6.1 Método da máxima verosimilhança (cont.)

Def.: X – v.a. com função massa/densidade de probabilidade $f(x, \theta)$, onde θ é um parâmetro desconhecido.

$(x_1, \ldots, x_n)$ – valores observados de uma a.a. $(X_1, \ldots, X_n)$ de dimensão n.

A função de verosimilhança da amostra é

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

A estimativa de máxima verosimilhança de θ é o valor $\hat{\theta}$ que maximiza L, i.e. tal que

$$\boxed{L(\hat{\theta} | x_1, \ldots, x_n) = \max_{\theta} L(\theta | x_1, \ldots, x_n)}$$

Observação:

- O estimador de MV de θ, denotado por $EMV(\theta)$, obtém-se substituindo $x_1, \ldots, x_n$ por $X_1, \ldots, X_n$, na expressão geral de $\hat{\theta}$ = estimativa de MV.

- Maximizar $L \iff$ Maximizar $\log L$

  $\log L$ – log-verosimilhança

Cálculos com $\log L$ são normalmente mais simples do que os cálculos com L.

• **Exemplo** (estimação por MV do parâmetro λ da Exp)

$X \sim Exp(\lambda)$, $\lambda > 0$, $f(x) = \lambda e^{-\lambda x}$, $x > 0$.

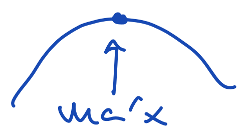
Dada uma amostra $(x_1, \dots, x_n)$, temos que

$$L(\lambda) \equiv L(\lambda \mid x_1, \dots, x_n) = f(x_1) \times \dots \times f(x_n) =$$

$$= \prod_{i=1}^{n} \left(\lambda e^{-\lambda x_i} \right) = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}, \quad \lambda > 0, \quad x_i > 0$$

$$\Rightarrow \log L(\lambda) = n \log(\lambda) - \lambda \left(\sum_{i=1}^{n} x_i \right)$$

$$\Rightarrow \begin{cases} \dfrac{d \log L}{d\lambda}(\lambda) = \dfrac{n}{\lambda} - \left(\sum_{i=1}^{n} x_i \right) = 0 \Rightarrow \lambda = \dfrac{n}{\sum_{i=1}^{n} x_i} \\[2ex] \dfrac{d^2 \log L}{d\lambda^2}(\lambda) = -\dfrac{n}{\lambda^2} < 0, \quad \forall \lambda > 0 \end{cases}$$



$$\Rightarrow \boxed{\begin{array}{ll} \hat{\lambda} = \dfrac{n}{\sum_{i=1}^{n} x_i} = \dfrac{1}{\bar{x}} = & \text{estimativa de MV de } \lambda \\ & \text{(valor numérico)} \\[2ex] EMV(\lambda) = \dfrac{n}{\sum_{i=1}^{n} X_i} = \dfrac{1}{\bar{X}} = & \text{estimador de MV de } \lambda \\ & \text{(variável aleatória)} \end{array}}$$

• **Propriedade da invariância dos estimadores de MV**

Se $EMV(\theta)$ é o estimador de MV do parâmetro θ, então $g(EMV(\theta))$ é o estimador de MV de $g(\theta)$, i.e.

$$\boxed{EMV(g(\theta)) = g(EMV(\theta))}.$$

Exemplo Seja $X_1, X_2, \ldots, X_n$ uma a.a. de uma população X com distribuição $Exp(\lambda)$. Determine o estimador de MV de $P(X>2)$.

$$P(X>2) = \int_2^{+\infty} f_X(x)\,dx = \int_2^{+\infty} \lambda e^{-\lambda x}\,dx = \left[-e^{-\lambda x}\right]_2^{+\infty} =$$

$$= -e^{-\infty} + e^{-2\lambda} = e^{-2\lambda}$$

$$\Rightarrow EMV\left(P(X>2)\right) = EMV\left(e^{-2\lambda}\right) =$$

$$= e^{-2EMV(\lambda)} = e^{-2/\overline{X}} \;//$$

- **Exercício 6.1**: Considere uma urna com bolas brancas e pretas na proporção de 3/1, desconhecendo-se qual a cor dominante. Seja p = prob. de sair bola preta numa extração. Qual a estimativa de MV de p se, ao extrairmos c/ reposição 3 bolas da urna, encontrássemos

  (a) 1 bola preta?

