

# **Estatística**

## **Intervalo de Confiança**

**Prof. Antonio Estanislau Sanches  
2017**

# Estatística Inferencial

Os dois principais procedimentos da estatística inferencial são:

- os estimadores por intervalo e
- os testes de hipóteses.

Normalmente os parâmetros de um processo ou produto (*temperatura média em um forno; comprimento médio de uma distância*) são desconhecidos, sendo necessário ESTIMAR o valor desses parâmetros e viabilizar testes para verificar se os valores, a eles arbitrados (*estimados*), podem ser considerados verdadeiros.

## Estimação

É um ramo da estatística inferencial que trata da estimativa de parâmetros desconhecidos da população, a partir dos dados de uma amostra. Tais estimativas podem ser:

**PONTUAIS** ou **INTERVALARES.**

## **ESTIMATIVA PONTUAL:**

o valor do parâmetro é IGUAL ao valor obtido da amostra;

## **ESTIMATIVA INTERVALAR:**

o valor do parâmetro será IGUAL ao valor obtido da amostra +/- uma variação;

Essa **variação** é diretamente proporcional à **DISPERSÃO** populacional, ou seja, quanto mais dispersa for a população, tanto maior será a variação entre os valores amostrais.

Por outro lado, o tamanho do intervalo de variação é inversamente proporcional a confiança de que ele contenha o parâmetro procurado, ou para se ter maior certeza (*confiança*) o tamanho da variação será maior.

Ou seja, quanto maior a amostra, maior será a precisão, porém, menor será a variação.

# ESTIMATIVA INTERVALAR

A estimação intervalar consiste na determinação de um intervalo onde, com uma certa confiança (*probabilidade*), pode-se prever que ali esteja o parâmetro  $\theta$  desconhecido, tendo-se em conta um estimador desse parâmetro.

Assim,  $P(L_1 < \theta < L_2) = 1 - \alpha \rightarrow$  significando que a probabilidade do intervalo aleatório  $L_1$  a  $L_2$  conter o valor exato  $\theta$  é de  $\alpha$ .

O intervalo é designado por um intervalo de confiança para o parâmetro  $\theta$ , com um nível de confiança  $\alpha$ .

