CE085 - Estatística Inferencial
Suficiência

Prof. Wagner Hugo Bonat

Laboratório de Estatística e Geoinformação - LEG
Curso de Bacharelado em Estatatística
Universidade Federal do Paraná - UFPR

8 de outubro de 2018

Conteúdo

Conteúdo

- Definição e motivação.
- Critério da fatorização de Fisher-Neyman.
- Teorema de Rao-Blackwell.
- Teorema de Lehmann-Scheffé.
- Exemplos.

Exemplo de motivação

- Seja $Y_i \sim N(\mu, \sigma^2)$ iid com log-verossimilhança

$$
\begin{aligned}
l(\mu, \sigma^2) &= -\frac{n}{2}\log(2\pi) - n\log(\sigma) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu)^2 \\
&= -\frac{n}{2}\log(2\pi) - n\log(\sigma) - \frac{1}{2\sigma^2}\left(\sum_{i=1}^{n} y_i^2 + n\mu^2 - 2\mu\sum_{i=1}^{n} y_i\right).
\end{aligned}
$$

- Note que a log-verossimilhança é determinada por

$$S_Y = \sum_{i=1}^{n} y_i \quad \text{e} \quad S_{YY} = \sum_{i=1}^{n} y_i^2,$$

a estatística $(S_Y, S_{YY})^\top$ é chamada **suficiente** para (μ, σ^2).

Exemplo de motivação

- Podemos escrever a log-verossimilhança apenas como função das estatísticas suficientes

$$l(\mu, \sigma^2) = -\frac{n}{2}\log(2\pi) - n\log(\sigma) - \frac{1}{2\sigma^2}(S_{YY} + n\mu^2 - 2\mu S_Y).$$

- Dimensão da estatística suficiente é igual ao número de parâmetros (neste caso).
- Suponha que $\sigma^2 = 1$ neste caso a log-verossimilhança é

$$l(\mu) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}S_{YY} - \frac{1}{2}n\mu^2 + \mu S_Y.$$

- A constante $-\frac{n}{2}\log(2\pi) - \frac{1}{2}S_{YY}$ não influência o corpo da log-verossimilhança.

Definição: Sufiência

- Seja $Y_1, \ldots, Y_n$ iid com fdp ou fp $f(y; \boldsymbol{\theta})$ para $\boldsymbol{\theta} \in \Omega$. Seja

$$X_1 = u_1(Y_1, \ldots, Y_n)$$

uma estatística com fdp ou fp $f_{X_1}(y; \boldsymbol{\theta})$.

- **Definição**: A estatística X_1 é chamada **suficiente** para o parâmetro θ se e somente se

$$\frac{f(y_1; \theta) \ldots f(y_n; \theta)}{f_{X_1}(u_1(y_1, \ldots, y_n))} = H(y_1, \ldots, y_n),$$

onde $H(y_1, \ldots, y_n)$ é uma função que não depende de θ.

Definição: Sufiência

- Em outras palavras (caso discreto)

$$f(y_1, \ldots, y_n | x_1; \theta) = \frac{f(y_1; \theta), \ldots, f(y_n; \theta)}{f_{X_1}(u_1(y_1, \ldots, y_n); \theta)}$$

- Caso contínuo

$$f(y_1, \ldots, y_n | x_1; \theta) \propto \frac{f(y_1; \theta), \ldots, f(y_n; \theta)}{f_{X_1}(u_1(y_1, \ldots, y_n); \theta)}$$

- Exemplo: Distribuição gamma.

