Exercícios: verossimilhança e suficiência

Disciplina: Inferência Estatística (MSc)
Instrutor: Luiz Carvalho

Setembro/2023

Motivação: A suficiência é um dos pilares dos chamados "princípios de redução de dados"; uma estatística suficiente encapsula toda a informação sobre um parâmetro e, portanto, deve ser preferida na busca por bons estimadores (veremos isso em detalhes em aulas posteriores). Nesta lista de exercícios vamos trabalhar o conceito de suficiência e a sua conexão com a função de verossimilhança – através do Teorema de Fatorização de Neyman-Fisher[1].

Notação: Como convenção adotamos $\mathbb{R} = (-\infty, \infty)$, $\mathbb{R}_+ = (0, \infty)$ e $\mathbb{N} = \{1, 2, \ldots\}$.

Dos livros-texto:

a) KN, Ch3: 4, 9a, 16a, 16b;

b) CB, Ch6: 6.1, 6.7.

c) * SV, Ch 2: 25, 29.

Extra:

1. Mostre que se $X_1, X_2, \ldots, X_n \sim P_\theta$ são amostra aleatória de uma família paramétrica qualquer, então

$$T(X_1, X_2, \ldots, X_n) = (X_{(1)}, X_{(2)}, \ldots, X_{(n)}),$$

é suficiente para θ.

2. Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória de P_θ e $x_1, x_2, \ldots, x_n$ uma amostra observada (realizada). Para cada um dos modelos abaixo, apresente a função de verossimilhança uma estatística suficiente para θ:

(a) (**Weibull**)

$$f_\theta(x) = \theta_1\theta_2 x^{\theta_2-1}\exp\left(-\theta_1 x^{-\theta_2}\right)\mathbb{I}(x \in \mathbb{R}_+), (\theta_1, \theta_2) \in \mathbb{R}_+ \times \mathbb{R}_+;$$

(b) (**Pareto**)

$$f_\theta(x) = \frac{\theta_1\theta_2^{\theta_1}}{x^{\theta_1+1}}\mathbb{I}\left(x \in (\theta_2, \infty)\right), (\theta_1, \theta_2) \in \mathbb{R}_+ \times \mathbb{R}_+;$$

[1]Ronald Aylmer Fisher (1890–1962), biólogo e estatístico inglês. Jerzy Neyman (1894–1981), matemático e estatístico polonês.

(c) (**Uniforme Discreta**)

$$P_\theta(X = x) = \frac{1}{\theta}\mathbb{I}(x \in \{1, 2, \ldots, \theta\}), \theta \in \mathbb{N};$$

(d) (**Binomial negativa**)

$$P_\theta(X = x) = \binom{x + \theta_1 - 1}{\theta_1 - 1}(1 - \theta_2)^x \theta_2^x, (\theta_1, \theta_2) \in \mathbb{N} \times (0, 1).$$

3. Dizemos que uma amostra aleatória tem distribuição gaussiana inversa se a densidade comum vale

$$f_\theta(x) = \sqrt{\frac{\theta_1}{2\pi x^3}} \exp\left(\frac{\theta_1(x - \theta_2)^2}{2\theta_2^2 x}\right) \mathbb{I}(x \in \mathbb{R}+),$$

para $(\theta_1, \theta_2) \in \mathbb{R}_+ \times \mathbb{R}_+$. Encontre uma estatística suficiente nos seguintes casos:

(a) $g(\theta) = \theta_1$ e θ_2 é conhecido;

(b) $g(\theta) = \theta_2$ e θ_1 é conhecido;

(c) $g(\theta) = (\theta_1, \theta_2)$, isto é, ambos são desconhecidos.

4. * Prove o 3.6 de KN (pág. 45);

5. * **Desafio**: Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma distribuição Uniforme$(0, \theta)$. Mostre que $T(X_1, X_2, \ldots, X_n) = \max(X_1, X_2, \ldots, X_n)$ é suficiente para θ **sem** usar o Teorema da Fatorização de Neyman-Fisher.

Dica: Seja $X_1, X_2, \ldots, X_n \sim P_\theta$ uma a.a. de uma família paramétrica qualquer. Se observamos a i-ésima estatística de ordem $X_{(i)} = a$, então:

i) Para todo $j < i$, $X_{(j)}|X_{(i)}$ tem distribuição com densidade $f_\theta(x)/F_\theta(x)$;

ii) Similarmente, para $k > i$, $X_{(k)}|X_{(i)}$ tem distribuição com densidade $f_\theta(x)/[1 - F_\theta(x)]$.