CC0288 - Inferê ncia Estatística I

Terceira Verificação de Aprendizagem - 03/05/2023.

Prof. Maurício

1. (Valor 6 pontos) Seja $X \sim Exp(\theta), \quad \theta > 0$. Baseado em $X_1, X_2, \ldots, X_n$ uma amostra aleatória de tamanho n de X responda ao que se pede.

   a. Qual o estimador pelo método dos momentos de θ?

   b. Qual o estimador pelo método dos mínimos quadrados de θ?

   c. Qual o estimador de Máxima Verossimilhança de θ?

   d. Mostre que X pertence à família exponencial de densidades e que a informação de Fisher é dada por $I_F(\theta) = \frac{1}{\theta^2}$.

   e. Mostre que $P(X > 2) = e^{-2\theta} = g(\theta)$.
   Encontre o estimador de Máxima Verossimilhança de $g(\theta)$ e sua distribuição aproximada em grandes amostras.

   f. Qual a estimativa de MV de θ e de $e^{-2\theta}$ baseado em :

```
> round(X,2)
[1]  0.02  0.38  1.23  0.92  3.20  1.04  0.69  2.48  0.63  1.96  0.45
[12]  1.09  1.52  1.00  1.20  2.74  0.49  0.95  1.58  0.64  0.45  2.54
[23]  1.30  0.61  1.07  0.08  3.15  1.19  0.19  0.14  1.30  3.66  1.08
[34]  4.17  1.24  0.48  7.51 13.82  1.35  0.95  1.18  2.89  0.89  3.87
[45]  4.11  0.15  1.35  1.42  1.43  2.93
>
> sum(X)
[1] 90.72231
>
> mean(X);1/mean(X);2/mean(X)
[1] 1.814446
[1] 0.5511323
[1] 1.102265
> exp(-mean(X))
[1] 0.1629281
>
> exp(-2*mean(X))
[1] 0.02654556
>
> exp(-mean(X))
[1] 0.1629281
>
> exp(-1/mean(X))
[1] 0.5762969
> exp(-2/mean(X))
[1] 0.3321181
>
```

Solução:

Vamos responder ao item **a**:

$$E_\theta(X) = \bar{X}.$$

$$\frac{1}{\theta} = \bar{X}.$$

$$\hat{\theta}_{MM} = \frac{1}{\bar{X}}.$$

Vamos responder ao item **b**:

$$S(\theta) = \sum_{i=1}^{n} [X_i - E(X_i)]^2 = \sum_{i=1}^{n} \left[X_i - \frac{1}{\theta}\right]^2$$

$$S(\theta) = \sum_{i=1}^{n} X_i^2 - \frac{2}{\theta} \sum_{i=1}^{n} X_i + \frac{n}{\theta^2}$$

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

Derivando em relação à θ temos:

$$S^{'}(\theta) = \frac{2n\bar{X}}{\theta^2} - \frac{2n}{\theta^3} = 2n\left[\frac{\bar{X}}{\theta^2} - \frac{1}{\theta^3}\right]$$

De $S^{'}(\theta) = 0$ temos:

$$\frac{\bar{X}}{\theta^2} - \frac{1}{\theta^3} = 0$$

temos

$$\theta = \frac{1}{\bar{X}}.$$

Vamos verificar se ele é ponto de mínimo.

A derivada segunda de $S(\theta)$ é dada por:

$$S^{''}(\theta) = 2n\left[-2\,\frac{\bar{X}}{\theta^3} + \frac{3}{\theta^4}\right]$$

Logo

$$S^{''}(1/\bar{X}) = 2n\left[-2\,\frac{\bar{X}}{1/\bar{X}^3} + \frac{3}{1/\bar{X}^4}\right] = 2n\left[-2\,\bar{X}^4 + 3\bar{X}^4\right] = 2n\bar{X}^4 > 0.$$

Portanto

$$\hat{\theta}_{MQ} = \frac{1}{\bar{X}}.$$

Vamos responder ao item **c**:

A função de verossimilhança de θ é dada por:

$$L(\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \theta \, e^{-\theta x_i} = \theta^n \, e^{-\theta \sum_{i=1}^{n} x_i}.$$

