/ /

Lista 2.5

① Seja $X_1, \ldots, X_n$ iid com pdf

$$f(x|\theta) = \theta x^{\theta-1}, \quad 0 \leq x \leq 1, \quad 0 < \theta < \infty.$$

ⓐ Encontre o MLE de θ, e mostre que a variancia $\to 0$ quando $n \to \infty$.

ⓑ Encontre o estimador de metodo dos momentos para θ.

$$f(x|\theta) = \prod \theta x_i^{\theta-1} = \theta^n \left(\prod x_i\right)^{\theta-1} = L(\theta|x)$$

$$\Rightarrow \frac{\partial}{\partial \theta} \log L = \frac{\partial}{\partial \theta}\left[n \log \theta + (\theta-1) \log \prod x_i\right] = \frac{n}{\theta} + \sum \log x_i$$

$$\Rightarrow \frac{\partial}{\partial \theta} \log L = \frac{n}{\theta} + \sum \log x_i = 0 \Rightarrow \theta = -\frac{n}{\sum \ln x_i}$$

E a segunda derivada é $-\frac{n}{\theta^2} < 0$, então este é o MLE. Para calcular a variancia de $\hat{\theta}$, note que, $Y_i = -\log X_i \sim Exp(1/\theta)$, então $-\sum \log X_i \sim Gamma(n, 1/\theta)$. Seja $\hat{\theta} = n/T$, onde $T \sim Gamma(n, 1/\theta)$. Temos

$$E\frac{1}{T} = \frac{\theta^n}{\Gamma(n)} \int_0^\infty \frac{1}{t} t^{n-1} e^{-\theta t} \, dt = \frac{\theta^n}{\Gamma(n)} \frac{\Gamma(n-1)}{\theta^{n-1}} = \frac{\theta}{n-1}$$

$$E\frac{1}{T^2} = \frac{\theta^n}{\Gamma(n)} \int_0^\infty \frac{1}{t^2} t^{n-1} e^{-\theta t} \, dt = \frac{\theta^n}{\Gamma(n)} \frac{\Gamma(n-2)}{\theta^{n-2}} = \frac{\theta^2}{(n-1)(n-2)}$$

e daí $\quad E\hat{\theta} = \frac{n}{n-1}\theta \quad$ e $\quad Var\,\hat{\theta} = \frac{n^2}{(n-1)^2(n-2)}\theta^2 \to 0 \quad (n \to \infty)$

/ /

(b)

$$EX_i = \int_0^1 x\,\theta x^{\theta-1}\,dx$$

$$= \theta \int_0^1 x^{\theta}\,dx = \theta\left[\frac{x^{\theta+1}}{\theta+1}\right]_0^1 = \frac{\theta}{\theta+1}$$

Daí $\quad \frac{1}{n}\sum X_i = \frac{\hat{\theta}}{\hat{\theta}+1} \Rightarrow \hat{\theta} = \frac{\sum X_i}{n-\sum X_i}$

/ /

② Seja $X_1, \dots, X_n$ uma amostra aleatória com pmf

$$P_\theta(X=x) = \theta^x (1-\theta)^{1-x}, \quad x = 0 \text{ ou } 1, \quad 0 \leq \theta \leq 1/2$$

ⓐ Encontre o estimador de método de momentos e MLE de θ.

ⓑ Encontre o MSE de cada estimador.

ⓒ Qual você prefere? Justifique.

ⓐ MM:

$$EX = \theta = \frac{1}{n}\sum X_i = \overline{X} \Rightarrow \tilde{\theta} = \overline{X}.$$

MLE:
$$f(X|\theta) = \prod_{i=1}^{n} \theta^{X_i}(1-\theta)^{1-X_i} = L(\theta|X)$$

$$\Rightarrow \frac{\partial}{\partial\theta}\left(\ln L(\theta|X)\right) = \frac{\partial}{\partial\theta}\left(\sum X_i \ln\theta + \sum(1-X_i)\ln(1-\theta)\right)$$

$$= \frac{\partial}{\partial\theta}\left(\ln\theta \sum X_i + \ln(1-\theta)n - \ln(1-\theta)\sum X_i\right)$$

$$= \frac{\sum X_i}{\theta} - \frac{n}{1-\theta} + \frac{\sum X_i}{1-\theta} = 0$$

$$\Rightarrow (1-\theta)\sum X_i - n\theta + \theta\sum X_i = 0 = \sum X_i - n\theta = 0$$

$$\Rightarrow \hat{\theta} = \frac{\sum X_i}{n}$$

quando $\overline{X} \leq 1/2$, pois $L(\theta|X)$ está considerando o intervalo do θ, logo $\hat{\theta} = \min\{\overline{X}, 1/2\}$

