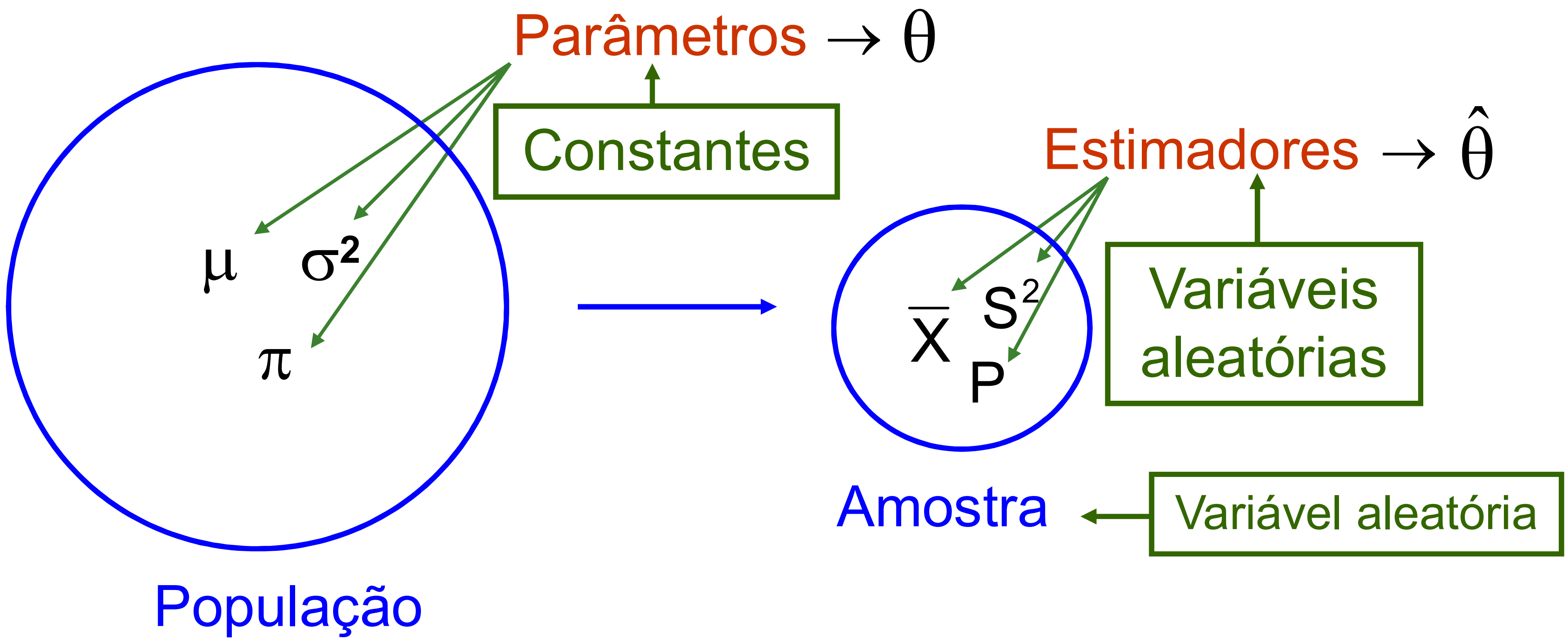
Inferência Estatística – Estimação

- Estimação de parâmetros
  - estimação por ponto ou por intervalo
- Intervalos de confiança
  - para média
  - para variância
  - para proporção
- Dimensionamento amostral

Estimação de parâmetros



⇨ O estimador é uma variável aleatória; portanto, pode assumir diferentes valores

Estimativa é um **valor particular** que o estimador assume

Processos de estimação

Estimação por ponto

É o processo através do qual obtemos um único ponto, ou seja, **um único valor** para estimar o parâmetro.

Exemplo: Amostra (1, 3, 2)

$$\bar{x} = \frac{\sum x_i}{n} = \frac{1+3+2}{3} = 2$$

estimativa por ponto de μ

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{(1-2)^2 + (3-2)^2 + (2-2)^2}{3-1} = 1$$

estimativa por ponto de σ^2

Estimação por intervalo

É um processo que permite obter os **limites de um intervalo** onde, com uma determinada probabilidade (**nível de confiança**), podemos esperar contenham o verdadeiro valor do parâmetro.

dois valores

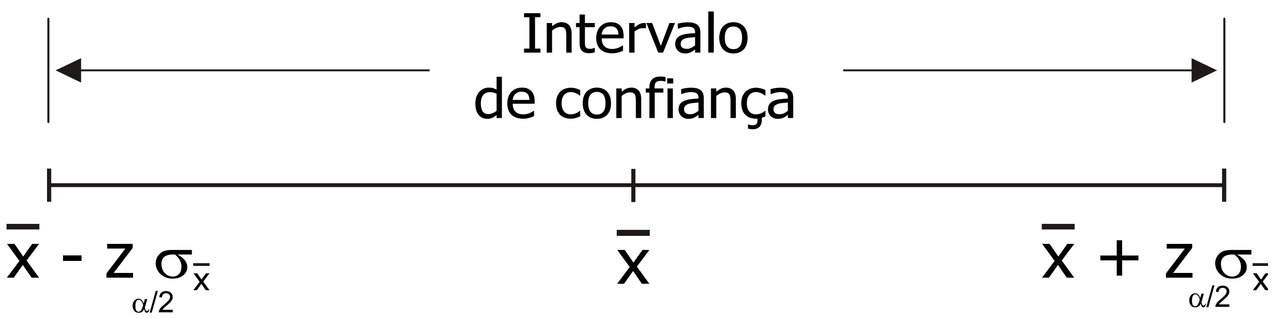
Limite inferior → $\underbrace{L_i < \theta < L_s}_{\text{Intervalo de confiança}}$ ← Limite superior

As estimativas por intervalo são preferíveis àquelas por ponto porque indicam a precisão, estabelecendo limites que, com uma determinada probabilidade, devem conter o parâmetro.

Estimação por intervalo

- Logo, para se ter confiança de estimar o verdadeiro parâmetro populacional, gera-se um intervalo de possíveis valores, a partir do valor (estimativa pontual) encontrado na amostra.

- Quanto maior a amplitude do intervalo, maior a confiança (probabilidade) de estimar corretamente o verdadeiro parâmetro populacional, porém menor será a precisão da estimação.



- Essa probabilidade $(1-\alpha)$ é chamada nível de confiança, sendo α o nível de significância, ou seja, a probabilidade do intervalo não conter o verdadeiro parâmetro populacional.

Intervalos para estimar...

- Média
- Variância
- Proporção

Como construir intervalos de confiança para μ?

1. Intervalo de confiança para a **média** de **uma** população (μ)

Duas situações

- **Conhecemos o valor de σ (<u>ou</u> $n > 30$)**
- **Não conhecemos o valor de σ (<u>e</u> $n \leq 30$)**

2. Intervalo de confiança para a **diferença entre médias** de **duas** populações (μ_1-μ_2)

1. Intervalo de confiança para a média de uma população (μ)

Situação 1: Quando conhecemos o valor de σ (ou $n > 30$)

Parâmetro $\rightarrow$ μ (média da população X)

Qual é o melhor estimador de μ?

