Cor dos olhos
- ...

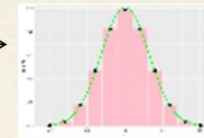
População



Variáveis Aleatórias

- Altura [m]
- Massa [kg]
- Concentração ferro no sangue [mg/L]
- Cor dos olhos
- ...

PDF - Altura



Parâmetros populacionais Θ

- μ = média de altura da população [m]
- σ = desvio padrão da média populacional [m]
- ...

Amostra



Estatísticas amostrais

- $\bar{x}$ = média amostral da altura
- s = desvio padrão amostral da altura
- ...

Inferência

Estimação de parâmetros populacionais a partir das estatísticas amostrais (estimadores)

Estimadores $\hat{\Theta}$

Estimação de Parâmetros

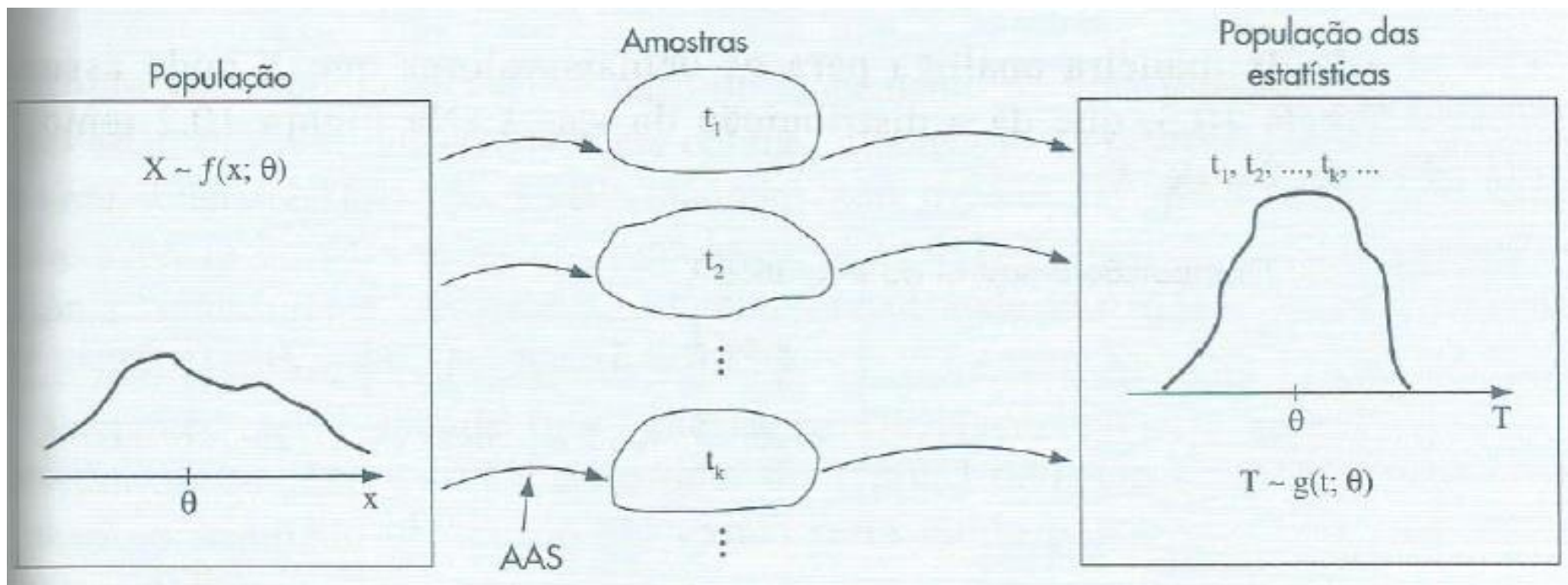
- Avaliar certas características dos elementos da população com base nos dados de uma amostra.
- Para avaliar essas características utilizamos, geralmente, a estimação dos parâmetros: estimação pontual (um valor numérico) ou uma estimação intervalar (um intervalo de valores, com determinada probabilidade).

Estimação de Parâmetros

Parâmetros (características na população)	Estatísticas (características da amostra)
μ = média de alguma variável quantitativa nos elementos da população	$\bar{x}$ = média da variável quantitativa, calculada com os elementos da amostra
σ = desvio padrão de uma variável, dentre os elementos da população	s = desvio padrão da variável, calculado com os elementos da amostra

Distribuição Amostral

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

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 da média ($\bar{X}$)

Distribuição Amostral da Média com σ^2 conhecida

Teorema Central do Limite:

Quando o tamanho da amostra (n) aumenta, independentemente da forma da distribuição da população, a distribuição amostral de $\bar{X}$ aproxima-se cada vez mais de uma distribuição Normal com média μ e variância $\frac{\sigma^2}{n}$.

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Distribuição Amostral da Média