Fazendo

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

temos:

$$L(\theta) = \theta^n \, e^{-\theta \, s}.$$

$$l(\theta) = \log(L(\theta)) = n \log(\theta) - \theta \, s.$$

Assim,

$$l^{'}(\theta) = \frac{n}{\theta} - s.$$

$$l^{''}(\theta) = -\frac{n}{\theta^2} < 0.$$

De $l^{'}(\theta) = 0$ temos:

$$\frac{n}{\theta} = s$$

$$\theta = \frac{n}{s} = \frac{n}{n\bar{x}} = \frac{1}{\bar{x}}.$$

Portanto

$$\hat{\theta}_{MV} = \frac{1}{\bar{X}}.$$

Vamos responder ao item **d**:

Sabemos que:

$$f(x|\theta) = \theta \, e^{-\theta \, x} \, I_A(x), \quad A = (0, \infty).$$

Note que o suporte não depende de θ e que:

$$\log[f(x|\theta)] = log(\theta) - \theta x = -\theta x + log(\theta) + 0.$$

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

Assim

$$c(\theta) = -\theta \quad ; T(x) = x \ \ ; d(\theta) = \log(\theta) \quad ; b(x) = 0.$$

Mostramos assim que pertence á família exponencial.

Vamos calcular agora a informação de Fisher:

$$f(X|\theta) = \theta \ e^{-\theta \ X}$$

$$\log\left(f(X|\theta)\right) = \log\left(\theta\right) - \theta X.$$

$$V = \frac{\partial \log f(X;\theta)}{\partial \theta} = \frac{1}{\theta} - X$$

$$W = \frac{\partial^2 \log f(X;\theta)}{\partial \theta^2} = -\frac{1}{\theta^2}.$$

Assim

$$I_F(\theta) = E(-W) = E\left(\frac{1}{\theta^2}\right) = \frac{1}{\theta^2}.$$

Vamos responder ao item **e**:

$$P(X > 2) = \int_2^{\infty} \theta \ e^{-\theta x} \ dx = \lim_{c \to \infty} \int_2^{c} \theta \ e^{-\theta x} \ dx$$

$$P(X > 2) = \lim_{c \to \infty} \ -e^{-\theta x} \ \Bigg|_0^c = \lim_{c \to \infty} \left[e^{-2\theta} - e^{-2c}\right]$$

$$g(\theta) = P(X > 2) = e^{-2\theta} - \lim_{c \to \infty} \ e^{-2c} = e^{-2\theta}.$$

Vamos usar a propriedade de invariância.

O estimador de MV de $g(\theta)$ é dado por:

$$\widehat{g(\theta)} = g\left(\hat{\theta}\right) = g\left(\frac{1}{\bar{X}}\right) = e^{-\frac{2}{\bar{X}}}.$$

A derivada de $g(\theta)$ é dada por:

$$g'(\theta) = -2 \ e^{-2\theta}.$$

$$\left[g'(\theta)\right]^2 = 4 \ e^{-4\theta}.$$

Assim

$$\frac{\left[g'(\theta)\right]^2}{I_F(\theta)} = 4 \ \theta^2 \ e^{-4\theta}.$$

Sabemos que se o tamanho da amostra é grande e as condições de regularidade estão satisfeitas temos:

$$\sqrt{n}\ \left(g\left(\hat{\theta}\right) - g(\theta)\right) \overset{a}{\sim} N\left(0, \frac{(g'(\theta))^2}{I_F(\theta)}\right) = N\left(0,\ \ 4\,\theta^2\, e^{-4\theta}\right).$$

Vamos responder ao item **f**:

A estimativa de MV de θ é dada por:

$$\theta_{est} = \frac{1}{\bar{x}} = \frac{1}{1,8144} = 0,55.$$

A estimativa de MV de $g(\theta) = exp(-2\theta)$ é dada por:

$$g(\theta_{est}) = exp\ \left(-\frac{2}{\bar{x}}\right) = exp\left(-\frac{2}{1,8144}\right) = 0,33.$$

2. (Valor 3 pontos) Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória da distribuição da variável aleatória X

$$f(x|\theta) = e^{-(x-\theta)} \; I_A(x), \quad A = (\theta, \infty), \theta > 0.$$

a. (Valor 1 ponto) Encontre uma estatística suficiente para θ.

b. (Valor 1 ponto) Mostre que a estatística obtida no item **a** é completa para θ.

c. (Valor 1 ponto) Baseado nesta estatística suficiente e completa , obtenha um estimador não viciado para θ. Mostre que ele é consistente. Este estimador pode ser melhorado?