Estimador $\rightarrow$ $\overline{X}$ (média aritmética simples da amostra)

De acordo com o **TCL**:

Se a população tem distribuição normal: $X \sim N(\mu, \sigma^2)$

então, $\overline{X} \sim N(\mu, \sigma^2/n)$

Padronizar a variável $\overline{X}$ $\rightarrow$ transformar $\overline{X}$ em Z

$$\overline{X} \sim N\,(\mu, \sigma^2/n)$$

$$\downarrow$$

$$Z \sim N\,(0,1)$$

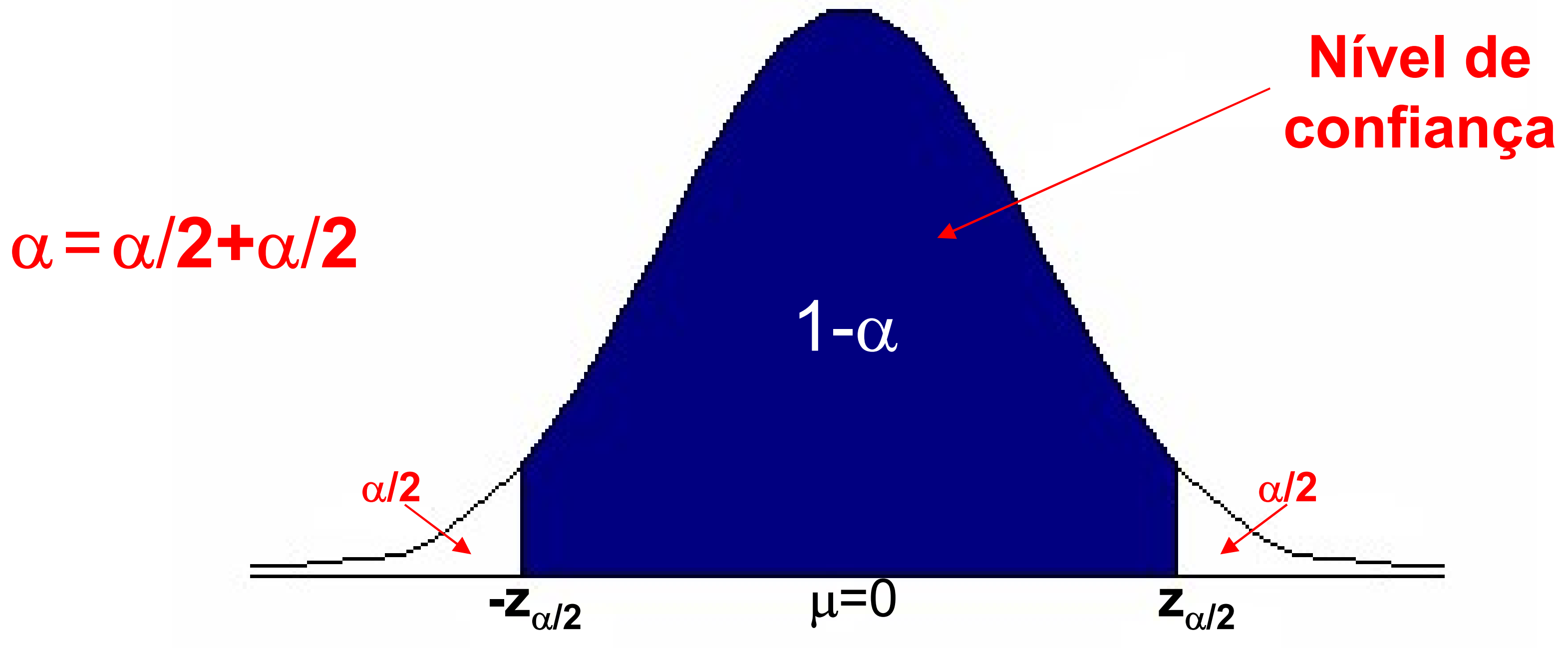
$$Z = \frac{X - \mu}{\sigma}$$

$$\Downarrow$$

$$Z = \frac{\overline{X} - \mu_{\overline{X}}}{\sigma_{\overline{X}}} \quad \Rightarrow \quad Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N\,(0,1)$$

Erro padrão da média

Curva normal padrão



$-z_{\alpha/2}$: valor de Z que delimita a área $\alpha/2$ a esquerda

$z_{\alpha/2}$: valor de Z que delimita a área $\alpha/2$ a direita

α: probabilidade de Z **não assumir** um valor entre $-z_{\alpha/2}$ e $z_{\alpha/2}$

$1-\alpha$: probabilidade de Z **assumir** um valor entre $-z_{\alpha/2}$ e $z_{\alpha/2}$

$P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1-\alpha \longleftarrow$ **nível de confiança**

$$P(-z_{\alpha/2}<Z<z_{\alpha/2})=1-\alpha \text{ , onde: } Z=\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$$

$$P\left(-z_{\alpha/2}<\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}<z_{\alpha/2}\right)=1-\alpha$$

$$P\left(-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}<\bar{X}-\mu<z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)=1-\alpha$$

$$P\left(-\bar{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}<-\mu<-\bar{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)=1-\alpha$$

$$P\left[\left(-\bar{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}<-\mu<-\bar{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)(-1)\right]=1-\alpha$$

$$P\left(\bar{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}>\mu>\bar{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)=1-\alpha$$

$$P\left(\bar{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}<\mu<\bar{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)=1-\alpha$$

1. Intervalo de confiança para a **média de uma** população (μ)

Situação 1: Quando conhecemos o valor de σ (<u>ou</u> n > 30)

$$P\Big(\underbrace{\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}}_{\text{Limite inferior}} < \mu < \underbrace{\overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}}_{\text{Limite superior}}\Big) = 1-\alpha$$

Nível de confiança

1-α é a probabilidade de que os limites contenham (ou cubram) o verdadeiro valor de μ.

$$IC\ (\mu;\ 1-\alpha):\quad \overline{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

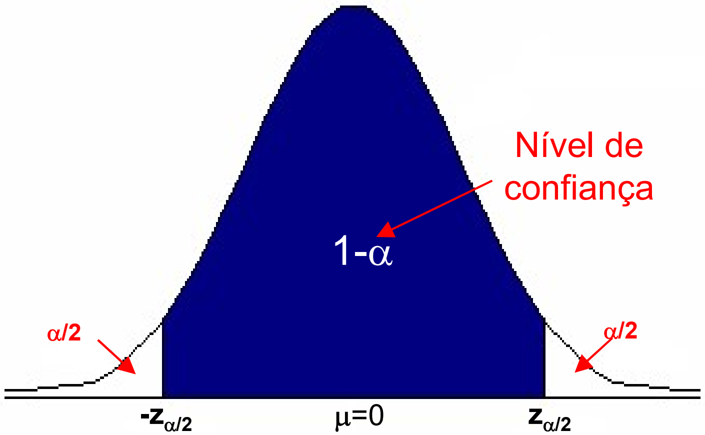
onde:

