Estimadores de máxima verosimilitud
Estadística. Grupo m3

Ejercicio 1. Sea $(X_1, \ldots, X_n)$ una muestra aleatoria simple de una variable aleatoria $X \sim Bernoulli(p)$.

1. Describir el espacio muestral $\mathcal{X}$ y el espacio paramétrico Θ.

   El espacio muestral es $\mathcal{X} = \{(x_1, \ldots, x_n) : x_j \in \{0,1\}\}$ y el espacio paramétrico es $\Theta = (0,1)$

2. Comprobar que la distribución pertenece a la familia exponencial. Encontrar un estadístico minimal suficiente.

   La función de densidad es

   $$f_p(x) = p^x(1-p)^{1-x} = \left(\frac{p}{1-p}\right)^x (1-p) = (1-p)\exp\left\{x \ln\left(\frac{p}{1-p}\right)\right\}$$

   para $x \in \{0,1\}$ y $p \in (0,1)$. El modelo pertenece a la familia exponencial tomando $c(p) = 1-p, h(x) = 1, q(p) = \ln\left(\frac{p}{1-p}\right)$ y $t(x) = x$. Entonces $T(X_1, \ldots, X_n) = \sum_{j=1}^{n} X_j$ es un estadístico minimal suficiente.

3. Hallar el estimador de máxima verosimilitud de p.

   La función de verosimilitud para la muestra observada es

   $$L(p) = f_p(x_1, \ldots, x_n) = \prod_{j=1}^{n} f_p(x_j) = p^t(1-p)^{n-t}$$

   donde $t = \sum_{j=1}^{n} x_j$, para $(x_1, \ldots, x_n) \in \mathcal{X}$ y $p \in (0,1)$. La función soporte es

   $$l(p) = \ln L(p) = t \ln p + (n-t)\ln(1-p).$$

   Derivamos e igualamos a cero

   $$\begin{aligned} l'(p) &= \frac{t}{p} - \frac{n-t}{1-p} = 0, \\ \hat{p} &= \frac{t}{n} = \bar{x}. \end{aligned}$$

Además

$$l''(\hat{p}) = \frac{-n}{\bar{x}(1-\bar{x})} < 0,$$

de donde $\hat{p}_{MV} = \bar{x}$ si $\bar{x} \neq \{0,1\}$.

Ejercicio 2. Sea $(X_1, \ldots, X_n)$ una muestra aleatoria simple de una variable aleatoria $X \sim N(0, \sigma^2)$. Encontrar el estimador de máxima verosimilitud para σ^2.

La función de verosimilitud es

$$\begin{aligned} L(\sigma^2) &= f_{\sigma^2}(x_1, \ldots, x_n) = \prod_{j=1}^{n} f_{\sigma^2}(x_j) = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2} x_j^2\right\} \\ &= \left(\frac{1}{\sqrt{2\pi}}\right)^n (\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{j=1}^{n} x_j^2\right\}, \end{aligned}$$

y la función soporte es

$$l(\sigma^2) = -n \ln\left(\sqrt{2\pi}\right) - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{j=1}^{n} x_j^2.$$

Derivamos e igualamos a cero

$$\begin{aligned} l'(\sigma^2) &= -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} \sum_{j=1}^{n} x_j^2 = 0, \\ \hat{\sigma}^2 &= \frac{1}{n} \sum_{j=1}^{n} x_j^2 \end{aligned}$$

Además

$$l''(\hat{\sigma}^2) = \frac{-1}{2(\hat{\sigma}^2)^3} \sum_{j=1}^{n} x_j^2 = \frac{-n^3}{2\left(\sum_{j=1}^{n} x_j^2\right)^2} < 0,$$

y entonces $\hat{\sigma}^2_{MV} = \frac{1}{n} \sum_{j=1}^{n} x_j^2 \in (0, \infty) = \Theta$.

Ejercicio 3. Sea $(X_1, \ldots, X_n)$ una muestra aleatoria simple de una variable aleatoria $X \sim N(\mu, \sigma^2)$. Encontrar el estimador de máxima verosimilitud para $\theta = (\mu, \sigma^2)$.