Digitalizado com CamScanner

/ /

(b) O MSE de $\tilde{\theta} = Var(\tilde{\theta}) + bias(\tilde{\theta})^2 = \left(\frac{(1-\theta)\theta}{n}\right) + \theta^2$.
O MSE de $\hat{\theta}$:

$$MSE(\hat{\theta}) = E(\hat{\theta}-\theta)^2 = \sum_{y=0}^{n} (\hat{\theta}-\theta)^2 \binom{n}{y} \theta^y (1-\theta)^{n-y}$$
$$= \sum_{y=0}^{\lfloor n/2 \rfloor} \left(\frac{y}{n}-\theta\right)^2 \binom{n}{y} \theta^y (1-\theta)^{n-y}$$
$$+ \sum_{y=\lfloor n/2 \rfloor+1} \left(\frac{1}{2}-\theta\right)^2 \binom{n}{y} \theta^y (1-\theta)^{n-y},$$

Onde $Y = \sum X_i \sim binomial(n, \theta)$.

(c) Considerando o MSE, temos que

$$MSE(\tilde{\theta}) = E(\bar{X}-\theta)^2 = \sum_{y=0}^{n} \left(\frac{y}{n}-\theta\right)^2 \binom{n}{y} \theta^y (1-\theta)^{n-y}.$$

Portanto, $MSE[\tilde{\theta}] - MSE(\hat{\theta}) = \sum_{y=\lfloor n/2 \rfloor+1}^{n} \left[\left(\frac{y}{n}-\theta\right)^2 - \left(\frac{1}{2}-\theta\right)^2\right] \binom{n}{y} \theta^y (1-\theta)^{n-y}$

$$= \sum_{y=\lfloor n/2 \rfloor+1} \left(\frac{y}{n}+\frac{1}{2}-2\theta\right)\left(\frac{y}{n}-\frac{1}{2}\right) \binom{n}{y} \theta^y (1-\theta)^{n-y}$$

Se $y/n > 1/2$ e $\theta \leq 1/2$ então os termos são positivos.
Daí, $MSE(\hat{\theta}) < MSE(\tilde{\theta})\ \forall\ 0 < \theta < 1/2$.

/ /

(3) Suponha que variáveis aleatórias $Y_1, \dots, Y_n$ satisfazem

$$Y_i = \beta X_i + \varepsilon_i$$

onde $X_1, \dots, X_n$ são constantes fixas, e $\varepsilon_i \sim N(0, \sigma^2)$, σ^2 desconhecido.

(a) Encontre uma estatística bi-dimensional suficiente para (β, σ^2)

(b) Encontre o MLE de β, e mostre que é não viesado.

(c) Encontre a distribuição do MLE de β.

(a)
$$L(\theta|Y) = \prod \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(Y_i - \beta X_i)^2\right\}$$

$$= (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum (Y_i^2 - 2\beta X_i Y_i + \beta^2 X_i^2)\right\}$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left\{-\frac{\beta^2 \sum X_i^2}{2\sigma^2}\right\} \exp\left\{-\frac{1}{2\sigma^2}\sum Y_i^2 + \frac{\beta}{\sigma^2}\sum X_i Y_i\right\}$$

$T(X_1, \dots, X_n) = \left(\sum Y_i^2, \sum X_i Y_i\right)$ é suficiente.

(b)
$$\ln L(\beta, \sigma^2 | Y) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum Y_i^2 + \frac{\beta}{\sigma^2}\sum X_i Y_i - \frac{\beta^2}{2\sigma^2}\sum X_i^2$$

Fixando σ^2,

$$\frac{\partial}{\partial\beta}\ln L = \frac{1}{\sigma^2}\sum X_i Y_i - \frac{\beta}{\sigma^2}\sum X_i^2 = 0 \Rightarrow \hat\beta = \frac{\sum X_i Y_i}{\sum X_i^2}$$

Digitalizado com CamScanner

/ /

Também,

$$\frac{\partial^2 \ell}{\partial \beta^2} = \frac{1}{\sigma^2} \sum x_i^2 < 0$$

então é maximo. Como $\hat{\beta}$ não depende de σ^2, é o MLE. ~~+~~ E $\hat{\beta}$ é não viesado

$$E\hat{\beta} = \frac{\sum x_i E_\beta Y_i}{\sum x_i^2} = \frac{\sum x_i \cdot \beta x_i}{\sum x_i^2} = \beta.$$

c) $\hat{\beta} = \sum a_i Y_i$, onde $a_i = x_i / \sum_j x_j^2$ constante. Pelo Corolario 4.6.10 [CB] $\hat{\beta}$ é uma normal com media β, e

$$\mathrm{Var}\,\hat{\beta} = \sum a_i^2 \,\mathrm{Var}\, Y_i = \sum \left(\frac{x_i}{\sum_j x_j^2}\right)^2 \sigma^2 = \frac{\sum x_i^2}{(\sum_j x_j^2)^2}\sigma^2 = \frac{\sigma^2}{\sum_j x_j^2}$$

