**CC0288 - Inferência Estatística I**

**Primeira Verificação de Aprendizagem - 24/03/2023.**

**Prof. Maurício**

1. (Valor 3 pontos) Seja $X$ uma única variável aleatória com distribuição de Bernoulli com parâmetro $\theta$. Sejam

$$\hat{\theta}_1 = X \quad e \quad \hat{\theta}_2 = \frac{1}{2}.$$

   a. Escreva a função de probabilidade de $X$ Identifique seu suporte, espaço paramétrico, sua média e variância.
   b. Verifique se $\hat{\theta}_1$ e $\hat{\theta}_2$ são não viciados para $\theta$.
   c. Compare os erros quadráticos médios dos dois estimadores. Faça um gráfico deles como função de $\theta$

**Solução:** $X \sim B(\theta)$

$$f(x|\theta) = \theta^x (1-\theta)^{1-x}\ I_A(x),\ A = \{0,1\}, 0 \leq \theta \leq 1.$$

O suporte é $A = \{0,1\}$ e o espaço paramétrico $\Theta = [0,1]$.

Além disso

$$E(X) = \theta \quad e \quad \sigma^2 = Var(X) = \theta(1-\theta).$$

Note que

$$E[\hat{\theta}_1] = E(X) = \theta.$$

Assim $\hat{\theta}_1$ é não viciado para $\theta$.

O erro quadrático médio de $\hat{\theta}_1$ é dado por:

$$EQM\left[\hat{\theta}_1\right] = E(X-\theta)^2 = Var(X) = \theta(1-\theta).$$

$$E[\hat{\theta}_2] = E\left(\frac{1}{2}\right) = \frac{1}{2} \neq \theta,$$

logo $\hat{\theta}_2$ é viciado para $\theta$.

Note que:

$$Var\left[\hat{\theta}_2\right] = Var\left[\frac{1}{2}\right] = 0.$$

O erro quadrático médio de $\hat{\theta}_2$ é dado por:

$$EQM\left[\hat{\theta}_2\right] = E(\frac{1}{2}-\theta)^2 = \left(\frac{1}{2}-\theta\right)^2.$$

Quando

$$EQM\left[\hat{\theta}_1\right] = EQM\left[\hat{\theta}_2\right]?$$

$$\theta(1-\theta) = \left(\frac{1}{2} - \theta\right)^2$$

$$\theta - \theta^2 = \frac{1}{4} - \theta + \theta^2$$

$$2\theta^2 - 2\theta + \frac{1}{4} = 0$$

$$\theta^2 - \theta + \frac{1}{8} = 0$$

Logo

$$\Delta = b^2 - 4ac = 1 - \frac{1}{2} = \frac{2}{4}.$$

$$\sqrt{\Delta} = \frac{\sqrt{2}}{2}.$$

Assim

$$\theta_1 = \frac{1 - \frac{\sqrt{2}}{2}}{2} = \frac{1}{2} - \frac{\sqrt{2}}{4}.$$

$$\theta_2 = \frac{1 + \frac{\sqrt{2}}{2}}{2} = \frac{1}{2} + \frac{\sqrt{2}}{4}.$$

Vamos fazer o gráfico no **R**:

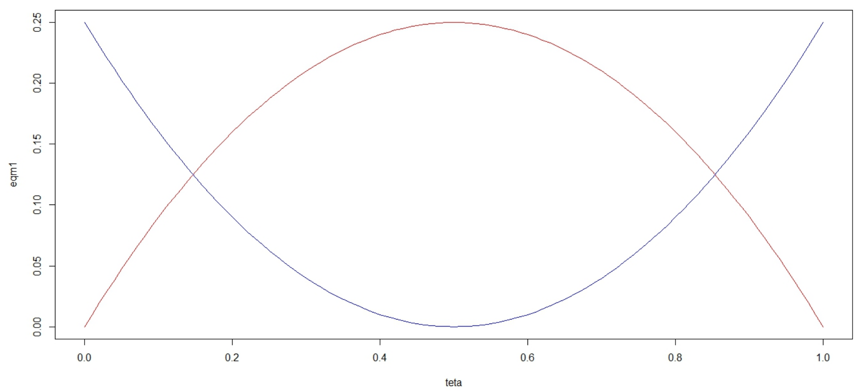


Figura 1:

Analisando o gráfico temos:

Para $\frac{1}{2} - \frac{\sqrt{2}}{4} < \theta < \frac{1}{2} + \frac{\sqrt{2}}{4}$ temos

$$EQM\left[\hat{\theta}_2\right] < EQM\left[\hat{\theta}_1\right]$$

Assim devemos escolher $\hat{\theta}_2$.

Para

$$0 < \theta < \frac{1}{2} - \frac{\sqrt{2}}{4} \text{ ou} \frac{1}{2} + \frac{\sqrt{2}}{4} \ \ \theta < 1$$

$$EQM\left[\hat{\theta}_1\right] < EQM\left[\hat{\theta}_2\right]$$

Assim devemos escolher $\hat{\theta}_1$.