Distribuições de X e Xbarra

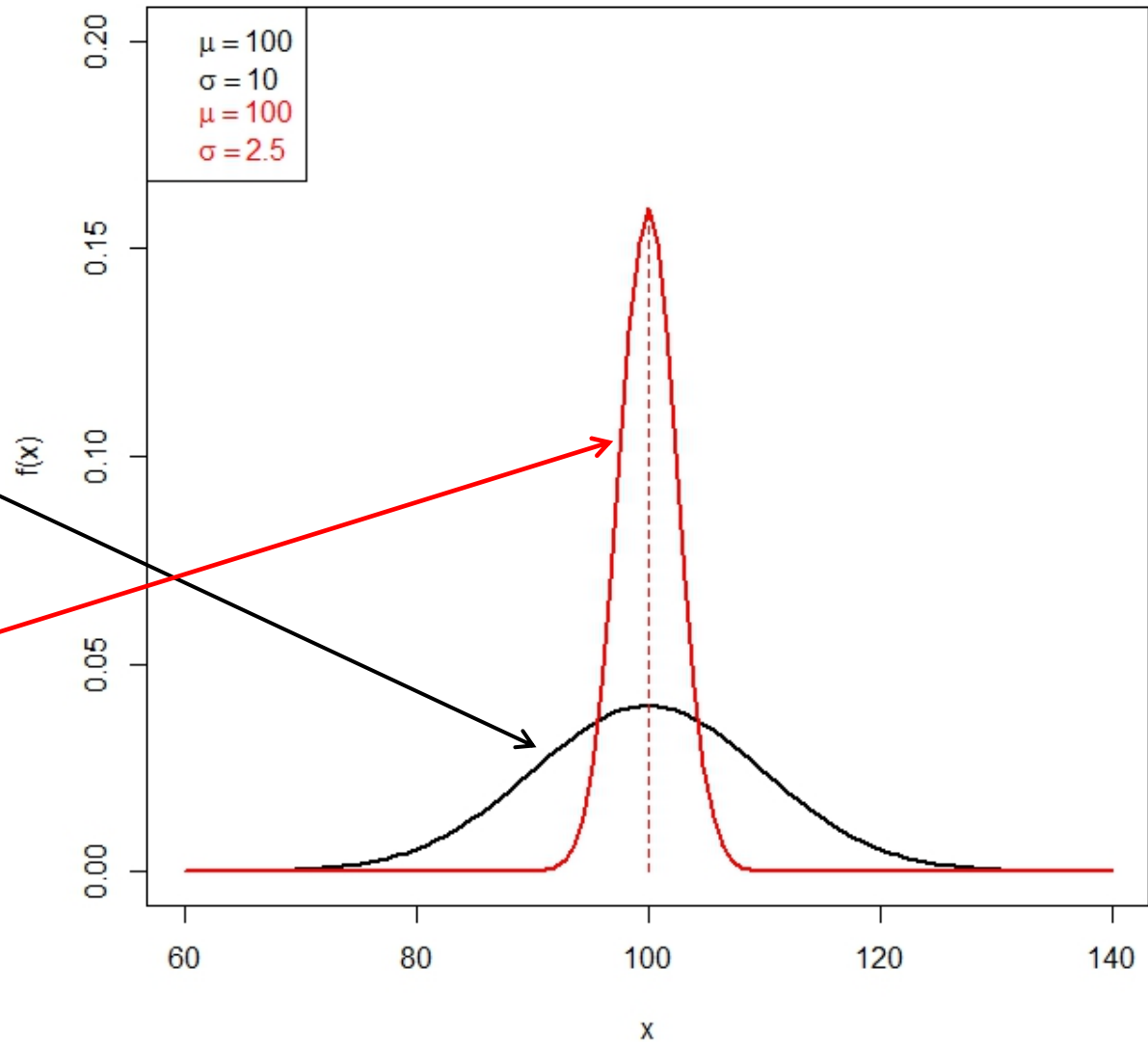
Teorema

Central do

Limite

$$X \sim N(\mu, \sigma^2)$$

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$



Distribuição Amostral da Média com σ^2 desconhecida

- Observe que, para utilização do teorema central do limite, temos que assumir que a variância populacional σ^2 é conhecida. Entretanto, é muito pouco provável que em uma situação real ela realmente seja conhecida. Assim, faz-se necessário o uso do seguinte Teorema:
- Seja uma população descrita por uma variável $X \sim N(\mu, \sigma^2)$ da qual coleta-se infinitas amostras de tamanho n a partir das quais são calculadas $\bar{X}$ e s^2 . Então a variável $T = \frac{\bar{X} - \mu}{\sqrt{\frac{s^2}{n}}}$, tem **distribuição t de Student**, que tem como único parâmetro a constante **$\nu = n - 1$** graus de liberdade.

Distribuição Amostral da Média com σ^2 desconhecida

- ***Distribuição t de Student***

- É semelhante à distribuição Normal, pois é simétrica em relação à média e tem forma campanular (sino);
- A distribuição t de Student é definida a partir de um único parâmetro, o número de graus de liberdade obtido por $\nu = n - 1$;
- Será bastante utilizada em testes de hipótese para média e diferença entre médias populacionais.

Estimativa por Ponto e por Intervalo

O **parâmetro populacional** é estimado por **dois valores**, obtidos por meio de cálculos com os dados amostrais e informações sobre a distribuição amostral da estatística, formando um intervalo denominado **intervalo de confiança**, dentro do qual se espera, com determinada probabilidade, encontrar o verdadeiro valor do parâmetro.

Estimativa por Intervalo

**1) Intervalo de confiança para média
(σ^2 conhecida e desconhecida);**

Estimativa por Intervalo

1) Intervalo de confiança para média (σ^2 conhecida)

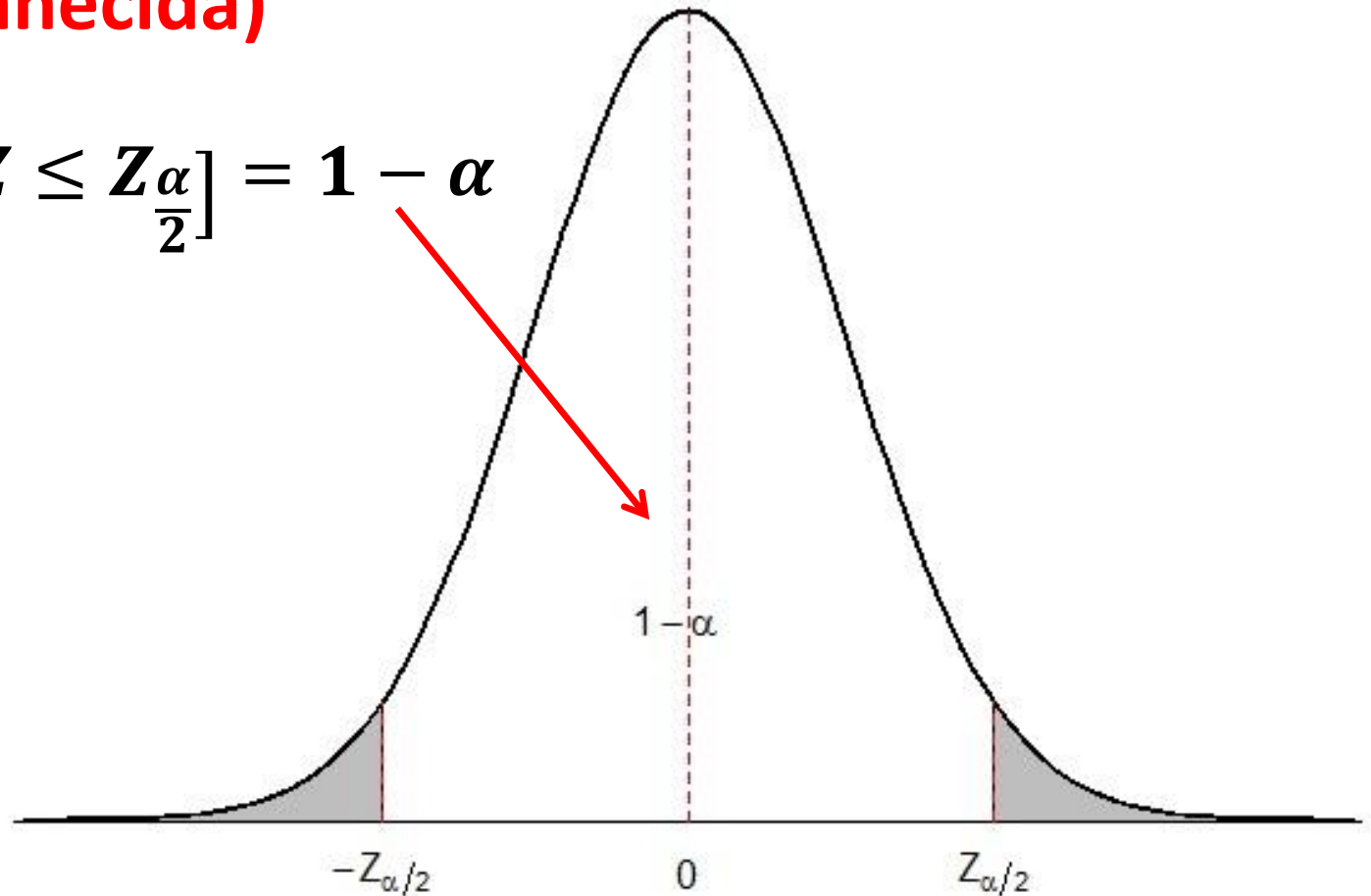
Definimos um **intervalo simétrico** ao redor da média amostral ($\bar{x}$) que contém valores “plausíveis” para a média populacional (μ)