# ESTIMAÇÃO INTERVALAR

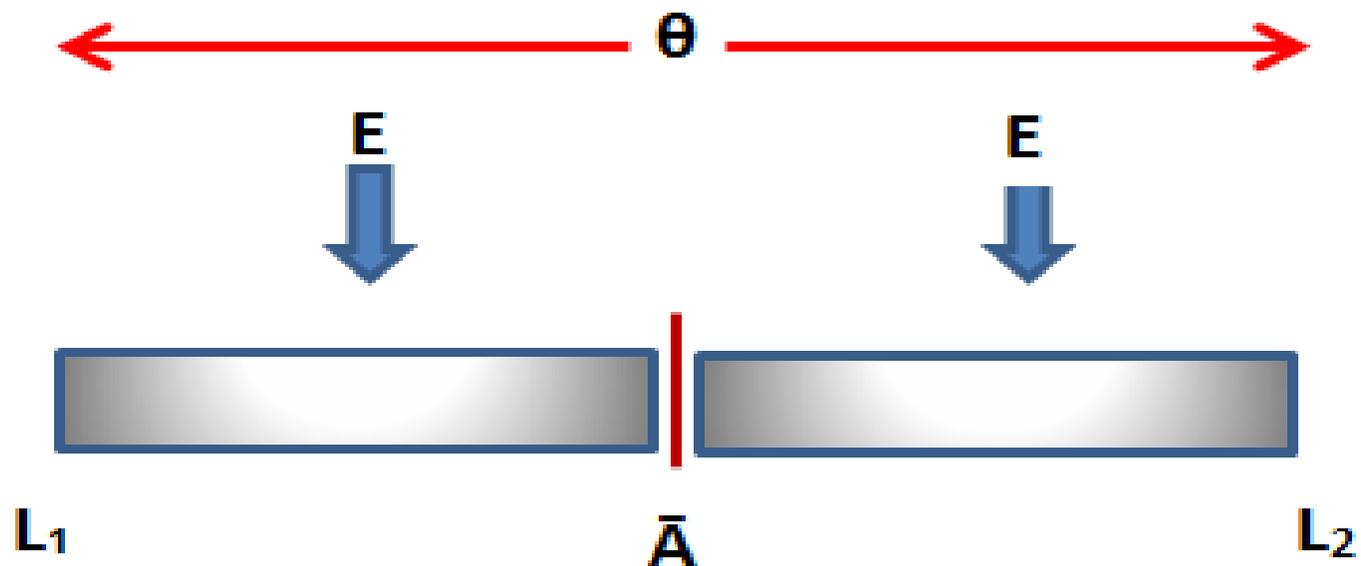
$L_1$  Limite inferior do intervalo

$L_2$  Limite superior do intervalo

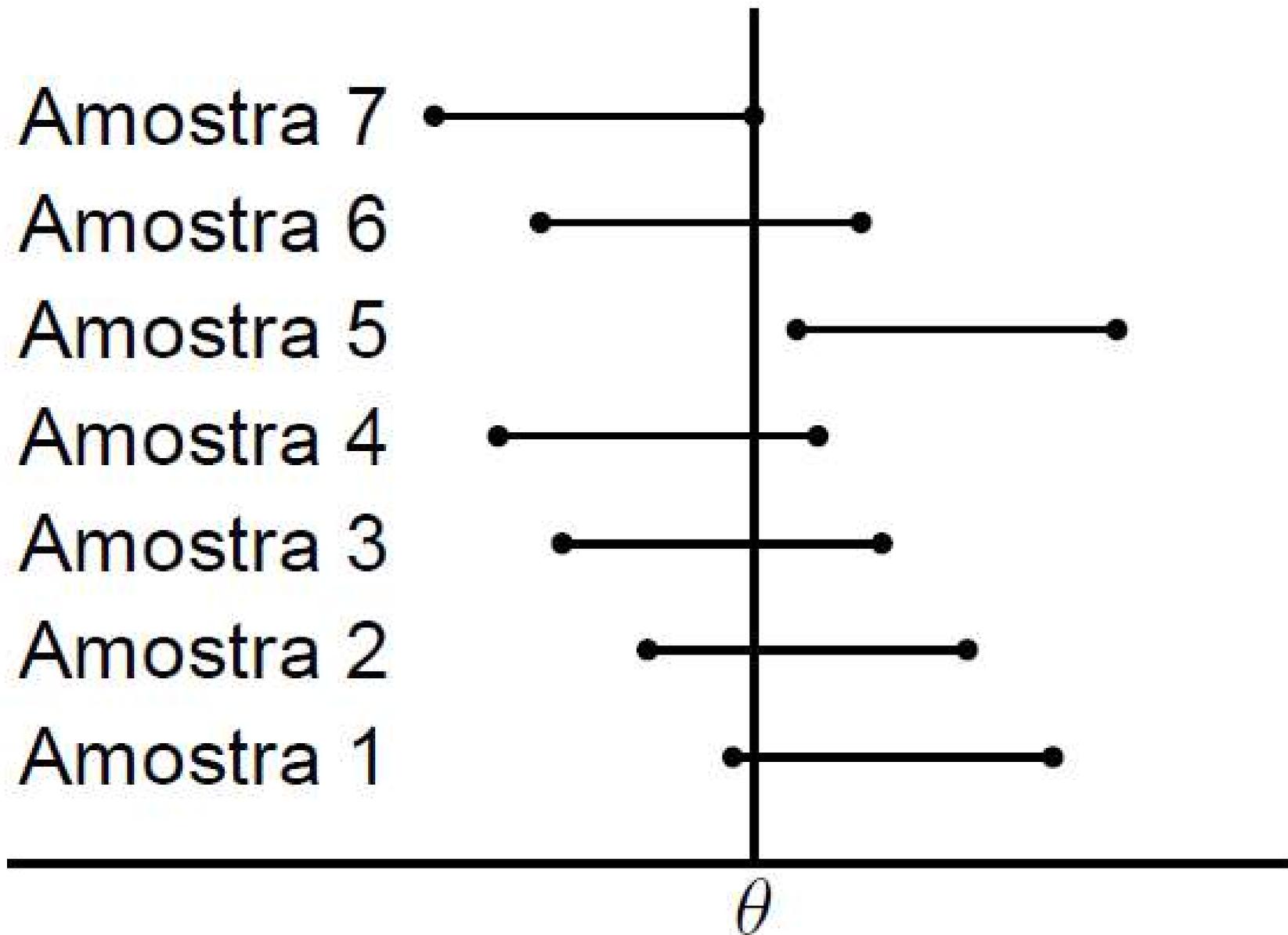
$E$  Comprimento do intervalo

$\bar{A}$  Média amostral

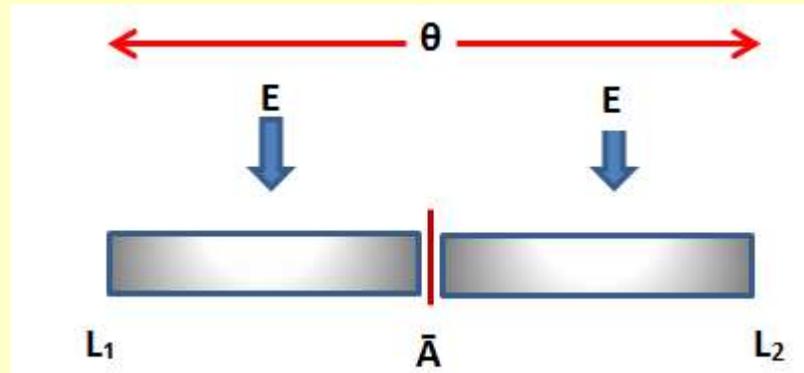
$\theta$  Média Populacional



# ESTIMAÇÃO INTERVALAR



# ESTIMAÇÃO INTERVALAR



## Propriedades dos I.C.:

- 1) Diminuindo-se o tamanho do intervalo ( $E$ ), aumenta-se o tamanho da amostra ( $n$ ) para um valor fixo de nível de confiança ( $\alpha$ );
- 2) Aumentando-se o nível de confiança ( $\alpha$ ), também aumenta-se o tamanho da amostra ( $n$ );

Na estatística, toda e qualquer afirmação deve vir acompanhada de um grau de confiança.

A estimativa por intervalos de confiança, consiste em determinar um INTERVALO definido por dois valores extremos  $L_1$  e  $L_2$ , tal que a probabilidade de que o parâmetro verdadeiro esteja contido nesse intervalo, seja  $(1-\alpha)$ , isto é:

$$P(L_1 < \theta < L_2) = 1 - \alpha$$

**Exemplo:** seja um I.C. com 90% de confiança, para estimar a média de uma população.

Embora se desconheça o verdadeiro valor da média populacional, se construirmos 100 intervalos, a partir de 100 amostras diferentes, espera-se que em 90 desses intervalos, estejam contidos o verdadeiro valor da média populacional.

$\alpha$  ► Nível de significância;

$1 - \alpha$  ► Nível de confiança

# EXEMPLO 1

A média e o desvio-padrão da resistência de um material sintético usado na confecção de carpete é, respectivamente, 100N e 20N. Se uma amostra de tamanho  $N = 40$ , for aleatoriamente selecionada, qual a probabilidade da média amostral ser menor que 105N?

$$s = \frac{\sigma}{\sqrt{N}} \quad \blacktriangleright \quad s = \frac{20}{\sqrt{40}} \quad \blacktriangleright \quad \underline{s = 3,162}$$

$$Z = \frac{X - \mu}{s} \quad \blacktriangleright \quad \underline{Z = \frac{105 - 100}{3,162}} \quad \blacktriangleright \quad Z = 1,58$$

`PADRONIZAR(105;100;20/RAIZ(40)) => Z = 1,58`

`DIST.NORMP.N(1,58;1) ▶ 94,29%`

$$P(X < 105) = 94,29\%$$

## IC (Intervalo de Confiança) da MÉDIA para grandes amostras (N ≥ 30):

Com reposição de eventos:

$$P\left[\bar{X} - Z \frac{\sigma}{\sqrt{N}} \leq \mu \leq \bar{X} + Z \frac{\sigma}{\sqrt{N}}\right]$$

