Ejercicios Tema 2 - Estimación SOLUCIONES. Taller 3

Ricardo Alberich, Juan Gabriel Gomila y Arnau Mir

Curso completo de estadística inferencial con R y Python

Contenidos

1 Estimación taller 3

1.1 Ejercicio 1

Supongamos que $X_1, X_2, \ldots, X_6$ es una muestra aleatoria de una variable aleatoria normal con media μ y varianza σ^2. Hallar la constante C tal que

$$C \cdot \left((X_1 - X_2)^2 + (X_3 - X_4)^2 + (X_5 - X_6)^2\right),$$

sea un estimador sin sesgo de σ^2.

1.2 Ejercicio 2

Supongamos que Θ_1 y Θ_2 son estimadores sin sesgo de un parámetro desconocido θ, con varianzas conocidas σ_1^2 y σ_2^2, respectivamente. Demostrar que $\Theta = (1-a)\cdot\Theta_1 + a\cdot\Theta_2$ también es insesgado para cualquier valor de $a \neq 0$.

1.3 Ejercicio 3

Sea $X_1, \ldots, X_{2n}$ una muestra aleatoria simple de una variable aleatoria $N(\mu, \sigma)$. Sea:

$$T = C\left(\left(\sum_{i=1}^{2n} X_i\right)^2 - 4n\sum_{i=1}^{n} X_{2i}X_{2i-1}\right)$$

un estimador del parámetro σ^2. ¿Cuál es el valor de C para que T sea un estimador insesgado?

1.4 Ejercicio 4

Una variable aleatoria X sigue la distribución de Rayleigh con parámetro $\theta > 0$ si es una variable aleatoria con valores $x > 0$ y función de densidad:

$$f(x) = \frac{x}{\theta} e^{-\frac{x^2}{2\theta}}.$$

Hallar el estimador máximo verosímil del parámetro θ.

2 Soluciones

2.1 Solución ejercicio 1

Recordemos que $E(X_i) = \mu$, y que $\sigma^2 = E(X_i^2) - (E(X_i))^2 = E(X_i^2) - \mu^2$ luego $E(X_i^2) = \sigma^2 + \mu^2$ para $i = 1, 2, 3$. Además al ser independientes
$E(X_1 \cdot X_2) = E(X_1) \cdot E(X_2)\mu \cdot \mu = \mu^2$ y lo mismo para $E(X_3 \cdot X_4) = E(X_5 \cdot X_6) = \mu^2$.

Ahora

$$\begin{aligned} E\left((X_1 - X_2)^2\right) &= E\left(X_1^2 - 2 \cdot X_1 \cdot X_2 + X_2^2\right) \\ &= E\left(X_1^2\right) - 2 \cdot E(X_1 \cdot X_2) + E\left(X_2^2\right) \\ &= (\sigma^2 + \mu^2) - 2 \cdot (\mu \cdot \mu) + (\sigma^2 + \mu^2) \\ &= 2 \cdot \sigma^2. \end{aligned}$$

Utilizando adecuadamente los cálculos anteriores

$$E\left(C \cdot \left((X_1 - X_2)^2 + (X_3 - X_4)^2 + (X_5 - X_6)^2\right)\right) = C \cdot 3 \cdot 2 \cdot \sigma^2.$$

Luego el valor de C para que el estimador sea insesgado es la solución de la ecuación

$$C \cdot 3 \cdot 2 \cdot \sigma^2 = \sigma^2$$

Así que el valor buscado es $C = \frac{1}{6}$.

2.2 Solución ejercicio 2

Supongamos que Θ_1 y Θ_2 son estimadores sin sesgo de un parámetro desconocido θ, con varianzas conocidas σ_1^2 y σ_2^2, respectivamente. Demostrar que $\Theta = (1 - a) \cdot \Theta_1 + a \cdot \Theta_2$ también es insesgado para cualquier valor de a.

Tenemos que $E(\Theta_1) = E(\Theta_2) = \theta$ y que $Var(\Theta_1) = \sigma_1$ y $Var(\Theta_2) = \sigma_2$.