$\overline{X}$: é o estimador de μ

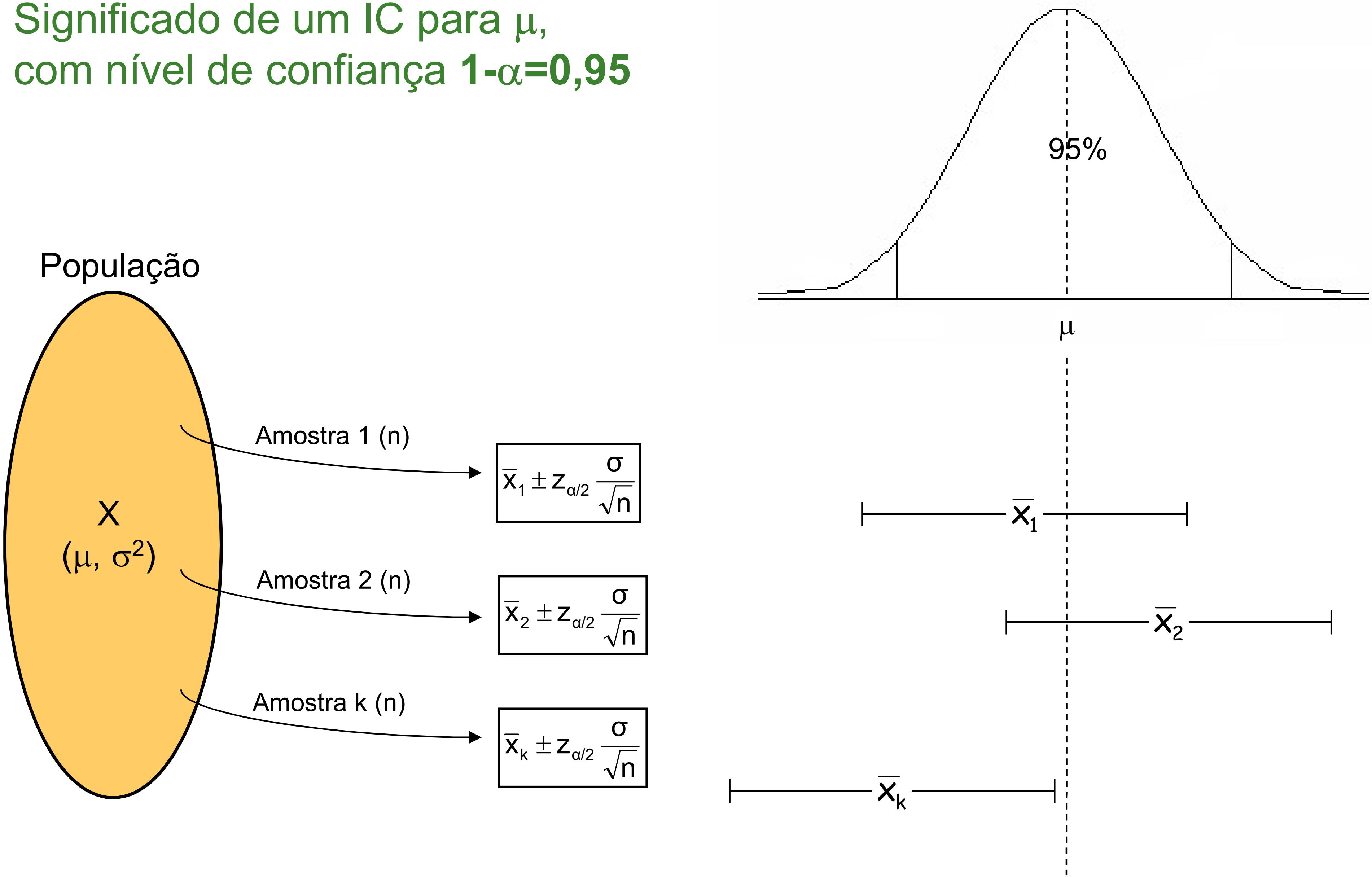
n : é o tamanho da amostra;

σ : é desvio padrão da população

$z_{\alpha/2}$: é o valor de Z que delimita a área $\alpha/2$ (tabela)



Significado de um IC para μ, com nível de confiança **$1-\alpha=0,95$**



95% dos intervalos contêm μ

⇨ Em geral, não conhecemos o parâmetro σ.

⇨ Por isso usamos uma estimativa desse parâmetro que é o **s** (desvio padrão obtido de uma amostra).

Em muitas situações, quando a amostra é grande (n > 30), a estimativa é considerada suficientemente próxima do parâmetro.

Duas pressuposições para a utilização desta metodologia:

1. A variável em estudo tem distribuição normal, $\mathbf{X \sim N(\mu, \sigma^2)}$

2. Conhecemos o valor de σ ou o tamanho da amostra é suficientemente grande para obtenção de uma estimativa aproximada da variação populacional (σ) $\rightarrow$ **n > 30**

Exemplo: Uma amostra de 100 funcionários de uma grande empresa apresentou nota média de 65,5 pontos e desvio padrão de 4,8 pontos para a satisfação com o salário. Obtenha o intervalo de confiança ao nível de 95%, para a verdadeira nota média de satisfação com o salário e conclua.

Variável em estudo: X = nota de satisfação com salário

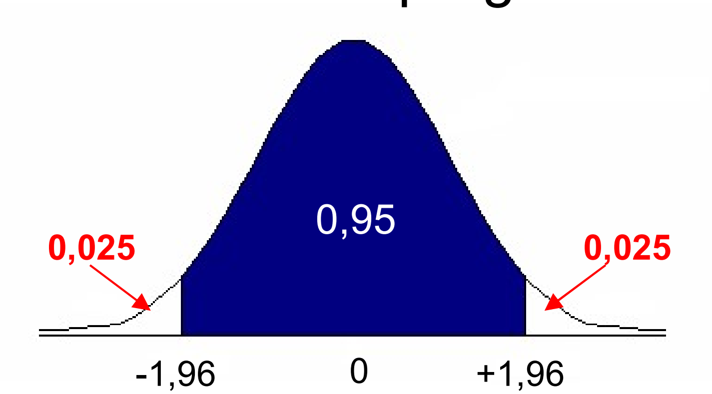
Pressuposição: A variável em estudo tem distribuição normal e n>30.

Estimativas:

$\bar{x} = 65,5$

$s = 4,8$

n = 100 empregados



IC (μ; 1–α): $\bar{X} \pm z_{\alpha/2} \dfrac{\sigma}{\sqrt{n}}$

IC (μ; 0,95): $65,5 \pm 1,96 \times \dfrac{4,8}{\sqrt{100}}$

IC (μ; 0,95): $65,5 \pm 0,941$

Limite inferior = 65,5 – 0,941 = 64,56

Limite superior = 65,5 + 0,941 = 66,44

$P(64,56 < \mu < 66,44) = 0,95$

Concluímos, com 95% de confiança, que a verdadeira nota média de satisfação dos funcionários com os salários deve estar entre 64,56 e 66,44 pontos.

Exercícios:

1) A variabilidade do tempo de atendimento em um consultório conhecida σ = 0,10 min. Uma amostragem com 20 pessoas indicou tempo médio de atendimento de 1,5 minutos. Construa um intervalo de confiança de 95% para o tempo médio de atendimento. **(1,46 ; 1,54)**

2) Suponha o desvio padrão da vida útil de uma determinada marca de tubo de imagem de TV é conhecido e é igual a σ = 500, mas que a média da vida útil é desconhecida. Supõe-se que a vida útil dos tubos de imagem tem uma distribuição aproximadamente normal. Para uma amostra de n = 15, a média da vida útil foi 8.900 horas de operação. Construir:

a) um intervalo de confiança de 95% para estimar a média da população. **(8647 ; 9153)**

b) um intervalo de confiança de 90% para estimar a média da população. **(8687 ; 9113)**

1. Intervalo de confiança para a média de uma população (μ)

Situação 2: Quando não conhecemos o valor de σ (e $n \leq 30$)

⇨ Quando a amostra é pequena, não podemos supor que o desvio padrão da amostra (**s**) seja uma estimativa suficientemente aproximada do parâmetro σ.

Portanto, não podemos utilizar a estatística **Z**, para construir o intervalo de confiança para μ.

$$IC\ (\mu;\ 1-\alpha):\quad \bar{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

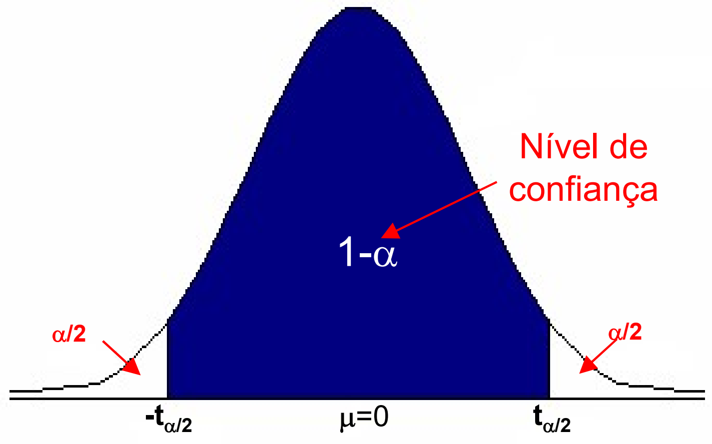
desconhecido

Nesse caso, em vez de **Z**, utilizamos a distribuição **t de Student**, com parâmetro ν.