Sem reposição de eventos:

$$P\left[\bar{X} - Z * \frac{\sigma}{\sqrt{n}} * \sqrt{\frac{N-n}{N-1}} \leq \mu \leq \bar{X} + Z * \frac{\sigma}{\sqrt{n}} * \sqrt{\frac{N-n}{N-1}}\right]$$

N = população

n = amostra

## IC da MÉDIA p/ amostras pequenas (N < 30):

Quando  $\sigma$  for conhecido:

$$P\left[\left(\bar{X} - Z \frac{\sigma}{\sqrt{n}}\right) < \mu < \left(\bar{X} + Z \frac{\sigma}{\sqrt{n}}\right)\right] \text{ p/ } \alpha/2 ; \text{ sendo n (amostra)}$$

Quando  $\sigma$  for desconhecido:

$$P\left[\left(\bar{X} - t \frac{\sigma}{\sqrt{n}}\right) < \mu < \left(\bar{X} + t \frac{\sigma}{\sqrt{n}}\right)\right] \text{ p/ } \alpha/2 ; \text{ sendo n (amostra) e } (t_{\alpha/2} ; n-1)$$

## Intervalo de confiança para a média:

Seja uma *amostra aleatória* (aa):  $X_1, X_2, \dots, X_n$ , com média  $\mu$  (desconhecida) e variância  $\sigma^2$  conhecida.

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

### Onde:

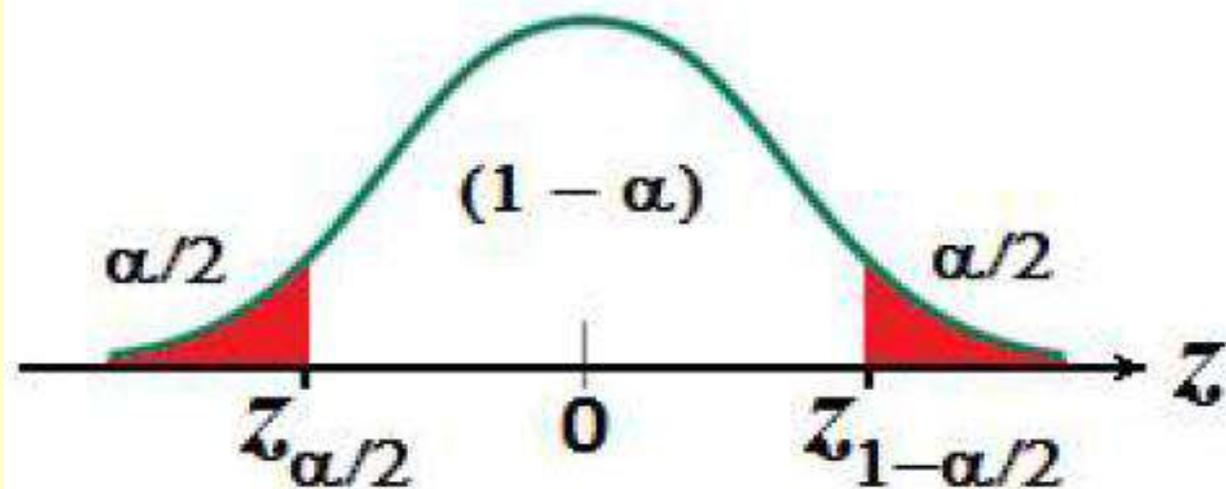
$\mu$  → **média populacional** (desconhecida);

$\bar{X}$  → **média amostral** (calculada);

$\sigma$  → **desvio padrão populacional** (conhecido);

$n$  → **tamanho da amostra**.

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



## Intervalo de confiança para a média:

Seja uma *amostra aleatória* (aa):  $X_1, X_2, \dots, X_n$ , com média  $\mu$  (desconhecida) e variância  $\sigma^2$  conhecida.

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

$$P\left(\bar{X} - Z_{(1-\alpha/2)} * \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + Z_{(1-\alpha/2)} * \frac{\sigma}{\sqrt{n}}\right)$$

Limite inferior

Limite superior

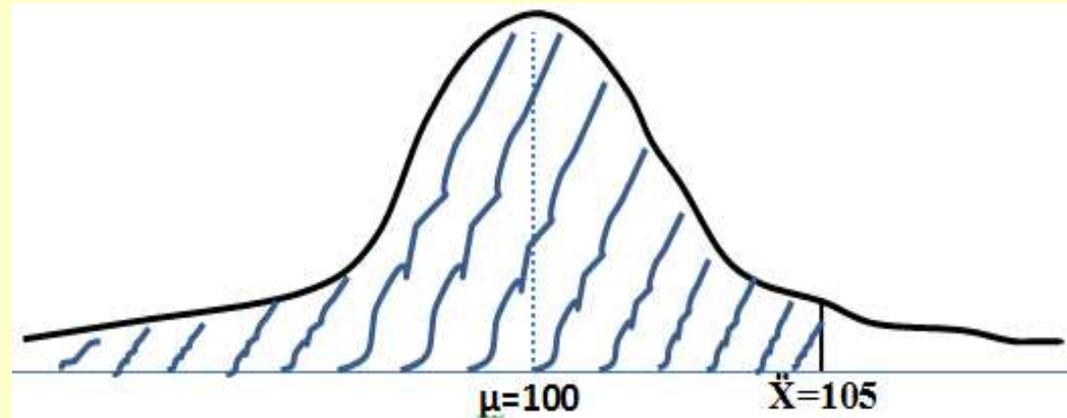
Se  $\alpha = 5\%$ , então:  $Z_{(1-\alpha/2)} = \text{INV.NORMP}(1-0,05/2) = 1,96$

Portanto, um *intervalo de confiança*  $(1 - \alpha) \times 100\%$  para  $\mu$ , com  $\sigma^2$  conhecido, é dado por:

$$\left[ \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} ; \bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right].$$

## EXEMPLO 2

A média e o desvio-padrão da resistência de um material sintético usado na confecção de carpete é, respectivamente, 100N e 20N. Se uma amostra de tamanho  $N = 40$ , for aleatoriamente selecionada, qual a probabilidade da média amostral ser menor que 105N?



