Cap. 7 Estimação intervalar

- Porquê estimação intervalar?

→ Uma estimativa pontual de um parâmetro não contém informação sobre a sua precisão.

→ Vamos construir estimativas na forma de intervalos e determinar a confiança do intervalo conter o real valor do parâmetro.

→ Em geral, o intervalo é da forma $(\hat{\theta}-\varepsilon, \hat{\theta}+\varepsilon)$, em que $\varepsilon > 0$ é uma medida da precisão da estimativa $\hat{\theta}$ do parâmetro θ.

- Que intervalos?

Def.: Seja $(X_1, \ldots, X_n)$ uma a.a. de uma pop. X com função de prob. $f(x;\theta)$, com $\theta \in \Theta$.
Sejam $T_1 = T_1(X_1, \ldots, X_n)$ e $T_2 = T_2(X_1, \ldots, X_n)$ duas estatísticas t. q. $T_1 < T_2$. Um intervalo aleatório de confiança $1-\alpha$ para θ é um intervalo (T_1, T_2) tal que

$$P(T_1 < \theta < T_2) = 1-\alpha =: \gamma, \quad \alpha \in]0,1[.$$

Dada uma amostra particular $(x_1, \ldots, x_n)$, a concretização do IAC para θ com grau de confiança $(\gamma \times 100)\%$ é denominada por intervalo de confiança a $(\gamma \times 100)\%$ para θ

e é dado por $(t_1 = T_1(x_1, \dots, x_n),\ t_2 = T_2(x_1, \dots, x_n))$

- Como obter intervalos de confiança?

Def.: Uma variável fulcral ou pivô é uma função de a.a. $(X_1, \dots, X_n)$ e do parâmetro θ com distribuição de prob. independente de θ.

Método da variável fulcral

$Z \equiv Z(X_1, \dots, X_n; \theta)$ = variável fulcral

1. Fixa-se o grau de confiança $\gamma = 1-\alpha$ e obtêm-se números c_α e d_α t.q. $P(c_\alpha < Z < d_\alpha) = \gamma = 1-\alpha$

2. Normalmente escolhem-se c_α e d_α t.q.
$$P(Z < c_\alpha) = \frac{\alpha}{2} = P(Z > d_\alpha)$$

3. Resolve-se a desigualdade $c_\alpha < Z < d_\alpha$ em ordem a θ de modo a obter
$$P(c_\alpha < Z < d_\alpha) = P(T_1 < \theta < T_2) = \gamma = 1-\alpha,$$
i.e. um IAC a γ para θ, (T_1, T_2).

4. Dada uma amostra particular $(x_1, \dots, x_n)$, (t_1, t_2) é o IC a $(\gamma \times 100)\%$ para θ.

Observação: $P(T_1 < \theta < T_2) = \gamma$ significa o seguinte: sendo N = # de observações repetidas de $(X_1, \dots, X_n)$ e $K(N)$ = # de intervalos (t_1, t_2) que contêm θ

então $\lim_{N \to +\infty} \frac{\kappa(N)}{N} = \gamma$.

No entanto: $\gamma \neq P(t_1 < \theta < t_2) = \begin{cases} 1, & \theta \in (t_1, t_2) \\ 0, & \text{c.c.} \end{cases}$

- Variáveis aleatórias fulcrais que vamos considerar

① IC para μ de uma normal, σ^2 conhecido

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

② IC para μ de uma normal, σ^2 desconhecido

$$\hat{\mu} = \overline{X}, \qquad T = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

t_{n-1} = distribuição t-Student com $n-1$ graus de liberdade

③ IC para σ^2 de uma normal, μ desconhecido

$$\hat{\sigma}^2 = S^2, \qquad Q = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

χ^2_{n-1} = distribuição qui-quadrado com $n-1$ graus de liberdade

④ IC aproximado para uma probabilidade de sucesso (proporção populacional)

$X \sim Ber(p)$, Y = # sucessos na amostra

$$\hat{p} = \bar{X} = \frac{Y}{n}, \quad Z = \frac{\bar{X} - p}{\sqrt{\frac{\bar{X}(1-\bar{X})}{n}}} \overset{a}{\sim} N(0,1)$$

7.1 IC para μ de uma normal, σ^2 conhecida

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

$V(X) = \sigma^2$ conhecida

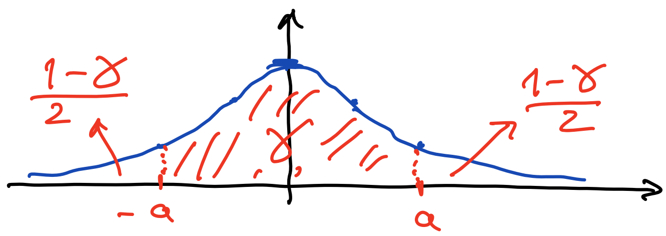
$(X_1, \dots, X_n)$ a.a., $\hat{\mu} = \bar{X}$ = estimador pontual de μ

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \leftarrow \text{independente de } \mu$$

$\nwarrow$ depende apenas de μ e da a.a.

