



Probabilidades e Estatística 2004/05

Colectânea de Exercícios

LEIC, LERCI, LEE

Capítulo 6 Estimação pontual

Nota: $EMV(a)$ designa o *estimador de máxima verosimilhança* de a e $emv(a)$ a respectiva estimativa. Além disso, L denota a função de verosimilhança.

Exercício 5.1 Seja $\{X_i, i = 1, \dots, 5\} \stackrel{\text{iid}}{\sim} X$, com valor $E[X] = \mu$ e $\text{Var}[X] = 2^2$.

(a) Como $E[X] = \mu$, $EM(\mu) = \bar{X}$. Uma vez que $L(\mu|x_1, \dots, x_5) = (\frac{1}{2\sqrt{2\pi^2}})^5 e^{-\sum_{i=1}^5 \frac{(x_i - \mu)^2}{8}}$, $emv(\mu) = \sum_{i=1}^5 x_i/5 = \bar{x}$, e logo $EMV(\mu) = \bar{X}$.

(b) $\bar{X} \sim N(\mu, \frac{4}{5})$ pelo que $P(\bar{X} > \mu + 1) = 1 - \Phi(\frac{\sqrt{5}}{2}) = 1 - 0.8682 = 0.1318$.

(c)

$$\begin{aligned} P(\min(X_1, \dots, X_5) < \mu - 2) &= 1 - P(\min(X_1, \dots, X_5) \geq \mu - 2) \\ &= 1 - P(X_1 \geq \mu - 2, \dots, X_5 \geq \mu - 2) \\ &= 1 - \prod_{i=1}^5 P(X_i \geq \mu - 2) \\ &= 1 - [P(X \geq \mu - 2)]^5 \\ &= 1 - [1 - \Phi(-1)]^5 = 0.5784. \end{aligned}$$

(d)

$$\begin{aligned} P(\max(X_1, \dots, X_5) > \mu + 3) &= 1 - P(\max(X_1, \dots, X_5) \leq \mu + 3) \\ &= 1 - P(X_1 \leq \mu + 3, \dots, X_5 \leq \mu + 3) \\ &= 1 - \prod_{i=1}^5 P(X_i \leq \mu + 3) \\ &= 1 - [P(X \leq \mu + 3)]^5 \\ &= 1 - \Phi^5(3/2) = 0.2923. \end{aligned}$$

Exercício 5.2

(a) Θ^2 é estimador centrado de θ^2 se e só se $E[\Theta^2] = \theta^2$. Como

$$E[\Theta^2] = \text{Var}[\Theta] + E^2[\Theta] = \text{Var}[\Theta] + \theta^2 > \theta^2$$

então Θ^2 não é estimador centrado de θ^2 .

(b) Usando o facto do valor esperado ser um operador linear,

$$\begin{aligned} E[(\Theta - \theta)^2] &= E[(\Theta - E[\Theta] + E[\Theta] - \theta)^2] \\ &= E[(\Theta - E[\Theta])^2] - 2(E[\Theta] - \theta)E[\Theta - E[\Theta]] + E[(E[\Theta] - \theta)^2] \\ &= \text{Var}(\Theta) - 0 + (E[\Theta] - \theta)^2 \\ &= \text{Var}(\Theta) + b^2. \end{aligned}$$

Exercício 5.3

(a) $E[w\bar{X}_1 + (1-w)\bar{X}_2] = wE[\bar{X}_1] + (1-w)E[\bar{X}_2] = wE[X_1] + (1-w)E[X_2] = w\mu + (1-w)\mu = \mu.$

(b) Como

$$\begin{aligned}\text{Var}(w\bar{X}_1 + (1-w)\bar{X}_2) &= w^2\text{Var}(\bar{X}_1) + (1-w)^2\text{Var}(\bar{X}_2) \\ &= \frac{w^2}{n}\text{Var}(X_1) + \frac{(1-w)^2}{n}\text{Var}(X_2) \\ &= \frac{w^2}{n}\sigma_1^2 + \frac{(1-w)^2}{n}\sigma_2^2\end{aligned}$$

a variância de $w\bar{X}_1 + (1-w)\bar{X}_2$ é mínima quando $w = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}.$

Exercício 5.4 Note-se que $E[\bar{X}] = E[X] = \mu$, $\text{Var}(\bar{X}) = \frac{\sigma^2}{3}$ e

$$\begin{aligned}E\left[\frac{X_1 + 2X_2 + X_3}{4}\right] &= \frac{E[X_1] + 2E[X_2] + E[X_3]}{4} = \mu \\ \text{Var}\left(\frac{X_1 + 2X_2 + X_3}{4}\right) &= \frac{\text{Var}(X_1) + 4\text{Var}(X_2) + \text{Var}(X_3)}{16} = \frac{6}{16}\sigma^2.\end{aligned}$$

Assim, a eficiência relativa de $\frac{X_1 + 2X_2 + X_3}{4}$ relativamente a $\bar{X}$ é $\frac{\text{Var}(\bar{X})}{\text{Var}(\frac{X_1 + 2X_2 + X_3}{4})} = \frac{16}{18}$, pelo que $\bar{X}$ é mais eficiente que $\frac{X_1 + 2X_2 + X_3}{4}$.