$$s = \frac{\sigma}{\sqrt{N}} \quad \blacktriangleright \quad s = \frac{20}{\sqrt{40}} \quad \blacktriangleright \quad s = 3,162$$

$$Z = \frac{X - \mu}{s} \quad \blacktriangleright \quad Z = \frac{105 - 100}{3,162} \quad \blacktriangleright \quad Z = 1,58$$

$$\text{DIST.NORMMP.N}(1,58;1) \quad \blacktriangleright \quad 94,29\%$$

$$P\left[\left(105 - 1,58 \frac{20}{\sqrt{40}}\right) < \mu < \left(105 + 1,58 \frac{20}{\sqrt{40}}\right)\right] = 94,29\%$$

$$P[(105 - 5) < \mu < (105 + 5)] = 94,29\%$$

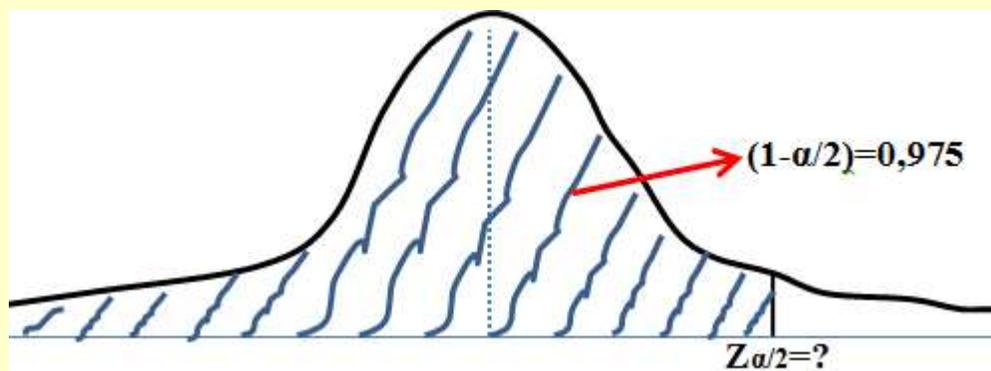
↑  
**E**

↑  
**E**

## EXEMPLO 3

Seja  $X$  a duração da vida de uma peça de aeronave, tal que  $\sigma=5$  horas. Admitindo-se que o ensaio de 100 peças, gerou uma vida média de  $\bar{X} = 500$  horas e que se deseja obter um IC de 95% para a média  $\mu$ .

Utilizar: IC da MÉDIA p/ amostras grandes ( $N \geq 30$ )



$$\text{INV.NORMP}(1-0,05/2) = 1,96$$

$$P\left[\left(500 - 1,96 \frac{5}{\sqrt{100}}\right) < \mu < \left(500 + 1,96 \frac{5}{\sqrt{100}}\right)\right]$$

Cálculo do IC  $\rightarrow 1,96 \frac{5}{\sqrt{100}}$  ou

$$\text{INT.CONFIANÇA.NORM}(0,05;5;100) = 0,98$$

$$\text{INV.NORM}(1-0,05/2;500;5/\text{RAIZ}(100)) = 500,98$$

$$P[(500 - 0,98) < \mu < (500 + 0,98)] = 95,0\%$$

↑  
**E**

↑  
**E**

# Problemas existentes na prática:

- 1) Estimar a média ( $\mu$ ) de uma população;
- 2) Estimar a diferença entre as médias de duas populações ( $\mu_A - \mu_B$ );
- 3) A proporção de uma população é conhecida por ( $p$ );
- 4) A diferença entre a proporção de duas populações é ( $p_1 - p_2$ );
- 5) A variância de uma população é conhecida por  $\sigma^2$ .

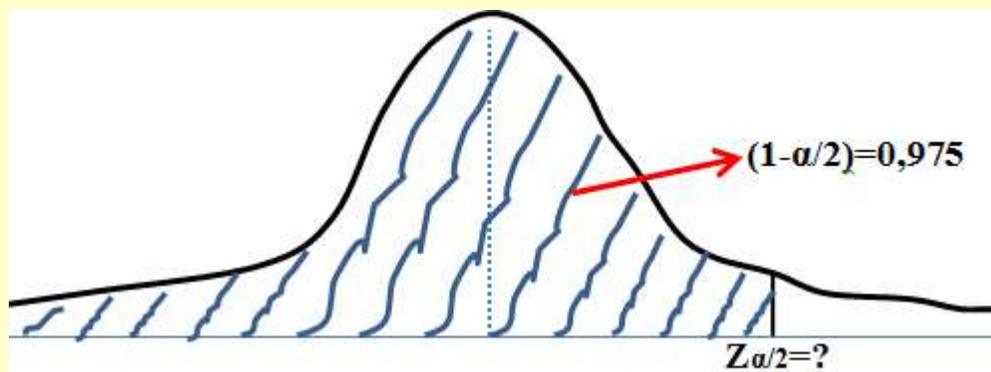
## A estrutura dos I.C. será sempre a mesma:

$$População = amostra \pm \left[ \frac{\text{Confiança} * \text{dispersão}}{\text{tamanho amostra}} \right]$$

## EXEMPLO 4

Historicamente sabe-se que o desvio-padrão da ddp de uma fonte de alimentação é de  $\sigma=10$  volts. Em uma inspeção no processo fabril da fonte, realizaram-se 50 medições, aleatoriamente, constatando-se uma média de  $\bar{X}=118$  volts. Determinar o IC de 95% para a tensão média de todas as fontes fabricadas

Utilizar: IC da MÉDIA p/ amostras grandes ( $N \geq 30$ )



$$\text{INV.NORMP.N}(0,975) = 1,96$$

$$P\left[\left(118 - 1,96 \frac{10}{\sqrt{50}}\right) < \mu < \left(118 + 1,96 \frac{10}{\sqrt{50}}\right)\right]$$

Cálculo do IC  $\rightarrow 1,96 \frac{10}{\sqrt{50}}$  ou  $\text{INT.CONFIANÇA.NORM}(0,05;10;50) = 2,772$

$$P\left[(118 - 2,772) < \mu < (118 + 2,772)\right] = 95,0\%$$

## EXEMPLO 5

Um processo transforma ferro gusa em vidro metálico, produzindo um metal mais resistente à tração e a corrosão, que o aço, porém, quebradiço e frágil às altas temperaturas.