Os intervalos são construídos de acordo com um dado “**nível de confiança ($1 - \alpha$)**”

Estimativa por Intervalo

1) Intervalo de confiança para média (σ^2 conhecida)

$$P\left[-Z_{\frac{\alpha}{2}} \leq Z \leq Z_{\frac{\alpha}{2}}\right] = 1 - \alpha$$



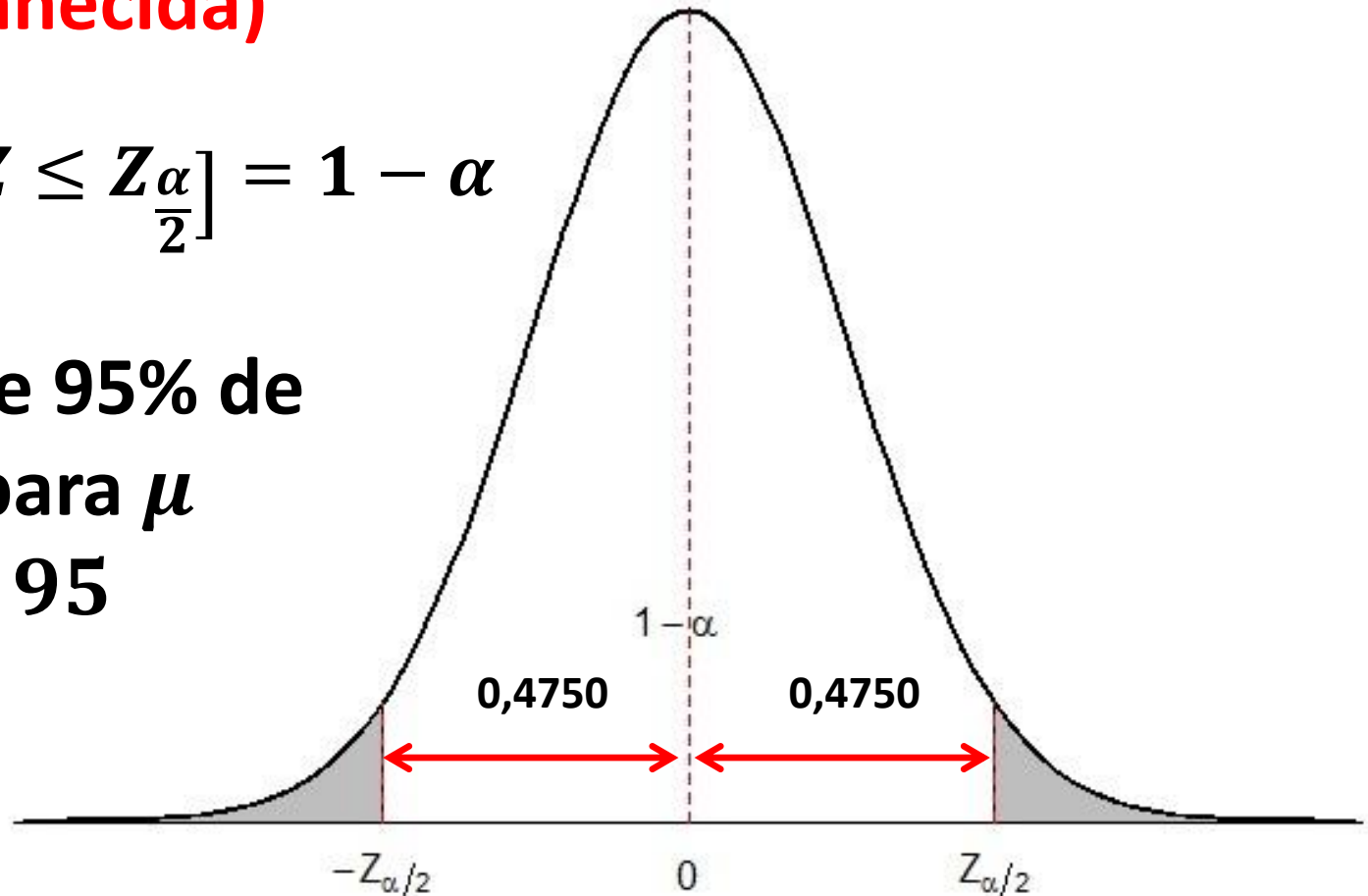
Estimativa por Intervalo

1) Intervalo de confiança para média (σ^2 conhecida)

$$P\left[-Z_{\frac{\alpha}{2}} \leq Z \leq Z_{\frac{\alpha}{2}}\right] = 1 - \alpha$$

Intervalo de 95% de
confiança para μ

$$1 - \alpha = 0,95$$



Estimativa por Intervalo

[illegible]

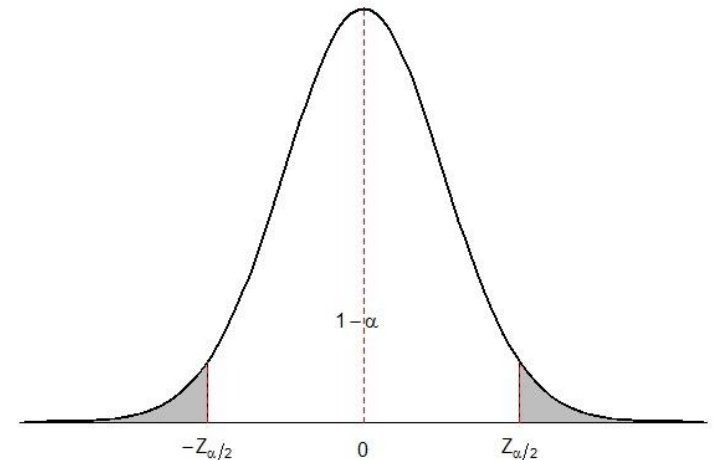
Estimativa por Intervalo

1) Intervalo de confiança para média (σ^2 conhecida)

$$P\left[-Z_{\frac{\alpha}{2}} \leq Z \leq Z_{\frac{\alpha}{2}}\right] = 0,95$$

$$P[-1,96 \leq Z \leq 1,96] = 0,95$$

$$P\left[-1,96 \leq \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \leq 1,96\right] = 0,95$$

