ACH2053 – Introdução à Estatística

Aula 11: Informação de Fisher

Valdinei Freire

valdinei.freire@usp.br

http://www.each.usp.br/valdinei

Escola de Artes, Ciências e Humanidades - USP

2025

Estatística e Estimadores

Seja um modelo paramétrico F_θ com espaço de parâmetros $\Omega \subseteq \mathbb{R}^d$ e seja $\mathbf{X} = (X_1, \ldots, X_n) \sim F_\theta$ para algum $\theta \in \Omega$. Um **estimador pontual** $\hat{\theta}$ de θ é uma estatística $r : \mathbb{R}^n \to \Omega$, com o propósito de estimar θ.

Exemplo: considere uma variável aleatória X e n amostras i.i.d. de X, $x_1, x_2, \ldots, x_n$. Construa um estimador para a esperança de X, isto é, $\mu = \mathsf{E}[X]$.

Considere o estimador:

$$T = \frac{\sum_{i=1}^n x_i}{n},$$

Lembre-se que o próprio estimador é uma variável aleatória:

- $\mathsf{E}[T] = \mu$, isto é, T tem viés nulo.
- $\mathsf{Var}[T] = \frac{\mathsf{Var}[X]}{n}$.

Teorema do Limite Central

Se as variáveis aleatórias $X_1, \ldots, X_n$ formam uma amostra aleatória de tamanho n de uma distribuição com média μ e desvio padrão σ $(0 < \sigma < \infty)$, então para cada x

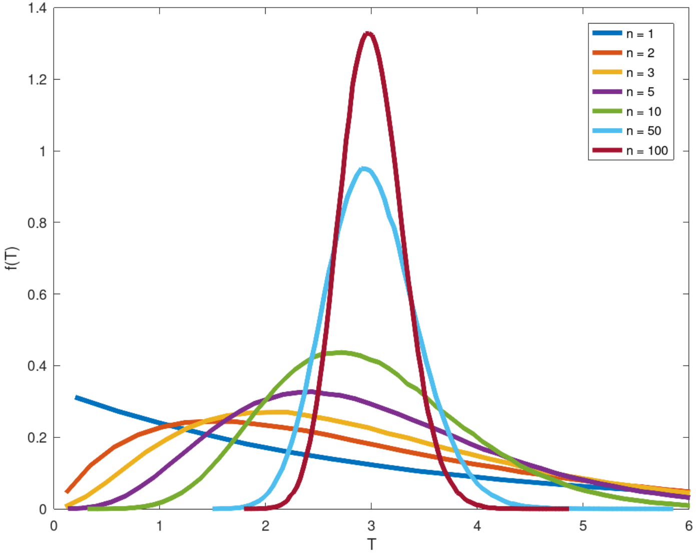
$$\lim_{n\to\infty} \Pr\left(\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \leq x\right) = \Phi(x),$$

e

$$\lim_{n\to\infty} \Pr\left(\frac{\sum_{i=1}^{n} X_n - n\mu}{\sqrt{n}\sigma} \leq x\right) = \Phi(x),$$

onde $\Phi(x)$ denota a c.d.f. da distribuição normal padrão.

Exemplo: X é exponencial



Informação de Fisher

- Nem todo estimador é o somatório de variáveis aleatórias, por exemplo, TRI.

- **Pergunta:** existe algo parecido com o Teorema do Limite Central para qualquer estimador?

- **Resposta:** Informação de Fisher
  1. quantidade de informação que uma amostra de dado tem sobre um parâmetro desconhecido
  2. intuitivamente, mais dado produz mais informação, mais dado informativo produz mais ainda informação
  3. a medida de informação pode ser utilizada para obter limites sobre a variância de estimadores
  4. a medida de informação permite aproximar a variância de estimadores para amostras grandes

Informação de Fisher: preliminares

Considere uma variável aleatória X tal que a p.d.f. é dada por $f(x|\theta)$, onde θ é um parâmetro desconhecido. Considere as seguintes condições:

