Métodos de estimación

Para obtener estimadores puntuales

Existen los siguientes métodos de estimación:

- Método de los Momentos
- Método de Máxima Verosimilitud
- Método Bayesiano
- Método de Mínimos Cuadrados

Método de los Momentos:

Consiste en igualar los momentos poblacionales con los muestrales, al resolver estas ecuaciones con valores de parámetros conocidos se obtienen los estimadores.

$\sum_{x \epsilon R_x} x^k f(x)$ v.a discreta

Momento poblacional de orden E(X^k)

$\int_{-\infty}^{+\infty} x^k f(x)dx$ v.a continua

Momentos muestrales de orden k: $M_k = \sum_{i=1}^{n} \frac{{X_i}^k}{n}$

- Entonces el método consiste en igualar los momentos muestrales con los poblacionales y plantear el sistema de ecuaciones:

E(X^k)= $\sum_{i=1}^{n} \frac{{X_i}^k}{n}$ k=1, …,q

Sea $X_1, X_2, \dots, X_n$ v.a iid con f(x_i, θ) con $\theta \epsilon \Theta \subseteq \mathbb{R}^q$

$E_\theta(X) = \sum_{i=1}^{n} \frac{X_i}{n}$

$E_\theta\ (X^2) = \sum_{i=1}^{n} \frac{{X_i}^2}{n}$

.

.

$E_\theta\ (X^{q+r}) = \sum_{i=1}^{n} \frac{{X_i}^{q+r}}{n}$

Sistema cuya solución $\hat{\theta}^{MM} = (\widehat{\theta_1}^{MM}, \widehat{\theta_2}^{MM}, \dots, \widehat{\theta_{q+r}}^{MM})$

Este sistema puede ser o no lineal, puede tener o no solución.

- **Ejemplo 1**:

Sea $X_i \sim \varepsilon(\lambda)$ con $f(x_i, \lambda)$=$\begin{cases} \lambda e^{-\lambda x}; \ x_i \geq 0 \\ 0 \ ; c.c \end{cases}$

$\lambda \epsilon \mathbb{R}^+$, en este caso θ=λ, q=1 $\Theta = \mathbb{R}^+$

$\theta \epsilon \ \Theta \subseteq \mathbb{R}$

E(X_i)=M_1

$\frac{1}{\lambda}$ = $\bar{X}$, despejo λ,

Por lo tanto el estimador de λ por el Método de los Momentos es:

$$\hat{\lambda}^{MM} = \frac{1}{\bar{X}}$$

- **Ejemplo 2**:
- Sea X_i ~U(- θ, θ) con $\theta \epsilon \mathbb{R}^+$ q=1

- $f(x_i)$=$\begin{cases} \frac{1}{2\theta}, -\theta \leq \ x_i \leq \theta \\ \quad 0, c.c \end{cases}$ E(X_i)= (- θ+ θ)/2=0

E(X_i)=0, no depende de θ, busco con otro momento

E(X^2)=$\sum_{i=1}^{n} \frac{{X_i}^2}{n}$

V(X)= E(X^2) $-$ $E(X)^2$ entonces, V(X)= E(X^2)= $\frac{(2\theta)^2}{12}$

$\frac{(2\theta)^2}{12}$= $\sum_{i=1}^{n} \frac{{X_i}^2}{n}$ despejo θ

θ^2= 3 $\sum_{i=1}^{n} \frac{{X_i}^2}{n}$ despejo θ, $\hat{\theta}^{MM}$=$\sqrt{\frac{3}{n}\sum_{i=1}^{n} X_i^2}$

Estimador de Máxima Verosimilitud (E.M.V)

El estimador de Máxima Verosimilitud será el valor del parámetro que maximiza la función de verosimilitud.

Definición:

Supongamos que X es una v.a con distribución de probabilidad f(x, θ) donde θ es un parámetro desconocido. Sean $x_1, x_2, \ldots, x_n$ los valores observados de una m.a de tamaño n. La función de verosimilitud de la muestra es

L(θ)= f(x_1, θ) ….. f(x_n, θ) función de distribución conjunta de la muestra

Por la Independencia es el producto de las función de densidad o de frecuencia según sea el caso

La función de verosimilitud es una función del parámetro desconocido θ. El estimador de máxima verosimilitud de θ es el valor de θ que maximiza la función de verosimilitud L(θ).

