Universidad de Chile
Facultad de Ciencias Físicas y Matemáticas
Departamento de Ingenieria industrial
IN-3242 – Estadistica

Estimador de Máxima Verosimilitud

MLE - EMV

Profesor: Raimundo Undurraga
Auxiliares: Brandon Galarza, Camila Jáuregui Leonardo Meneses, Francisca Monetta, Matias Reyes, Bastian Urzúa, Antonia Villegas

P1.-

1. El tiempo de realización en minutos de una determinada tarea dentro de un proceso industrial es una VA con función de densidad:

$$f(x) = \frac{x}{\theta^2} e^{\frac{-x}{\theta}}, x > 0$$

Suponga que tiene una muestra aleatoria simple de N observaciones. Buscamos estimar el parámetro θ a partir del método de méxima verosimilitud.

a) Plantee la función de verosimilitud. Intuitivamente, ¿Qué significa esta función?

b) Obtenga una expresión para la función de log-verosimilitud.

c) Determine el estimador máximo verosímil de θ. Explique en qué consiste obtener un estimador a través de este método.

2. Considere la siguiente función de distribución:

$$f(x) = \begin{cases} \frac{\alpha}{x^{\alpha-1}} & si \quad 0 \leq x \leq 1 \\ \\ 0 & si \qquad \sim \end{cases}$$

Encuentre el EMV de α para una muestra aleatoria de tamaño N .

$\boxed{\textbf{P2.-}}$ Considere una muestra que proviene de una variable aleatoria D que distribuye Exponencial con un parámetro poblacional dado por λ, tal que su densidad corresponde a:

$$Pr(D = d_i) = -\lambda e^{-\lambda d_i}$$

Usted, a su vez, observa una muestra de 8 observaciones, dada por {4.8, 5, 3.6, 10, 4.7, 6.2, 7.3, 5}.

a) Obtenga el Estimador por Máxima Verosimilitud de λ, describiendo detenidamente los pasos y corroborando que sea el que maximice la probabilidad de observar la muestra

b) ¿Cuál es el valor estimado, dada la muestra, de $\widehat{\lambda}_{EMV}$?

Resumen

- **Función de Verosimilitud *(Likelihood)***

  Sea $X = (x_1, x_2, \cdots, x_n)$ variables aleatorias con densidad conjunta $f(X|\theta) = f(x_1, x_2, \cdots, x_n, \theta)$, con probabilidad conjunta con una función que depende de θ.

  Por lo tanto la función verosimilitud se define como:

$$L(\theta|X) = L(\theta, x_1, x_2, \cdots, x_n) = f(X|\theta)$$

  Likelihood : Probable que ocurra

- **Estimador de Máxima Verosimilitud (*MLE o EMV*)**

  Maximizamos $L(\theta|X)$ para encontrar aquel $\widehat{\theta}$ que maximiza la probabilidad de ocurrencia de lo que observamos en los datos de X

$$\widehat{\theta}_{EMV} = \underset{\theta}{\text{argmax}} L(\theta|X)$$

- **Para estimar el estimador se requiere:**

  i. Conocer o asumir la distribución de los datos X (Recordando la materia vista en la primera parte del curso).

  ii. Considerar que la muestra aletaoria $X = (x_1, x_2, \cdots, x_n)$ sean i.i.d. Mediante esto, ya que la función se puede descomponer de la siguiente forma

$$L(\theta|X) = f(X|\theta) = f(x_1, x_2, \cdots, x_n|\theta) = \prod_{i=1}^{N} f(X_i|\theta)$$

  Para facilitar el método anterior, podemos trabajar con la log-verosimilud, obteniendo:

$$ln(L(\theta|X)) = ln\left(\prod_{i=1}^{N} f(X_i|\theta)\right)$$

  Llegando a lo siguiente:

$$\sum_{i=1}^{N} ln(f(X_i|\theta)) \tag{1}$$

  (1) Podemos obtner esta trasformación ya que ln() es una función monótona y creciente, por lo tanto :

$$\widehat{\theta}_{EMV} = \underset{\theta}{\text{argmax}} L(\theta|X) = \underset{\theta}{\text{argmax}} ln(L(\theta|X))$$

- **Propiedades de EMV para <u>Muestras Grandes</u>:**

  Bajo condiciones de regularidad:

  - EMV es **consistente**
  - EMV es **asintóticamente eficiente**
  - EMV es **asintóticamente distribuido normal**

- **Propiedades de EMV para <u>Muestras Finitas</u>:**

  - EMV es consistente, pero pueden estar severamente sesgado en muestras finitas
  - La estimación matriz varianza-covarianza puede ser muy dudosa en muestras finitas

<u>"Recetario para obtener una EMV/MLE"</u>

a) Escribir la función verosimilitud $\boxed{\mathrm{L}(\theta|X) = f(x_1, x_2, \cdots, x_n, \theta)}$

b) Escribir la función log-verosimilitud $\boxed{\ln(\mathrm{L}(\theta|X)}$

c) Maximizar la función usando la Condición de Primer Orden (CPO) $\frac{\partial ln(L(\theta|X))}{\partial\theta} = 0$

d) Despejar θ denotándolo como $\widehat{\theta}_{EMV}$

e) Comprobar que sea un máximo, o sea, que se cumpla la Condición Segundo Orden (CSO) $\frac{\partial^2 ln(L(\theta|X)}{\partial\theta^2} < 0$