Formulário: Você pode usar sem provar:

$$\mu = E(X) = \theta + 1 \quad \sigma^2 = 1.$$

$$F(x) = 0 \text{ para } x \leq \theta \;\; F(x) = 1 - e^{-(x-\theta)} \text{ para } \; x > \theta.$$

A f.d.p. do Máximo é dada por:

$$g_{Y_n}(y) = n \; [F(y)]^{n-1} \; f(y).$$

A f.d.p. do Mínimo é dada por:

$$g_{Y_1}(y) = n \; [1 - F(y)]^{n-1} \; f(y).$$

Solução:

A distribuição conjunta da amostra é dada por:

$$f(x_1, x_2, \ldots, x_n|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} e^{-(x_i-\theta)} \; I_A(x_i).$$

$$f(x_1, x_2, \ldots, x_n|\theta) = e^{-s} \, e^{n\theta} \prod_{i=1}^{n} I_A(x_i)$$

Quando

$$\prod_{i=1}^{n} I_A(x_i) = 1.$$

Isto é equivalente a

$$I_A(x_i) = 1 \;\; i = 1, 2, \ldots, n.$$

Assim

$$x_i > \theta; \; i = 1, 2, \ldots, n.$$

$$x_1 > \theta, x_2 > \theta, \ldots, x_n > \theta$$

ou

$$y_1 = min(x_1, x_2, \ldots, x_n) > \theta$$

Logo

$$I_{(\theta,\infty)}(y_1) = 1.$$

Assim

$$f(x_1, x_2, \ldots, x_n|\theta) = e^{-s}\, e^{n\theta} \qquad I_A(y_1), \quad A = (\theta, \infty).$$

Fazendo

$$h(x_1, x_2, \ldots, x_n) = e^{-s} = e^{-\sum_{i=1}^{n} x_i}$$

$$g(y_1, \theta) = e^{n\theta} \qquad I_A(y_1), \quad A = (\theta, \infty).$$

$$f(x_1, x_2, \ldots, x_n|\theta) = h(x_1, x_2, \ldots, x_n) \times g(y_1, \theta)$$

temos que

$$Y_1 = min(X_1, X_2, \ldots, X_n)$$

é uma estatística suficiente para θ.

A função de sobrevivência de X para $x > \theta$ é dada por:

$$S(x) = e^{-(x-\theta)}.$$

A f.d.p. do mínimo é dada por:

$$g_{Y_1}(y) = n\ [S(y)]^{n-1}\ f(y)$$

$$g_{Y_1}(y) = n\ \left[e^{-(y-\theta)}\right]^{n-1}\ e^{-(y-\theta)}\ I_A(y)$$

$$g_{Y_1}(y) = n\ e^{-n(y-\theta)}\ I_A(y)$$

Vamos responder ao item **b**:

Vamos supor que

$$E[h(Y_1)] = 0, \quad \forall\, \theta > 0.$$

Assim,

$$E[h(Y_1)] = \int_{\theta}^{\infty} h(y)\ n\ e^{-n(y-\theta)}\ dy = 0, \quad \forall\, \theta > 0.$$

Diferenciado em relação à θ temos:

$$-n\ h(\theta)\ \ e^{-n(\theta-\theta)}\ (1) = 0.$$

$$h(\theta) = 0\ \forall\ \theta > 0.$$

Finalmente

De

$$E[h(Y_1)] = 0$$

temos

$$Y_1 = 0.$$

Logo Y_1 também é completa!!!!

