UNIVERSIDADE DE SÃO PAULO

Faculdade de Zootecnia e Engenharia de Alimentos

ZAB1111 – Estatística Básica

Aula 12

INTERVALO DE CONFIANÇA PARA A MÉDIA

Material preparado pelo Prof. Dr. César Gonçalves de Lima – FZEA/USP

3.2. ESTIMAÇÃO POR INTERVALO

- A estimativa pontual de um parâmetro θ é bastante útil, mas não fornece qualquer indicação da precisão a ela associada.
- Precisão é o grau de variação de resultados de uma medição e tem como base o desvio-padrão de uma série de repetições da mesma análise.
- É desejável que uma estimativa pontual esteja acompanhada por alguma medida do erro da estimativa.

Intervalo de Confiança é um estimador que envolve a determinação de um intervalo a respeito da estimativa do parâmetro, juntamente com alguma **medida de confiança** de que o verdadeiro valor do parâmetro esteja neste intervalo.

A probabilidade de que o I.C. contenha o verdadeiro valor do parâmetro é chamada de **coeficiente** ou **nível de confiança**, sendo denotado pela letra grega γ (gama).

Objetivo: Encontrar um I.C. de pequena amplitude que inclua o verdadeiro valor do parâmetro com uma confiança (γ) alta.

A amplitude de um I.C. é sempre calculada como a diferença entre os seus limites:

$$H = lim_{sup} - lim_{inf}$$

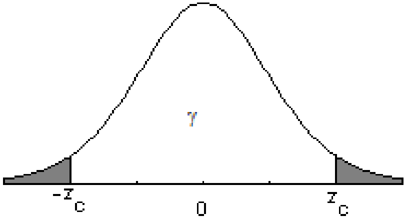
3.2.1. Intervalo de confiança para a média populacional $IC(\mu)$

O intervalo de confiança para a média populacional com um coeficiente de confiança $\gamma = 1 - \alpha$, **quando a variância populacional** (σ^2) **é conhecida**, é dado por:

$$IC(\mu; 100\gamma\%) = \left[\bar{x} - z_c \frac{\sigma}{\sqrt{n}}; \bar{x} + z_c \frac{\sigma}{\sqrt{n}}\right]$$

Onde z_c é o valor crítico, tabelado, da distribuição normal padrão, tal que:

$$\gamma = P(-z_c \leq Z \leq z_c)$$

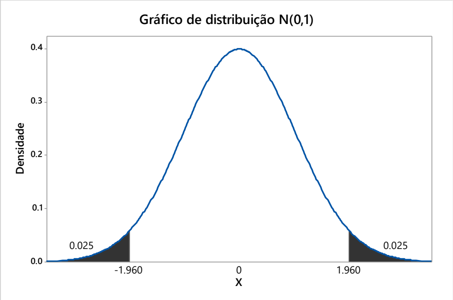


Nota: z_c também é denotado como $z_{\alpha/2}$. Por exemplo: se queremos construir um $IC(\mu; \gamma = 0{,}95)$, o valor z_c tal que:

$$0{,}95 = P(-z_c \leq Z \leq z_c)$$

é o valor tabelado $z_{0{,}025} = 1{,}96$, tal que

$$P(Z > 1{,}96) = P(Z < -1{,}96)$$
$$= 0{,}025$$



Prova: Sabemos que se tirarmos uma amostra de n elementos de uma população $X \sim N(\mu, \sigma^2)$, a média amostral $\bar{x} \sim N(\mu, \sigma^2/n)$ e que

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

Seja z_c um valor tal que:

$$\gamma = 1 - \alpha = P(-z_c < Z < z_c)$$

Então:

$$\gamma = P\left(-z_c < \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} < z_c\right) = P\left(-z_c \frac{\sigma}{\sqrt{n}} < \bar{x} - \mu < z_c \frac{\sigma}{\sqrt{n}}\right)$$

$$= P\left(-\bar{x} - z_c \frac{\sigma}{\sqrt{n}} < -\mu < -\bar{x} + z_c \frac{\sigma}{\sqrt{n}}\right)$$

$$= P\left(\bar{x} - z_c \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_c \frac{\sigma}{\sqrt{n}}\right)$$

$\Rightarrow$ Podemos admitir que

$$IC(\mu; 100\gamma\%) = \left[\bar{x} - z_c \frac{\sigma}{\sqrt{n}}; \bar{x} + z_c \frac{\sigma}{\sqrt{n}}\right]$$

Note que:

γ	0,80	0,85	0,90	0,95	0,99
z_c	1,28	1,44	1,64	1,96	2,58

- A amplitude H do $IC(\mu)$ está diretamente relacionada com o nível de confiança γ: se quero uma maior confiança na estimativa da média, vou construir um IC com maior amplitude.
- Se desejarmos um $IC(\mu)$ mais curto e que tenha uma confiança adequada, precisamos aumentar o tamanho da amostra.

