Practica 11(2021)

Preámbulo:

Para una distribución binomial:

$$\begin{aligned} \text{Parameters:} & n \in \{0,1,2,\ldots\}, \text{number of trials} \\ \text{Support:} & k \in \{0,1,\ldots,n\}, \text{number of successes} \\ \text{PMF:} & \binom{n}{k} p^k q^{n-k} \end{aligned}$$

$$\begin{aligned} \mathrm{E}(X) &= \sum_{k=0}^{n} k \binom{n}{k} p^k q^{n-k} \\ &= \sum_{k=1}^{n} k \binom{n}{k} p^k q^{n-k} \left(k = 0, k \binom{n}{k} p^k q^{n-k} = 0 \right) \\ &= \sum_{k=1}^{n} n \binom{n-1}{k-1} p^k q^{n-k} \left(k \binom{n}{k} = n \binom{n-1}{k-1} \right) \\ &= np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} q^{(n-1)-(k-1)} \\ & (\text{sacar } np \text{ y } (n-1)-(k-1) = n-k) \\ &= np \sum_{j=0}^{m} \binom{m}{j} p^j q^{m-j} \\ & (\text{putting } m = n-1, j = k-1) \\ &= np \,(\text{ Binomial Theorem and } p+q=1) \end{aligned}$$

Problema 1

Enunciado: Resolución:

a)

$$\begin{aligned} V &= \frac{\partial}{\partial \theta} \ln \left(\frac{1}{\sqrt{2\pi\theta}} e^{\frac{-x^2}{2\theta}} \right) \\ &= \frac{\partial}{\partial \theta} \left\{ \frac{-x^2}{2\theta} \ln(e) - \ln(\sqrt{2\pi\theta}) \right\} \\ &= \frac{\partial}{\partial \theta} \left\{ \frac{-x^2}{2\theta} - \frac{1}{2} \ln \theta + \frac{1}{2} \ln(2\pi) \right\} \\ &= \frac{-x^2}{2} \cdot \frac{\partial}{\partial \theta} \frac{1}{\theta} - \frac{1}{2} \frac{\partial}{\partial \theta} \ln \theta \\ &= \frac{1}{\theta^2} \left(x^2 - \theta \right) \end{aligned}$$

$$\langle V \rangle = 0$$

Inciso b

$$\begin{aligned} V = \frac{\partial}{\partial \theta} \ln \left(\frac{x^\theta}{1+2^\theta} \right) &= \partial_\theta \left(\ln x^\theta - \ln(1+2\theta) \right) \\ &= \ln x - \partial_\theta \ln \left(1+2^\theta\right) = \ln x - \frac{1}{1+2^\theta} \cdot \partial_\theta (1+2^\theta) \\ &= \ln x - \frac{1}{1+2^\theta} \partial_\theta e^{\theta \ln 2} = \ln x - \frac{1}{1+2^\theta} \cdot \ln 2 \; e^{\theta \ln 2} \\ &= \ln x - \frac{\ln 2}{1+2^\theta} \; 2^\theta \end{aligned}$$

$$\langle V \rangle = \left(\frac{1+2^\theta}{1+2^\theta} \right) = 1$$

Problema 2

Enunciado:

Resolución: Demostrar:

$$\mathcal{I}(\theta) = -\mathrm{E} \left[\frac{\partial^2}{\partial \theta^2} \log f(X;\theta) \mid \theta \right]$$

se tiene que

$$\begin{aligned} \frac{\partial^2}{\partial \theta^2} \log f(X;\theta) &= \frac{\frac{\partial^2}{\partial \theta^2} f(X;\theta)}{f(X;\theta)} - \left(\frac{\frac{\partial}{\partial \theta} f(X;\theta)}{f(X;\theta)} \right)^2 \\ &= \frac{\frac{\partial^2}{\partial \theta^2} f(X;\theta)}{f(X;\theta)} - \left(\frac{\partial}{\partial \theta} \log f(X;\theta) \right)^2 \end{aligned}$$

y que

$$\mathrm{E} \left[\frac{\frac{\partial^2}{\partial \theta^2} f(X;\theta)}{f(X;\theta)} \mid \theta \right] = \frac{\partial^2}{\partial \theta^2} \int f(x;\theta) dx = \frac{\partial^2}{\partial \theta^2} \; 1 = 0$$

por lo que

$$\begin{aligned} \mathrm{E} \left[\frac{\partial^2}{\partial \theta^2} \log f(X;\theta) \right] &= \mathrm{E} \left[\frac{\frac{\partial^2}{\partial \theta^2} f(X;\theta)}{f(X;\theta)} \right] - \mathrm{E} \left[\left(\frac{\partial}{\partial \theta} \log f(X;\theta) \right)^2 \right] \\ &= -\mathrm{E} \left[\left(\frac{\partial}{\partial \theta} \log f(X;\theta) \right)^2 \right] = -J(\theta) \end{aligned}$$