$\Rightarrow$ Z é uma v.a. fulcral

Dado $\gamma = 1 - \alpha$ $(\gamma \approx 1, \alpha \approx 0)$ e tirando partido da simetria da normal, determinamos $a > 0$ tal que



$$P(-a \leq Z \leq a) = \gamma \Leftrightarrow \Phi(a) = 1 - \frac{1-\gamma}{2} = \frac{1+\gamma}{2}$$

$$\Leftrightarrow \quad a = \Phi^{-1}\left(\frac{1+\gamma}{2}\right)$$

Com este valor de a temos então que:

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

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

$$\Leftrightarrow P\left(\mu \in \left[\bar{X} - \frac{a\sigma}{\sqrt{n}}, \bar{X} + \frac{a\sigma}{\sqrt{n}}\right]\right) = \gamma$$

Ou seja $IAC_\gamma(\mu) = \left[\underbrace{\overline{X} - \frac{a\sigma}{\sqrt{n}}}_{= T_1}, \underbrace{\overline{X} + \frac{a\sigma}{\sqrt{n}}}_{= T_2}\right]$ e

$$\boxed{IC_{(\gamma\times100)\%}(\mu) = \left[\bar{x} - \frac{a\sigma}{\sqrt{n}}, \bar{x} + \frac{a\sigma}{\sqrt{n}}\right]}$$

Exemplo 1: $X \sim N(\mu, 1)$, amostra com $n=10$ e valores observados $(8.7, 9.1, 10.0, 11.9, 11.7, 8.9, 10.4, 11.2, 10.2, 8.9)$

(a) Determine um IC a 95% para μ.

$$\gamma = 0.95 \Rightarrow a = \Phi^{-1}\left(\frac{1+\gamma}{2}\right) = \Phi^{-1}(0.975) \overset{\text{tabela}}{=} 1.96$$

$$\Rightarrow IAC_{95\%}(\mu) = \left[\overline{X} - \frac{1.96}{\sqrt{10}}, \overline{X} + \frac{1.96}{\sqrt{10}}\right] \simeq \left[\overline{X} - 0.62, \overline{X} + 0.62\right]$$

Como $\bar{x} = 10.1$, temos que

$$IC_{95\%}(\mu) = [10.1 - 0.62, 10.1 + 0.62] = [9.48, 10.72]$$

(b) Determine a menor dimensão n de uma amostra que permite garantir com 95% de confiança que $|\bar{x} - \mu| \leq 0.25$.

- $IC_{(\gamma\times100)\%}(\mu) = \left[\bar{x} - \frac{a\sigma}{\sqrt{n}}, \bar{x} + \frac{a\sigma}{\sqrt{n}}\right] \Leftrightarrow |\bar{x} - \mu| \leq \frac{a\sigma}{\sqrt{n}}$

$$\frac{a\sigma}{\sqrt{n}} \leq 0.25 \Leftrightarrow \sqrt{n} \geq \frac{a\sigma}{0.25} = \frac{1.96}{0.25} = 7.84$$

$$\Leftrightarrow n \geq 61.47$$

$$\Rightarrow \boxed{n = 62}$$

<u>Exemplo 2</u>: $(X_1, \dots, X_n)$ a.a. de $X \sim N(\mu, 16)$.

Se $IAC_\gamma(\mu) = \left[\overline{X} - \frac{12}{\sqrt{n}}, \overline{X} + \frac{12}{\sqrt{n}}\right]$, qual é o nível de confiança γ deste intervalo.

- $\frac{a\sigma}{\sqrt{n}} = \frac{12}{\sqrt{n}} \Leftrightarrow \frac{4a}{\sqrt{n}} = \frac{12}{\sqrt{n}} \Leftrightarrow a = 3$

- $\Phi^{-1}\left(\frac{1+\gamma}{2}\right) = 3 \Leftrightarrow \frac{1+\gamma}{2} = \Phi(3) \overset{\text{tabela}}{=} 0.998650$

$\Rightarrow \gamma = 2(0.998650) - 1 = 0.9973$ //