Vamos calcular a esperança de Y_1:

$$E(Y_1) = \int_{\theta}^{\infty}\ y\ n\ e^{-n(y-\theta)}\ dy$$

Fazendo a mudança de variável

$$u = n(y-\theta) \quad ; du = n\ dy.$$

$$y = \frac{u}{n} + \theta.$$

$$E(Y_1) = \int_{0}^{\infty}\ (\frac{u}{n} + \theta)\ e^{-u}\ du$$

$$E(Y_1) = \frac{1}{n}\ \int_{0}^{\infty} u\ e^{-u}\ du + \theta\ \int_{0}^{\infty}\ e^{-u}\ du$$

$$E(Y_1) = \frac{1}{n}\ \Gamma(2) + \theta\ \Gamma(1) = \frac{1}{n} \times 1! + \theta\ 0! = \frac{1}{n} + \theta$$

Logo

$$E(Y_1) - \frac{1}{n} = E\left(Y_1 - \frac{1}{n}\right) = \theta.$$

Assim

$$T = Y_1 - \frac{1}{n}$$

é o nosso estimador procurado. Ele não pode ser melhorado pelo Teorema de Lehmann-Scheffé pois T é um estimador não viciado função de uma estatística suficiente e completa .

Calcule a variância de T:

Note que

$$Var(T) = Var\left(Y_1 - \frac{1}{n}\right) = Var(Y_1).$$

Vamos calcular a esperança de Y_1^2:

$$E(Y_1^2) = \int_\theta^\infty y^2\ n\ e^{-n(y-\theta)}\ dy$$

Fazendo a mesma mudança de variável temos:

$$E(Y_1^2) = \int_0^\infty (\frac{u}{n} + \theta)^2\ e^{-u}\ du$$

$$E(Y_1^2) = \frac{1}{n^2}\ \int_0^\infty u^2\ e^{-u}\ du + \frac{2\theta}{n}\ \int_0^\infty u\ e^{-u}\ du + \theta^2\ \int_0^\infty\ e^{-u}\ du$$

$$E(Y_1^2) = \frac{1}{n^2}\ \Gamma(3) + \frac{2\theta}{n}\ \Gamma(2) + \theta^2.$$

$$E(Y_1^2) = \frac{2}{n^2} + \frac{2\theta}{n} + \theta^2 = \frac{1}{n^2} + \frac{1}{n^2} + \frac{2\theta}{n} + \theta^2$$

$$E(Y_1^2) = \frac{1}{n^2} + \left[\theta + \frac{1}{n}\right]^2$$

A variância de Y_1 é dada por:

$$Var(Y_1) = \frac{1}{n^2} + \left[\theta + \frac{1}{n}\right]^2 - \left[\theta + \frac{1}{n}\right]^2 = \frac{1}{n^2}.$$

Mostre Y_1 é um estimador assintoticamente não viciado de θ.

$$\lim_{n\to\infty} E(Y_1) = \lim_{n\to\infty} \left[\theta + \frac{1}{n}\right] = \theta.$$

Mostre Y_1 é um estimador consistente para θ.
Note que

$$\lim_{n\to\infty} Var(Y_1) = \lim_{n\to\infty} \frac{1}{n^2} = 0.$$

Assim como ele é assintoticamente não viciado e $\lim\limits_{n\to\infty}\ Var(Y_1) = 0$, ele é consistente.

3. (Valor 1 ponto) Seja $\sim U(0, \theta)$. Uma amostra aleatória $X_1, X_2, \ldots, X_n$ é retirada.

   Sabemos que $Y_n = Max(X_1, X_2, \ldots, X_n)$ é uma estatística suficiente e completa para θ.

   Além disso seja

$$T^* = \frac{n+1}{n}\, Y_n$$

   Provamos em sala de aula que $E(T^*) = \theta$.

   Quem é o estimador não viciado de variância uniformemente mínima de θ? Diga o nome do resultado que garante suas resposta.

   Agora sabemos

$$T = 2\bar{X}$$

   é um estimador não viciado de θ.

   Seja um novo estimador para θ definido por:

$$T_1 = E\left(2\bar{X}\ |Y_n\right).$$

   Identifique T_1. Diga o nome do resultado que garante suas resposta.

   Solução: Como T^* é um estimador não viciado função de uma estatística suficiente e completa ele é o ENVVUM para θ usando Lehmann-Scheffé.

   Sabemos que se T é uma estatística suficiente e completa para θ e S um estimador não viciado para θ então

$$\hat{\theta} = E(S|T)$$

   é o ENVVUM para θ.

   Assim

$$T_1 = E\left(2\bar{X}\ |Y_n\right) = \frac{n+1}{n}\, Y_n$$

   pelo Teorema de Lehmann-Scheffé ou Rao-Blackwell.