1. Intervalo de confiança para a média de uma população (μ)

Situação 2: Quando não conhecemos o valor de σ (ou n ≤ 30)

$$P(\underbrace{\overline{X} - t_{\alpha/2}\frac{S}{\sqrt{n}}} < \mu < \underbrace{\overline{X} + t_{\alpha/2}\frac{S}{\sqrt{n}}}) = 1-\alpha$$



$$IC\ (\mu;\ 1-\alpha):\ \overline{X} \pm t_{\alpha/2}\frac{S}{\sqrt{n}}$$

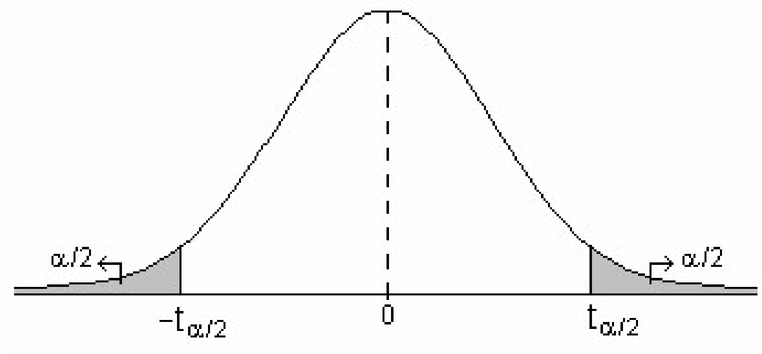
onde: $t_{\alpha/2}$ é o valor da estatística **T** que delimita a área $\alpha/2$

Este valor é encontrado na tabela da **distribuição t** de Student, a partir dos valores de ν e de α.

Pressuposição

A variável em estudo tem distribuição normal ➠ **X ~ N (μ,σ^2)**

Tabela II. Limites da distribuição t de Student.



Graus de Liberdade (ν)	**Limites bilaterais**: $P(\lvert t \rvert > t_{\alpha/2})$							
	Nível de Significância (α)							
	0,50	0,20	0,10	0,05	0,025	0,02	0,01	0,005
1	1,000	3,078	6,314	12,706	25,542	31,821	63,657	127,320
2	0,816	1,886	2,920	4,303	6,205	6,965	9,925	14,089
3	0,715	1,638	2,353	3,183	4,177	4,541	5,841	7,453
4	0,741	1,533	2,132	2,776	3,495	3,747	4,604	5,598
5	0,727	1,476	2,015	2,571	3,163	3,365	4,032	4,773
6	0,718	1,440	1,943	2,447	2,969	3,143	3,707	4,317
7	0,711	1,415	1,895	2,365	2,841	2,998	3,500	4,029
8	0,706	1,397	1,860	2,306	2,752	2,896	3,355	3,833
9	0,703	1,383	1,833	2,262	2,685	2,821	3,250	3,690
10	0,700	1,372	1,813	2,228	2,634	2,764	3,169	3,581

...

Exemplo:

Através da seguinte amostra de tamanho 15, procura-se **estimar a verdadeira potência média de aparelhos eletrônicos** de alta sensibilidade, medida em microwatts, utilizando um intervalo de confiança de 95%:

26,7; 25,8; 24,0; 24,9; 26,4; 25,9; 24,4; 21,7; 24,1; 25,9; 27,3; 26,9; 27,3; 24,8; 23,6

Resolução:

Variável em estudo: X = potência de aparelhos eletrônicos (microwatts)

Pressuposição: A variável em estudo tem distribuição normal.

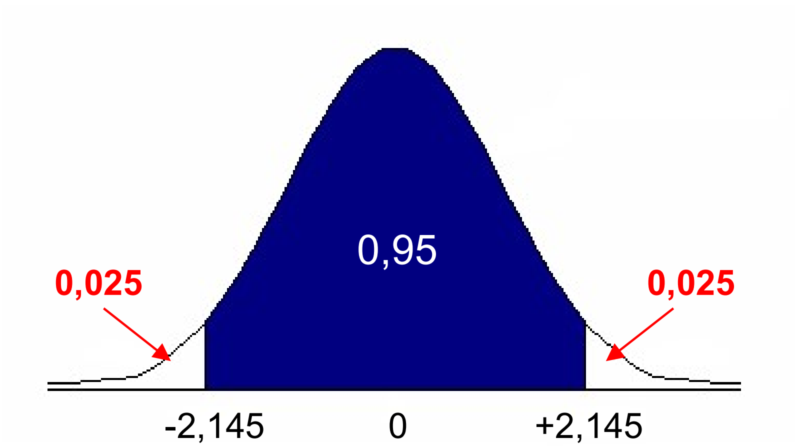
$$IC(\mu;\, 1-\alpha):\ \bar{X} \pm t_{\alpha/2}\,\frac{S}{\sqrt{n}}$$

Estimativas:

$\bar{x}=25{,}31$ $\alpha=0{,}05$

$s=1{,}579$ $\nu=15-1=14$

$n=15$ $t_{0,025;14}=2{,}145$



Construção do intervalo:

$$IC(\mu;1-\alpha): \bar{X} \pm t_{\alpha/2}\frac{S}{\sqrt{n}}$$

$$IC\ (\mu;\ 0{,}95):\ 25{,}31 \pm 2{,}145 \times \frac{1{,}58}{\sqrt{15}}$$

$IC\ (\mu;\ 0{,}95):\ 25{,}31 \pm 0{,}874$

Limite inferior: $25{,}31 - 0{,}874 = 24{,}44$

Limite superior: $25{,}31 + 0{,}874 = 26{,}18$

$P(24{,}44 < \mu < 26{,}18) = 0{,}95$

Conclusão: Com confiança de 95%, pode-se dizer que a verdadeira potência média dos aparelhos eletrônicos deve estar entre 24,44 a 26,18 microwatts.

Exercício:

Um engenheiro de desenvolvimento de um fabricante de pneus está investigando a vida do pneu em relação a um novo componente de borracha.

Ele fabricou 40 pneus e testou-os até o fim da vida em um teste na estrada. A média e o desvio padrão da amostra são 61.492 e 6.085 km, respectivamente. O engenheiro acredita que a vida média desse novo pneu é maior em relação a 60.000 km, que é a vida média do pneu antigo.

Obtenha o intervalo de confiança, ao nível de 99%, para a vida média do pneu e conclua a respeito da suposição do engenheiro.