/ /

(4) Considere $Y_1,\ldots,Y_n$ como no exercício (3)

(a) Mostre que $\sum Y_i / \sum X_i$ é um estimador não viesado para β

(b) Calcule a variancia de $\sum Y_i / \sum X_i$ e compare com a variancia do MLE.

(a) Não é viesado pois

$$E_\beta\left[\frac{\sum Y_i}{\sum X_i}\right] = \frac{\sum X_i E_\beta \beta}{\sum X_i} = \beta$$

(b) E para a variancia

$$\mathrm{Var}\left(\frac{\sum Y_i}{\sum X_i}\right) = \frac{1}{(\sum X_i)^2} \sum \mathrm{Var}\, Y_i = \frac{\sum \sigma^2}{n^2 \bar{x}^2} = \frac{\sigma^2}{n \bar{X}^2}$$

Como $\sum x_i^2 - n\bar{x}^2 = \sum (x_i - \bar{x})^2 \geq 0$, $\sum x_i^2 \geq n\bar{x}^2$. Daí

$$\mathrm{Var}\,\hat{\beta} = \frac{\sigma^2}{\sum x_i^2} \leq \frac{\sigma^2}{n\bar{x}^2} = \mathrm{Var}\left(\frac{\sum Y_i}{\sum X_i}\right).$$

/ /

(5) Considere $Y_1, \dots, Y_n$ como no exercício (3)

(a) Mostre que $[\sum(Y_i/x_i)]/n$ é não viesado β.

(b) Calcule a variância do estimador e compare com os exercícios (3) e (4).

(a) $$\frac{1}{n} E_\beta \sum \frac{Y_i}{x_i} = \frac{1}{n}\sum \frac{EY_i}{x_i} = \frac{1}{n}\sum \frac{\beta x_i}{x_i} = \beta$$

(b) $$\frac{1}{n^2} Var \sum \frac{Y_i}{x_i} = \frac{1}{n}\sum \frac{Var Y_i}{x_i^2} = \frac{\sigma^2}{n^2}\sum \frac{1}{x_i^2}$$

Temos que, $\frac{1}{n}\sum\frac{1}{x_i^2} \geq \frac{n}{\sum x_i^2}$. Daí

$$Var\,\hat{\beta} = \frac{\sigma^2}{\sum x_i^2} \leq \frac{\sigma^2}{n^2}\sum\frac{1}{x_i^2} = Var\frac{1}{n}\sum\frac{Y_i}{x_i}$$

E usando a desigualdade de Jensen, temos que

$$\frac{1}{\bar{x}^2} \leq \frac{1}{n}\sum\frac{1}{x_i^2}$$

Daí,

$$Var\left(\frac{\sum Y_i}{\sum x_i}\right) = \frac{\sigma^2}{n\bar{x}^2} \leq \frac{\sigma^2}{n^2}\sum\frac{1}{x_i^2} = Var\frac{1}{n}\sum\frac{Y_i}{x_i}$$

/ /

(6) Seja $X_1, \ldots, X_n \overset{iid}{\sim} Bernoulli(p)$. Mostre que a variancia de $\bar{X}$ satisfaz o Limite Inferior de Cramér-Rao, e por isso $\bar{X}$ é o melhor estimador não viesado de p.

Assunto não coberto.

/ /

(7) Seja $X_1, \dots, X_n$ uma amostra aleatoria de uma população com media μ e variancia σ^2.

(a) Mostre que o estimador $\sum a_i X_i$ é não viesado para μ se $\sum a_i = 1$

(b) Sobre todos os estimadores da forma $\sum a_i X_i$ encontre o com variancia minima e calcule ela.

(a) $$E_\mu \sum a_i X_i = \sum E a_i X_i = \sum a_i \mu = \mu . 1 = \mu$$

(b) $$\mathrm{Var} \sum a_i X_i \overset{iid}{=} \sum \mathrm{Var}\, a_i X_i = \sum a_i^2 \mathrm{Var}\, X_i = \sigma^2 \sum a_i^2$$

Então, queremos minimizar $\sum a_i^2$ com sujeito a $\sum a_i = 1$. Somando e subtraindo a media, $1/n$, temos

$$\sum a_i^2 = \sum_i \left[\left(a_i - \frac{1}{n}\right) + \frac{1}{n}\right]^2 = \sum \left(a_i - \frac{1}{n}\right)^2 + \frac{1}{n}.$$

Aqui, $\sum a_i^2$ é minimizado com $a_i = 1/n, \forall i$. Então $\frac{\sum X_i}{n} = \overline{X}$ tem variancia minima, com variancia de $\frac{\sigma^2}{n}$.

