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

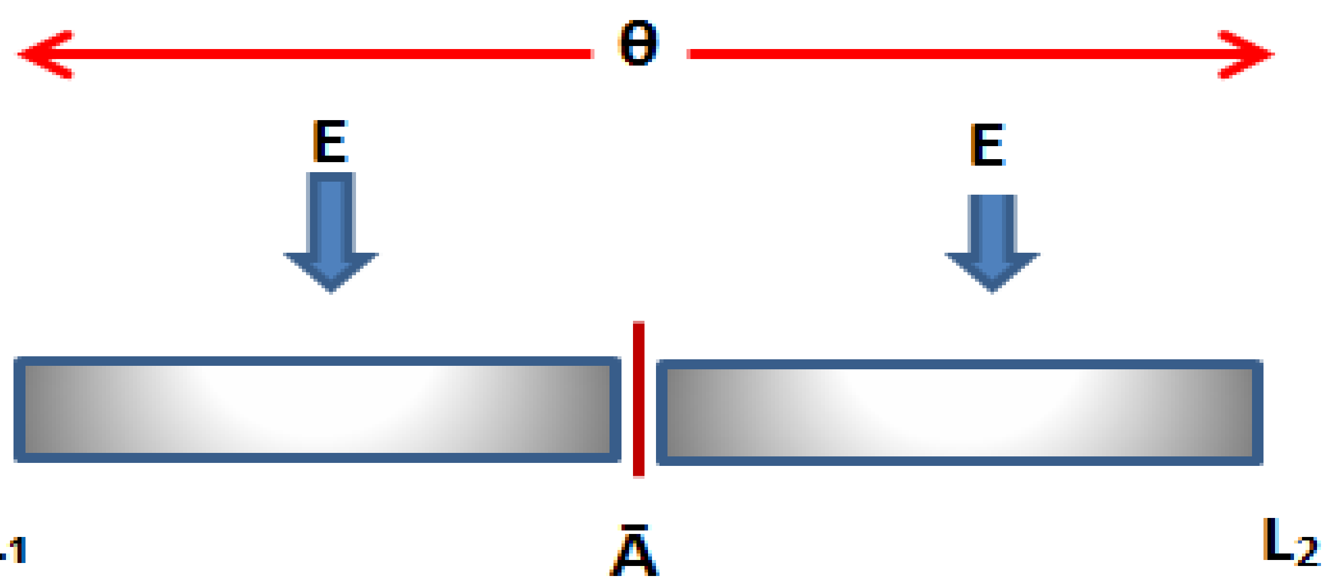
A estimação intervalar consiste na determinação de um intervalo onde, com uma certa confiança (*probabilidade*), pode-se prever que ali esteja o parâmetro θ desconhecido, tendo-se em conta um estimador desse parâmetro.