Resolução:

Variável em estudo: X = vida do pneu (km)

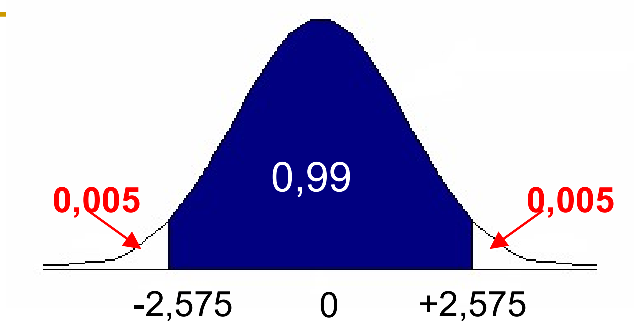
Pressuposições:

1. A variável em estudo tem distribuição normal.
2. A amostra é grande ($n>30$).

Estimativas: $\bar{x}$ = 61.492 km

s = 6.085 km

n = 40 pneus



Construção do intervalo: $IC(\mu; 1-\alpha): \bar{X} \pm z_{\alpha/2} \dfrac{\sigma}{\sqrt{n}}$

$$IC\ (\mu;\ 0{,}99):\ 61.492 \pm 2{,}575 \times \frac{6.085}{\sqrt{40}}$$

$$IC\ (\mu;\ 0{,}99):\ 61.492 \pm 2.477$$

Limite inferior: 61.492 – 2.477 = 59.015

Limite superior: 61.492 + 2.477 = 63.969

$P(59.015 < \mu < 63.969) = 0{,}99$

Conclusão: O intervalo de confiança, ao nível de 99%, para a verdadeira vida média do novo pneu é de 59.015 a 63.969 km.

Como o valor 60.000 km **está** coberto pelo intervalo, a vida média do novo pneu **não** é maior que a do pneu antigo.

Exercício: A quantidade mensal de produtos entregues por uma empresa segue uma distribuição Normal com média e variância desconhecidas. Analise os dados a seguir, que representam uma amostra de 20 meses e construa um intervalo de 95% para a média.

17,4	18,2	18,3	18,8	19,0	19,2	19,3	19,6	19,6	19,9
20,2	20,2	20,5	20,7	20,9	21,0	21,3	21,5	21,9	22,6

$$\overline{X} = 20{,}01$$

$$S = 1{,}34$$

$$19{,}38 \leq \mu \leq 20{,}64$$

Exercício: Como encarregado de compras de um supermercado, suponha que você toma uma amostra aleatória de 12 latas de vagens em conserva de um setor de enlatados. O peso líquido de cada lata de vagens está apresentado na tabela: Supondo que o peso líquido médio por lata seja normalmente distribuído, estimar o peso líquido médio das vagens enlatadas usando um intervalo de confiança de 95%.

Peso por lata (em 10 gramas)	15,7	15,8	15,9	16,0	16,1	16,2
Número de latas	1	2	2	3	3	1

média = 15,97 e desvio-padrão = 0,15 R: (15,88 ; 16,06)

Exercício: A vida média de operação para uma amostra de n = 10 lâmpadas foi de 4.000 horas, com desvio padrão de 200 horas. Supõe-se que o tempo de operação das lâmpadas em geral tenha distribuição aproximadamente normal. Estima a vida média de operação para a população de lâmpadas da qual foi extraída a amostra, usando um intervalo de confiança de 95%. R: (3.857 a 4.143 h)

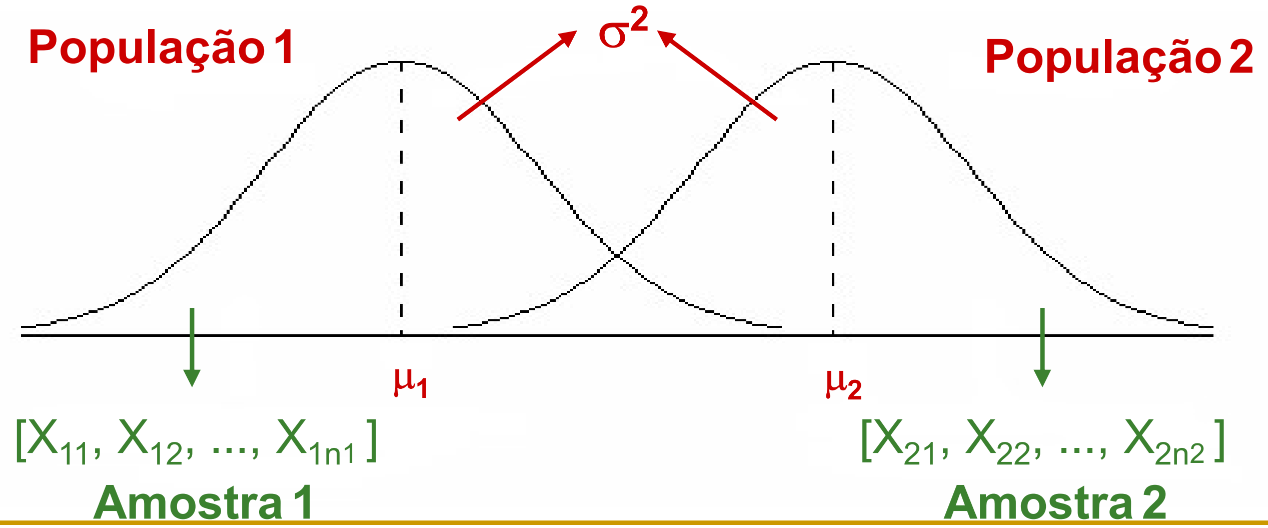
2. Intervalo de confiança para a diferença entre médias de duas populações (μ_1-μ_2)

Há três **pressuposições** que devem ser atendidas para o uso desse procedimento:

1. A variável em estudo tem distribuição normal: **X ~ N (μ,σ^2)**

2. As variâncias das populações são iguais $(\sigma_1^2 = \sigma_2^2 = \sigma)$

3. As amostras retiradas das populações são **independentes**



Atendidas as pressuposições, desejamos comparar as médias das populações, estimando por intervalo, a diferença $\mu_1 - \mu_2$.

Utilizamos, a seguinte expressão

$$IC\ (\mu_1 - \mu_2\ ;\ 1-\alpha):\ \bar{X}_1 - \bar{X}_2 \pm t_{\alpha/2} \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) S^2}$$

$$S^2 = \frac{S_1^2(n_1 - 1) + S_2^2(n_2 - 1)}{(n_1 - 1) + (n_2 - 1)}$$

Variância combinada

onde: $t_{\alpha/2}$ é o valor da estatística **T** que delimita a área $\alpha/2$

Este valor é encontrado na tabela da **distribuição t** de Student a partir dos valores de ν e de α.

Grau de liberdade combinado

$$\nu = (n_1 - 1) + (n_2 - 1)$$

Exemplo:

Dez cobaias adultas criadas em laboratório, foram separadas, aleatoriamente, em dois grupos: um foi tratado com ração normalmente usada no laboratório (padrão) e o outro grupo foi submetido a uma nova ração (experimental). As cobaias foram pesadas no início e no final do período de duração do experimento. Os ganhos de peso (em gramas) observados foram os seguintes:

Ração experimental	220	200	210	220	210
Ração padrão	200	180	190	190	180

Construa o intervalo de confiança, ao nível de 99%, para a diferença entre as médias das duas populações.

Resolução:

Variável em estudo: X = ganho de peso (g)

Pressuposições:

1. A variável em estudo tem distribuição normal.
2. As variâncias das populações são iguais ($\sigma_1^2 = \sigma_2^2 = \sigma^2$).
3. As amostras retiradas das populações são independentes.

