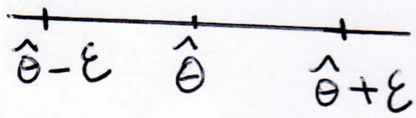
# Estimação por Intervalo

$\theta$ parâmetro desconhecido

$\hat{\theta}$ - estimador pontual de $\theta$, fornece um único valor numérico como estimativa de $\theta$

$$\hat{\theta} - \varepsilon \leftarrow \hat{\theta} \rightarrow \hat{\theta} + \varepsilon$$




$[\hat{\theta} - \varepsilon, \hat{\theta} + \varepsilon]$ conjunto de estimativas, que será um intervalo de confiança para $\theta$.

O procedimento é denominado estimação por intervalo.

Nas pesquisas eleitorais, $p$ - real proporção de eleitores que vota no candidato, $\hat{p}$ - proporção amostral de pessoas que afirmam que votarão no candidato

$$\hat{p} - 0{,}02 \; , \; \hat{p} + 0{,}02$$

$\varepsilon = 0{,}02$ margem de erro

A Construção do Intervalo de Confiança depende da distribuição amostral do estimador

1º Caso - Intervalo de Confiança para a média $\mu$ de uma distribuição normal com variância $\sigma^2$ conhecida

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

$(X_1, X_2 \dots X_n)$ amostra aleatória dessa distribuição

Objetivo: Estimar $\mu$ (média da variável $X$)

Estimador pontual de $\mu$: $\hat{\mu} = \overline{X}$

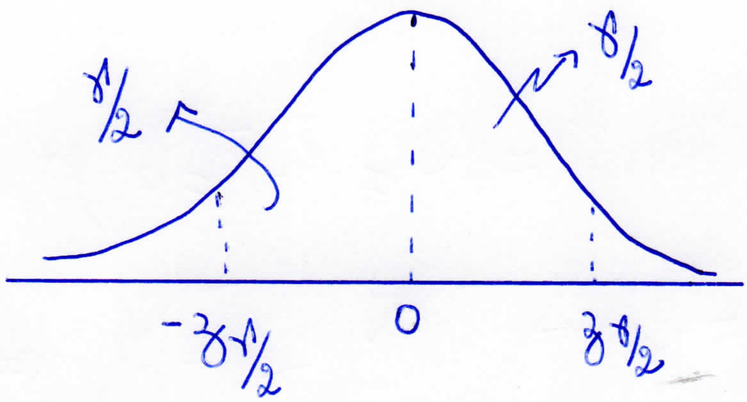
Resultado:

$\overline{X} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$ esta é a distribuição amostral de $\overline{X}$

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

Fixado $0 < \gamma < 1$, podemos determinar $z_{\gamma/2}$ tal que

$$P\left(-z_{\gamma/2} \leq Z \leq z_{\gamma/2}\right) = \gamma \qquad Z \sim N(0,1)$$




$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

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

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

$$IC(\mu, \gamma) = \left[\bar{X} - z_{\gamma/2}\frac{\sigma}{\sqrt{n}} \,;\, \bar{X} + z_{\gamma/2}\frac{\sigma}{\sqrt{n}}\right]$$

é um intervalo de confiança para $\mu$ com coeficiente de confiança $\gamma$.

Obs:

- Os extremos do IC são variáveis aleatórias e a probabilidade desse IC (com extremos aleatórios) conter $\mu$ é $\gamma$. Na prática, $\gamma$ é fixado em valores altos: 0,90, 0,95 e 0,99.

- Observada a estimativa $\bar{x}$ (valor numérico de $\bar{X}$) o intervalo se torna

$$\left[\bar{x} - z_{\gamma/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\gamma/2}\frac{\sigma}{\sqrt{n}}\right]$$

Os extremos são numéricos e não podemos mais fazer afirmações probabilísticas, já que esses elementos não são variáveis aleatórias.