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

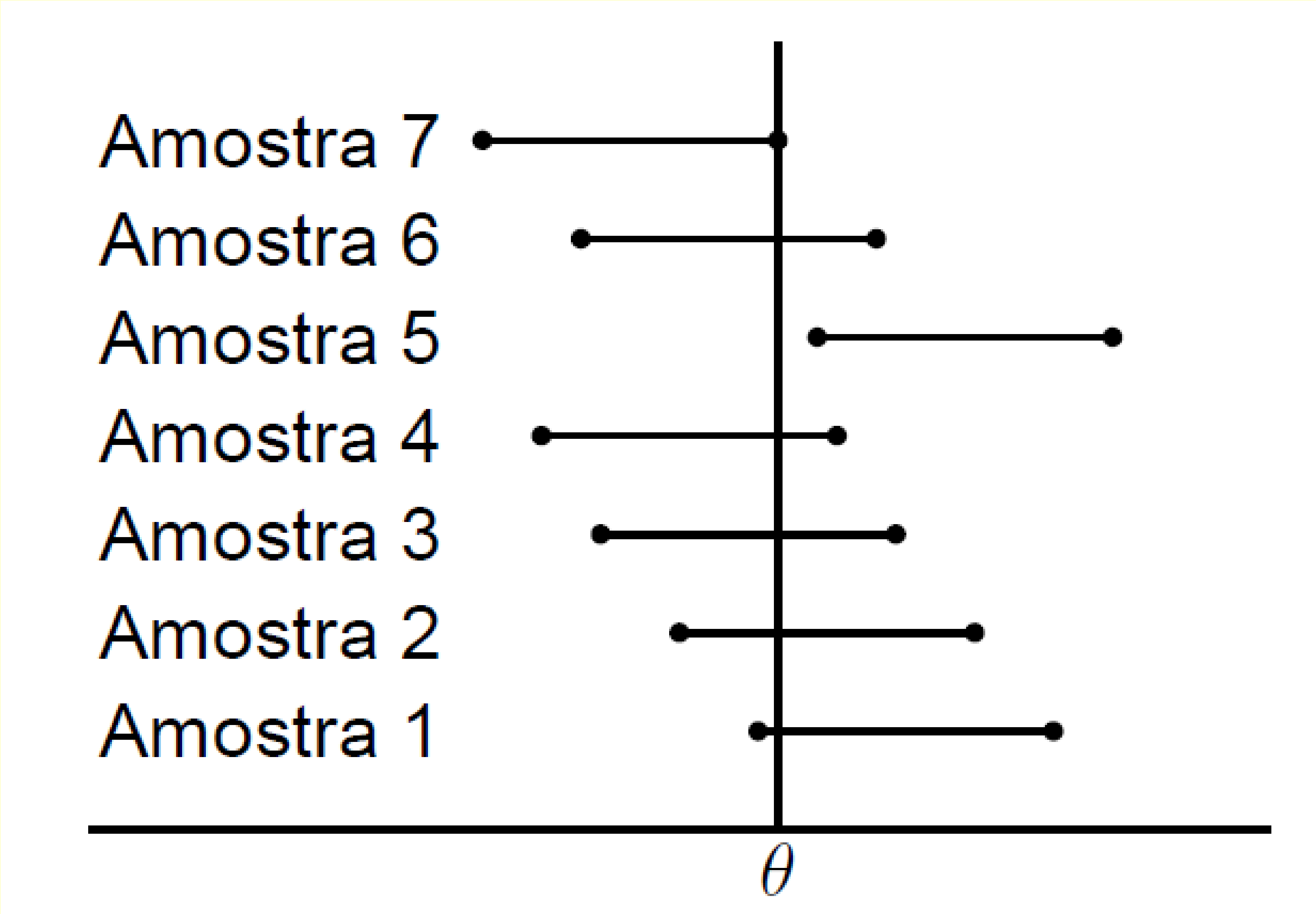
O intervalo é designado por um intervalo de confiança para o parâmetro $\boldsymbol{\theta}$, com um nível de confiança α.

ESTIMAÇÃO INTERVALAR

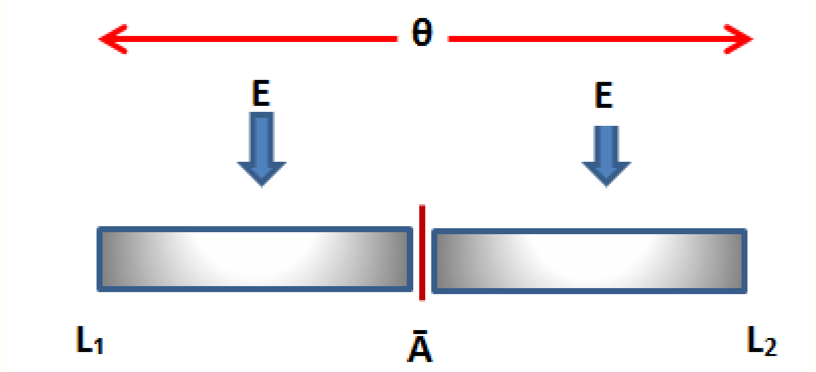
L_1	Limite inferior do intervalo
L_2	Limite superior do intervalo
E	Comprimento do intervalo
$\bar{A}$	Média amostral
θ	Média Populacional



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 (**α**);

2) Aumentando-se o nível de confiança (**α**), 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.

α ► 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?