Exercício 5.5

(a) A mediana de uma distribuição normal de valor esperado μ é igual a μ . Logo

$$\begin{aligned}E[\tilde{X}] &= \mu & \text{Var}(\tilde{X}) &= [2\sqrt{n}f(\mu)]^{-2} = \frac{\pi\sigma^2}{2n} \\ E[\bar{X}] &= \mu & \text{Var}(\bar{X}) &= \frac{\sigma^2}{n}\end{aligned}$$

Logo

$$\text{ef}(\bar{X}, \tilde{X}) = \frac{\text{Var}(\tilde{X})}{\text{Var}(\bar{X})} = \frac{\pi}{2} > 1$$

pelo que $\bar{X}$ é um estimador de μ mais eficiente que $\tilde{X}$.

(b) Dado que $f_X(\cdot)$ é uma função simétrica em torno de μ , a mediana é igual a μ . Logo

$$\begin{aligned}E[\tilde{X}] &= \mu & \text{Var}(\tilde{X}) &= [2\sqrt{n}f(\mu)]^{-1} = \frac{\sigma^2}{2n} \\ E[\bar{X}] &= \mu & \text{Var}(\bar{X}) &= \frac{\sigma^2}{n}\end{aligned}$$

Logo

$$\text{ef}(\bar{X}, \tilde{X}) = \frac{\text{Var}(\tilde{X})}{\text{Var}(\bar{X})} = \frac{1}{2} < 1$$

pelo que $\bar{X}$ é um estimador de μ menos eficiente que $\tilde{X}$.

- (c) Os resultados anteriores sugerem que para populações normais a média amostral é mais eficiente que a mediana amostral como estimador do valor esperado da população, mas que há outras populações simétricas em que a mediana amostral é um estimador mais eficiente do valor esperado que a média amostral.

Exercício 5.6 Como

$$\text{ef}(T_1, T_2) = \frac{\text{Var}(T_2) + (E[T_2] - \theta)^2}{\text{Var}(T_1)} = \frac{3 + 4\theta^2}{9}$$

T_1 é mais eficiente que T_2 se $3 + 4\theta^2 > 9$, ou seja $\theta > \sqrt{3}/2$; assim a escolha entre os estimadores T_1 e T_2 deve basear-se na informação disponível sobre o intervalo de valores de θ .

Exercício 5.7

- (a) Seja X_i a v.a. que toma o valor 1 se a bola que sai na i -ésima extracção é preta e o valor 0 caso contrário. Tem-se $\{X_i, i = 1, 2, 3\} \stackrel{\text{iid}}{\sim} X \sim \text{Bernoulli}(p)$, pelo que $X = \sum_{i=1}^3 X_i \sim \text{Bin}(3, p)$. Logo

$$L(p; X = k) = \binom{3}{k} p^k (1-p)^{3-k}, \quad k = 0, 1, 2, 3.$$

Como $L(\frac{1}{4}; X = 1) > L(\frac{3}{4}; X = 1)$, então $\text{emv}(p) = \frac{1}{4}$.

- (b) Como $L(\frac{1}{4}; X = 2) < L(\frac{3}{4}; X = 2)$, então $\text{emv}(p) = \frac{3}{4}$.

- (c) $L(p; X = 2) = \binom{3}{2} p^2 (1-p)^1$, pelo que $\text{emv}(p) = \frac{2}{3}$.

Exercício 5.8 $P(\text{bola branca}) = \frac{r}{r+1} = p$. Como $\text{emv}(p) = \frac{k}{n}$, utilizando a propriedade da invariância dos estimadores de máxima verosimilhança, vem que $\text{emv}(r) = \frac{k}{N-k}$.

Exercício 5.9 Seja X a v.a. que contabiliza o número de andares vendidos num dia, $X \sim \text{Poisson}(\lambda)$, e $\{X_i, i = 1, \dots, n\} \stackrel{\text{iid}}{\sim} X$ uma a.a. com média amostral $\bar{X}$.

- (a) Como $E[X] = \mu$ e $L(\mu; x_1, \dots, x_n) = \frac{e^{-\lambda n} \lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}$, $EMV(\lambda) = \bar{X}$. Como $E[\bar{X}] = \frac{1}{n} \sum_{i=1}^n E[X_i] = \frac{n\mu}{n} = \mu$, o estimador de máxima verosimilhança é centrado.
- (b) Como $\text{Var}[X] = \lambda$, um estimador centrado de $\text{Var}[X]$ é $\bar{X}$.
- (c) $\text{emv}(\lambda) = \frac{8}{20}$.
- (d) Como $L(\lambda; X_1 = 0, \dots, X_{15} = 0, X_{16} \geq 1, X_{17} \geq 1) = e^{-15\lambda} (1 - e^{-\lambda})^2$, $\text{emv}(\lambda) = \ln(17/15) = 0.1252$.