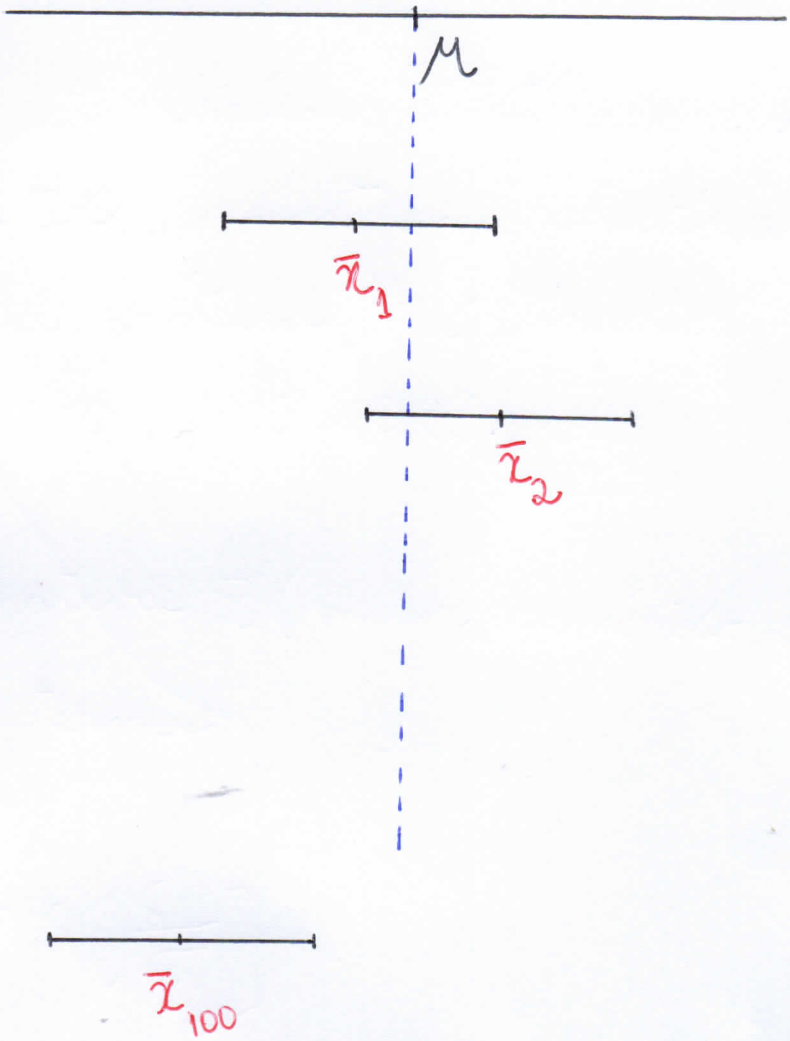
Interpretação:

Tomando-se 100 amostras de mesmo tamanho e construindo-se 100 intervalos pelo mesmo procedimento e com mesmo coeficiente de confiança $\gamma$, então aproximadamente $100\gamma$ desses intervalos conterão o valor de $\mu$.

IC $\bar{x}_1 \pm z_{\gamma/2}\frac{\sigma}{\sqrt{n}}$

$\bar{x}_2 \pm z_{\gamma/2}\frac{\sigma}{\sqrt{n}}$

$\bar{x}_{100} \pm z_{\gamma/2}\frac{\sigma}{\sqrt{n}}$




Se $\gamma = 0{,}95$, aproximad/ 95 dos 100 intervalos conteriam $\mu$

Na prática, construímos apenas um intervalo e como não conhecemos $\mu$, não sabemos se ele contem ou não $\mu$, mas "acreditamos que estamos dentre os 95 que contem $\mu$".

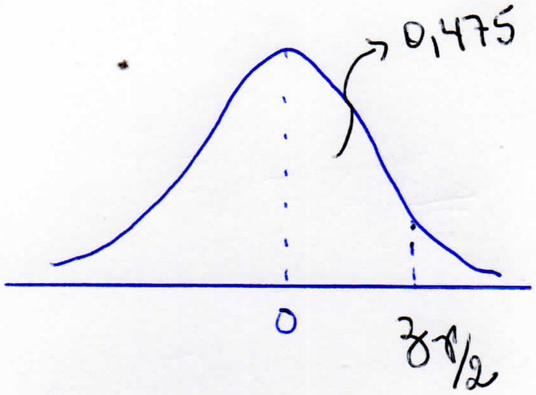
Exemplo 7.18 pag 246 Magalhães e Lima

X - comprimento de jacarés (em m)

$X \sim N(\mu ; 0{,}01)$ $\sigma = \sqrt{0{,}01}$

Estimar $\mu$ por intervalo com coeficiente de confiança $\gamma = 0{,}95$ com base em uma amostra de tamanho $n = 10$ que forneceu média 1,69 m.

$\bar{x} = 1{,}69$




$z_{\gamma/2} = z_{0{,}475} = 1{,}96$

$IC(\mu, 0{,}95)$:

$$\left[1{,}69 - 1{,}96 \frac{\sqrt{0{,}01}}{\sqrt{10}} ; 1{,}69 + 1{,}96 \frac{\sqrt{0{,}01}}{\sqrt{10}}\right] = [1{,}63 ; 1{,}75]$$

Obs:

Amplitude do IC $= 2 z_{\gamma/2} \frac{\sigma}{\sqrt{n}}$

- Aumentando $n$, a amplitude diminui, implicando em intervalos mais informativos.

- Quanto maior $\gamma$, maior $z_{\gamma/2}$, maior a amplitude (mantendo os demais fixos).