Seleciona-se uma amostra aleatória de 20 pedaços de vidro metálico, registrando-se a temperatura na qual se torna quebradiço. Obtendo-se uma média,  $\bar{X}=315$  °C e  $s=8$  °C.

Determinar um IC com 90% de confiança para a temperatura média na qual o vidro metálico se torna frágil. Utilizar: IC da MÉDIA p/ amostras pequenas (N<30)

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

$$n = 20 \rightarrow gl = n - 1 \rightarrow \boxed{gl = 19} \text{ sendo } \alpha = 90\%$$

$$\begin{aligned} \text{P/ condição UNICAUDAL} \quad t_{0,1} &= \text{INV.T}(1-0,1;19) = 1,328 \\ &\text{ou } \text{INVT}(2*0,1;19) = 1,328 \end{aligned}$$

Para condição UNICAUDAL

$$\begin{aligned} \text{P/ condição BICAUDAL} \quad t_{0,1} &= \text{INV.T.BC}(1-0,1;19) = 1,729 \\ &\text{ou } \text{INVT}(0,1;19) = 1,729 \end{aligned}$$

Para condição BICAUDAL

$$\text{ou } P\left[\left(315 - 1,729 * \frac{8}{\sqrt{20}}\right) < \mu < \left(315 + 1,729 * \frac{8}{\sqrt{20}}\right)\right] \quad E=3,09$$

**IC: P[311,907 <  $\mu$  < 318,093] com 90% de confiança**

## EXEMPLO 6

Construa um IC ao nível de 90% de confiança, para a amostra:

**9, 8, 12, 7, 9, 6, 11, 6, 10, 9**

Utilizar: **IC da MÉDIA n/ amostras pequenas (N<30)**

$$P\left[\left(\bar{X} - t_{1-\alpha} \cdot \frac{s}{\sqrt{n}}\right) < \mu < \left(\bar{X} + t_{1-\alpha} \cdot \frac{s}{\sqrt{n}}\right)\right]$$

	J	K	L	
24	<b>Amostra</b>	<b>Parâmetros</b>	<b>Valores</b>	<b>Descrição</b>
25	9	Média =	8,70	=MÉDIA(J25:J34)
26	8	Variância =	4,01	=VARA(J25:J34)
27	12	Desv Pad =	2,00	=DESVPADAJ(J25:J34)
28	7	n =	10,0	=CONT.VALORES(J25:J34)
29	9	gl =	9,0	=L28-1
30	6	alfa =	0,90	
31	11	t <sub>α</sub> =	1,833	=INV.T.BC(1-90%;9)
32	6	E =	1,161	=L31*L27/RAIZ(L28)
33	10	L <sub>Sup</sub> =	9,861	=L25+L32
34	9	L <sub>Inf</sub> =	7,539	=L25-L32

**BICAUDAL**

$$E = t_{\text{alfa}} * \frac{s}{\sqrt{n}}$$

# RESUMÃO:

Cálculo do IC  $\rightarrow P[(\bar{X}-E) < \mu < (\bar{X}+E)]$

Para o valor de E:

Amostras Grandes $N \geq 30$	<b>Z</b>	$E = Z_{\alpha} * \frac{\sigma}{\sqrt{n}}$
Amostras Pequenas $N < 30$	<b>t</b>	$E = t_{\alpha} * \frac{s}{\sqrt{n}}$

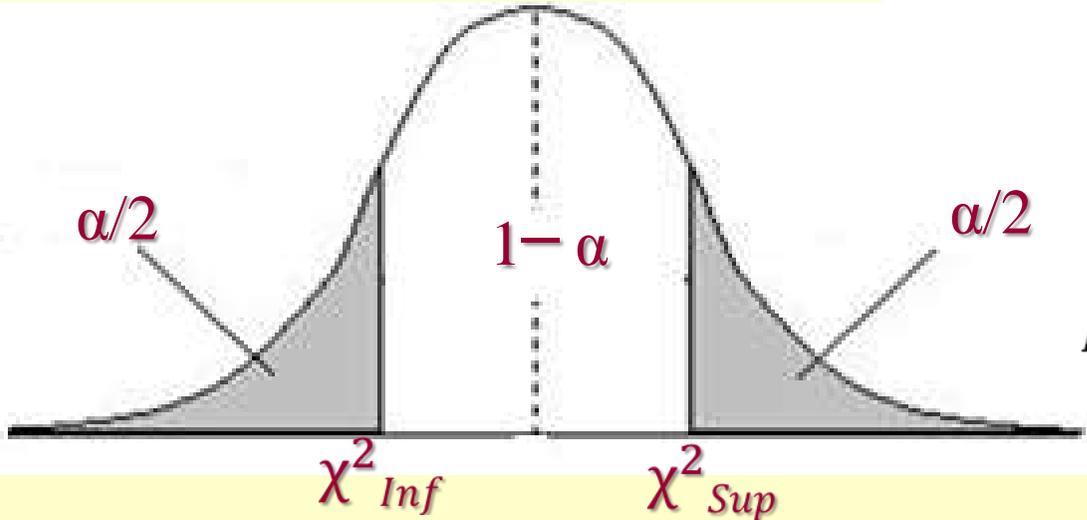
$$Z_{\alpha} = \text{INV.NORMP.N}(\alpha/2)$$

Sendo:

$$t_{\alpha} = \text{INV.T}(\alpha ; g|)$$

# IC da VARIÂNCIA de uma população Normal:

$$\text{Teorema de Fisher} \rightarrow \chi_{n-1}^2 = \frac{(n-1) * s^2}{\sigma^2}$$



$$P[\sigma_{Inf}^2 < \sigma^2 < \sigma_{Sup}^2] \text{ com } 1-\alpha \text{ de confiança}$$

Substituindo  $\sigma$  por Fisher:

$$P\left[\frac{(n-1)*s^2}{\chi_{Sup}^2} < \sigma^2 < \frac{(n-1)*s^2}{\chi_{Inf}^2}\right]$$

