PREMIOS JORGE JUAN 2004
ESTADÍSTICA

Sean X_1, X_2, ..., X_n variables aleatorias independientes, con función de densidad común:

$$f(x) = \lambda e^{-\lambda x}, \quad x > 0, \quad \lambda > 0$$

Supongamos que estas variables no son observables, y que en su lugar, podemos observar las variables aleatorias Y_1, Y_2, ..., Y_n, donde

$$Y_i = kd \quad \text{si } kd = X_i < (k+1)d$$

para k = 0, 1, 2, ... y d > 0 es conocido.

a) Demostrar que las variables aleatorias Y_1, Y_2, ..., Y_n son independientes, y calcular la distribución de cada una de ellas.

b) Demostrar que la función de probabilidad conjunta de Y_1, Y_2, ..., Y_n es:

$$P(Y_1 = y_1, ..., Y_n = y_n) = \exp\left\{-\lambda \sum_{i=1}^{n} y_i\right\}\left(1 - e^{-\lambda\delta}\right)^n$$

c) Demostrar que el estimador de máxima verosimilitud de ?, basado en Y_1, Y_2, ..., Y_n , es

$$\hat{\lambda} = \frac{1}{\delta} \operatorname{Log}\left(\frac{n\delta + \sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} Y_i}\right)$$

d) Demostrar que si $\left\{\hat{\lambda}_n\right\}$ es la sucesión de estimadores de máxima verosimilitud obtenida en c), entonces

$$\sqrt{n}\left(\hat{\lambda}_n - \lambda\right) \xrightarrow{d} N\left(0, \sigma^2(\lambda, \delta)\right)$$

donde $\sigma^2(\lambda, \delta) \to \lambda^2$ cuando $\delta \to 0$.

SOLUCION

a) Cada v.a. Y_i es función de la v.a. X_i, $\forall$i=1,...,n, y X_1, X_2, ..., X_n son variables aleatorias independientes, por lo tanto también lo son Y_1, Y_2, ..., Y_n.

$$P(Y_i = y) = P(k\delta \leq X_i < (k+1)\delta) = \int_{k\delta}^{(k+1)\delta} \lambda \exp(-\lambda x) dx = e^{-\lambda y}\left(1 - e^{-\lambda\delta}\right)$$

$$y = k\delta,\ k = 0,1,\ldots\ \delta > 0$$

b) Inmediato, pues $P(Y_1 = y_1,\ldots,Y_n = y_n) = \prod_{i=1}^{n} P(Y_i = y_i)$

c) La función de verosimilitud de ?, dada una realización $(y_1,\ldots,y_n)$:

$$L(y_1,\ldots,y_n,\lambda) = \exp\left\{-\lambda \sum_{i=1}^{n} y_i\right\}\left(1 - e^{-\lambda\delta}\right)^n, \quad \lambda > 0$$

Maximizamos su logaritmo:

$$\log L(y_1,\ldots,y_n,\lambda) = \left\{-\lambda \sum_{i=1}^{n} y_i\right\} + n\log\left(1 - e^{-\lambda\delta}\right)$$

$$\frac{\partial}{\partial\lambda}\log L(y_1,\ldots,y_n,\lambda) = \left\{-\sum_{i=1}^{n} y_i\right\} + \frac{n}{1 - e^{-\lambda\delta}}\delta e^{-\lambda\delta} = 0 \Leftrightarrow$$

$$\hat{\lambda} = \frac{1}{\delta}\operatorname{Log}\left(\frac{n\delta + \sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} y_i}\right).$$

Se comprueba fácilmente que $\frac{\partial^2}{\partial\lambda^2}\log L(y_1,\ldots,y_n,\lambda) < 0 \quad \forall\lambda$

Por lo tanto se trata de un máximo (absoluto), y será el estimador de máxima verosimilitud del parámetro.

d) Sabiendo que se verifica :

$$\left(I_n(\lambda)\right)^{1/2}\left(\hat{\lambda}_n - \lambda\right) \xrightarrow{d} N(0,1)$$

donde $I_n(\lambda) = -E\left(\frac{\partial^2}{\partial\lambda^2}\log f(Y_1,\ldots,Y_n)\right)$, siendo $f(y_1,\ldots,y_n)$ la función de probabilidad conjunta de Y_1, Y_2, ..., Y_n, tendremos:

$$I_n(\lambda) = -E\left(\frac{\partial^2}{\partial\lambda^2}\log f(Y_1,\ldots,Y_n)\right) = -E\left(\frac{-n\delta^2 e^{-\lambda\delta}}{(1-e^{-\lambda\delta})^2}\right) = \frac{n\delta^2 e^{-\lambda\delta}}{(1-e^{-\lambda\delta})^2}$$

Entonces $(I_n(\lambda))^{1/2} = \frac{\sqrt{n}\delta e^{-\frac{\lambda\delta}{2}}}{(1-e^{-\lambda\delta})}$, pues $\lambda\delta > 0 \Rightarrow e^{\lambda\delta} > 1$.

Por lo tanto:

$$\frac{\sqrt{n}\delta e^{-\frac{\lambda\delta}{2}}}{(1-e^{-\lambda\delta})}\left(\hat{\lambda}_n - \lambda\right) \xrightarrow{d} Z \Rightarrow \sqrt{n}\left(\hat{\lambda}_n - \lambda\right) \xrightarrow{d} \frac{(1-e^{-\lambda\delta})}{\delta e^{-\frac{\lambda\delta}{2}}} Z$$

siendo Z una v.a. con distribución N(0,1), de manera que

$\frac{(1-e^{-\lambda\delta})}{\delta e^{-\frac{\lambda\delta}{2}}} Z$ tiene distribución $N\left(0, \sigma^2(\lambda,\delta)\right)$, siendo

$$\sigma^2(\lambda,\delta) = \frac{(1-e^{-\lambda\delta})^2}{\delta^2 e^{-\lambda\delta}}.$$

Por último, se comprueba que $\lim_{\delta\to 0}\sigma^2(\lambda,\delta) = \lambda^2$, aplicando 2 veces la regla de L´Hôpital.