Para

$$\theta = \frac{1}{2} - \frac{\sqrt{2}}{4} \text{ ou } \theta = \frac{1}{2} + \frac{\sqrt{2}}{4}$$

é indiferente a escolha.

2. (Valor 2 pontos) Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória de tamanho $n$ de $X$ com $E(X) = \mu$ e $Var(X) = \sigma^2$.

Sejam

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \quad e \quad S^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}.$$

a. Mostre que $\bar{X}$ é um estimador não viciado para $\mu$ e que sua variância é dada por $\frac{\sigma^2}{n}$.
b. Mostre que $S^2$ é um estimador não viciado para $\sigma^2$.

**Solução:**

De acordo com o enunciado temos para $i = 1, 2, \ldots, n$ :

$$E(X_i) = \mu, \ \ Var(X_i) = \sigma^2 \quad e \ \ E(X_i^2) = Var(X_i) + E^2(X_i) = \sigma^2 + \mu^2.$$

Além disso temos que as variáveis são independentes.

Vamos mostrar que $\bar{X}$um estimador não viciado para $\mu$.

$$E\left(\bar{X}\right) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$$

Colocando a constante para fora do operador esperança:

$$E\left(\bar{X}\right) = \frac{1}{n}\, E\left(\sum_{i=1}^{n} X_i\right)$$

$$E\left(\bar{X}\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} \mu = \frac{1}{n}\, n\, \mu = \mu.$$

Por outro lado sabemos que se as variáveis são independentes temos:

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) = \sum_{i=1}^{n} \sigma^2 = n\sigma^2.$$

$$Var\left(\bar{X}\right) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\, Var\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\ n\sigma^2 = \frac{\sigma^2}{n}.$$

Lembrando que

$$(n-1)S^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2 = \sum_{i=1}^{n} X_i^2 - n\bar{X}^2.$$

Aplicando o operador esperança temos:

$$(n-1)E(S^2) = \sum_{i=1}^{n} E(X_i^2) - nE(\bar{X}^2)$$

Lembrando que

$$E(\bar{X}^2) = Var(\bar{X}) + \mu^2 = \mu^2 + \frac{\sigma^2}{n}.$$

$$(n-1)E(S^2) = \sum_{i=1}^{n} (\sigma^2 + \mu^2) - n(\mu^2 + \frac{\sigma^2}{n}) =$$

$$(n-1)E(S^2) = n\mu^2 + n\sigma^2 - \sigma^2 - n\mu^2 = (n-1)\sigma^2.$$

Finalmente

$$E(S^2) = \sigma^2.$$

3. (Valor 5 pontos) Seja $X \sim Gama(3, \theta)$.

   a. Mostre que

$$f(x|\theta) = \frac{\theta^3}{2}\, x^2\, e^{-\theta x} \;\; I_{(0,\infty)}\,(x),\; \theta > 0.$$

   Identifique seu suporte e espaço paramétrico.

   b. Mostre que

$$\log(f(x|\theta)) = [c(\theta)T(x) + d(\theta) + b(x)]\; I_A(x),$$

   e $A$ não depende de $\theta$.

   c Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória de tamanho $n$ de $X$. Qual a distribuição conjunta da amostra?

   d. Qual a função de verossimilhança de $\theta$?

   e. Qual a lei de $S = \sum_{i=1}^{n} X_i$ ? Baseado em $S$ proponha um estimador não viciado para $g(\theta) = \frac{1}{\theta}$.

**Solução:**

No formulário temos para $X \sim Gama(r, \lambda)$.:

$$f(x) = \frac{\lambda^r}{\Gamma(r)}\, x^{r-1}\, e^{-\lambda x} \;\; I_A(x),\; A = (0, \infty), r > 0, \lambda > 0$$

Temos $r = 3 > 0$ e $\lambda = \theta > 0$.

Sabemos que $\Gamma(3) = 2! = 2$. Logo

$$f(x|\theta) = \frac{\theta^3}{2}\, x^2\, e^{-\theta x} \;\; I_{(0,\infty)}\,(x),\; \theta > 0.$$

O suporte de $X$

$$A(x) = \{x \in R | f(x \mid \theta > 0)\} = (0, \infty).$$

O espaço paramétrico

$$\Theta = (0, \infty)$$

e ele não depende do parâmetro $\theta$.