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

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

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:

$$P[\ddot{X} - Z\frac{\sigma}{\sqrt{N}} \leqq \mu \leqq \ddot{X} + Z\frac{\sigma}{\sqrt{N}}]$$

Sem reposição de eventos:

$$P[\ddot{X} - Z * \frac{\sigma}{\sqrt{n}} * \sqrt{\frac{N-n}{N-1}} \leqq \mu \leqq \ddot{X} + Z * \frac{\sigma}{\sqrt{n}} * \sqrt{\frac{N-n}{N-1}}]$$

N = população

n = amostra

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

Quando σ for conhecido:

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

Quando σ for desconhecido:

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

Intervalo de confiança para a média:

Seja uma *amostra aleatória (aa)*: *X1*, *X2*, . . . , *Xn*, com média **μ** (desconhecida) e variância **σ²** conhecida.

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

Onde:

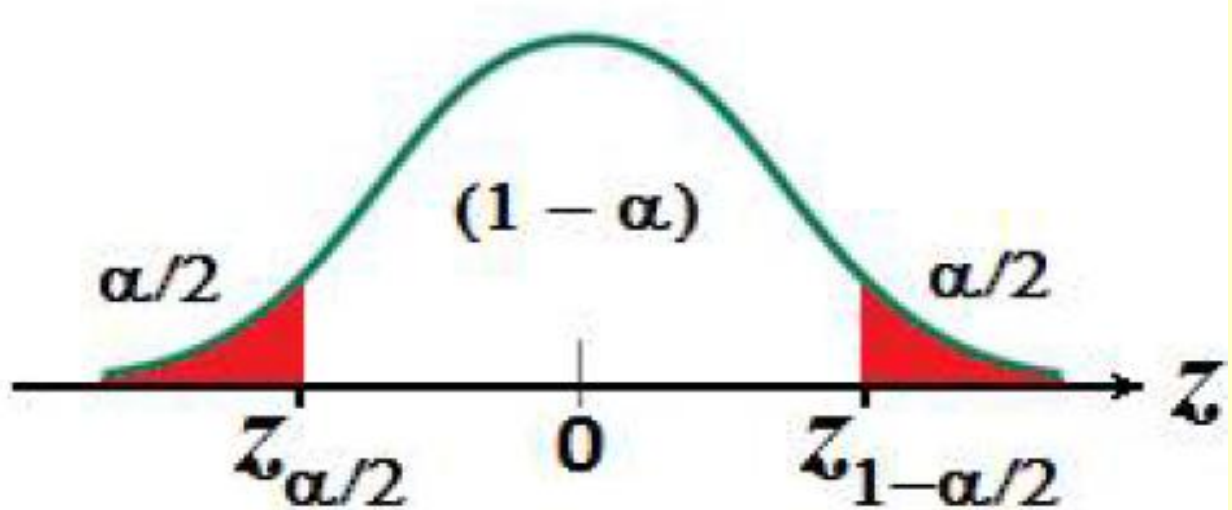
μ → média populacional (*desconhecida*);

Ẍ → média amostral (*calculada*);

σ → desvio padrão populacional (*conhecido*);

n → tamanho da amostra.

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



Intervalo de confiança para a média:

Seja uma *amostra aleatória (aa)*: *X1*, *X2*, . . . , *Xn*, com média **μ** (desconhecida) e variância **σ²** conhecida.

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

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

Limite inferior **Limite superior**

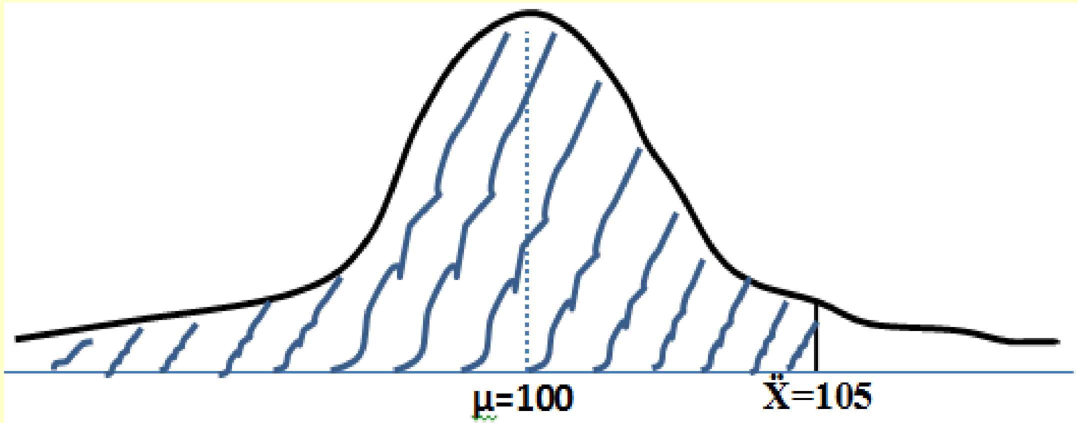
Se α = 5%, então: $Z_{(1-\alpha/2)}$ = INV.NORMP(1-0,05/2) = 1,96

Portanto, um *intervalo de confiança* **(1 - α)×100%** para μ, com **σ²** *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?



