CC-0288 - Inferência Estatística I

Exemplo 1.3.8 - 17/03/2023

Prof. Maurício Mota

1. Vamos comentar o Exemplo 1.3.8 da página 14 do livro do Heleno e da Mônica.
   Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de $X \sim U(0, \theta)$.
   Sejam

$$\hat{\theta}_1 = \bar{X} \quad e \quad \hat{\theta}_2 = Y_n = max(X_1, X_2, X_n).$$

Escolha o melhor estimador para θ baseado no critério do menor erro quadrático médio.
A função densidade de probabilidade de X é:

$$f(x|\theta) = \frac{1}{\theta}\, I_A(x), \quad A = (0, \theta).$$

Note que o suporte A depende de θ. O espaço paramétrico é dado por:

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

O r-ésimo momento em relação à origem é dado por:

$$E(X^r) = \int_0^\theta x^r \frac{dx}{\theta} = \frac{1}{\theta} \quad \frac{x^{r+1}}{r+1}\Big|_0^\theta = \frac{\theta^r}{r+1}.$$

Assim,

$$\mu = E(X) = \frac{\theta}{2}.$$

$$E(X^2) = \frac{\theta^2}{3}.$$

$$\sigma^2 = V(X) = E(X^2) - E^2(X) = \frac{\theta^2}{3} - \frac{\theta^2}{4} = \frac{\theta^2}{12}.$$

Vamos estudar as propriedades de $\hat{\theta}_1 = \bar{X}$:

$$E[\hat{\theta}_1] = E[\bar{X}] = \mu = \frac{\theta}{2},$$

portanto é um estimador viciado.
Seu viés é dado por:

$$B[\hat{\theta}_1] = B[\hat{\theta}_1] - \theta = \frac{\theta}{2} - \theta = -\frac{\theta}{2}.$$

Sua variância é dada por:

$$Var[\hat{\theta}_1] = Var(\bar{X}) = \frac{\sigma^2}{n} = \frac{\frac{\theta^2}{12}}{n} = \frac{\theta^2}{12n}.$$

Seu erro médio quadrático é dado por:

$$EQM[\hat{\theta}_1] = Var[\hat{\theta}_1] + B^2[\hat{\theta}_1] = \frac{\theta^2}{12n} + \frac{\theta^2}{4} = \frac{1+3n}{12n}\,\theta^2$$

$$EQM[\hat{\theta}_1] = (\frac{1}{4} + \frac{1}{12n})\theta^2.$$

Note que

$$\lim_{n\to\infty} EQM[\hat{\theta}_1] = \frac{\theta^2}{4}.$$

Vamos calcular a f.d.p. de $Y_n = max(X_1, X_2, X_n)$:
Seja $G(y)$ a acumulada de Y_n. Assim

$$G(y) = P(Y_n \leq y) = P(max(X_1, X_2, X_n) \leq y) = P(X_1 \leq y, X_2 \leq y, \ldots, X_n \leq y)$$

Como as variáveis são independentes temos:

$$G(y) = P(X_1 \leq y)P(X_2 \leq y) \ldots P(X_n \leq y)$$

Como elas são identicamente distribuídas

$$G(y) = [F(y)]^n$$

Vamos obter a f.d.p. de Y_n:

$$g(y) = n[F(y)]^{n-1}\,F'(y) = n[F(y)]^{n-1}\,f(y).$$

A acumulada de X para $0 < x < \theta$ temos:

$$F(x) = \frac{x}{\theta}.$$

Finalmente:

$$g(y) = n[\frac{y}{\theta}]^{n-1}\,\frac{1}{\theta}\,I_A(y), \;\; A = (0, \theta)$$

$$g(y) = n\frac{y^{n-1}}{\theta^n} I_A(y), \;\; A = (0, \theta).$$

O r-ésimo momento em relação à origem é dado por:

$$E(Y_n^r) = \int_0^\theta y^r\; n\,\frac{y^{n-1}x}{\theta^n}\,dy = n\int_0^\theta \frac{y^{r+n-1}x}{\theta^n}\,dy = \frac{n}{\theta^n}\quad\frac{x^{r+n}}{r+n}\Big|_0^\theta = \frac{n\,\theta^r}{r+n}.$$

Assim,

$$\mu = E(Y_n) = \frac{n\theta}{n+1} \neq \theta.$$

Y_n é um estimador viciado para θ. Ele é assintoticamente não viciado.