Exercício 5.10 Dado que $X \sim N(\mu, 2^2)$, então a função de verosimilhança de μ com base numa amostra $x_1, x_2, \dots, x_n$ é dada por

$$L(\mu; x_1, x_2, \dots, x_n) = \left[\frac{1}{2\sqrt{2\pi}} \right]^n e^{-\frac{1}{8} \sum_{i=1}^n (x_i - \mu)^2}$$

Considerando o logaritmo da função de verosimilhança, $\ln L(\mu; x_1, x_2, \dots, x_n) = -n \ln(2\sqrt{2\pi}) - \frac{1}{8} \sum_{i=1}^n (x_i - \mu)^2$, e maximizando (faça-o!) em função em ordem a μ , vem que

$$\text{emv}(\mu) = \bar{x}, \quad EMV(\mu) = \bar{X}.$$

Exercício 5.11 Seja $X \sim \text{Normal}(\mu, \sigma^2)$ a v.a. que designa a voltagem que um cabo eléctrico pode suportar, $\{X_i, i = 1, \dots, 12\} \stackrel{\text{iid}}{\sim} X$ e $\bar{X}$ e S_n^2 a média e variância amostral não corrigida, respectivamente. Como

$$L(\mu, \sigma; x_1, \dots, x_{12}) = (\sigma\sqrt{2\pi})^{-n} e^{-\frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2}}$$

tem-se $EMV(\mu) = \bar{X}$ e $EMV(\sigma) = S_n = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2}$. Assim,

$emv(\mu)$	$\bar{x}$	55.833
$emv(\sigma^2)$	s_n^2	101.639
$emv(\sigma)$	s_n	10.082
$emv(P(X > 70))$	$1 - \Phi(\frac{70 - \bar{x}}{s_n})$	0.0793

sendo que na estimação de σ^2 e de $P(X > 70)$ foi utilizada a propriedade de invariância dos estimadores de máxima verosimilhança.

Exercício 5.12

(a) Tem-se

$$\begin{aligned} E[T_1] &= \frac{1}{n^2(n-1)} E \left[\sum_{i=1}^n (nX_i - n\bar{X})^2 \right] \\ &= \frac{1}{n^2(n-1)} n^2 E \left[\sum_{i=1}^n [(X_i - \mu) + (\mu - \bar{X})]^2 \right] \\ &= \frac{1}{(n-1)} E \left[\sum_{i=1}^n [(X_i - \mu)^2 + 2(\mu - \bar{X})(X_i - \mu) + (\mu - \bar{X})^2] \right] \\ &= \frac{1}{(n-1)} E \left[\sum_{i=1}^n (X_i - \mu)^2 - 2n(\bar{X} - \mu)^2 + n(\bar{X} - \mu)^2 \right] \\ &= \frac{1}{(n-1)} \left[\sum_{i=1}^n E[(X_i - \mu)^2] - nE[(\bar{X} - \mu)^2] \right] \\ &= \frac{1}{(n-1)} [nE[(X - \mu)^2] - \text{Var}(X)] \\ &= \text{Var}(X) = \sigma^2 \end{aligned}$$

e

$$E[T_2] = E \left[\frac{n-1}{n} T_1 \right] = \frac{n-1}{n} E[T_1] = \frac{n-1}{n} \sigma^2.$$

(b) Do exercício anterior, caso X possua distribuição normal, o estimador de máxima verosimilhança de σ^2 é S_n^2 , pelo que T_2 é o estimador de máxima verosimilhança de σ^2 .

Exercício 5.13 Seja $X \sim \text{Exponencial}(\mu^{-1})$ a v.a. que designa a duração (em horas) de uma pilha, $\{X_i, i = 1, \dots, 10\} \stackrel{\text{iid}}{\sim} X$ e $\bar{X}$ é a média amostral respectiva. Como

$$L(\mu; x_1, \dots, x_n) = \mu^{-n} e^{-\frac{\sum_{i=1}^n x_i}{\mu}} I_{[0, +\infty]}(\min x_i)$$

(onde $I_{[0, +\infty]}(\min x_i) = 1$ se e só se $\min x_i > 0$, valendo 0 caso contrário) tem-se que $EMV(\mu) = \bar{X}$, $emv(\mu) = \bar{x} = 174$ e $emv(P(X > 200)) = e^{-\frac{200}{\bar{x}}} = 0.3168$, pelo princípio da invariância dos estimadores de máxima verosimilhança.

Exercício 5.14

(a) Uma vez que

$$L(\alpha; x_1, \dots, x_n) = \frac{\prod_{i=1}^n x_i}{\alpha^{2n}} e^{-\frac{\sum_{i=1}^n x_i^2}{2\alpha^2}} I_{[0, +\infty[}(\min x_i)$$

$$\text{tem-se } EMV(\alpha) = \sqrt{\frac{\sum_{i=1}^n X_i^2}{2n}} \text{ e logo } emv(\alpha) = \sqrt{\frac{98.13}{24}} = 2.022.$$

(b) Usando o princípio da invariância dos estimadores de máxima verossimilhança, tem-se

$$emv(E[X]) = 2.022\sqrt{\frac{\pi}{2}} = 2.5342 \text{ e } emv(\text{Var}[X]) = (2 - \frac{\pi}{2})2.022^2 = 1.7548.$$