/ /

(8) Prove o que foi dito no exemplo 7.3.24 [CB]: Seja T é uma estatística suficiente completa para o parametro θ, e $h(X_1, \ldots, X_n)$ é algum estimador não viesado de $\tau(\theta)$, então $\phi(T) = E[h(X_1, \ldots X_n) | T]$ é o melhor estimador não viesado de $\tau(\theta)$.

(Teorema de Rao-Blackwell)

$$\tau(\theta) = E_\theta h = E_\theta [h | T] = E_\theta \phi(T)$$

logo $\phi(T)$ é não viesado para $\tau(\theta)$. Também

$$\begin{aligned} Var_\theta h &= Var_\theta [E(h|T)] + E_\theta [Var(h|T)] \\ &= Var_\theta \phi(T) + E_\theta [Var(h|T)] \\ &\geq Var_\theta (T) \end{aligned}$$

Logo, $\phi(T)$ é o melhor estimador não viesado de $\tau(\theta)$.

/ /

9) Seja $X_1, \ldots, X_{n+1} \overset{iid}{\sim}$ Bernoulli(p), e defina a função $h(p)$ por

$$h(p) = P\left(\sum_{i=1}^{n} X_i > X_{n+1} \mid p\right)$$

a probabilidade as das primeiras n observações excederem $(n+1)^a$ observação.

a) Mostre que

$$T(X_1, \ldots, X_n) = \begin{cases} 1 & \text{se } \sum_{i=1}^{n} X_i > X_{n+1} \\ 0 & \text{c.c.} \end{cases}$$

é um estimador não viesado de $h(p)$.

b) Encontre o melhor estimador não viesado de $h(p)$

a) T é uma v.a bernoulli. Daí

$$E_p T = P_p(T=1) = P_p\left(\sum_{i=1}^{n} X_i > X_{n+1} \mid p\right) = h(p).$$

b) $\sum_{i=1}^{n+1} X_i$ é uma estatística suficiente completa para θ, então $E[T \mid \sum X_i]$ é o melhor estimador não viesado de $h(\theta)$

Temos

$$E\left[T \mid \sum X_i = y\right] = P\left(\sum_{i=1}^{n} X_i > X_{n+1} \mid \sum X_i = y\right)$$

$$= P\left(\sum_{i=1}^{n} X_i > X_{n+1}, \sum_{i=1}^{n+1} X_i = y\right) \Big/ P\left(\sum_{i=1}^{n+1} X_i = y\right)$$

O denominador é igual a $\binom{n+1}{y} p^y (1-p)^{n+1-y}$. Se $y = 0$ o numerador é

$$P\left(\sum_{i=1}^{n} X_i > X_{n+1}, \sum_{i=1}^{n+1} X_i = 0\right) = 0$$

Digitalizado com CamScanner

/ /

Se $y>0$ o numerador é

$$P\left(\sum_{i=1}^{n} X_i \geq 0, \sum_{i=1}^{n} X_i = y\right) P(X_{n+1}=0) + P\left(\sum_{i=1}^{n} X_i \geq 1, \sum_{i=1}^{n} X_i = y-1\right) P(X_{n+1}=1)$$

Para todo $y>0$,

$$P\left(\sum_{i=1}^{n} X_i \geq 0, \sum_{i=1}^{n} X_i = y\right) = P\left(\sum_{i=1}^{n} X_i = y\right) = \binom{n}{y} p^{y}(1-p)^{n-y}.$$

Se $y = 1$ ou 2, então

$$P\left(\sum_{i=1}^{n} X_i \geq 1, \sum_{i=1}^{n} X_i = y-1\right) = 0$$

E se $y>1$, então

$$P\left(\sum_{i=1}^{n} X_i \geq 1, \sum_{i=1}^{n} X_i = y-1\right) = P\left(\sum X_i = y-1\right) = \binom{n}{y-1} p^{y-1}(1-p)^{n-y+1}.$$

Portanto, o UMVUE é

$$E\left(T \mid \sum_{i=1}^{n} X_i = y\right) = \begin{cases} 0, & y=0 \\ \dfrac{1}{(n+1)(n+1-y)}, & y = 1 \text{ ou } 2 \\ 1, & y>2 \end{cases}$$