Critério da fatorização de Fisher-Neyman

- A estatística $X_1 = u_1(Y_1, \ldots, Y_n)$ é suficiente para θ se e somente se

$$f(y_1; \theta) \ldots f(y_n; \theta) = k_1(u_1(y_1, \ldots, y_n); \theta) k_2(y_1, \ldots, y_n),$$

  onde $k_2(y_1, \ldots, y_n)$ não depende de θ.
- Note que o lado esquerdo é a verossimilhança, assim

$$L(\theta) \propto k_1(X_1; \theta).$$

- Além disso, a log-verossimilhança é

$$l(\theta) = const + \log k_1(X_1; \theta).$$

- De onde segue que a função escore

$$U(\theta) = \frac{\partial}{\partial \theta} \log k_1(X_1; \theta).$$

Exemplos: Distribuição power

- Considere $Y_1, \ldots, Y_n$ iid da distribuição power

$$f(y;\theta) = \theta y^{\theta-1} \quad \text{para} \quad 0 < y < 1,$$

onde $\theta > 0$. Considere a estatística $\prod_{i=1}^{n} Y_i$. Então,

$$f(y_1;\theta)\ldots, f(y_n;\theta) = \theta^n \prod_{i=1}^{n} x_i^{\theta-1}.$$

Note que esta função depende dos dados apenas através de $x_1 = \prod_{i=1}^{n} y_i$, assim pelo Teorema da fatorização x_1 é suficiente para θ.

Exemplos: Distribuição Weibull

- Considere a distribuição Weibull

$$f(y;\theta) = \theta y^{\theta-1} e^{-y^{\theta}}, \quad \text{for} \quad y > 0.$$

- Log-verossimilhança

$$l(\theta) = n\log(\theta) + (\theta - 1)\sum_{i=1}^{n} \log y_i - \sum_{i=1}^{n} e^{\theta \log y_i},$$

- Função escore

$$U(\theta) = \frac{n}{\theta} + \sum_{i=1}^{n} \log y_i - \sum_{i=1}^{n} \log(y_i) e^{\theta \log y_i}.$$

- Neste caso a amostra completa é a menor estatística suficiente.

Esperança condicional

- Se X e Y são v.a e X tem esperança então

$$E(E(X|Y)) = E(X)$$

e se X tem variância, então

$$Var(X) = E(Var(X|Y)) + Var(E(X|Y)).$$

- Disto, segue que

$$Var(X) \geq V(E(X|Y)).$$

Teorema de Rao-Blackwell

- Seja $X_1 = u_1(Y_1, \ldots, Y_n)$ uma estatística suficiente para θ e seja $X_2 = u_2(Y_1, \ldots, Y_n)$ um estimador não viciado para θ. Então,
$$\tilde{\theta} = E(Y_2|Y1),$$
é também um estimador não viciado para θ e $Var(\tilde{\theta}) \leq Var(Y_2)$ para todo $\theta \in \Omega$.
- Demonstração.

Estimador não-viciado de variância minima

- **Definição**: A estatística $X_2 = u_2(Y_1, \ldots, Y_n)$ é chamada estimador não-viciado de variância minima (MVUE) para θ se X_2 é não-viciado para θ e se a variância de X_2 é menor ou igual a variância de qualquer outro estimador não viciado para θ.
- Rao-Blackwell teorema diz que o MVUE é sempre função de uma estatística suficiente.
- A existência de MVUE não é única e nem sempre existe.

Teorema: EMV é função de uma estatística suficiente

- Se o EMV $\hat{\theta}$ é unicamente determinado apartir de $Y_1, \ldots, Y_n$ e $X_1 = u_1(Y_1, \ldots, Y_n)$ é uma estatística suficiente, então $\hat{\theta}$ é uma função de Y_1.
- Demonstração.
- Exemplo: Distribuição exponencial.

Familias completas

- **Definição**: A familia $\{f_{X_1}(\cdot;\theta)\}$ é chamada completa se a condição
$$E(u(X_1)) = 0, \quad \text{para todo} \theta \in \Omega.$$
- Também podemos dizer que a estatística X_1 é completa.
- Exemplo: Distribuição exponencial, Bernoulli e normal.

Teorema de Lehmann-Scheffé

- Seja $\tilde{\theta}$ um estimador não-viciado para θ, tal que $\tilde{\theta}$ é função de uma estatística suficiente e completa X_1. Então, $\tilde{\theta}$ é o único MVUE de θ.
- Demonstração.
- Exemplos: Família exponencial.