  (b) 2 bolas pretas?

  Suponha agora que desconhecíamos qq relação entre o nº de bolas brancas e pretas.

  (c) Qual a estimativa de MV de p se, ao extrairmos 3 bolas c/ reposição, encontrássemos 2 bolas pretas?

$X =$ v.a. # bolas pretas em 3 extrações c/ reposição

$\Rightarrow X \sim Bin(3, p)$

(a) A.a. $X_1 \sim Bin(3,p)$ c/ valor $X_1 = x_1 = 1$

$$\Rightarrow L(p) \equiv L(p \mid x_1 = 1) = P(X_1 = 1) = \binom{3}{1} p^1 (1-p)^{3-1}$$

$$= 3\, p\, (1-p)^2$$

Tendo em conta o racio 3/1, só há 2 possibilidades para p:

(1) $p = 1/4 \Rightarrow L(1/4) = 3 \times \frac{1}{4} \cdot \left(\frac{3}{4}\right)^2 = \frac{27}{64}$

(2) $p = 3/4 \Rightarrow L(3/4) = 3 \times \frac{3}{4} \times \left(\frac{1}{4}\right)^2 = 9/64$

$\Rightarrow$ Máx em $p = 1/4$

$$\Rightarrow \boxed{\hat{p} = 1/4}$$

(b) A.a. $X_1 \sim Bin(3,p)$ c/ valor $X_1 = x_1 = 2$

$$\Rightarrow L(p) \equiv L(p \mid x_1 = 2) = P(X_1 = 2) = \binom{3}{2} p^2 (1-p)^{3-2}$$

$$= 3\, p^2 (1-p)$$

Tendo em conta o racio 3/1, só há 2 possibilidades para p:

(1) $p = 1/4 \Rightarrow L(1/4) = 3 \times \left(\frac{1}{4}\right)^2 \frac{3}{4} = \frac{9}{64}$

(2) $p = 3/4 \Rightarrow L(3/4) = 3 \times \left(\frac{3}{4}\right)^2 \frac{1}{4} = 27/64$

$\Rightarrow$ Máx em $p = 3/4$

$$\Rightarrow \boxed{\hat{p} = 3/4}$$

(c) A função de verosimilhança é igual à da alínea b), i.e.

$$L(p) = 3p^2(1-p),$$

mas agora queremos encontrar o ponto de máximo de L sem a restrição do racio 3/1.

$$\Rightarrow \frac{\partial L}{dp} = 6p(1-p) - 3p^2 = 3p(2-2p-p) = 0$$

$$\Rightarrow p = 0 \vee (2-3p=0 \Leftrightarrow p = 2/3).$$

$$\frac{\partial^2 L}{\partial p^2} = 3(2-3p) + 3p(-3) = 6 - 9p - 9p =$$
$$= 6 - 18p \Rightarrow \frac{\partial^2 L}{\partial p^2}(2/3) = 6 - \frac{18\times 2}{3} =$$
$$= -6 < 0$$

max

$$\Rightarrow \boxed{\hat{p} = 2/3}$$

- 1º Sem. 21/22, Exame 2 de 24/02/2022 às 10:30

(6) X = v.a. com $f_X(x) = \begin{cases} 2\beta x(1-x^2)^{\beta-1}, & x \in (0,1) \\ 0, & \text{c.c.} \end{cases}$

$\beta > 0$ desconhecido.