Problema 3

Enunciado:

Resolución: Sabiendo que

$$P\left(x_1, \ldots, x_n \mid \theta\right)=\prod_{i=0}^n P\left(x_i \mid \theta\right)$$

$$\log \prod_{i=0}^n P\left(x_i \mid \theta\right)=\sum_{i=0}^n \log P\left(x_i \mid \theta\right)$$

Se tiene que

$$\begin{aligned}
J(\theta) & =-E\left[\frac{\partial^2}{\partial \theta^2} \log P\left(x_1, \ldots, x_n \mid \theta\right)\right] \\
& =-E\left[\sum_{i=0}^n \frac{\partial^2}{\partial \theta^2} \log P\left(x_i \mid \theta\right)\right] \\
& =\sum_{i=0}^n-E\left[\frac{\partial^2}{\partial \theta^2} \log P\left(x_i \mid \theta\right)\right] \\
& =\sum_{i=0}^n J_i(\theta) \\
& =n J(\theta)
\end{aligned}$$

Problema 4

Enunciado:

Resolución: Primer caso

$$\partial_\theta^2 \log(\theta e^{-\theta x})=\partial_\theta\left(\frac{1}{\theta}-x\right)=\frac{-1}{\theta^2}$$

$$\begin{aligned}
J_1(\theta) & =-E[\partial_\theta^2 \log(\theta e^{-\theta x})]=-E[\frac{-1}{\theta^2}] \\
& =\frac{1}{\theta^2} E[1]=\frac{1}{\theta^2}
\end{aligned}$$

Segundo caso:

Usando que:

$$\partial_\theta^2 \log(\theta^2 e^{-\theta^2 x})=-\frac{2}{\theta^2}-2x$$

$$E[x]=\int_{\mathbb{R}^+} x e^{-\theta x} dx=\frac{1}{\theta^2}$$

Se tiene que:

$$\begin{aligned}
J_2(\theta) & =-E[\partial_\theta^2(\theta^2 e^{-\theta^2 x})] \\
& =-E[-\frac{2}{\theta^2}-2x] \\
& =\frac{2}{\theta^2}+2E[x]=\frac{4}{\theta^2}
\end{aligned}$$

Problema 5

Enunciado:

Resolución: Si hay dos formas de parametrizar:

$$X \sim f(x \mid \theta)=g(x \mid \eta), \text{ con } \theta=\phi(\eta), \quad \eta=\psi(\theta)$$

se pueden relacionar las informaciones de Fisher.

$$\begin{aligned}
J^\theta(\theta) & =\mathrm{E}\left\{\left(\frac{\partial}{\partial \theta} \log f(x \mid \theta)\right)^2\right\} \\
& =\mathrm{E}\left\{\left(\frac{\partial}{\partial \theta} \log g(x \mid \psi(\theta))\right)^2\right\} \\
& =\mathrm{E}\left\{\left(\frac{\partial}{\partial \eta} \log g(x \mid \eta) \frac{\partial \eta}{\partial \theta}\right)^2\right\} \\
& =\mathrm{E}\left\{\left(\frac{\partial}{\partial \eta} \log g(x \mid \eta)\right)^2\right\}\left(\frac{\partial \eta}{\partial \theta}\right)^2 \\
& =J^\eta(\eta)\left(J_\theta^\eta\right)^2 \\
& =J^\eta(\eta)\left(\frac{\partial \eta}{\partial \theta}\right)^2 (1\mathrm{D})
\end{aligned}$$

Se puede del otro lado:

$$I^\eta(\eta)=I^\theta(\theta)\left(\frac{\partial \theta}{\partial \eta}\right)^2 (1\mathrm{D})$$

Problema 6

Enunciado:

Resolución: Statement

La cota Cramér-Rao va a ser incrementalmente generalizada. Todas las versiones de la cota requieren ciertas condiciones de regularidad que se mencionan al final.