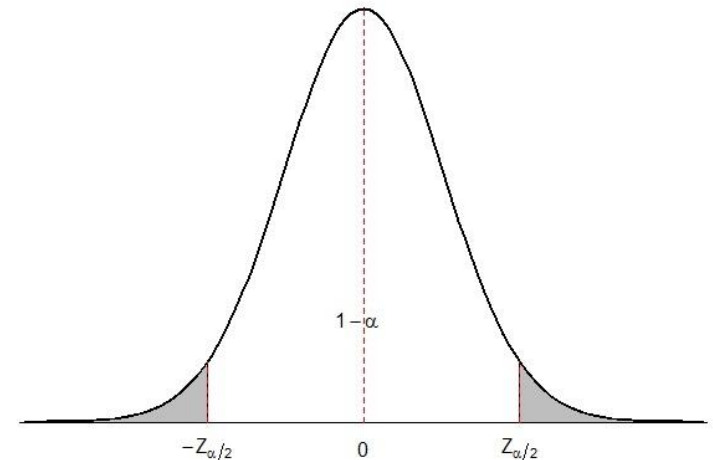


Estimativa por Intervalo

1) Intervalo de confiança para média (σ^2 conhecida)

$$P \left[-1,96 \leq \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \leq 1,96 \right] = 0,95$$

$$P \left[\bar{X} - 1,96 \sqrt{\frac{\sigma^2}{n}} \leq \mu \leq \bar{X} + 1,96 \sqrt{\frac{\sigma^2}{n}} \right] = 0,95$$



$$IC(\mu)_{1-\alpha}: \bar{X} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma^2}{n}}$$

Estimativa por Intervalo

Exemplo: Um pesquisador está estudando a resistência de um certo material sob determinadas condições. Ele sabe que essa variável tem distribuição Normal com $\rightarrow \sigma^2 = 4 \text{ un}^2$. Foi extraída uma amostra aleatória de tamanho 10 ($n=10$) obtendo-se os seguintes valores: 7,9; 6,8; 5,4; 7,5; 7,9; 6,4; 8,0; 6,3; 4,4; 5,9.

- (a) Calcule a estimativa pontual ($\bar{X}$) da média populacional (μ), com base nesta amostra.
- (b) Determine o intervalo de confiança para a resistência média (μ) com um coeficiente de confiança de 95%

Estimativa por Intervalo

Exemplo: Um pesquisador está estudando a resistência de um certo material sob determinadas condições. Ele sabe que essa variável tem distribuição Normal com $\sigma^2 = 4 \text{ un}^2$. Foi extraída uma amostra aleatória de tamanho 10 ($n=10$) obtendo-se os seguintes valores: 7,9; 6,8; 5,4; 7,5; 7,9; 6,4; 8,0; 6,3; 4,4; 5,9.

(a) Calcule a estimativa pontual ($\bar{X}$) da média populacional (μ), com base nesta amostra.

$$\bar{X} = \frac{7,9 + 6,8 + \dots + 5,9}{10} = 6,65 \text{ unidades}$$

Estimativa por Intervalo

Exemplo: Um pesquisador está estudando a resistência de um certo material sob determinadas condições. Ele sabe que essa variável tem distribuição Normal com $\sigma^2 = 4 \text{ un}^2$. Foi extraída uma amostra aleatória de tamanho 10 ($n=10$) obtendo-se os seguintes valores: 7,9; 6,8; 5,4; 7,5; 7,9; 6,4; 8,0; 6,3; 4,4; 5,9.

(b) Determine o intervalo de confiança para a resistência média (μ) com um coeficiente de confiança de 95%

$$IC(\mu)_{1-\alpha}: \bar{X} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma^2}{n}}$$

$$\bar{X} = 6,65$$

$$Z_{\alpha/2} = 1,96$$

$$\sigma^2 = 4$$

$$n = 10$$

Estimativa por Intervalo

Exemplo: Um pesquisador está estudando a resistência de um certo material sob determinadas condições. Ele sabe que essa variável tem distribuição Normal com $\sigma^2 = 4 \text{ un}^2$. Foi extraída uma amostra aleatória de tamanho 10 ($n=10$) obtendo-se os seguintes valores: 7,9; 6,8; 5,4; 7,5; 7,9; 6,4; 8,0; 6,3; 4,4; 5,9.

$$5,41 < \mu < 7,89$$

Se forem retiradas amostras repetidas vezes nesta população aproximadamente em 95% das vezes a resistência média do material (μ) estará no intervalo encontrado (entre 5,41 unidades e 7,89 unidades).

Estimativa por Intervalo

Exemplo: Um pesquisador está estudando a resistência de um certo material sob determinadas condições. Ele sabe que essa variável tem distribuição Normal com $\sigma^2 = 4 \text{ un}^2$. Foi extraída uma amostra aleatória de tamanho 10 ($n=10$) obtendo-se os seguintes valores: 7,9; 6,8; 5,4; 7,5; 7,9; 6,4; 8,0; 6,3; 4,4; 5,9.

(c) **Determine o intervalo de confiança para a resistência média (μ) com um coeficiente de confiança de 99%.**