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

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

DIST.NORMP.N(1,58;1) ► 94,29%

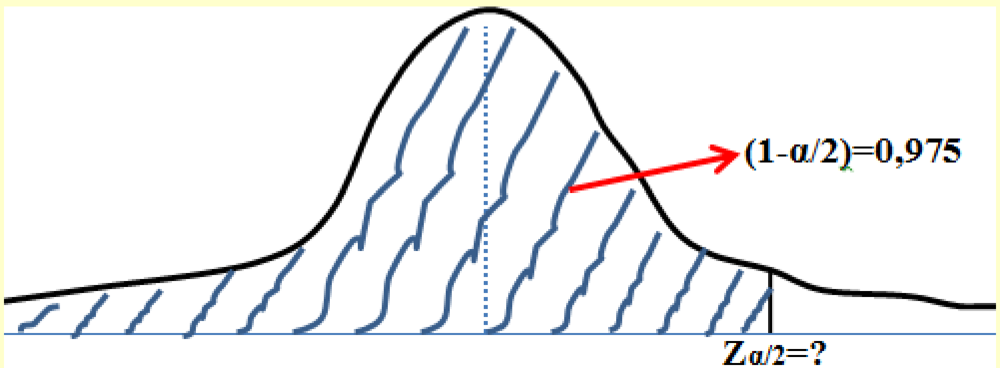
$$P\left[\left(105 - 1,58\,\frac{20}{\sqrt{40}}\right) < \mu < \left(105 + 1,58\,\frac{20}{\sqrt{40}}\right)\right] = 94{,}29\%$$

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

EXEMPLO 3

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

Utilizar: **IC da MÉDIA p/ amostras grandes (N>=30)**



INV.NORMP(1-0,05/2) = 1,96

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

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

INT.CONFIANÇA.NORM(0,05;5;100) = 0,98

INV.NORM(1-0,05/2;500;5/RAIZ(100)) = 500,98

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

E E

Problemas existentes na prática:

1) Estimar a média (**μ**) 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 σ^2.

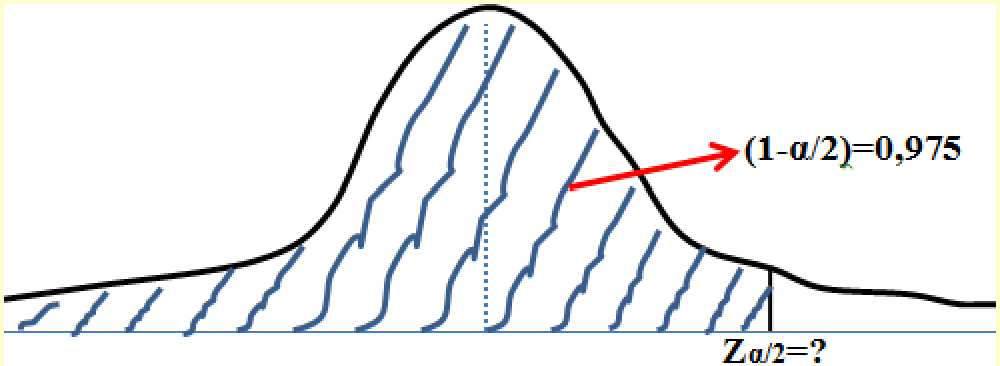
A estrutura dos I.C. será sempre a mesma:

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

EXEMPLO 4

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

Utilizar: **IC da MÉDIA p/ amostras grandes (N>=30)**



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** INT.CONFIANÇA.NORM(0,05;10;50) = 2,772

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

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, $\ddot{X}$=315 ºC e s=8 ºC. 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)**