# IC do DESV PAD de uma população Normal:

$$P\left[s * \sqrt{\frac{(n-1)}{\chi_{Sup}^2}} < \sigma < s * \sqrt{\frac{(n-1)}{\chi_{Inf}^2}}\right]$$

# IC da VARIÂNCIA de uma população Normal:

## Exemplo 7

Calcular o IC para  $n=10$  ;  $s^2=4$  e  $\alpha=10\%$     ✧ Logo, se  $n=10 \rightarrow gl=9$

$$\chi^2_{Sup} \rightarrow \text{INV.QUIQUA}(0,95;9) = 16,92$$

$$\chi^2_{Inf} \rightarrow \text{INV.QUIQUA}(0,05;9) = 3,325$$

$$P\left[\frac{(n-1)*s^2}{\chi^2_{Sup}} < \sigma^2 < \frac{(n-1)*s^2}{\chi^2_{Inf}}\right]$$

$$P\left[\frac{9*4}{16,92} < \sigma^2 < \frac{9*4}{3,33}\right]$$

$$P[2,13 < \sigma^2 < 10,83] \text{ com } 90\% \text{ de confiança}$$

# IC da PROPORÇÃO ou probabilidade (p) - Binomial:

$$Z = \frac{f - p}{\sqrt{\frac{p * q}{n}}}$$

$f \rightarrow$  frequência relativa  $\rightarrow f_r = \frac{X_i}{n}$   
 $p \rightarrow$  proporção de sucesso  
 $q \rightarrow$  não sucesso  $\rightarrow 1 - p$

$$P[ p_{\text{inf}} < p < p_{\text{sup}} ]$$

$$P \left[ f - Z_{\frac{\alpha}{2}} * \sqrt{\frac{p * q}{n}} < p < f + Z_{\frac{\alpha}{2}} * \sqrt{\frac{p * q}{n}} \right] \text{ para } (1 - \alpha) \text{ de confiança}$$

Porém, como “p” é um valor desconhecido, substitui-se pela frequência Relativa,  $f_r$

$$P \left[ f - Z_{\frac{\alpha}{2}} * \sqrt{\frac{f * (1 - f)}{n}} < p < f + Z_{\frac{\alpha}{2}} * \sqrt{\frac{f * (1 - f)}{n}} \right] \text{ para } (1 - \alpha) \text{ de confiança}$$

# IC da PROPORÇÃO ou probabilidade (p) - Binomial:

$$P \left[ f - Z_{\frac{\alpha}{2}} * \sqrt{\frac{f * (1 - f)}{n}} < p < f + Z_{\frac{\alpha}{2}} * \sqrt{\frac{f * (1 - f)}{n}} \right] \text{ para } (1 - \alpha) \text{ de confiança}$$

## Exemplo 8

Numa pesquisa eleitoral com 500 eleitores, 260 se mostraram favoráveis ao candidato “Y”. Calcular o IC com 90% de confiança para o candidato “Y”.  $n=500$  ;  $X=260$  ;  $\alpha=0,1$  ;  $1 - \alpha=0,9$  ;  $f_r = \frac{260}{500} = 0,52$  ;  $Z_{\frac{\alpha}{2}} = ?$

$$Z_{\frac{\alpha}{2}} = \text{INV.NORMP.N}(0,95) = 1,64$$

$$P[f - E < p < f + E]$$

$$E = 1,64 * \sqrt{\frac{0,52 * (1 - 0,52)}{500}} = 0,037$$

$$P[0,52 - 0,037 < p < 0,52 + 0,037]$$

$$P[0,483 < p < 0,557] \text{ com } 90\% \text{ de certeza}$$

# IC da diferença entre as MÉDIAS de duas populações NÃO INDEPENDENTES entre si:

Para determinar o **IC** da diferença das médias de duas populações, precisamos que essas duas populações sejam independentes .

Porém, na prática, temos algumas **situações onde as populações não são independentes**.

**Por exemplo:** numa situação de comparação inter laboratorial onde dois laboratórios medem a mesma peça, as medidas entre os laboratórios não são independentes. Neste caso, utilizamos o **IC de dados pareados**.

Consideremos duas amostras dependentes  $X_1 \dots X_n$  e  $Y_1 \dots Y_n$ . Neste caso teremos observações pareadas, isto é, podemos considerar na realidade uma amostra de pares  $(X_1, Y_1), \dots (X_n, Y_n)$ . Vamos definir  $D_i = X_i - Y_i$ , para  $i = 1, \dots n$ . Assim obteremos a amostra  $D_1 \dots D_n$ , resultante das diferenças entre os valores de cada par. Aqui, apesar das amostras serem dependentes, vamos considerar que  $D_i \approx N(\mu_D; \sigma^2_D)$ .

O parâmetro  $\mu_D$  será estimado pela média amostral das diferenças, ou seja,  $\bar{D}$ , O parâmetro  $\sigma^2_D$  será estimado pela variância amostral das diferenças, ou seja:

$$t_{(1-\alpha)} = INV.T(1 - \alpha; gl) \quad \text{unicaudal} \quad gl = n - 1$$

$$s_D^2 = \frac{\sum_{i=1}^n (D_i - \bar{D})^2}{n - 1}$$

$$IC(\mu_D, 1 - \alpha) = \left( \bar{D} - t_{\alpha/2} \frac{s_D}{\sqrt{n}}; \bar{D} + t_{\alpha/2} \frac{s_D}{\sqrt{n}} \right)$$

bicaudal

# IC da diferença entre as MÉDIAS de duas populações NÃO INDEPENDENTES entre si:

Determinar o IC da diferença de médias do resultados dos testes de um grupo de alunos sobre um vídeo instrutivo. Seleciona-se uma amostra aleatória simples de 15 alunos, para assistirem ao vídeo, em seguida, aplica-se um outro teste similar. São fornecidas suas notas de aproveitamento para o cálculo do IC da diferença das médias, ao nível 5% de significância.