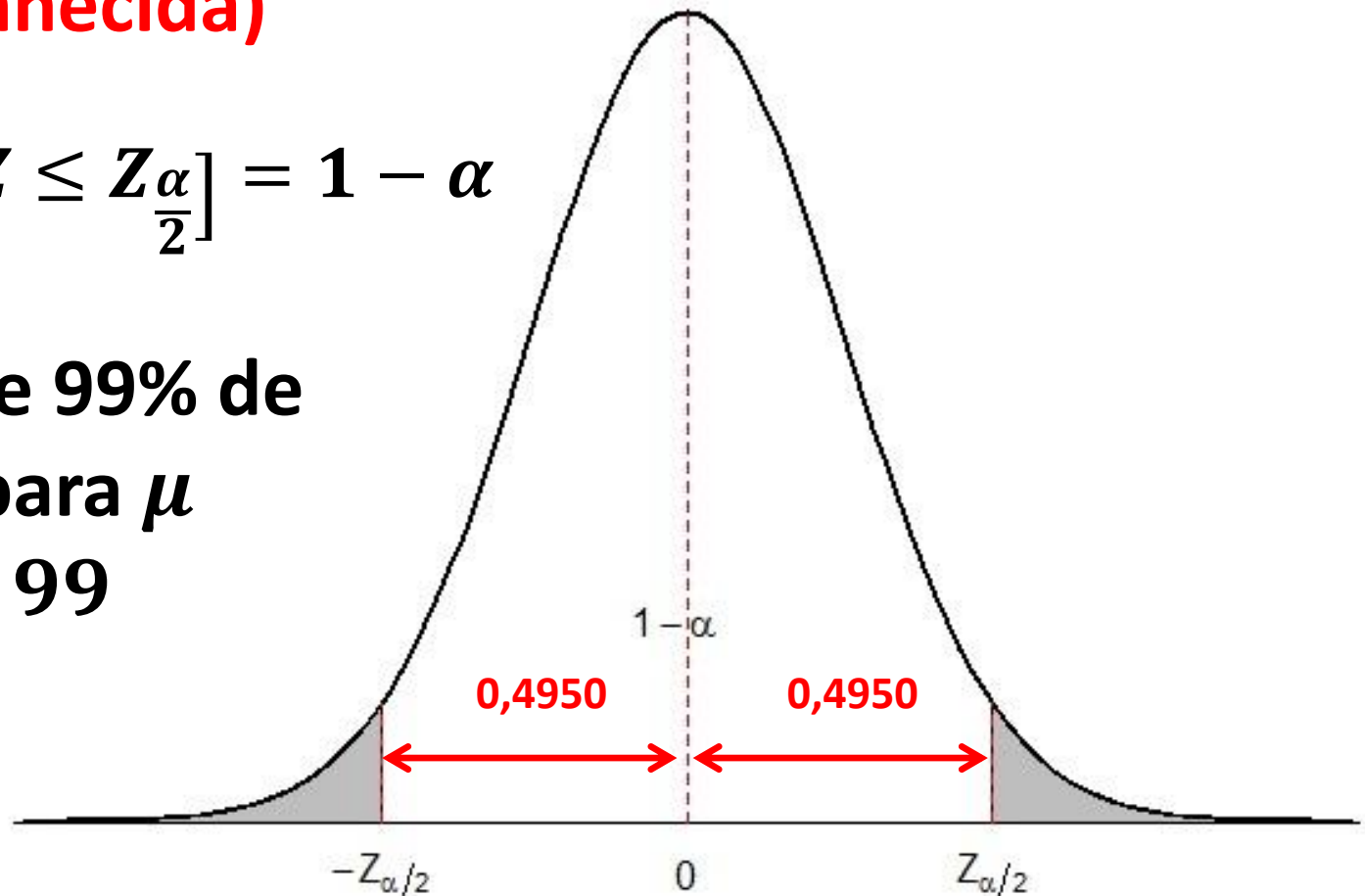
Estimativa por Intervalo

1) Intervalo de confiança para média (σ^2 conhecida)

$$P\left[-Z_{\frac{\alpha}{2}} \leq Z \leq Z_{\frac{\alpha}{2}}\right] = 1 - \alpha$$

Intervalo de 99% de
confiança para μ

$$1 - \alpha = 0,99$$



2,57 2,58

also $2,575 \rightarrow 7 \frac{d}{2}$

[illegible]

Estimativa por Intervalo

Exemplo: Um pesquisador está estudando a resistência de um certo material sob determinadas condições. Ele sabe que essa variável tem distribuição Normal com $\sigma^2 = 4 \text{ un}^2$. Foi extraída uma amostra aleatória de tamanho 10 ($n=10$) obtendo-se os seguintes valores: 7,9; 6,8; 5,4; 7,5; 7,9; 6,4; 8,0; 6,3; 4,4; 5,9.

(c) Determine o intervalo de confiança para a resistência média (μ) com um coeficiente de confiança de 99%.

$$IC(\mu)_{1-\alpha}: \bar{X} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma^2}{n}}$$

$$\bar{X} = 6,65$$

$$Z_{\alpha/2} = 2,575$$

$$\sigma^2 = 4$$

$$n = 10$$

Estimativa por Intervalo

Exemplo: Um pesquisador está estudando a resistência de um certo material sob determinadas condições. Ele sabe que essa variável tem distribuição Normal com $\sigma^2 = 4 \text{ un}^2$. Foi extraída uma amostra aleatória de tamanho 10 ($n=10$) obtendo-se os seguintes valores: 7,9; 6,8; 5,4; 7,5; 7,9; 6,4; 8,0; 6,3; 4,4; 5,9.

(c) Determine o intervalo de confiança para a resistência média (μ) com um coeficiente de confiança de 99%

$$IC(\mu)_{0,99}: 6,65 \pm 2,575 \sqrt{\frac{4}{10}} \rightarrow 5,02 < \mu < 8,28$$

Estimativa por Intervalo

Exemplo: Um pesquisador está estudando a resistência de um certo material sob determinadas condições. Ele sabe que essa variável tem distribuição Normal com $\sigma^2 = 4 \text{ un}^2$. Foi extraída uma amostra aleatória de tamanho 10 ($n=10$) obtendo-se os seguintes valores: 7,9; 6,8; 5,4; 7,5; 7,9; 6,4; 8,0; 6,3; 4,4; 5,9.

(d) Comparar os dois intervalos construídos.

- O coeficiente de confiança aumentou (95% para 99%);
- O intervalo de confiança aumentou

$$5,41 < \mu < 7,89 \quad \text{para} \quad 5,02 < \mu < 8,28$$

Estimativa por Intervalo

1) Intervalo de confiança para média (σ^2 desconhecida);

Relembrando

A distribuição amostral de uma estatística.

- Distribuição amostral da média ($\bar{X}$) com σ^2 desconhecida

Teorema do Limite Central não se verifica

Seja uma população descrita por uma variável $X \sim N(\mu, \sigma^2)$ da qual coleta-se infinitas amostras de tamanho n a partir das quais são calculadas $\bar{X}$ e s^2 . Então a variável

$$T = \frac{\bar{X} - \mu}{\sqrt{\frac{s^2}{n}}},$$

tem **distribuição t de Student com $\nu = n - 1$** graus de liberdade.