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

n = 20 → gl = n - 1 →	$\boxed{gl = 19}$	sendo α =	90%
P/ condição UNICAUDAL	$t_{0,1}$=	INV.T(1-0,1;19)=	1,328
	ou	INVT(2*0,1;19)=	1,328
P/ condição BICAUDAL	$t_{0,1}$=	INV.T.BC(1-0,1;19)=	1,729
	ou	INVT(0,1;19)=	1,729

Para condição UNICAUDAL

Para condição BICAUDAL

$$\text{ou } P[(315 - 1{,}729 * \frac{8}{\sqrt{20}}) < \mu < (315 + 1{,}729 * \frac{8}{\sqrt{20}})]$$

E=3,09

IC: P[311,907 < μ < 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 p/ amostras pequenas (N<30)**

$$P[(\ddot{X} - t_{1-\alpha} * \frac{s}{\sqrt{n}}) < \mu < (\ddot{X} + t_{1-\alpha} * \frac{s}{\sqrt{n}})]$$

	J	K	L	
24	Amostra	Parâmetros	Valores	Descrição
25	9	Média =	8,70	=MÉDIA(J25:J34)
26	8	Variância =	4,01	=VARA(J25:J34)
27	12	Desv Pad =	2,00	=DESVPADA(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_α =	1,833	=INV.T.BC(1-90%;9)
32	6	E =	1,161	=L31*L27/RAIZ(L28)
33	10	L_{Sup} =	9,861	=L25+L32
34	9	L_{Inf} =	7,539	=L25-L32

BICAUDAL

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

RESUMÃO:

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

Para o valor de E:

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

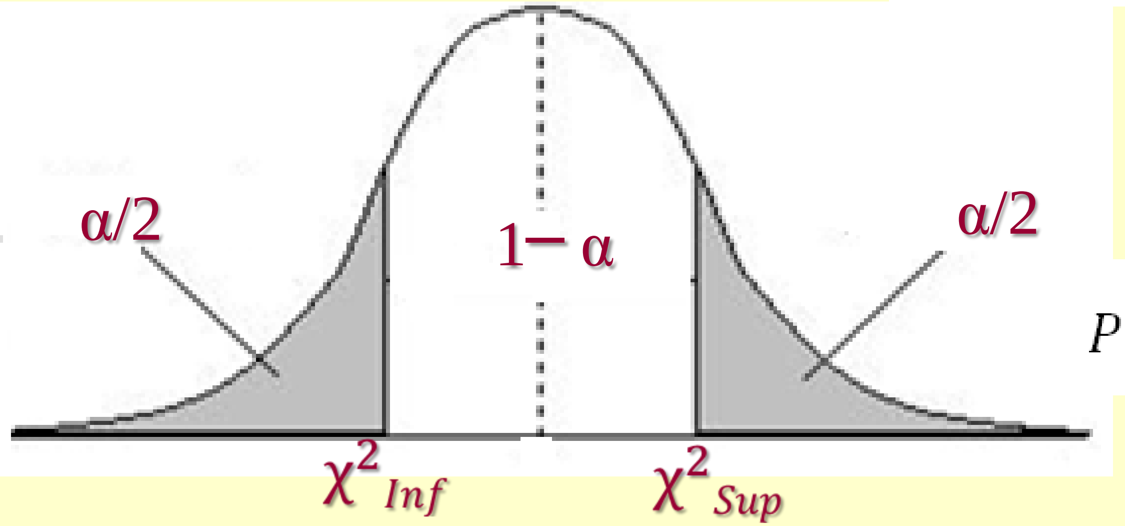
Z_α = INV.NORMP.N(α/2)

Sendo:

t_α = INV.T(α ; gl)

IC da VARIÂNCIA de uma população Normal:

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



$P[\sigma^2_{Inf} < \sigma^2 < \sigma^2_{Sup}]$ com 1-α de confiança

Substituindo σ por Fisher:

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

IC do DESV PAD de uma população Normal:

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

IC da VARIÂNCIA de uma população Normal:

Exemplo 7

Calcular o IC para n=10 ; s^2=4 e α=10%   ❖ Logo, se n=10 → gl=9

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

$\chi^2_{Inf} \rightarrow$ 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 → frequência relativa → $f_r = \frac{x_i}{n}$

p → proporção de sucesso

q → não sucesso → 1 – p

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

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

Porém, como " p" é um valor desconhecido, substitui-se pela frequência Relativa, f_r