- sabe-se que $\theta \in \Omega$, e Ω é um intervalo aberto.
- $x \in S$, sendo S conhecido e $f(x|\theta) > 0$ para todo $x \in S$ e $\theta \in \Omega$.
- defina $\lambda(x|\theta) = \log f(x|\theta)$, isto é, o loglikelihood.
- assuma que $\lambda(x|\theta)$ possui segunda derivada e defina:

$$\lambda'(x|\theta) = \frac{\partial}{\partial\theta}\lambda(x|\theta) \qquad \text{e} \qquad \lambda''(x|\theta) = \frac{\partial^2}{\partial\theta^2}\lambda(x|\theta)$$

Informação de Fisher: definição

Seja X uma variável aleatória com p.d.f. $f(x|\theta)$ e que todas as condições anteriores seja atendida. A Informação de Fisher $I(\theta)$ na variável aleatória X é definida por:

$$I(\theta) = \mathsf{E}_\theta[\lambda'(X|\theta)^2]$$

A Informação de Fisher também pode ser calculada por:

$$I(\theta) = -\mathsf{E}_\theta[\lambda''(X|\theta)]$$

ou

$$I(\theta) = \mathsf{Var}_\theta[\lambda'(X|\theta)]$$

Informação de Fisher: sobre amostras

Seja $X_1, \ldots, X_n$ variáveis aleatórias com p.d.f. conjunta $f(x_1, \ldots, x_n|\theta)$ e que todas as condições anteriores seja atendida. Defina $\lambda(x_1, \ldots, x_n|\theta) = \log f(x_1, \ldots, x_n|\theta)$.

A Informação de Fisher $I(\theta)$ na amostra aleatória $X_1, \ldots, X_n$ é definida por:

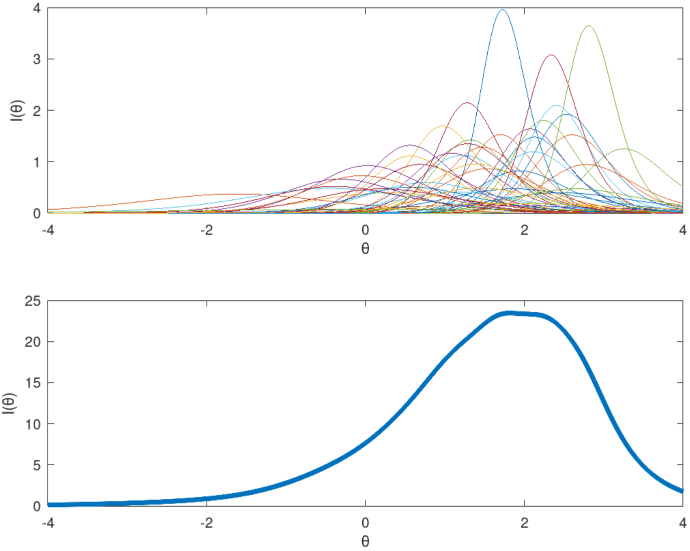
$$I(\theta) = \mathsf{E}_\theta[\lambda'(X_1, \ldots, X_n|\theta)^2]$$

Quando as variáveis aleatórias são independentes condicionadas em θ tem-se que:

$$\begin{aligned} I(\theta) &= -\mathsf{E}_\theta[\lambda''(X_1, \ldots, X_n|\theta)] \\ &= -\sum_{i=1}^{n} \mathsf{E}_\theta[\lambda''(X_i|\theta)] = \sum_{i=1}^{n} I_i(\theta) \end{aligned}$$

onde $I_i(\theta)$ é a Informação de Fisher na variável aleatória X_i.

Exemplo: TRI



Limite de Cramér-Rao

Theorem

Seja $X_1, \ldots, X_n$ *variáveis aleatórias com p.d.f. conjunta* $f(x_1, \ldots, x_n|\theta)$ *e que todas as condições anteriores seja atendida. Seja* T *um estimador para* θ *sem viés. Então:*

$$\mathit{Var}_\theta[T] \geq \frac{1}{\sum_{i=1}^{n} I_i(\theta)}$$

Theorem

Suponha que em um problema arbitrário o M.L.E. $\hat{\theta}_n$ *é determinado resolvendo a equação* $\lambda'(x_1, \ldots, x_n|\theta) = 0$ *e que todas as condições anteriores seja atendida. Então, a distribuição de* $\hat{\theta}_n$ *pode ser aproximada por uma distribuição normal com média* θ *e variância* $\frac{1}{\sum_{i=1}^{n} I_i(\theta)}$.