Estimativas:

$$IC(\mu_1 - \mu_2; 1-\alpha): \bar{X}_1 - \bar{X}_2 \pm t_{\alpha/2}\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)S^2}$$

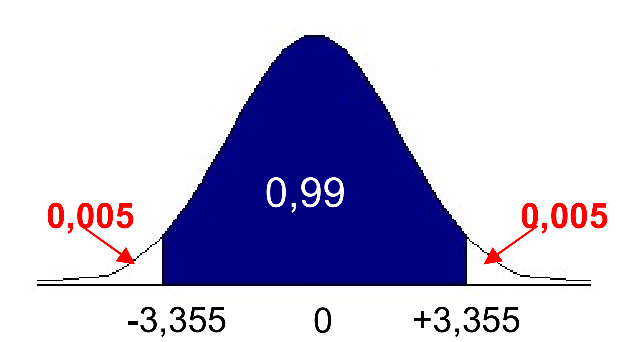
Experimental	Padrão
$\bar{x}_1 = 212$	$\bar{x}_2 = 188$
$s_1^2 = 70$	$s_2^2 = 70$
$n_1 = 5$	$n_2 = 5$

$$s^2 = \frac{70 \times (5-1) + 70 \times (5-1)}{(5-1)+(5-1)} = 70$$

$\alpha = 0,01$

$\nu = (5 - 1) + (5 - 1) = 8$

$t_{0,005;8} = 3,355$



Construção do intervalo:

$$IC(\mu_1 - \mu_2; 1-\alpha): \overline{X}_1 - \overline{X}_2 \pm t_{\alpha/2}\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)S^2}$$

$$IC(\mu_1 - \mu_2; 0{,}99): (212 - 188) \pm 3{,}355\sqrt{\left(\frac{1}{5}+\frac{1}{5}\right)70}$$

$$IC(\mu_1 - \mu_2; 0{,}99): 24 \pm 17{,}78$$

Limite inferior: 24 – 17,78 = 6,22

Limite superior: 24 + 17,78 = 41,78

$$P(6{,}22 < \mu_1 - \mu_2 < 41{,}78) = 0{,}99$$

Conclusão: A probabilidade de o intervalo de 6,22 a 41,78 conter a diferença entre a média de ganho de peso da população que recebeu ração experimental e a média de ganho de peso da população que recebeu a ração padrão é de 0,99.

Como o zero (0) está **fora** do intervalo, podemos concluir que as médias são iguais **diferentes**, sendo a ração experimental melhor que a padrão.

Exercício:

Duas máquinas são usada para encher pacotes de leite. O volume segue aproximadamente o modelo normal. Duas amostras de 16 pacotes de cada máquina indicou:

Máquina A

1021	1016	1012	1011	1014	1018	1022	1027
1008	1015	1013	1013	1017	1019	1007	1003

Máquina B

1011	1015	1017	1015	1021	1021	1010	1007
1022	1018	1016	1015	1020	1022	1025	1030

Construa um intervalo de 95% para a diferença entre as duas médias das máquinas. Baseado nos resultados desses cálculos você concluiria que as duas máquinas fornecem o mesmo volume médio?

média =	1014,75	1017,813
desvio =	6,049793	5,84487
variância =	36,6	34,1625
n=	16	16

S_p^2= 35,38125 S^2(diff) = 4,422656

média 2 - média 1= -3,0625 (diferença das médias amostrais)

valor t = 2,042272

S(diff) = 2,103011

lim.inf = **-7,3574**

lim.sup= **1,2324**

Sim, pois o valor zero está contido no intervalo

Exercício:

Um eixo deve ser montado no interior de um rolamento. Uma amostra de doze unidades indicou para o diâmetro interno do rolamento $\overline{X}_1 = 2{,}538\text{cm}$ e $S_1 = 0{,}008$; e para o diâmetro do eixo $\overline{X}_2 = 2{,}520\text{cm}$ e $S_2 = 0{,}006$. Supondo variâncias iguais, calcule o intervalo de confiança de 99% para a folga de montagem.

$$S_p^{\,2} = \frac{(11)0{,}008^2 + (11)0{,}006^2}{12+12-2} = 0{,}000050$$

$$S = \sqrt{S_p^{\,2}\left(\frac{1}{12}+\frac{1}{12}\right)} = 0{,}00289$$

$$\nu = 12+12-2 = 22$$

$$t_{0{,}005;22} = 2{,}819$$

$$(2{,}538 - 2{,}52) - 2{,}819(0{,}00289) \le \text{folga} \le (2{,}538 - 2{,}52) + 2{,}819(0{,}00289)$$

$$0{,}00985 \le \text{folga} \le 0{,}02615$$

Intervalo de confiança para a variância

Outra distribuição importante, definida a partir da distribuição Normal é a distribuição qui-quadrado χ^2.

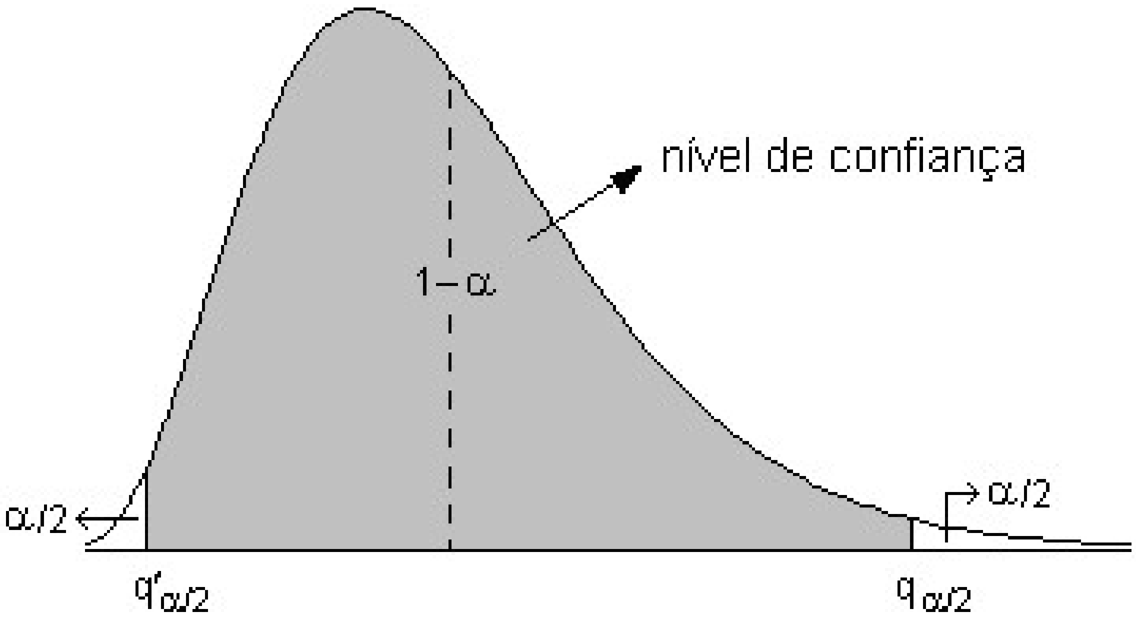
A distribuição χ^2 é a base para inferências a respeito da variância σ^2.

Pressuposição:

X é uma variável com Distribuição Normal com média e variância desconhecidas.

Então, um intervalo de confiança é obtido usando-se a seguinte expressão:

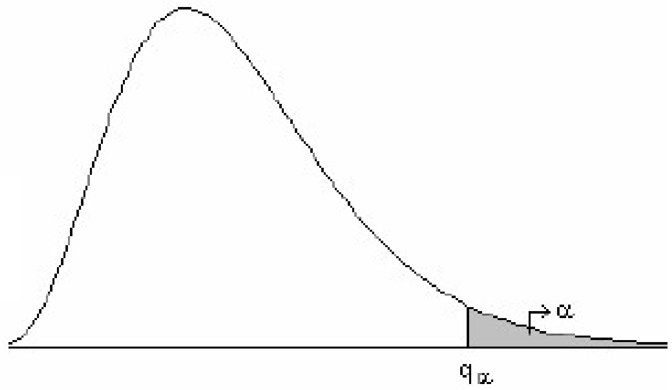
$$\boxed{P\left(\frac{(n-1)S^2}{q_{\alpha/2,n-1}} \leq \sigma^2 \leq \frac{(n-1)S^2}{q'_{\alpha/2,n-1}}\right) = 1-\alpha}$$



$$\left(\frac{(n-1)S^2}{q_{\alpha/2,n-1}} ; \frac{(n-1)S^2}{q'_{\alpha/2,n-1}} \right)$$

- S^2 é o estimador da variância populacional σ^2;
- n é o tamanho da amostra;
- $\nu = n-1$ é o número de graus de liberdade associado à variância;
- $q'_{\alpha/2}$ é o valor da distribuição qui-quadrado, com ν=n-1 graus de liberdade, que delimita a área $\alpha/2$ à esquerda;
- $q_{\alpha/2}$ é o valor da distribuição qui-quadrado com ν=n-1 graus de liberdade que delimita a área $\alpha/2$ à direita.