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

IC da PROPORÇÃO ou probabilidade (p) - Binomial:

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

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". $n=500 \; ; \; X=260 \; ; \; \alpha=0,1 \; ; \; 1-\alpha=0,9 \; ; \; f_r=\frac{260}{500}=0,52 \; ; \; Z_{\frac{\alpha}{2}}=?$

$Z_{\frac{\alpha}{2}}=$ 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] com 90% de certeza

IC da diferença entre as MÉDIAS de duas populações NÃO INDENPENDENTES 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 μ_D será estimado pela média amostral das diferenças, ou seja, $\overline{D}$, O parâmetro σ^2_D será estimado pela variância amostral das diferenças, ou seja:

$$t_{(1-\alpha)} = INV.T(1-\alpha\,;\mathrm{gl})$$ unicaudal

gl = n - 1

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

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

bicaudal

IC da diferença entre as MÉDIAS de duas populações NÃO INDENPENDENTES 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	**Antes**	**Depois**
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 INDENPENDENTES 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	$(di - \overline{D})^2$
2	74	80	-6	1,777778
3	64	74	-10	28,44444
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	53,77778
12	70	75	-5	0,111111
13	60	59	1	32,11111
14	59	61	-2	7,111111
15	67	70	-3	2,777778
16	69	74	-5	0,111111
17		$\sum$ =	-70	147,3333
18		$\overline{D}$ =	-4,66667	
19			S_D=	3,244042

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

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

$$S_D = 3{,}244$$

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

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

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

$$\boxed{IC = (-5{,}396\,;\,-3{,}937)}$$

IC da diferença entre as MÉDIAS de populações normais, com σ_1 e σ_2 conhecidos (n>=30):

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

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

IC da diferença entre as MÉDIAS de populações normais, com σ_1 e σ_2 conhecidos (n>=30):

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

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 $\ddot{X}_1=42$; de P_2 um amostra de 16 elementos e $\ddot{X}_2=35$. Construir um IC ao nível de 90% para a diferença das médias.

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

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

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

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

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

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

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

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

Sendo: $\boxed{gl = n_1 + n_2 - 2}$

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

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

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

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 :

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

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

Exemplo 10

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

$\ddot{X}_1 = 15$	$s_1 = 3$	$n_1 = 10$	$\alpha = 0{,}05$	$1 - \alpha = 0{,}95$
$\ddot{X}_2 = 12$	$s_2 = 2$	$n_2 = 12$		

gl = 20

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

condição UNICAUDAL

$S_p^2 = 6{,}25$

$E = 2{,}233$

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

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

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

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

Exemplo 11

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

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_1, obteve-se uma amostra com $n_1=6$; $\ddot{X}_1=5,1$; $s_1=0,8$ e do Op_2 uma amostra com $n_2=10$; $\ddot{X}_2=4,2$; $s_2=0,5$.

$n_1 =$	6	$n_2 =$	10
$\ddot{X}_1 =$	5,1	$\ddot{X}_2 =$	4,2
$s_1 =$	0,8	$s_2 =$	0,5
alfa =	0,05	1-alfa =	0,95

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

gl = 7,393 ou gl = 7

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

condição UNICAUDAL

E = 0,858

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

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

$L_{Sup} =$ 1,758

$L_{inf} =$ 0,042

IC da diferença entre PROPORÇÕES de populações normais :

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

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_1, obteve-se uma amostra de tamanho n_1=100 forneceu 6 peças "*não conformes*"; Do Op_2, obteve-se uma amostra de tamanho n_1=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_1 =	100	n_2 =	200
x_1 =	5,1	x_2 =	4,2
α =	0,1	α/2 =	0,05
(1 - α) =	0,9	(1 - α)/2 =	0,95

$$p_1=\frac{x_1}{n_1}=\frac{6}{100}=0{,}06$$

$$p_2=\frac{x_2}{n_2}=\frac{8}{200}=0{,}04$$

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

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

$Z_{\alpha/2}$ =	INV.NORMP(0,95) =		1,645
$\dot{y}_1 - \dot{y}_2$ =	0,02	L_{sup} =	-0,025
E =	0,045	L_{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