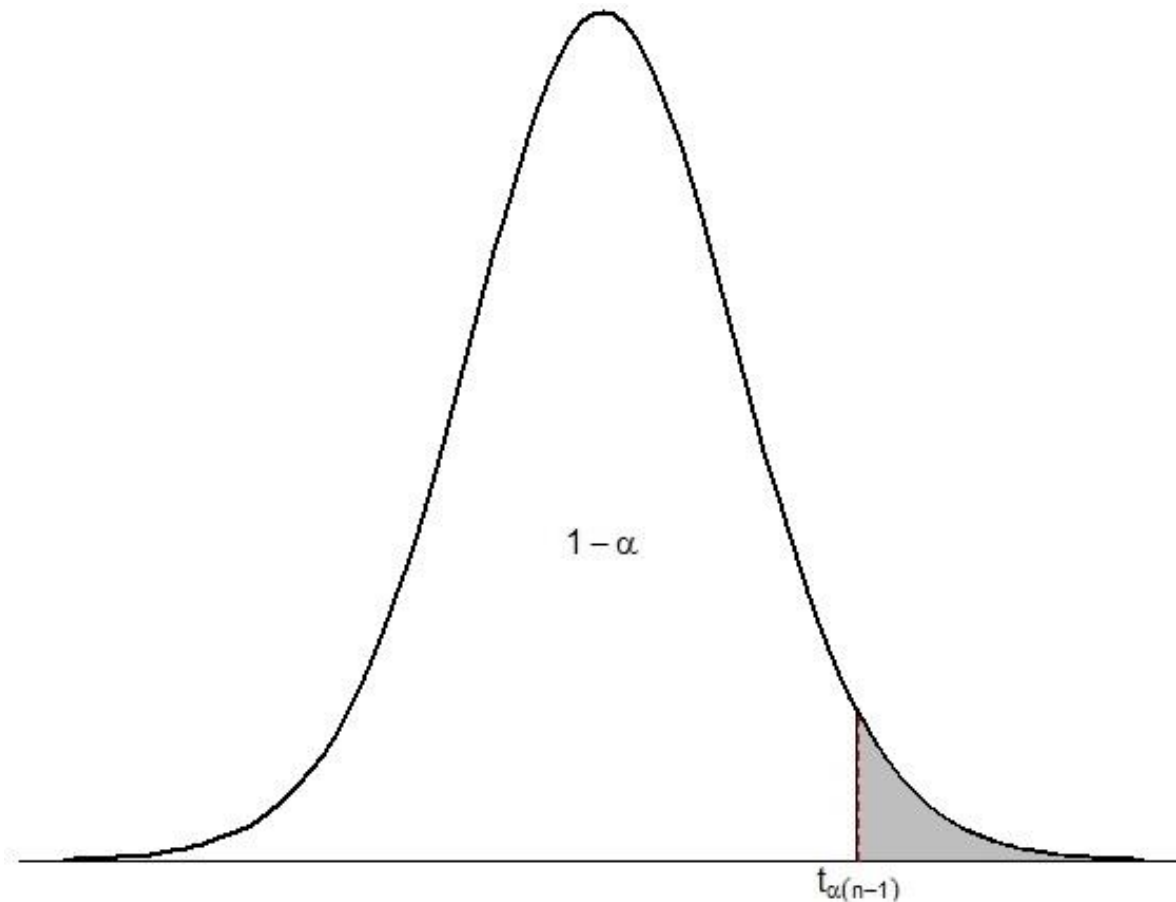
Distribuição t Student

Tabela A.4. Quantis superior da distribuição t de Student (t_α) com ν graus de liberdade (n-1) e para diferentes valores de α , de acordo com a seguinte afirmação probabilística: $P[T > t_\alpha] = \alpha$.

	α								
ν	0,250	0,200	0,150	0,100	0,050	0,025	0,010	0,005	0,001
1	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	636,619
2	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,599
3	0,765	0,979	1,250	1,638	2,353	3,182	4,541	5,841	12,924
4	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,500	5,408
8	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073

Distribuição t Student

Tabela A.4. Quantis superior da distribuição t de Student (t_α) com ν graus de liberdade e para diferentes valores de α , de acordo com a seguinte afirmação probabilística: $P[T > t_\alpha] = \alpha$.



Estimativa por Intervalo

1) Intervalo de confiança para média (σ^2 desconhecida)

$$P\left[-t_{\frac{\alpha}{2}} \leq T \leq t_{\frac{\alpha}{2}}\right] = 1 - \alpha$$

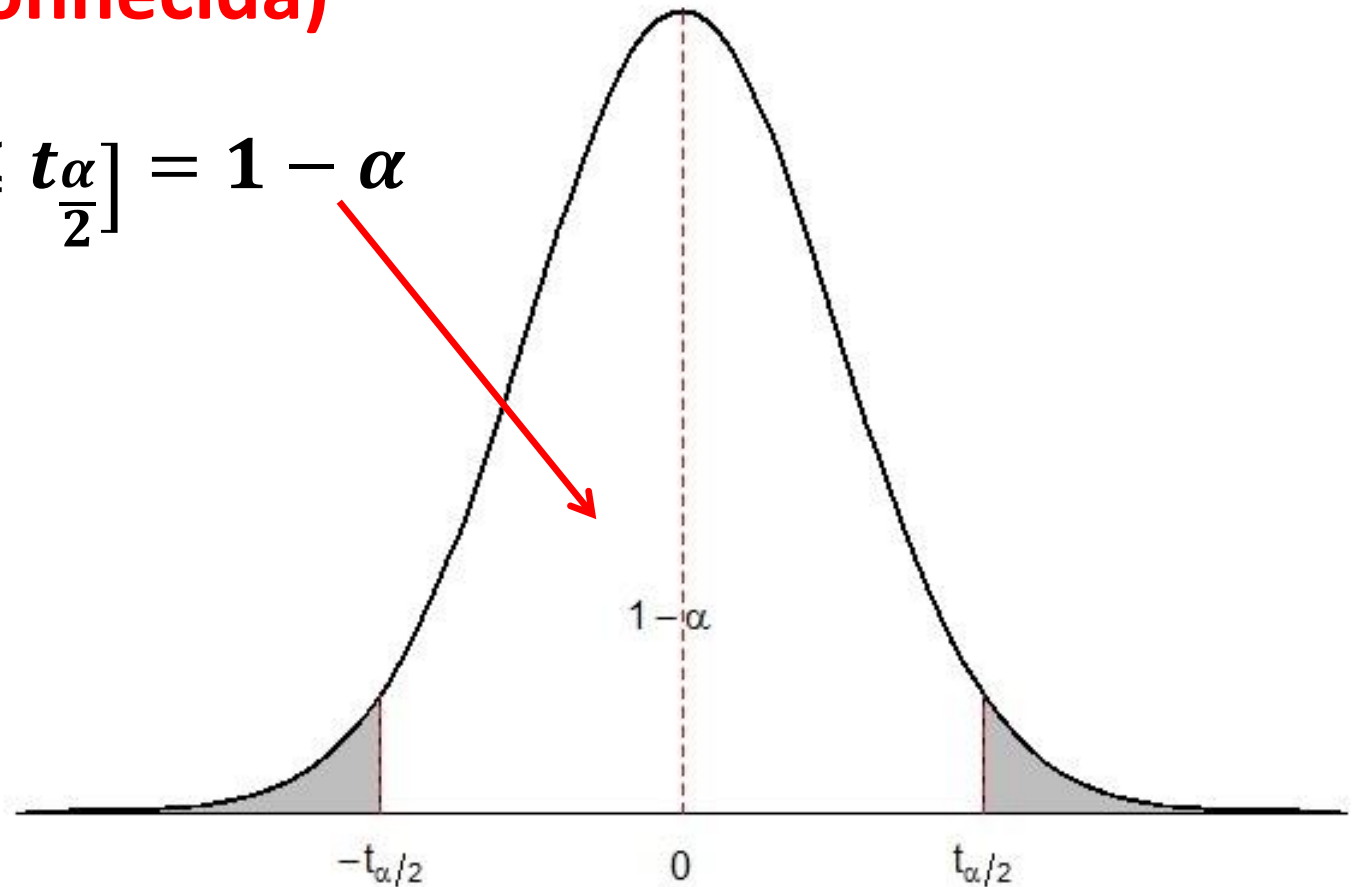
Intervalo de 99% de confiança para μ com σ^2 desconhecida: $1 - \alpha = 0,99 \rightarrow \alpha = 0,01$

Intervalo de 95% de confiança para μ com σ^2 desconhecida: $1 - \alpha = 0,95 \rightarrow \alpha = 0,05$