Tabela III. Limites unilaterais da distribuição qui-quadrado (χ^2).



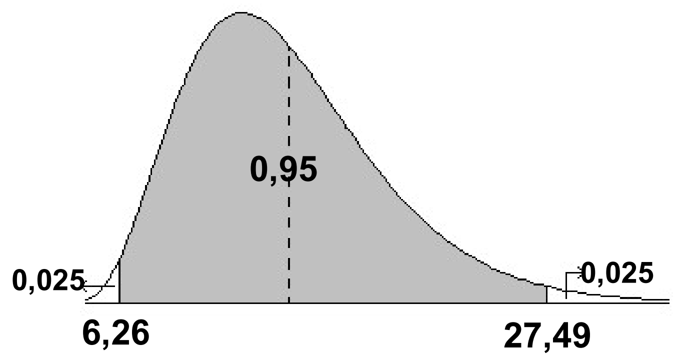
Graus de Liberdade	Nível de Significância (α)									
	Esquerda (q')					Direita (q)				
(ν)	0,005	0,01	0,025	0,05	0,1	0,1	0,05	0,025	0,01	0,005
1	0,00	0,00	0,00	0,00	0,02	2,71	3,84	5,02	6,63	7,88
2	0,01	0,02	0,05	0,10	0,21	4,61	5,99	7,38	9,21	10,60
3	0,07	0,11	0,22	0,35	0,58	6,25	7,81	9,35	11,34	12,84
4	0,21	0,30	0,48	0,71	1,06	7,78	9,49	11,14	13,28	14,86
5	0,41	0,55	0,83	1,15	1,61	9,24	11,07	12,83	15,09	16,75
6	0,68	0,87	1,24	1,64	2,20	10,64	12,59	14,45	16,81	18,55
7	0,99	1,24	1,69	2,17	2,83	12,02	14,07	16,01	18,48	20,28
8	1,34	1,65	2,18	2,73	3,49	13,36	15,51	17,53	20,09	21,95
9	1,73	2,09	2,70	3,33	4,17	14,68	16,92	19,02	21,67	23,59
10	2,16	2,56	3,25	3,94	4,87	15,99	18,31	20,48	23,21	25,19

...

Exemplo: Uma das maneiras de manter sob controle a qualidade de um produto é controlar sua variabilidade. Uma máquina de encher pacotes de café está regulada para enchê-los conforme uma distribuição normal com média de 500g. Colheu-se uma amostra de 16 pacotes e observou-se uma variância de 169g^2. Com esse resultado encontrar o intervalo de confiança de 95% para σ^2.

$s^2 = 169 \qquad n = 16$

$q'_{0,025;15} = 6,26 \qquad q_{0,025;15} = 27,49$



$$\boxed{\frac{(n-1)S^2}{q_{\alpha/2,n-1}} \le \sigma^2 \le \frac{(n-1)S^2}{q'_{\alpha/2,n-1}}}$$

$$\frac{15(169)}{27,49} \le \sigma^2 \le \frac{15(169)}{6,26}$$

$$92,22 \le \sigma^2 \le 404,95$$

ou

$$9,60 \le \sigma \le 20,12$$

Exercício: A quantidade mensal de produtos entregues por uma empresa segue uma distribuição Normal com média e variância desconhecidas. Analise os dados a seguir, que representam uma amostra de 20 meses e construa um intervalo de 95% para a variância da quantidade mensal de produtos entregues.

17,4	18,2	18,3	18,8	19,0	19,2	19,3	19,6	19,6	19,9
20,2	20,2	20,5	20,7	20,9	21,0	21,3	21,5	21,9	22,6

$$\overline{X} = 20{,}01$$

$$S = 1{,}34$$

$$1{,}04 \leq \sigma^2 \leq 3{,}84$$

ou

$$1{,}02 \leq \sigma \leq 1{,}96$$

Intervalo de confiança para a proporção (π)

Exemplos: proporção de produtos defeituosos em um lote ou proporção de eleitores que votarão em um candidato.

Pressuposição:
n é grande $\Rightarrow$ np > 5 e n(1-p) > 5 (p é o estimador de π) $p = \frac{x}{n}$

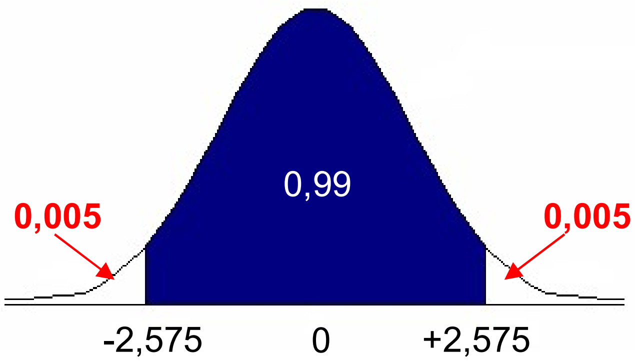
Assim, a aproximação Normal pode ser usada, resultando no seguinte intervalo de confiança:

$$\boxed{P\left(p - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \leq \pi \leq p + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right) = 1-\alpha}$$

$$\Downarrow$$

$$IC(\pi; 1\text{-}\alpha) = \left(p \pm z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right)$$

Exemplo: Foi realizada uma pesquisa de mercado para verificar a preferência da população de em relação ao consumo de determinado produto. Para isso, colheu-se uma amostra de 300 consumidores e, destes, 180 disseram consumir o produto. Encontre o intervalo de confiança de 99% para a proporção de consumidores do produto na população.



$n = 300 \qquad p = {}^{180}\!/_{300} = 0,6$

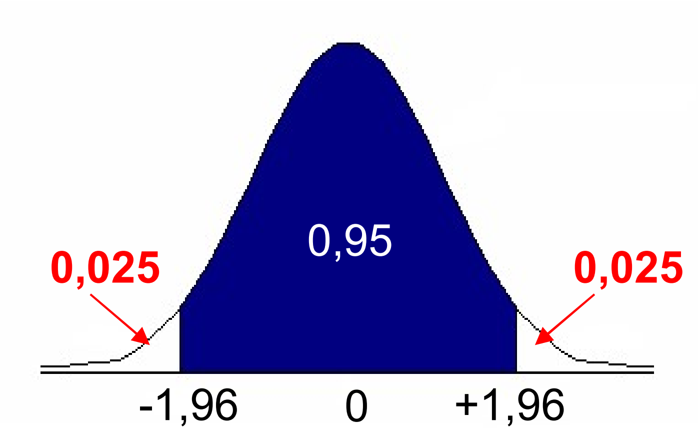
$$IC(\pi; 1\text{-}\alpha) = \left(p \pm z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \right)$$

$$\left(0,6 \pm 2,575\sqrt{\frac{0,6(1-0,6)}{300}} \right)$$

$$(0,6 \pm 0,0728) = (0,5272; 0,6728)$$

Com uma confiança de 99%, a proporção de consumidores que preferem o produto pesquisado deve estar entre 52,72% e 67,28%.