	A	B
1	<b>Antes</b>	<b>Depois</b>
2	74	80
3	64	74
4	79	83
5	90	92
6	89	96
7	94	98
8	55	59
9	75	77
10	88	93
11	66	78
12	70	75
13	60	59
14	59	61
15	67	70
16	69	74

# IC da diferença entre as MÉDIAS de duas populações NÃO INDEPENDENTES entre si:

Determinar o IC da diferença de médias do resultados dos testes de um grupo de alunos sobre um vídeo instrutivo. Seleciona-se uma amostra aleatória simples de 15 alunos, para assistirem ao vídeo, em seguida, aplica-se um outro teste similar. São fornecidas suas notas de aproveitamento para o cálculo do IC da diferença das médias, ao nível 5% de significância.

	A	B	C	D
1	Antes	Depois	D=di	(di - $\bar{D}$ ) <sup>2</sup>
2	74	80	-6	1,777778
3	64	74	-10	28,444444
4	79	83	-4	0,444444
5	90	92	-2	7,111111
6	89	96	-7	5,444444
7	94	98	-4	0,444444
8	55	59	-4	0,444444
9	75	77	-2	7,111111
10	88	93	-5	0,111111
11	66	78	-12	53,777778
12	70	75	-5	0,111111
13	60	59	1	32,111111
14	59	61	-2	7,111111
15	67	70	-3	2,777778
16	69	74	-5	0,111111
17		$\Sigma =$	-70	147,3333
18		$\bar{D} =$	-4,66667	
19			$S_D =$	3,244042

$$\bar{D} = \frac{\sum di}{n} \rightarrow -70/15 = -4,667$$

$$s_D^2 = \frac{\sum_{i=1}^n (D_i - \bar{D})^2}{n - 1} \quad S_D = \sqrt{\frac{\Sigma[(-6 - 4,667)^2 \dots (-5 - 4,667)^2]}{15 - 1}}$$

$$S_D = 3,244$$

$$t_{(1-\alpha)} = INV.T(1 - \alpha ; gl) = INV.T(0.95 ; 14) = 1,761$$

$$IC(\mu_D, 1 - \alpha) = \left( \bar{D} - t_{\alpha/2} \frac{s_D}{\sqrt{n}} ; \bar{D} + t_{\alpha/2} \frac{s_D}{\sqrt{n}} \right)$$

$$IC = \left( -4,667 - 1,761 * \left( \frac{3,244}{\sqrt{15}} \right) ; -4,667 + 1,761 * \left( \frac{3,244}{\sqrt{15}} \right) \right)$$

$$IC = ( -5,396 ; -3,937 )$$

**IC da diferença entre as MÉDIAS de populações normais, com  $\sigma_1$  e  $\sigma_2$  conhecidos ( $n \geq 30$ ):**

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)}}$$

$$P\left[(\bar{X}_1 - \bar{X}_2) - Z_{\frac{\alpha}{2}} * \sqrt{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)} < (\mu_1 - \mu_2) < (\bar{X}_1 - \bar{X}_2) + Z_{\frac{\alpha}{2}} * \sqrt{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)}\right]$$

# IC da diferença entre as MÉDIAS de populações normais, com $\sigma_1$ e $\sigma_2$ conhecidos ( $n \geq 30$ ):

$$P\left[\left(\bar{X}_1 - \bar{X}_2\right) - Z_{\frac{\alpha}{2}} * \sqrt{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)} < (\mu_1 - \mu_2) < \left(\bar{X}_1 - \bar{X}_2\right) + Z_{\frac{\alpha}{2}} * \sqrt{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)}\right]$$

## Exemplo 9

Sejam as populações  $P_1 \rightarrow N(\mu_1 ; 2,89)$  e  $P_2 \rightarrow N(\mu_2 ; 1,69)$ . De  $P_1$  foi extraída uma amostra de 25 elementos com  $\bar{X}_1=42$ ; de  $P_2$  um amostra de 16 elementos e  $\bar{X}_2=35$ . Construir um IC ao nível de 90% para a diferença das médias.

$\bar{X}_1 = 42$	$\sigma_1 = 2,89$	$n_1 = 25$	$\alpha = 0,1$	$1 - \alpha = 0,9$
$\bar{X}_2 = 35$	$\sigma_2 = 1,69$	$n_2 = 16$		

$$Z_{\frac{\alpha}{2}} = \text{INV.NORMP.N}(0,95) = 1,64$$

$$P\left[\left(\bar{X}_1 - \bar{X}_2\right) - E < (\mu_1 - \mu_2) < \left(\bar{X}_1 - \bar{X}_2\right) + E\right]$$

$$E = 1,64 * \sqrt{\left(\frac{2,89}{25} + \frac{1,69}{16}\right)} = 0,77$$

$$P\left[(42 - 35) - 0,77 < (\mu_1 - \mu_2) < (42 - 35) + 0,77\right]$$

$$P\left[6,23 < (\mu_1 - \mu_2) < 7,77\right] \text{ com } 90\% \text{ de certeza}$$

**IC da diferença entre as MÉDIAS de populações normais, com VARIÂNCIAS IGUAIS e desconhecidas :**

$$P\left[\left(\bar{X}_1 - \bar{X}_2\right) - t_{\frac{\alpha}{2}} * S_p * \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} < (\mu_1 - \mu_2) < \left(\bar{X}_1 - \bar{X}_2\right) + t_{\frac{\alpha}{2}} * S_p * \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}\right]$$

Sendo:  $gl = n_1 + n_2 - 2$

$$S_p^2 = \left\{ \frac{[(n_1 - 1) * s_1^2] + [(n_2 - 1) * s_2^2]}{(n_1 + n_2 - 2)} \right\}$$

**IC da diferença entre as MÉDIAS de populações normais, com VARIÂNCIAS DIFERENTES e desconhecidas :**

$$P\left[\left(\bar{X}_1 - \bar{X}_2\right) - t_{\frac{\alpha}{2}} * \sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)} < (\mu_1 - \mu_2) < \left(\bar{X}_1 - \bar{X}_2\right) + t_{\frac{\alpha}{2}} * \sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}\right]$$

Sendo:  $gl = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{(n_1-1)} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{(n_2-1)}}$