Seu viés é dado por:

$$B[\hat{\theta}_2] = B[\hat{\theta}_2] - \theta = \frac{n\theta}{n+1} - \theta = -\frac{\theta}{n+1}.$$

$$E(Y_n^2) = \frac{n\ \theta^2}{n+2}.$$

$$V(Y_n) = E(Y_n^2) - E^2(Y_n) = \frac{n\ \theta^2}{n+2} - \frac{n^2\theta^2}{(n+1)^2} = n\ \theta^2 \left[\frac{1}{n+2} - \frac{n}{(n+1)^2}\right] = .$$

$$V(Y_n) = n\ \theta^2 \left[\frac{(n+1)^2 - n(n+2)}{(n+2)(n+1)^2}\right] = \frac{n\ \theta^2}{(n+2)(n+1)^2}$$

O erro médio quadrado de Yn é dado por:

$$EQM(Y_n) = Var(Y_n) + B^2(Y_n) = \frac{n\theta^2}{(n+2)(n+1)^2} + \frac{\theta^2}{(n+1)^2} = \frac{2(n+1)\theta^2}{(n+2)(n+1)^2} + \frac{\theta^2}{(n+1)^2}$$

$$EQM(Y_n) = \frac{2\theta^2}{(n+2)(n+1)}.$$

Para $n = 3$ temos:

$$EQM(Y_3) = \frac{2\theta^2}{(3+2)(3+1)} = \frac{2\theta^2}{20} = \frac{\theta^2}{10}.$$

$$EQM(\bar{X}_3) = \frac{1+9}{36}\ \theta^2 = \frac{5}{18}\theta^2.$$

$$\frac{EQM(Y_3)}{EQM(\bar{X}_3)} = 0,36.$$

Para $n = 5$ temos:

$$EQM(Y_5) = \frac{2\theta^2}{(5+2)(5+1)} = \frac{2\theta^2}{42} = \frac{\theta^2}{21}.$$

$$EQM(\bar{X}_5) = \frac{1+15}{60}\ \theta^2 = \frac{4}{15}\theta^2.$$

$$\frac{EQM(Y_5)}{EQM(\bar{X}_5)} = \frac{5}{28} = 0,1786.$$

Para $n = 10$ temos:

$$EQM(Y_{10}) = \frac{2\theta^2}{(10+2)(10+1)} = \frac{2\theta^2}{42} = \frac{\theta^2}{66}.$$

$$EQM(\bar{X}_{10}) = \frac{1+30}{120}\ \theta^2 = \frac{31}{120}\theta^2.$$

$$\frac{EQM(Y_{10})}{EQM(\bar{X}_{10})} = \frac{20}{343} = 0,0853.$$

Para $n = 20$ temos:

$$EQM(Y_{20}) = \frac{2\theta^2}{(20+2)(20+1)} = \frac{2\theta^2}{42} = \frac{\theta^2}{231}.$$

$$EQM(\bar{X}_{20}) = \frac{1+60}{240}\,\theta^2 = \frac{61}{240}\theta^2.$$

$$\frac{EQM(Y_{20})}{EQM(\bar{X}_{20})} = \frac{240}{13791} = 0,017.$$

No geral temos:

$$\frac{EQM(Y_n)}{EQM(\bar{X}_n)} = \frac{24n}{(n+1)(n+2)(3n+1)} < 1,$$

para $n > 1$.

Y_n é melhor que $\bar{X}$ para todo θ e $n > 1$.

Não se esqueça de olhar o comentário 01.