Caso 1

Se empieza por que el parámetro es escalar y tiene un estimador sin bias, se supone que θ es un desconocido parámetro determinista que se estima desde n observaciones independientes de x, cada uno desde una distribución de acuerdo a una densidad de probabilidad $f(x;\theta)$.

$$\mathrm{var}(\hat{\theta}) \geq \frac{1}{J(\theta)}$$

donde la $J(\theta)$ viene dada por

$$J(\theta) = n\mathrm{E}_\theta \left[\left(\frac{\partial \ell(X;\theta)}{\partial \theta} \right)^2 \right]$$

y $\ell(x;\theta) = \log(f(x;\theta))$. Bajo ciertas condiciones es posible:

$$J(\theta) = -n\mathrm{E}_\theta \left[\frac{\partial^2 \ell(X;\theta)}{\partial \theta^2} \right]$$

Definimos la eficiencia como:

$$e(\hat{\theta}) = \frac{J(\theta)^{-1}}{\mathrm{var}(\hat{\theta})}$$

y la cota se puede reescribir como:

$$e(\hat{\theta}) \leq 1$$

Caso 2

Considerando que hay bias, y el estimador con bias $T(X)$, no tiene valor de expectación θ sino que es una $\psi(\theta)$. Se tiene $E\{T(X)\} - \theta = \psi(\theta) - \theta$ que puede ser no nula. Esto implica que

$$\mathrm{var}(T) \geq \frac{[\psi'(\theta)]^2}{J(\theta)}$$

donde $\psi'(\theta)$ es la derivada de $\psi(\theta)$ (respecto a θ). Bound on the variance of biased estimators Apart from being a bound on estimators of functions of the parameter, this approach can be used to derive a bound on the variance of biased estimators with a given bias, as follows. Consider an estimator $\hat{\theta}$ with bias $b(\theta) = E\{\hat{\theta}\} - \theta$, and let $\psi(\theta) = b(\theta) + \theta$. By the result above, any unbiased estimator whose expectation is $\psi(\theta)$ has variance greater than or equal to $(\psi'(\theta))^2 / I(\theta)$. Thus, any estimator $\hat{\theta}$ whose bias is given by a function $b(\theta)$ satisfies

$$\mathrm{var}(\hat{\theta}) \geq \frac{[1 + b'(\theta)]^2}{I(\theta)}$$

The unbiased version of the bound is a special case of this result, with $b(\theta) = 0$.

It's trivial to have a small variance - an "estimator" that is constant has a variance of zero. But from the above equation we find that the mean squared error of a biased estimator is bounded by

$$\mathrm{E}\left((\hat{\theta} - \theta)^2 \right) \geq \frac{[1 + b'(\theta)]^2}{I(\theta)} + b(\theta)^2$$

using the standard decomposition of the MSE. Note, however, that if $1 + b'(\theta) < 1$ this bound might be less than the unbiased Cramér-Rao bound $1/I(\theta)$. For instance, in the example of estimating variance below, $1 + b'(\theta) = \frac{n}{n+2} < 1$. Multivariate case Extending the Cramér-Rao bound to multiple parameters, define a parameter column vector

$$\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_d]^T \in \mathbb{R}^d$$

with probability density function $f(x;\boldsymbol{\theta})$ which satisfies the two regularity conditions below. The Fisher information matrix is a $d \times d$ matrix with element $I_{m,k}$ defined as

$$I_{m,k} = \mathrm{E}\left[\frac{\partial}{\partial \theta_m} \log f(x;\boldsymbol{\theta}) \frac{\partial}{\partial \theta_k} \log f(x;\boldsymbol{\theta}) \right] = -\mathrm{E}\left[\frac{\partial^2}{\partial \theta_m \partial \theta_k} \log f(x;\boldsymbol{\theta}) \right]$$

Let $\boldsymbol{T}(X)$ be an estimator of any vector function of parameters, $\boldsymbol{T}(X) = (T_1(X), \ldots, T_d(X))^T$, and denote its expectation vector $\mathrm{E}[\boldsymbol{T}(X)]$ by $\boldsymbol{\psi}(\boldsymbol{\theta})$. The Cramér-Rao bound then states that the covariance matrix of $\boldsymbol{T}(X)$ satisfies

$$\mathrm{cov}_\theta(\boldsymbol{T}(X)) \geq \frac{\partial \boldsymbol{\psi}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} [I(\boldsymbol{\theta})]^{-1} \left(\frac{\partial \boldsymbol{\psi}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^T$$

Problema 8

Enunciado: Demostrar:

$$J_{ik}(\vec{\theta}) = E\left[\frac{\partial}{\partial\theta_i}\log p(x \mid \vec{\theta})\frac{\partial}{\partial\theta_k}\log p(x \mid \vec{\theta})\right] = -E\left[\frac{\partial^2}{\partial\theta_i\partial\theta_k}\log p(x \mid \vec{\theta})\right]$$