Determine a estimativa de MV de $P(X > 0.5)$, baseada na amostra $(x_1, \dots, x_6)$ de X t.q.

$$\sum_{i=1}^{6} \ln(1-x_i^2) \simeq -3.698.$$

- $P(X>0.5) = \int_{0.5}^{+\infty} f_X(x) = \int_{0.5}^{1} 2\beta x (1-x^2)^{\beta-1} dx$

$$= \left[-(1-x^2)^{\beta}\right]_{0.5}^{1} = 0.75^{\beta}$$

$$\Rightarrow \quad \widehat{P(X>0.5)} = 0.75^{\hat{\beta}}$$

- Estimativa de MV de β:

$$L(\beta) \equiv L(\beta | x_1, \dots, x_6) = \prod_{i=1}^{6} 2\beta x_i (1-x_i^2)^{\beta-1} =$$

$$= 2^6 \beta^6 \left(\prod_{i=1}^{6} x_i\right)\left(\prod_{i=1}^{6}(1-x_i)^2\right)^{\beta-1}, \quad \beta>0$$

$$(\log L)(\beta) = 6\log(2) + 6\log(\beta) + \sum_{i=1}^{6}\log(x_i) + (\beta-1)\sum_{i=1}^{6}\log(1-x_i^2)$$

$$\Rightarrow \frac{d\log L}{d\beta} = \frac{6}{\beta} + \sum_{i=1}^{6}\log(1-x_i^2) = 0 \Rightarrow \beta = \frac{-6}{\sum_{i=1}^{6}\log(1-x_i^2)}$$

$$\frac{d^2\log L}{d\beta} = -\frac{6}{\beta^2} < 0 \qquad \text{máx.}$$

$$\Rightarrow \hat{\beta} = \frac{-6}{-3.698} = 1.6225$$

- $\widehat{P(X>0.5)} = 0.75^{\hat{\beta}} = 0.75^{1.6225} = 0.6270$ //

- <u>Exercício 6.2</u>

$X \sim Exp(\lambda)$, $f_X(x) = \lambda e^{-\lambda x}, x>0$, $x_1 + \dots + x_{10} = 1740$

$\widehat{P(X>200)} = ?$

- $EMV(\lambda) = \frac{1}{\bar{X}} \Rightarrow \hat{\lambda} = \frac{1}{\frac{1740}{10}} = \frac{1}{174}$

$$\cdot\ P(X>200) = \int_{200}^{+\infty} \lambda e^{-\lambda x}\,dx = \left[-e^{-\lambda x}\right]_{200}^{+\infty} =$$

$$= e^{-200\lambda}$$

$$\Rightarrow \widehat{P(X>200)} = e^{-200\hat{\lambda}} = e^{-\frac{200}{174}}$$

$$= 0.3168$$

- <u>O estimador de MV</u> <u>pode não ser único — exemplo</u>

$X \sim Unif(\theta, \theta+1)$, θ = parâmetro desconhecido

$(X_1, \dots, X_n)$ = a.a. de X

$$f_X(x,\theta) = \begin{cases} 1, & \text{qd } \theta \le x \le \theta+1 \\ 0, & \text{c.c.} \end{cases}$$

$$L(\theta) = L(\theta \mid x_1, \dots, x_n) = \prod_{i=1}^{n} f_X(x_i, \theta) =$$

$$= \begin{cases} 1, & \text{qd } \theta \le x_i \le \theta+1,\ \forall i=1,\dots,n \\ 0, & \text{c.c} \end{cases}$$

$$= \begin{cases} 1, & \text{qd } \max\{x_1,\dots,x_n\}-1 \le \theta \le \min\{x_1,\dots,x_n\} \\ 0, & \text{c.c.} \end{cases}$$

$\Rightarrow \hat{\theta}$ = qq valor no intervalo

$[\max\{x_1,\dots,x_n\}-1, \min\{x_1,\dots,x_n\}]$