- $z_{\gamma/2} \frac{\sigma}{\sqrt{n}} = \varepsilon$ de modo que IC: $[\bar{X} - \varepsilon, \bar{X} + \varepsilon]$

Isolando $n \Rightarrow n = \left(\frac{z}{\varepsilon}\right)^2 \sigma^2$

↳ obtido anteriormente

- Ideal: intervalos de confiança de baixa amplitude com alto coeficiente de confiança. Isto só é obtido com o aumento do tamanho da amostra.

Exemplo 7.19 Pag 247 Magalhães e Lima

X - vida de baterias automotivas em meses

Com base em estudos anteriores admite-se que

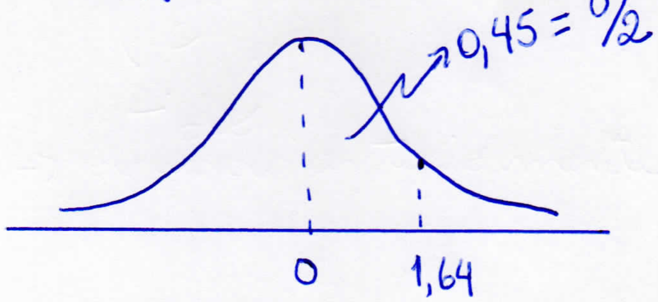
$X \sim N(\mu, 4{,}5^2)$

$\mu$ : vida média desconhecida

Qual deve ser o tamanho da amostra para que a amplitude do intervalo de 90% de confiança para a vida média seja de 3 meses?

$$\text{Amplitude do IC} = 2\, z_{\gamma/2} \frac{\sigma}{\sqrt{n}} = 3$$

$\sigma = 4{,}5$




$z_{\gamma/2} = 1{,}64$

$$2 \cdot 1{,}64 \cdot \frac{4{,}5}{\sqrt{n}} = 3$$

$\sqrt{n} = 4{,}92 \qquad n \simeq 25$

a margem de erro é $\varepsilon = z_{\gamma/2} \dfrac{\sigma}{\sqrt{n}} = 1{,}5$

Amplitude : $2\varepsilon = 3{,}0$ $\qquad$ IC : $[\bar{X} - 1{,}5 \,;\, \bar{X} + 1{,}5]$

Qual é o tamanho de amostra necessário para que o erro na estimação de $\mu$ seja no máximo 1 com probabilidade 0,95?

$$P(|\bar{X} - \mu| \leq 1) = 0{,}95$$

$$\gamma = 0{,}95 \Rightarrow z_{\gamma/2} = 1{,}96 \qquad \varepsilon = 1$$

$$n = \left(\frac{z_{\gamma}}{\varepsilon}\right)^2 \sigma^2 \Rightarrow n = \left(\frac{1{,}96}{1}\right)^2 \cdot 4{,}5^2 = 77{,}79 \simeq 78$$

$$IC: [\bar{X} - \varepsilon \,;\, \bar{X} + \varepsilon]$$

$$IC: [\bar{X} - 1 , \bar{X} + 1]$$

Aumento no coeficiente de Confiança
Redução no erro $\Big\} \Rightarrow$

$\Rightarrow$ aumento em $n$

2.º Caso - Intervalo de Confiança para a média $\mu$ de uma variável com variância $\sigma^2$ conhecida (não há a suposição de distribuição normal)

Pelo Teorema do Limite Central,

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \xrightarrow[n \to \infty]{} N(0,1)$$

Portanto, fixado $\gamma$, determinamos $z_{\gamma/2}$ tal que

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

$Z \sim N(0,1)$

$\therefore \left[\bar{X} - z_{\gamma/2}\frac{\sigma}{\sqrt{n}} ; \bar{X} + z_{\gamma/2}\frac{\sigma}{\sqrt{n}}\right]$ é um IC aproximado para $\mu$, com coeficiente de confiança $\gamma$.

Exemplo

Consideremos a variável aleatória

X - tempo de conexão à Internet em uma certa população, com distribuição de probabilidades desconhecida

$E(X) = \mu \rightarrow$ desconhecido

$\mu$: tempo médio de conexão populacional

Por experiências anteriores $\sigma = \sqrt{50}$ minutos

Uma amostra aleatória de 500 conexões resultou em tempo médio de 25 minutos

Construir um IC para $\mu$ com coeficiente de confiança 0,92.

$\bar{x} = 25 \quad \sigma = \sqrt{50} \quad n = 500 \quad \gamma = 0{,}92$


$$\left[\bar{X} - z_{\gamma/2}\frac{\sigma}{\sqrt{n}} \,;\, \bar{X} + z_{\gamma/2}\frac{\sigma}{\sqrt{n}}\right]$$

$$\left[25 - 1{,}75\frac{\sqrt{50}}{\sqrt{500}} \,;\, 25 + 1{,}75\frac{\sqrt{50}}{\sqrt{500}}\right] = \left[24{,}45 \,;\, 25{,}55\right]$$