Resolución: se tiene que

$$\frac{\partial^2}{\partial\theta_i\partial\theta_k}\log p(x \mid \vec{\theta}) = \frac{\partial_{\theta_i}\partial_{\theta_k}p(x \mid \vec{\theta})}{p(x \mid \vec{\theta})} - \left(\frac{\partial_{\theta_i}p(x \mid \vec{\theta})}{p(x \mid \vec{\theta})}\right)\left(\frac{\partial_{\theta_k}p(x \mid \vec{\theta})}{p(x \mid \vec{\theta})}\right)$$

y que

$$\mathrm{E}\left[\frac{\partial_{\theta_i}\partial_{\theta_k}p(x \mid \vec{\theta})}{p(x \mid \vec{\theta})}\right] = \partial_{\theta_i}\partial_{\theta_k}\int_{\mathbb{R}} p(x \mid \vec{\theta})dx = \partial_{\theta_i}\partial_{\theta_k}1 = 0$$

por lo que

$$\begin{aligned}\mathrm{E}\left[\frac{\partial^2}{\partial\theta_i\partial\theta_k}\log p(x \mid \vec{\theta})\right] &= \mathrm{E}\left[\frac{\partial_{\theta_i}\partial_{\theta_k}p(x \mid \vec{\theta})}{p(x \mid \vec{\theta})}\right] - \mathrm{E}\left[\left(\frac{\partial_{\theta_i}p(x \mid \vec{\theta})}{p(x \mid \vec{\theta})}\right)\left(\frac{\partial_{\theta_k}p(x \mid \vec{\theta})}{p(x \mid \vec{\theta})}\right)\right]\\ &= -\mathrm{E}\left[\left(\frac{\partial_{\theta_i}p(x \mid \vec{\theta})}{p(x \mid \vec{\theta})}\right)\left(\frac{\partial_{\theta_k}p(x \mid \vec{\theta})}{p(x \mid \vec{\theta})}\right)\right]\\ &= -\mathrm{E}\left[\left(\partial_{\theta_i}\log p(x \mid \vec{\theta})\right)\left(\partial_{\theta_k}\log p(x \mid \vec{\theta})\right)\right]\\ &= -J_{ik}(\theta)\end{aligned}$$

Problema 9

Enunciado:

Resolución:

$$\vec{\vartheta}(t) = \begin{pmatrix} q_1(t) \\ q_2(t) \end{pmatrix} = \begin{pmatrix} \frac{t}{2} \\ \frac{t}{2} \end{pmatrix}, \quad \dot{\vec{\vartheta}}(t) = \begin{pmatrix} \dot{q}_1(t) \\ \dot{q}_2(t) \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

$$Dist(\vec{\vartheta}^1, \vec{\vartheta}^2) = \int_0^1 \sqrt{\dot{\vec{\vartheta}}^T J(\vec{\vartheta})\dot{\vec{\vartheta}}}dt$$

$$J_{ik}(\vartheta) = -E\left[\partial_{q_1}\log P(x \mid q_1, q_2)\partial_{q_2}\log P(x \mid q_1, q_2)\right]$$

$$\partial_{q_1}\log P(x \mid q_1, q_2) = \frac{\partial_{q_1}P(x \mid q_1, q_2)}{P(x \mid q_1, q_2)} = \begin{cases} \dfrac{-1}{1-q_1-q_2}, x = 0 \\ \dfrac{1}{q_1}, x = 1 \\ 0, x = 2 \end{cases}$$

$$\partial_{q_2}\log P(x \mid q_1, q_2) = \frac{\partial_{q_2}P(x \mid q_1, q_2)}{P(x \mid q_1, q_2)} = \begin{cases} \dfrac{-1}{1-q_1-q_2}, x = 0 \\ 0, x = 1 \\ \dfrac{1}{q_2}, x = 2 \end{cases}$$

$$\begin{aligned}J_{11}(\vartheta) &= \frac{1}{1-q_1-q_2} + \frac{1}{q_1}\\ J_{22}(\vartheta) &= \frac{1}{1-q_1-q_2} + \frac{1}{q_2}\\ J_{12}(\vartheta) &= J_{21}(\vartheta) = \frac{1}{1-q_1-q_2}\end{aligned}$$

$$J\begin{pmatrix} q_1(t) \\ q_2(t) \end{pmatrix} = \begin{bmatrix} \frac{1}{1-q_1-q_2} + \frac{1}{q_1} & \frac{1}{1-q_1-q_2} \\ \frac{1}{1-q_1-q_2} & \frac{1}{1-q_1-q_2} + \frac{1}{q_2} \end{bmatrix}$$

$$dist = \frac{1}{4}\sqrt{\frac{4}{1-q_1-q_2} + \frac{1}{q_1} + \frac{1}{q_2}}$$