También se lo puede definir como:

Dado X=$(X_1, X_2, \dots, X_n)$ con X$\sim F(x,\theta)$ $\theta \epsilon \Theta \subseteq \mathbb{R}^k$

$\hat{\theta}^{MV}$= $\underset{\theta\epsilon\Theta}{\operatorname{Argmax}} f(x,\theta)$

Pasos para obtener estimadores de Máxima Verosimilitud:

1) Encontrar la función de Máxima Verosimilitud

2)Aplicar logaritmo en base e, (función continua, creciente y conserva los máximos)

3) Derivo respecto al parámetro, e igualo a 0 para buscar puntos críticos

4) Despejo el parámetro

5) Compruebo que es máximo derivando nuevamente (no se pide)

Ejemplo caso discreto

Ejemplo (caso discreto)

Sea X $\sim$P(λ) su función de frecuencia es p(x)=$e^{-\lambda}\frac{\lambda^{x}}{x!}$ con x=0, 1,….

La función de verosimilitud es:

L(λ)=$_{\text{Indep}}$p(x_1).p(x_2)….p(x_n)= $e^{-\lambda}\frac{\lambda^{x_1}}{x_1!}$. $e^{-\lambda}\frac{\lambda^{x_2}}{x_2!}$…. $e^{-\lambda}\frac{\lambda^{x_n}}{x_n!}$=

Agrupando queda:

$$=e^{-n\lambda}\frac{\lambda^{\left(\sum_{i=1}^{n} x_i\right)}}{\prod_{i=1}^{n} x_i!}$$

Ahora hay que maximizar, aplico ln

$\ln(L(\lambda)) = \ln(e^{-n\lambda}\frac{\lambda^{(\sum_{i=1}^{n} x_i)}}{\prod_{i=1}^{n} x_i!}) = \ln(e^{-n\lambda}) + \ln(\lambda^{(\sum_{i=1}^{n} x_i)}) - \ln(\prod_{i=1}^{n} x_i!) =$

$-n\lambda \ln(e) + (\sum_{i=1}^{n} x_i)\ln(\lambda) - \ln(\prod_{i=1}^{n} x_i!) =$

$-n\lambda \cdot 1 \quad + (\sum_{i=1}^{n} x_i)\ln(\lambda) - \ln(\prod_{i=1}^{n} x_i!) =$

Ahora derivo con respecto a λ:

$\frac{d\ln(L(\lambda))}{d\lambda} = -n + \sum_{i=1}^{n} x_i / \lambda = 0$ despejo λ

$\lambda = \frac{1}{n}\sum_{i=1}^{n} x_i$, entonces $\hat{\lambda}^{MV} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$

¿Cómo sé que es máximo?

Verifico con la 2º derivada (no lo comprobamos)

- **Ejemplo caso continuo**: X $\sim\epsilon(\lambda)$ $f(x) = \begin{cases} \lambda e^{-\lambda x}, si\ x \geq 0 \\ 0,\ c.c \end{cases}$

L(λ)= $\prod_{i=1}^{n} f(x_i)$= $\prod_{i=1}^{n} \lambda e^{-\lambda x_i}$= $\lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}$; $\forall i, x_i \geq 0$

Aplico Ln; ln(L(λ))= ln($\lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}$)=$_{\text{prop}}$ ln(λ^n) + ln($e^{-\lambda \sum_{i=1}^{n} x_i}$)=

$n$ln(λ)+ $(-\lambda \sum_{i=1}^{n} x_i)$ ln(e) = $n$ln(λ)+ $(-\lambda \sum_{i=1}^{n} x_i)$. 1

Ahora derivo con respecto a λ:

$\frac{d\ln(L(\lambda))}{d\lambda}$= n/ λ $- \sum_{i=1}^{n} x_i = 0$ despejo λ,

λ= $\frac{n}{\sum_{i=1}^{n} x_i}$ con lo cual $\hat{\lambda}^{MV} = \frac{1}{\bar{X}}$

Observación: notar que con ambos métodos el estimador es el mismo, esto no siempre ocurre