# IC da diferença entre as MÉDIAS de populações normais, com VARIÂNCIAS IGUAIS e desconhecidas :

$$P\left[\left(\bar{X}_1 - \bar{X}_2\right) - t_{\frac{\alpha}{2}} * S_p * \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} < (\mu_1 - \mu_2) < \left(\bar{X}_1 - \bar{X}_2\right) + t_{\frac{\alpha}{2}} * S_p * \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}\right]$$

$$\text{Sendo: } \underline{gl = n_1 + n_2 - 2} \quad S_p = \sqrt{\frac{[(n_1 - 1) * s_1^2] + [(n_2 - 1) * s_2^2]}{(n_1 + n_2 - 2)}}$$

## Exemplo 10

Das populações  $P_1$  foi extraída uma amostra de  $n_1=10$  ;  $\bar{X}_1=15$  ;  $s_1=3$  e de  $P_2$  uma amostra de  $n_2=12$  ;  $\bar{X}_2=12$  ;  $s_2=2$  . Construir um IC ao nível de 95% para a diferença das médias. *Considere a mesma variância populacional para ambas as populações.*

$$\begin{array}{llllll} \bar{X}_1 = 15 & s_1 = 3 & n_1 = 10 & \alpha = 0,05 & 1 - \alpha = 0,95 \\ \bar{X}_2 = 12 & s_2 = 2 & n_2 = 12 & & \end{array}$$

$$gl = 20$$

$$t_{\text{alfa}} = \text{INV.T}(0,05 ; 20) = 2,086$$

condição UNICAUDAL

$$P\left[\left(\bar{X}_1 - \bar{X}_2\right) - E < (\mu_1 - \mu_2) < \left(\bar{X}_1 - \bar{X}_2\right) + E\right]$$

$$S_p^2 = 6,25$$

$$E = 2,233$$

$$(\bar{X}_1 - \bar{X}_2) = 3$$

$$P[0,767 < (\mu_1 - \mu_2) < 5,233] \text{ com } 95\% \text{ de certeza}$$

# IC da diferença entre as MÉDIAS de populações normais, com VARIÂNCIAS DIFERENTES e desconhecidas :

## Exemplo 11

$$P[(\bar{X}_1 - \bar{X}_2) - t_{\frac{\alpha}{2}} * \sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)} < (\mu_1 - \mu_2) < (\bar{X}_1 - \bar{X}_2) + t_{\frac{\alpha}{2}} * \sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}]$$

Dois trabalhadores realizam a mesma operação em uma máquina. Deseja-se estimar a diferença entre o tempo médio de operação de cada um. Considerando a variabilidade da máquina como igual para ambos e normal a distribuição dos tempos de operação para cada operador, determine um IC de 95% de certeza para a diferença entre os tempos médios de operação de cada operador. Do Op<sub>1</sub>, obteve-se uma amostra com n<sub>1</sub>=6 ;  $\bar{X}_1=5,1$  ; s<sub>1</sub>=0,8 e do Op<sub>2</sub> uma amostra com n<sub>2</sub>=10 ;  $\bar{X}_2=4,2$  ; s<sub>2</sub>=0,5 .

$$t_{\text{alfa}} = \text{INV.T}(0,05 ; 7) = 2,365$$

condição UNICAUDAL

$$E = 0,858$$

$$(\bar{X}_1 - \bar{X}_2) = 0,9$$

n <sub>1</sub> = 6	n <sub>2</sub> = 10
$\bar{X}_1 = 5,1$	$\bar{X}_2 = 4,2$
s <sub>1</sub> = 0,8	s <sub>2</sub> = 0,5
alfa = 0,05	1-alfa = 0,95

$$\text{Sendo: } gl = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{(n_1-1)} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{(n_2-1)}}$$

$$gl = 7,393 \quad \text{ou} \quad gl = 7$$

$$P[(\bar{X}_1 - \bar{X}_2) - E < (\mu_1 - \mu_2) < (\bar{X}_1 - \bar{X}_2) + E]$$

$$L_{\text{sup}} = 1,758$$

$$L_{\text{inf}} = 0,042$$

# IC da diferença entre PROPORÇÕES de populações normais :

$$P\left[(\hat{p}_1 - \hat{p}_2) - Z_{\frac{\alpha}{2}} * \sqrt{\left(\frac{\hat{p}_1 * (1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2 * (1 - \hat{p}_2)}{n_2}\right)} < (P_1 - P_2) < (\hat{p}_1 - \hat{p}_2) + Z_{\frac{\alpha}{2}} * \sqrt{\left(\frac{\hat{p}_1 * (1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2 * (1 - \hat{p}_2)}{n_2}\right)}\right]$$

## Exemplo 12

Dois trabalhadores realizam a mesma operação em uma máquina. Deseja-se estimar a diferença na proporção das peças “*não conformes*” produzidas pelos dois operários. Do Op<sub>1</sub>, obteve-se uma amostra de tamanho n<sub>1</sub>=100 forneceu 6 peças “*não conformes*”; Do Op<sub>2</sub>, obteve-se uma amostra de tamanho n<sub>1</sub>=200 com 8 peças “*não conformes*”; Determinar com 90% de certeza, um IC para a diferença de proporção de peças “*não conformes*” produzidas pelos operários.

n <sub>1</sub> =	100	n <sub>2</sub> =	200
x <sub>1</sub> =	5,1	x <sub>2</sub> =	4,2
α =	0,1	α/2 =	0,05
(1 - α) =	0,9	(1 - α)/2 =	0,95

$P_1 = \frac{x_1}{n_1} = \frac{6}{100} = 0,06$
$P_2 = \frac{x_2}{n_2} = \frac{8}{200} = 0,04$

$$E = Z_{\frac{\alpha}{2}} * \sqrt{\left(\frac{\hat{p}_1 * (1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2 * (1 - \hat{p}_2)}{n_2}\right)}$$

$$P[(\hat{y}_1 - \hat{y}_2) - E < (P_1 - P_2) < (\hat{y}_1 - \hat{y}_2) + E]$$

$$Z_{\alpha/2} = \text{INV.NORMP}(0,95) = 1,645$$

$$\hat{y}_1 - \hat{y}_2 = 0,02 \quad L_{\text{sup}} = -0,025$$

$$E = 0,045 \quad L_{\text{inf}} = 0,065$$

*Não há como concluir se existe diferença entre o trabalho dos operários, visto que o ZERO consta do IC*

**F I M**