O tamanho de amostra necessário para termos uma confiança (γ) de que o erro na estimação da média populacional seja inferior a $\varepsilon_{\bar{x}}$, é calculado por:

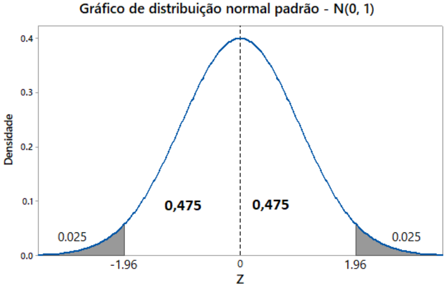
$$n = \left(\frac{z_c \sigma}{\varepsilon_{\bar{x}}}\right)^2$$

Exemplo 3.2. O peso de bovinos Nelore aos 210 dias de idade tem distribuição normal com variância $400kg^2$. Baseado numa amostra de 30 animais, cujo peso médio foi de 186 kg pede-se:

a) Construir um I.C. para o peso médio dos bovinos com uma confiança $\gamma = 0{,}95$ e outro para $\gamma = 0{,}99$.

b) Qual a confiança em afirmar que o verdadeiro peso médio é de $[180; 192]kg$?

$a)$ Para $\gamma = 0{,}95$, $z_c = 1{,}96$

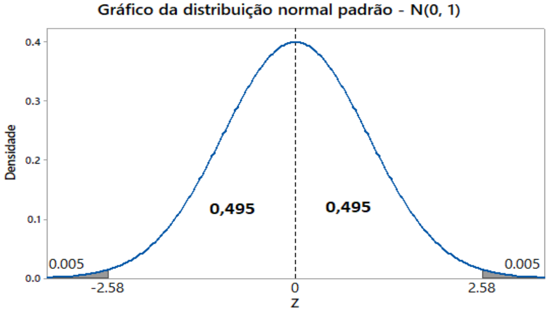


$$IC(\mu, 95\%) = \left[186 - 1{,}96\frac{20}{\sqrt{30}}; 186 + 1{,}96\frac{20}{\sqrt{30}}\right]$$

$$= [178{,}84; 193{,}16]kg \qquad \text{(amplitude 14,32kg)}$$

Conclusão: Este IC contém o verdadeiro valor do peso médio dos bezerros aos 210 dias de idade com 95% de confiança.

Para $\gamma = 0{,}99\ z_c = 2{,}58$



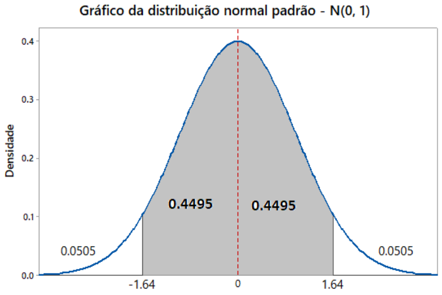
$$IC(\mu; 99\%) = \left[186 - 2{,}58\frac{20}{\sqrt{30}}; 186 + 2{,}58\frac{20}{\sqrt{30}}\right]$$

$$= [176{,}58;\ 195{,}42]kg$$

Conclusão: Este intervalo de amplitude 18,84kg contém o verdadeiro valor do peso médio dos bezerros com 99% de confiança.

Note que a amplitude do $IC(\mu; 99\%)$ é maior que $IC(\mu; 95\%)$, ou seja quanto maior a confiança, maior a amplitude do IC.

$b)$ Se $H = 192 - 180 = 12\ kg \Rightarrow 12 = 2 * z_c \dfrac{20}{\sqrt{30}} \Rightarrow z_c = 1{,}64$



$$\gamma = P(-1{,}64 < Z < 1{,}64)$$
$$= 2P(0 < Z < 1{,}64)$$
$$= 2(0{,}4495) = 0{,}8990$$

ou seja, o IC contém o verdadeiro peso médio dos bezerros com 89,9% de confiança.