Note que:

$$\log(f(x|\theta)) = \log\left(\frac{\theta^3}{2}\right) + \log(x^2) + \log\left(e^{-\theta x}\right)$$

$$= \log(\theta^3) - \log(2) + 2\log(x) - \theta x$$
$$= -\theta x + 3\log(\theta) + 2\log(x) - \log(2)$$

Assim

$$c(\theta) = -\theta \ \ , \ T(x) = x \ \ \ , d(\theta) = 3\log(\theta), b(x) = 2\log(x) - \log(2).$$

A distribuição conjunta da amostra é dada por:

$$f(x_1, x_2, \ldots, x_n|\theta) = \prod_{i=1}^{n} f(x|\theta)$$

$$= \prod_{i=1}^{n} \frac{\theta^3}{2}\, x_i^2 \ e^{-\theta x_i} \ \ I_{(0,\infty)}\ (x_i)$$

$$= \frac{\theta^{3n}}{2^n} \ \ \prod_{i=1}^{n} x_i^2 \ \ e^{-\theta s} \ \prod_{i=1}^{n};\ I_{(0,\infty)}\ (x_i)$$

em que $s = \sum_{i=1}^{n} \ x_i$.

A função de verossimilhança é dada por:

$$L(\theta) = \frac{\theta^{3n}}{2^n} \ \ \prod_{i=1}^{n} x_i^2 \ \ e^{-\theta s}\ I_{(0,\infty)}\ (\theta).$$

A estimação de $\theta$ será baseado na Estatística:

$$S = \sum_{i=1}^{n} \ X_i.$$

Por isso vamos achar sua distribuição amostral.

A função geradora de momentos de $X$ é dada por:

$$M_X(t) = \left[\frac{\theta}{\theta - t}\right]^3, \ t < \theta.$$

A função geradora de momentos de $S = \sum_{i=1}^{n} X_i$, independentes, é dada por:

$$M_S(t) = [M_X(t)]^n$$

$$M_S(t) = \left[\left[\frac{\theta}{\theta - t}\right]^3\right]^n, t < \theta.$$

$$M_S(t) = \left[\frac{\theta}{\theta - t}\right]^{3n}, t < \theta,$$

Logo

$$S \sim Gama(3n, \theta).$$

Sabemos que

$$E(S) = \frac{r}{\lambda} = \frac{3n}{\theta}.$$

$$\frac{1}{3n}\, E(S) = \frac{1}{\theta} = g(\theta).$$

$$E\left(\frac{S}{3n}\right) = \frac{1}{\theta}$$

Assim

$$T = \frac{S}{3n} = \frac{n\bar{X}}{3n} = \frac{\bar{X}}{3},$$

é o nosso estimador procurado.

4. (Valor 1 ponto) Mostre que entre os estimadores lineares para $E(X) = \mu$ baseado em uma amostra aleatória $X_1, X_2, \ldots, X_n$ $T = \bar{X}$ tem a menor variância.

**Solução:**

Seja

$$T = \sum_{i=1}^{n} a_i \; X_i$$

,

Com $E(T) = \theta$.

Note que:

$$E(T) = \sum_{i=1}^{n} a_i \; E(X_i) = \sum_{i=1}^{n} a_i \;\; \theta = \theta$$

Logo

$$\sum_{i=1}^{n} a_i = 1$$

A variância de $T$ é dada por:

$$V(T) = \sum_{i=1}^{n} Var(a_i \; X_i) = \sum_{i=1}^{n} a_i^2 \; Var(X_i) = \sigma^2 \; \sum_{i=1}^{n} a_i^2.$$

Para minimizar a varIância de $T$ basta minimizar $\sum_{i=1}^{n} a_i^2$ sujeito à restrição $\sum_{i=1}^{n} a_i = 1$.

Note que:

$$\sum_{i=1}^{n} \left(a_i - \frac{1}{n}\right)^2 = \sum_{i=1}^{n} \left(a_i^2 - \frac{2}{n} \; a_i + \frac{1}{n^2}\right)$$

$$\sum_{i=1}^{n} \left(a_i - \frac{1}{n}\right)^2 = \sum_{i=1}^{n} \; a_i^2 - \frac{2}{n} \; \sum_{i=1}^{n} \; a_i + n \; \frac{1}{n^2}$$

$$\sum_{i=1}^{n} \left(a_i - \frac{1}{n}\right)^2 = \sum_{i=1}^{n} \; a_i^2 - \frac{2}{n} \quad \times 1 + \frac{1}{n}$$

$$\sum_{i=1}^{n} \left(a_i - \frac{1}{n}\right)^2 = \sum_{i=1}^{n} \; a_i^2 - \frac{1}{n}$$

Portanto

$$\sum_{i=1}^{n} \; a_i^2 = \sum_{i=1}^{n} \left(a_i - \frac{1}{n}\right)^2 + \frac{1}{n}.$$

O menor valor ocorre quando:

$$\sum_{i=1}^{n} \left( a_i - \frac{1}{n} \right)^2 = 0$$

isto é,

$$a_i - \frac{1}{n} = 0; \ \ ; a_i = \frac{1}{n}, i = 1, 2, \ldots, n$$

Logo

$$T_0 = \sum_{i=1}^{n} a_i \ X_i = \sum_{i=1}^{n} \ \frac{1}{n} \ X_i = \frac{\displaystyle\sum_{i=1}^{n} X_i}{n} = \bar{X}.$$

Você poderia provar por multiplicadores de Lagrange.