Intervalos de confiança

- Intervalos de confiança;
- Caso normal: média;
- Intervalos de confiança unilaterais;
- Estatística pivotal.

Lembremos que

$$U = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sqrt{\frac{\Delta^2}{n-1}}} \sim \mathrm{T}(n-1). \tag{26}$$

Para $c > 0$, podemos computar $\Pr(-c < U < c) = \gamma$:

$$\begin{aligned}
&\Pr\left(-c < \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sqrt{\frac{\Delta^2}{n-1}}} < c\right) = \gamma,\\
&\Pr\left(\bar{X}_n - \frac{c\sigma'}{\sqrt{n}} < \mu < \bar{X}_n + \frac{c\sigma'}{\sqrt{n}}\right) = \gamma,\\
&T_{n-1}(c) - T_{n-1}(-c) = 2T_{n-1}(c) - 1 = \gamma.
\end{aligned}$$

Concluímos que $c = F_T^{-1}\left(\frac{1+\gamma}{2}; n-1\right)$.

Definição de intervalo de confiança

O conceito de **intervalo de confiança** é fundamental em Estatística e nas aplicações em Ciência.

Definição 37 (Intervalo de confiança)

Seja $\boldsymbol{X} = \{X_1, X_2, \ldots, X_n\}$ *uma amostra aleatória, cada variável aleatória com p.d.f.* $f(x \mid \theta)$, *e considere uma função real* $g(\theta)$. *Sejam* $A(\boldsymbol{X})$ *e* $B(\boldsymbol{X})$ *duas estatísticas de modo que valha*

$$\Pr\{A(\boldsymbol{X}) < g(\theta) < B(\boldsymbol{X})\} \geq \gamma. \tag{27}$$

Dizemos que $I(\boldsymbol{X}) = (A(\boldsymbol{X}), B(\boldsymbol{X}))$ *é um* ***intervalo de confiança*** *de* $100\gamma\%$ *para* $g(\theta)$. *Se a desigualdade for uma igualdade para todo* $\theta \in \Omega$, *dizemos que o intervalo é* ***exato***.

Revisitando o caso Normal

No caso do intervalo de confiança para o parâmetro de média, temos

$$\Pr\left\{A(\boldsymbol{X}) < g(\mu) < B(\boldsymbol{X})\right\} \geq \gamma,$$

com $g(\mu) = \mu$ e

$$A(\boldsymbol{X}) = \bar{X}_n - \frac{c\sigma'}{\sqrt{n}} = \bar{X}_n - \frac{c\sqrt{\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2}}{\sqrt{n(n-1)}},$$

$$B(\boldsymbol{X}) = \bar{X}_n + \frac{c\sigma'}{\sqrt{n}} = \bar{X}_n + \frac{c\sqrt{\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2}}{\sqrt{n(n-1)}}.$$

Interpretação de um intervalo de confiança

ATENÇÃO: a interpretação de um intervalo é crucial. Muita gente confunde o que um intervalo de confiança significa!

Observação 19 (Um intervalo de confiança não é uma afirmação sobre o(s) parâmetro(s)!)

A afirmação probabilística da forma $\operatorname{Pr}\{A(\boldsymbol{X}) < g(\theta) < B(\boldsymbol{X})\} = \gamma$ *diz respeito à distribuição conjunta das variáveis aleatórias* $A(\boldsymbol{X})$ *e* $B(\boldsymbol{X})$ *para um valor fixo de* θ *– e, portanto, de* $g(\theta)$*.*

Ideia 3 (Intervalos de confiança são procedimentos)

Como de costume na teoria ortodoxa (frequentista), o foco da construção de um intervalo confiança está em dar garantias probabilísticas ***com relação à distribuição dos dados****. Dizer que* $\operatorname{Pr}\{A(\boldsymbol{X}) < g(\theta) < B(\boldsymbol{X})\} = \gamma$ *é dizer que, se eu gerasse* M *grande amostras aleatórias* $\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \ldots, \boldsymbol{X}^{(M)}$ *de tamanho* n *e construisse* M *intervalos* $I(\boldsymbol{X}^{(1)}), I(\boldsymbol{X}^{(2)}), \ldots, I(\boldsymbol{X}^{(M)})$*, eu esperaria encontrar:*

$$\frac{1}{M} \sum_{i=1}^{M} \mathbb{I}\left(g(\theta) \in I(\boldsymbol{X}^{(1)})\right) \approx \gamma.$$

O que aprendemos?

- Intervalos de confiança;
  "Um intervalo $(A(\boldsymbol{X}), B(\boldsymbol{X}))$ de confiança de 100γ% para $g(\theta)$ é tal que $\Pr\left[A(\boldsymbol{X}) < g(\theta) < B(\boldsymbol{X})\right] \geq \gamma$";
- Um intervalo de confiança é uma afirmação probabilística sobre **as estatísticas** $A(\boldsymbol{X})$ e $B(\boldsymbol{X})$ a partir da **distribuição conjunta dos dados**;

Leitura recomendada

- De Groot seção 8.5;
- Próxima aula: De Groot, seção 9.1;
- **Exercícios recomendados**
  - De Groot.

    Seção 8.5: 1, 4 e 6.