Utilizando a fórmula

$$n = \left(\frac{z_c\sigma}{\varepsilon_{\bar{x}}}\right)^2$$

podemos calcular o tamanho da amostra a ser usado para estimar o peso de bovinos Nelore aos 210 dias ($\sigma = 20\ kg$), admitindo diferentes erros na estimativa da média ($\varepsilon_{\bar{x}}$):

	Erros ($\varepsilon_{\bar{x}}$) em kg					
γ	10	8	6	4	2	1
0,95	16	25	43	97	385	1537
0,99	27	42	74	167	666	2663

Note que o tamanho da amostra cresce com o aumento do coeficiente de confiança (γ) ou com a diminuição da margem de erro.

Problema: Geralmente nós não conhecemos a variância populacio-nal (σ^2) dos dados. Neste caso precisamos estimá-la com os dados de uma amostra, usando o estimador s^2.

Como calcular um $IC(\mu; 100\gamma\%)$ se desconhecemos σ^2?

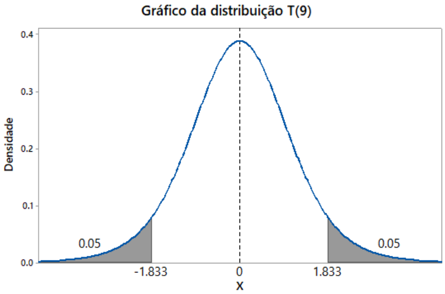
O intervalo de confiança para a média quando a variância populacional (σ^2) é desconhecida é calculado por:

$$IC(\mu; 100\gamma\%) = \left[\bar{x} - t_c \frac{s}{\sqrt{n}}; \bar{x} + t_c \frac{s}{\sqrt{n}}\right]$$

onde $s =$ é o desvio padrão amostral e t_c é obtido da Tabela III (t-Stu-dent) com $n - 1$ graus de liberdade, tal que $\gamma = P(-t_c \leq T \leq t_c)$.

Exemplo 3.3. Dez animais foram alimentados com certa ração durante 15 dias e os seus ganhos de peso foram: 2,71; 2,93; 3,10; 3,12; 3,23; 3,76; 3,89; 4,01; 4,16 e 4,23 kg. Construir um I.C. para o ganho médio de peso com $\gamma = 0{,}90$.

Resolução: Da amostra tem-se: $\bar{x} = 3{,}51$ e $s^2 = 0{,}3081$.



Como $\gamma = 0{,}90$ entramos na Tábua III com $p = \frac{1-0{,}90}{2} = 0{,}05$ e 9 g.l. e obtemos $t_c = 1{,}833$.

Então:

$$IC(\mu; 90\%): 3{,}51 \pm 1{,}833\ \sqrt{\frac{0{,}3081}{10}} = 3{,}51 \pm 0{,}32$$

$$\Rightarrow IC(\mu; 90\%) = [3{,}19; 3{,}83]\ kg$$

Este IC (de amplitude 0,64 kg) contém o verdadeiro ganho médio de peso de bovinos alimentados com certa ração durante 15 dias com 90% de confiança.

Resumo: Se queremos calcular um intervalo de confiança para a média populacional, μ, e a variância populacional, σ^2, é

- **conhecida**, usamos:

$$IC(\mu; 100\gamma\%) = \left[\bar{x} - z_c \frac{\sigma}{\sqrt{n}}; \bar{x} + z_c \frac{\sigma}{\sqrt{n}}\right]$$

- **desconhecida**, usamos:

$$IC(\mu; 100\gamma\%) = \left[\bar{x} - t_c \frac{s}{\sqrt{n}}; \bar{x} + t_c \frac{s}{\sqrt{n}}\right]$$

Em que s^2 é a variância amostral.

EXERCÍCIO Um provedor de acesso à internet está monitorando a duração do tempo de conexões de seus clientes com o intuito de dimensionar seus equipamentos. Embora não se conheça a média desse tempo, sabe-se que o desvio-padrão, por analogia a outros serviços, é considerado igual a 17,8 minutos. Uma amostra de 80 conexões resultou num valor médio $\bar{x} = 25{,}2$ minutos.

a) Comente sobre o tempo médio de conexão, baseando-se num intervalo com confiança de 96% para a média.

b) Calcule o tamanho de amostra necessário para diminuir a amplitude deste intervalo de confiança pela metade, com a mesma confiança.