La función de verosimilitud es

$$\begin{aligned} L(\theta) &= f_\theta(x_1, \ldots, x_n) = \prod_{j=1}^{n} f_\theta(x_j) = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x_j-\mu)^2\right\} \\ &= \left(\frac{1}{\sqrt{2\pi}}\right)^n (\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{j=1}^{n}(x_j-\mu)^2\right\}, \end{aligned}$$

y la función soporte es

$$l(\theta) = -n\ln\left(\sqrt{2\pi}\right) - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{j=1}^{n}(x_j-\mu)^2.$$

Derivamos con respecto a μ e igualamos a cero

$$\begin{aligned} \frac{\partial l(\theta)}{\partial \mu} &= \frac{1}{\sigma^2}\sum_{j=1}^{n}(x_j-\mu) = \frac{1}{\sigma^2}\left(\sum_{j=1}^{n} x_j - n\mu\right) = 0, \\ \hat{\mu} &= \bar{x}. \end{aligned}$$

Derivando ahora con respecto a σ^2 e igualando a cero

$$\begin{aligned} \frac{\partial l(\theta)}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{j=1}^{n}(x_j-\mu)^2 = 0, \\ \hat{\sigma}^2 &= \frac{1}{n}\sum_{j=1}^{n}(x_j-\mu)^2. \end{aligned}$$

Si se sustituye μ por $\hat{\mu}$, obtenemos como estimador de la varianza la varianza muestral $\hat{\sigma}^2 = \frac{1}{n}\sum_{j=1}^{n}(x_j-\bar{x})^2$.

La matriz hessiana en $(\hat{\mu}, \hat{\sigma}^2)$

$$H = \begin{pmatrix} -\frac{n}{\hat{\sigma}^2} & 0 \\ 0 & -\frac{n(\hat{\sigma}^2)^2}{2} \end{pmatrix}$$

es definida negativa, por lo tanto el estimador de máxima verosimilitud es $\hat{\theta}_{MV} = (\hat{\mu}, \hat{\sigma}^2) = \left(\bar{x}, \frac{1}{n}\sum_{j=1}^{n}(x_j-\bar{x})^2\right) \in \Theta.$

Ejercicio 4. Sea $(X_1, \ldots, X_n)$ una muestra aleatoria simple de una variable aleatoria $X \sim U(0, \theta)$.

1. Comprobar que θ es un parámetro de escala y hallar un estadístico minimal suficiente y completo para θ.

   El parámetro θ es un parámetro de escala porque la distribución de la variable $Y = \frac{X}{\theta} \sim U(0,1)$ no depende de θ.

   La función de densidad conjunta es

$$f_\theta(x_1, \ldots, x_n) = \prod_{j=1}^{n} f_\theta(x_j) = \prod_{j=1}^{n} \frac{1}{\theta} I_{(0,\theta)}(x_j) = \frac{1}{\theta^n} I_{(0,\theta)}(x_{(n)}).$$

   Por el teorema de factorización $T(x_1, \ldots, x_n) = x_{(n)}$ es suficiente.

   Además, el cociente

$$\frac{f_\theta(x_1, \ldots, x_n)}{f_\theta(y_1, \ldots, y_n)} = \frac{\frac{1}{\theta^n} I_{(0,\theta)}(x_{(n)})}{\frac{1}{\theta^n} I_{(0,\theta)}(y_{(n)})} = \frac{I_{(0,\theta)}(x_{(n)})}{I_{(0,\theta)}(y_{(n)})}$$

   no depende de θ cuando $x_{(n)} = y_{(n)}$, entonces el estadístico es minimal suficiente.

   Para comprobar que es completo, se calcula

$$E(g(T(X_1, \ldots, X_n))) = E(g(X_{(n)})) = \int_0^\theta g(x) f_{X_{(n)}}(x) dx.$$

   La densidad de la variable $X_{(n)}$ es

$$f_{X_{(n)}}(x) = n f_\theta(x) (F(x))^{n-1} = \frac{n}{\theta^n} x^{n-1} I_{(0,\theta)}(x),$$

   y, por tanto

$$E(g(X_{(n)})) = \int_0^\theta g(x) \frac{n}{\theta^n} x^{n-1} dx = 0,$$

   implica que

$$\begin{aligned} \int_0^\theta g(x) x^{n-1} dx &= 0 \\ \frac{d}{d\theta} \int_0^\theta g(x) x^{n-1} dx &= 0 \end{aligned}$$

de donde se concluye que $g(\theta)\theta^{n-1} = 0$ y $g = 0$ c.s.

2. Hallar el estimador de máxima verosimilitud para θ.

   La función de verosimilitud es

$$L(\theta) = \frac{1}{\theta^n} I_{(x_{(n)},\infty)}(\theta)$$

   Como $\frac{1}{\theta^n}$ es decreciente, el máximo se alcanza en $\theta = x_{(n)}$. Si tomamos como espacio paramétrico $\Theta = (0, \theta]$, entonces

$$L(\theta) = \frac{1}{\theta^n} I_{[x_{(n)},\infty)}(\theta)$$

   y el estimador de máxima verosimilitud es $\hat{\theta}_{MV} = x_{(n)}$.