Nos piden que demostremos que para cualquier $a \neq 0$

$$E(\Theta) = E\left((1 - a) \cdot \Theta_1 + a \cdot \Theta_2\right) = \theta.$$

efectivamente

$E(\Theta) = (1 - a) \cdot E(\Theta_1) + a \cdot E(\Theta_2) = (1 - a) \cdot \theta + a \cdot \theta = \theta.$

2.3 Solución ejercicio 3

Como $X_1, \ldots, X_{2n}$ una muestra aleatoria simple de una variable aleatoria $N(\mu, \sigma)$. Sabemos que $\sum_{i=1}^{2\cdot n} X_i$ sigue una ley $N\left(2 \cdot n \cdot \mu, \sqrt{2 \cdot n \cdot \sigma^2}\right)$ luego $E\left(\sum_{i=1}^{2\cdot n} X_i\right) = 2 \cdot n \cdot \mu$, $Var(\sum_{i=1}^{2\cdot n} X_i) = 2 \cdot n \cdot \sigma^2$ y $E\left(\left(\sum_{i=1}^{2\cdot n} X_i\right)^2\right) = 2 \cdot n \cdot \sigma^2 + 2 \cdot n \cdot \mu^2$. Ademas como $X_{2\cdot i}$ y $X_{2\cdot i-1}$ son independientes $E(X_{2\cdot i} \cdot X_{2\cdot i-1}) = E(X_{2\cdot i}) \cdot E(X_{2\cdot i-1}) = \mu \cdot \mu = \mu^2$.

Entonces

$$\begin{aligned} E(T) &= E\left(C \cdot \left(\left(\sum_{i=1}^{2n} X_i\right)^2 - 4n\sum_{i=1}^{n} X_{2i}X_{2i-1}\right)\right) \\ &= C \cdot \left(2 \cdot n \cdot \sigma^2 + (2 \cdot n \cdot \mu)^2 - 4 \cdot n \cdot n \cdot \mu^2\right) \\ &= C \cdot 2 \cdot n \cdot \sigma^2 \end{aligned}$$

Por lo que el valor de C pedido es la solución de la ecuación

$$C \cdot 2 \cdot n \cdot \sigma^2 = \sigma^2$$

despejando C obtenemos que el valor buscado es $C = \frac{1}{2\cdot n}$.

2.4 Solución ejercicio 4

Sea $x_1, x_2, \ldots, x_n$ la realización de una más de una variable aleatoria X con distribución Rayleigh de parámetro $\theta > 0$. Su función de verosimilitud es

$$\begin{aligned} L(\theta|x_1, , x_2, \ldots, x_n) &= f_X(x_1;\theta) \cdot f_X(x_1;\theta) \cdot \ldots \cdot f_x(x_n;\theta) \\ &= \tfrac{x_1}{\theta} \mathrm{e}^{-\frac{x_1^2}{2\theta}} \cdot \tfrac{x_2}{\theta} \mathrm{e}^{-\frac{x_2^2}{2\theta}} \cdot \ldots \cdot \tfrac{x_n}{\theta} \mathrm{e}^{-\frac{x_n^2}{2\theta}} \\ &= \tfrac{\prod_{i=1}^n x_i}{\theta^n} \cdot \mathrm{e}^{-\frac{\sum_{i=1}^n x_i^2}{2\theta}}. \end{aligned}$$

Buscamos el θ que maximiza $L(\theta|x_1, , x_2, \ldots, x_n)$ que es el mismo que maximiza su logaritmo, es decir

$$\ln(L(\theta|x_1, , x_2, \ldots, x_n)) = \ln\left(\frac{\prod_{i=1}^n x_i}{\theta^n} \cdot \mathrm{e}^{-\frac{\sum_{i=1}^n x_i^2}{2\theta}}\right) = \ln\left(\prod_{i=1}^n x_i\right) - \ln(\theta^n) - \frac{\sum_{i=1}^n x_i^2}{2\theta}.$$

Ahora derivamos la expresión respecto de θ, la igualamos a cero y despejamos θ

$$\ln(L(\theta|x_1, , x_2, \ldots, x_n))' = \left(\ln\left(\prod_{i=1}^n x_i\right) - \ln(\theta^n) - \frac{\sum_{i=1}^n x_i^2}{2\theta}\right)' = -\frac{n}{\theta} + \frac{\sum_{i=1}^n x_i^2}{2\theta^2} = 0.$$

de donde multiplicando los dos términos de la última igualdad por θ^2.

$-n \cdot \theta + \frac{\sum_{i=1}^n x_i^2}{2} = 0$

resolviendo la ecuación obtenemos que el estimador máximo verosímil de θ es

$$\hat{\theta} = \frac{\sum_{i=1}^n x_i^2}{2 \cdot n}.$$