Estimativa por Intervalo

1) Intervalo de confiança para média (σ^2 desconhecida)

$$P\left[-t_{\frac{\alpha}{2}} \leq T \leq t_{\frac{\alpha}{2}}\right] = 1 - \alpha$$



Estimativa por Intervalo

1) Intervalo de confiança para média (σ^2 desconhecida)

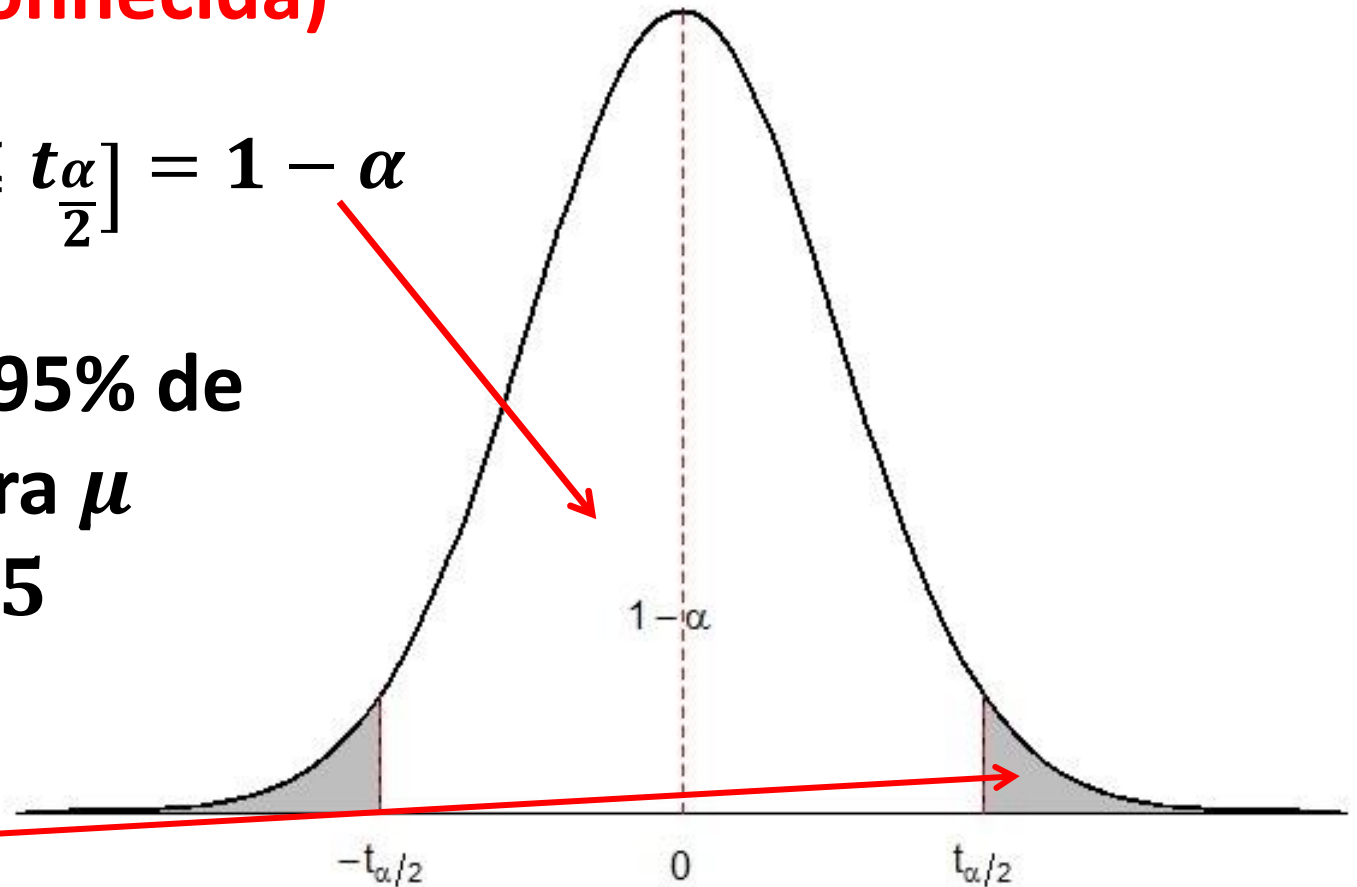
$$P\left[-t_{\frac{\alpha}{2}} \leq T \leq t_{\frac{\alpha}{2}}\right] = 1 - \alpha$$

Intervalo de 95% de
confiança para μ

$$1 - \alpha = 0,95$$

$$\alpha = 0,05$$

$$\frac{\alpha}{2} = 0,025$$



Distribuição t Student

Tabela A.4. Quantis superior da distribuição t de Student (t_α) com ν graus de liberdade (n-1) e para diferentes valores de α , de acordo com a seguinte afirmação probabilística: $P[T > t_\alpha] = \alpha$.

	α								
ν	0,250	0,200	0,150	0,100	0,050	0,025	0,010	0,005	0,001
1	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	636,619
2	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,599
3	0,765	0,979	1,250	1,638	2,353	3,182	4,541	5,841	12,924
4	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,500	5,408
8	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073

Estimativa por Intervalo

1) Intervalo de confiança para média (σ^2 desconhecida)

$$P\left[-t_{\frac{\alpha}{2}} \leq T \leq t_{\frac{\alpha}{2}}\right] = 1 - \alpha$$

$$P\left[-t_{\frac{\alpha}{2}} \leq \frac{\bar{X} - \mu}{\sqrt{\frac{S^2}{n}}} \leq t_{\frac{\alpha}{2}}\right] = 1 - \alpha$$

$$P\left[\bar{X} - t_{\frac{\alpha}{2}} \sqrt{\frac{S^2}{n}} \leq \mu \leq \bar{X} + t_{\frac{\alpha}{2}} \sqrt{\frac{S^2}{n}}\right] = 1 - \alpha$$

Estimativa por Intervalo

1) Intervalo de confiança para média (σ^2 desconhecida)

$$IC(\mu)_{1-\alpha}: \bar{X} \pm t_{\frac{\alpha}{2}} \sqrt{\frac{s^2}{n}}$$

Exemplo: Suponha que o consumo diário de água de uma comunidade seja uma variável com distribuição normal e que uma amostra de 15 observações (n) produziu as seguintes estatísticas $\bar{X} = 63m^3$ e $s^2 = 121m^6$. Vamos obter um intervalo de 95% de confiança para a média populacional.