Exercício: Um empresário deseja conhecer a satisfação de seus clientes em relação aos serviços prestados por sua empresa. Em uma amostra aleatória de n=100 clientes entrevistados, 12 pessoas demonstraram insatisfação com os serviços prestados. Construa um intervalo de 95% de confiança para a proporção de clientes insatisfeitos. R: (5,63% ; 18,37%)



$$n = 100 \qquad p = {}^{12}\!/_{100} = 0{,}12$$

$$IC(\pi; 1-\alpha) = \left(p \pm z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \right)$$

$$\left(0{,}12 \pm 1{,}96\sqrt{\frac{0{,}12(1-0{,}12)}{100}} \right) = (0{,}0563; 0{,}1887)$$

Com 95% de confiança, a proporção de pessoas insatisfeitas com os serviços prestados deve estar entre 5,63% e 18,37%.

Exercícios:

1) Numa pesquisa de mercado, 400 pessoas foram entrevistadas sobre sua preferência por determinado produto. Destas 400 pessoas, 240 disseram preferir o produto. Determinar um intervalo de confiança de 95% de probabilidade para o percentual de preferência dos consumidores em geral para este produto. R: (55,20% ; 64,80%)

2) Um administrador de uma universidade coleta dados sobre uma amostra aleatória de âmbito nacional de 230 alunos de cursos de Administração de Empresas e encontra que 54 de tais estudantes têm diplomas de Técnico em Contabilidade. Usando um intervalo de confiança de 90%, estimar a proporção nacional de estudantes que possuem diplomas de Técnico de Contabilidade. R: (19% ; 28%)

Dimensionamento de amostra

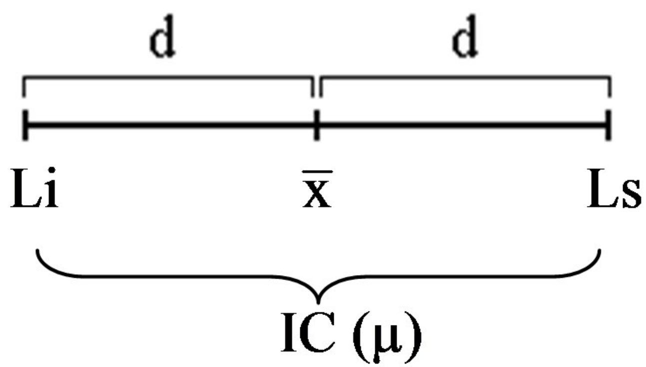
Os intervalos de confiança para média tem as formas:

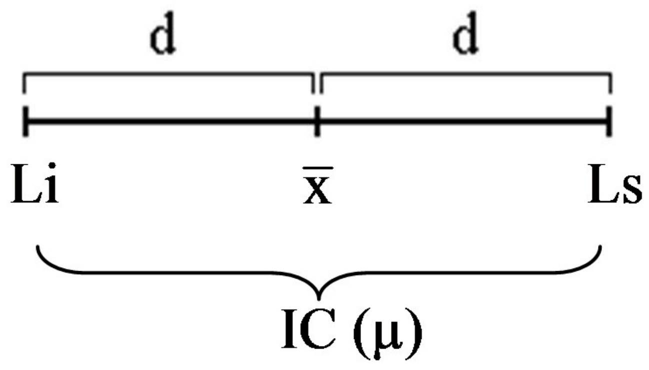
$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \qquad\qquad \bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

A semi-amplitude do intervalo de confiança, que é a precisão da estimação, é dada por

$$d = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$d = t_{\alpha/2} \frac{s}{\sqrt{n}}$$





- O grau de precisão é utilizado para estabelecer a semi-amplitude desejada para o intervalo de confiança, e é comum que ele seja expresso em percentual.
- Tomando d = $\gamma \bar{x}$, com γ entre 0 e 1, estabelecemos o grau de precisão como uma percentagem da média.
- Por exemplo, tomar γ = 0,10 significa que pretendemos obter um tamanho de amostra n tal que, no intervalo de confiança para μ, tenhamos uma semi-amplitude máxima que corresponda a 10 % do valor da média amostral.
- Então d será a magnitude (na unidade de medida que a variável em estudo é medida) dessa semi-amplitude máxima desejada.

➢Da expressão, isolando n temos

$$d = z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \Rightarrow \boxed{n = z^2_{\alpha/2} \frac{\sigma^2}{d^2}}$$

Variância populacional

$$d = t_{\alpha/2} \frac{s}{\sqrt{n}} \Rightarrow \boxed{n = t^2_{\alpha/2} \frac{s^2}{d^2}}$$

Estimativa da variância através de amostra piloto

Exemplo: Qual o tamanho da amostra necessário para estimar a média populacional de uma característica dimensional de um processo com 95% de confiança e precisão de 0,5 cm ? Sem conhecimento da variabilidade populacional, estima-se o desvio padrão populacional através de uma amostra piloto.

7	11	12	11	13	8	15	8	11	16
10	12	9	6	11	10	11	10	12	9

Solução:

$\alpha = 0,05$

$t_{0.025,19} = 2,093$

$d = 0,5$ cm

$s_1 = 2,46$

$$n = t^2_{\alpha/2}\frac{s^2}{d^2} = 2,093^2\frac{2,46^2}{0,5^2} = 106$$

Logo, é necessária uma amostra de pelos menos 106 observações, devendo ser coletadas mais 86.

$$IC(\pi;1\text{-}\alpha)=\left(p\pm \underbrace{z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}}_{\text{precisão (d)}}\right) \Rightarrow \boxed{n=\frac{z_{\alpha/2}^{2}\,p(1-p)}{d^{2}}}$$

Exemplo: O fornecedor alega que entrega 10% de produtos defeituosos. Qual o tamanho de amostra suficiente para estimar a proporção de produtos defeituosos entregues por este fornecedor com precisão de 0,03 e 95% de confiança?

Solução:

$\alpha = 0{,}05 \Rightarrow z_{0,025} = 1{,}96$; $p = 0{,}10$; $d = 0{,}03$

$$n=\frac{1{,}96^{2}\,.\,0{,}10\,.\,(1-0{,}10)}{0{,}03^{2}}=384{,}16$$

Logo, é necessária uma amostra de 385 produtos.

Exercícios:

1) Um analista do departamento de pessoal deseja estimar o número médio de horas de treinamento anual para os capatazes de uma divisão da companhia, com um erro de 3,0 horas e com 90% de confiança. Baseado em dados de outras divisões, ele estima o desvio padrão das horas de treinamento em $\sigma = 20{,}0$ horas. Qual o tamanho mínimo necessário da amostra? (R: n = 121)

2) Um comprador potencial deseja estimar o valor médio das compras por cliente em uma loja de brinquedos em um aeroporto. Com base em uma amostra piloto de 40 vendas, o desvio padrão dos valores de vendas foi estimado em s = $0,80. Qual o tamanho mínimo que deveria ter uma amostra se ele deseja estimar a média das vendas dentro de $0,25 com confiança de 99%? (R: n = 75)

Exercícios:

3) Qual o tamanho da amostra necessário para estimar a média populacional de uma característica dimensional de um processo com 95% de confiança e precisão de 0,5cm ?

Sem conhecimento da variabilidade populacional, estima-se o desvio-padrão populacional através de uma amostra piloto.

7	11	12	11	13	8	15	8	11	16
10	12	9	6	11	10	11	10	12	9

(R: n = 106, logo é necessário coletar mais 86 (106-20) peças)

4) Quer-se estimar a proporção de porto-alegrenses maiores de 16 anos que são favoráveis à flexibilização das leis trabalhistas. Qual o tamanho mínimo da amostra necessário para um erro absoluto de estimação de no máximo 2 pontos percentuais, com uma confiança de 95%? (R: n = 2401)