Estimativa por Intervalo

$$IC(\mu)_{1-\alpha}: \bar{X} \pm t_{\frac{\alpha}{2}} \sqrt{\frac{s^2}{n}}$$

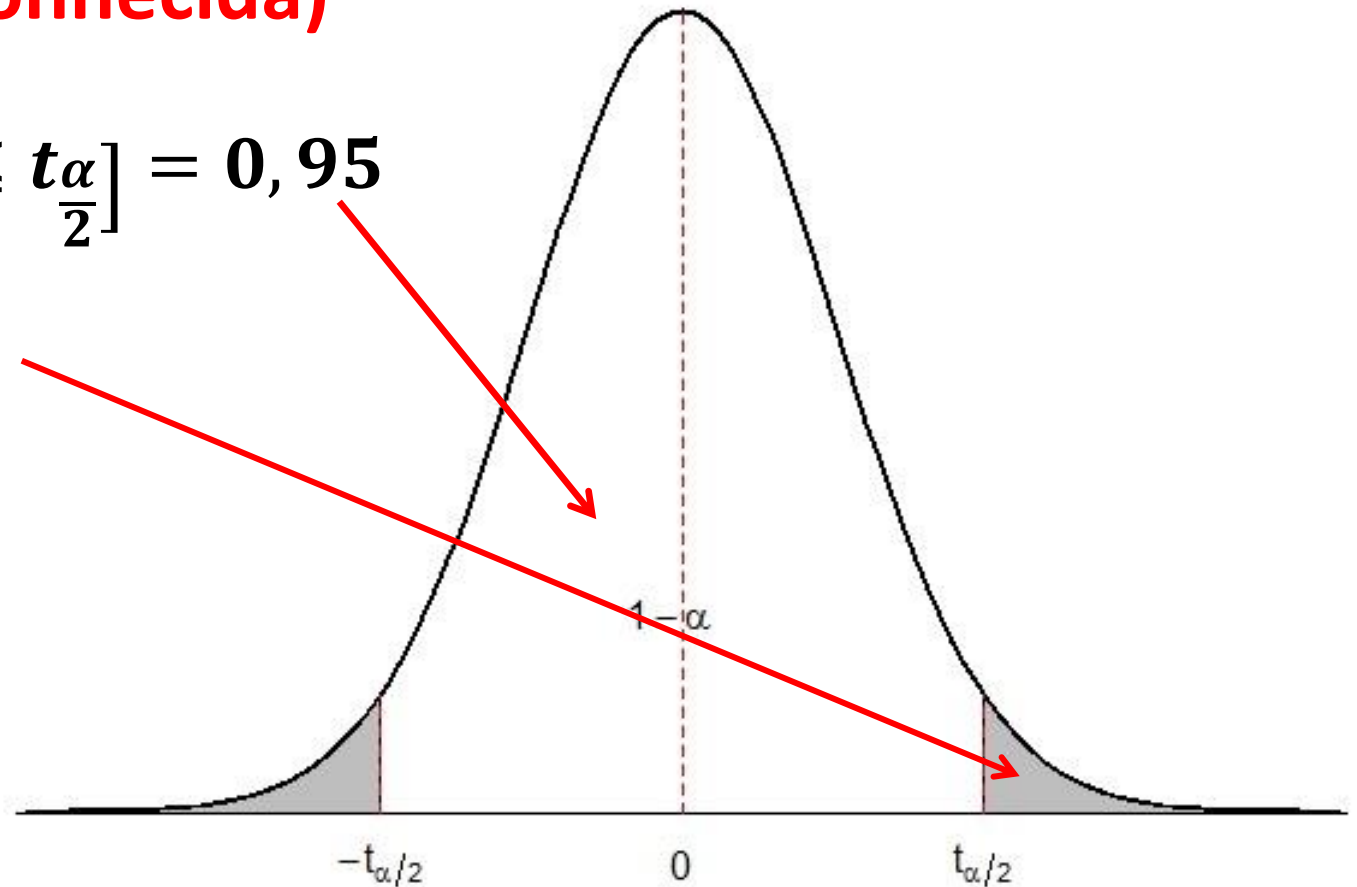
Exemplo: Suponha que o consumo diário de água de uma comunidade seja uma variável com distribuição normal e que uma amostra de **15 observações (n)** produziu as seguintes estatísticas $\bar{X} = 63m^3$ e $s^2 = 121m^6$. Vamos obter um intervalo de **95% de confiança** para a média populacional.

Estimativa por Intervalo

1) Intervalo de confiança para média (σ^2 desconhecida)

$$P\left[-t_{\frac{\alpha}{2}} \leq T \leq t_{\frac{\alpha}{2}}\right] = 0,95$$

$$\frac{\alpha}{2} = 0,025$$



Distribuição t Student

Tabela A.4. Quantis superior da distribuição t de Student (t_α) com ν graus de liberdade (n-1) e para diferentes valores de α , de acordo com a seguinte afirmação probabilística: $P[T > t_\alpha] = \alpha$.

	α								
ν	0,250	0,200	0,150	0,100	0,050	0,025	0,010	0,005	0,001
1	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	636,619
2	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,599
3	0,765	0,979	1,250	1,638	2,353	3,182	4,541	5,841	12,924
4	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,500	5,408
8	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073

Estimativa por Intervalo

$$IC(\mu)_{1-\alpha}: \bar{X} \pm t_{\frac{\alpha}{2}} \sqrt{\frac{s^2}{n}}$$

Exemplo: Suponha que o consumo diário de água de uma comunidade seja uma variável com distribuição normal e que uma amostra de **15 observações (n)** produziu as seguintes estatísticas $\bar{X} = 63m^3$ e $s^2 = 121m^6$. Vamos obter um intervalo de **95% de confiança** para a média populacional.

$$IC(\mu)_{0,95}: 63 \pm 2,145 \sqrt{\frac{121}{15}}$$

Estimativa por Intervalo

$$IC(\mu)_{1-\alpha}: \bar{X} \pm t_{\frac{\alpha}{2}} \sqrt{\frac{s^2}{n}}$$

Exemplo: Suponha que o consumo diário de água de uma comunidade seja uma variável com distribuição normal e que uma amostra de **15 observações (n)** produziu as seguintes estatísticas $\bar{X} = 63m^3$ e $s^2 = 121m^6$. Vamos obter um intervalo de **95% de confiança** para a média populacional.

$$IC(\mu)_{0,95}: 63 \pm 6,09$$

$$56,91m^3 < \mu < 69,09m^3$$

Estimativa por Intervalo

$$56,91m^3 < \mu < 69,09m^3$$

Se forem retiradas amostras repetidas vezes nesta população aproximadamente em 95% das vezes o consumo médio de água (μ) estará entre $56,91m^3$ e $69,09m^3$.

Resumo

Como estimar ?

Parâmetro (μ)
desconhecido

Por ponto
(pontual)

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n}$$

Por intervalo
(de confiança)

Variância σ^2
(conhecida)

Variância σ^2
(desconhecida)

- Distribuição Normal Padronizada: $N(0,1)$

$$IC(\mu)_{1-\alpha}: \bar{X} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma^2}{n}}$$

- Distribuição t-Student

- Estimar σ^2 (ponto) $\rightarrow s^2$

$$IC(\mu)_{1-\alpha}: \bar{X} \pm t_{\frac{\alpha}{2}} \sqrt{\frac{s^2}{n}}$$