- **Ejemplo:**

Sea $X_1, X_2, \dots, X_n$ m.a de una v.a U(0,θ),entonces $f(x_i) = \begin{cases} \frac{1}{\theta}, x_i \epsilon (0,\theta) \\ 0, c.c \end{cases}$

L(θ)= $\prod_{i=1}^{n} f(x_i)$= $\prod_{i=1}^{n} \frac{1}{\theta} \quad \forall i, x_i \epsilon (0,\theta)$

L(θ) no es una función continua del parámetro

L(θ) $\underbrace{=}_{(*)} \quad \frac{1}{\theta^n} \prod_{i=1}^{n} I_{(0,\theta)}(x_i) \underbrace{=}_{(**)}$

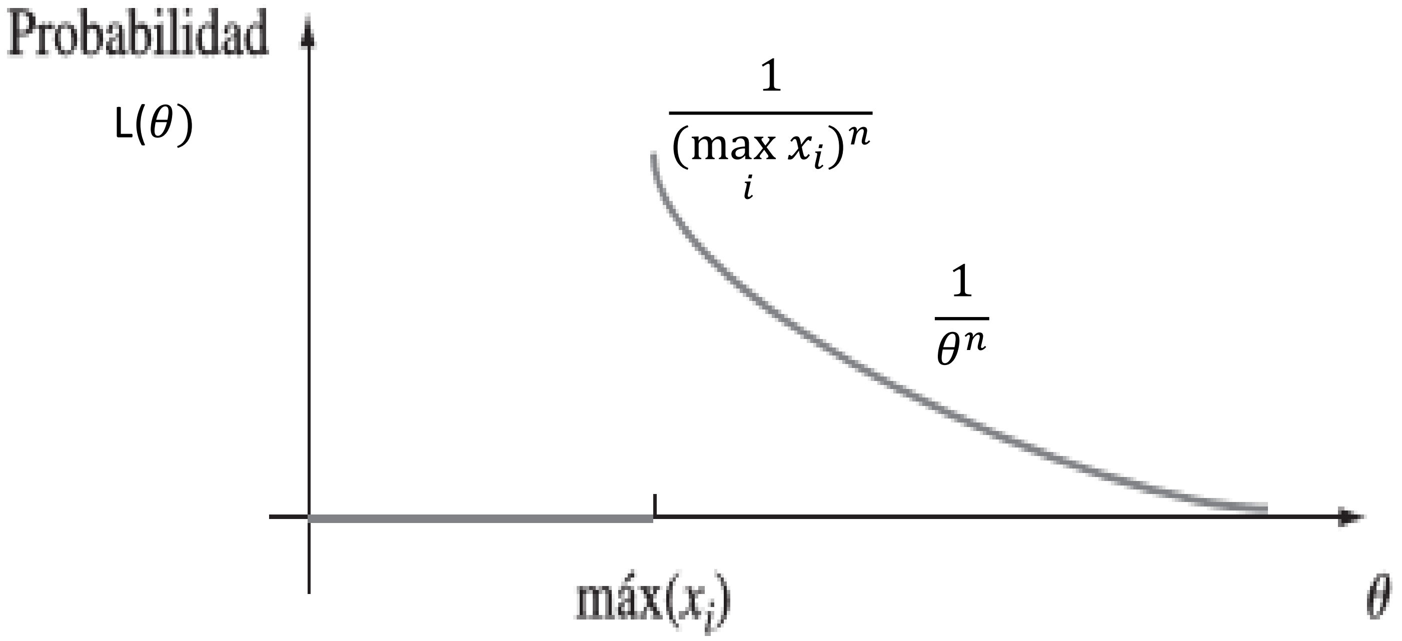
= $\frac{1}{\theta^n} I_{(-\infty,\theta)}(\max_i x_i)\, I_{(0,\infty)}(\min_i x_i)$

Como la función es decreciente, se maximiza la seleccionar θtan pequeño como sea posible sujeto a $x_i \epsilon (0,\theta)$.

$\hat{\theta}^{MV} = \max_i X_i$ =$X_{(n)}$

(*) defino la función indicadora I como: $I_{(a,b)}(z) = \begin{cases} 1, z \epsilon [a,b] \\ 0, \quad c.c \end{cases}$

(**)Notar que $\forall i, i = 1, \dots, n \; x_i \epsilon (0,\theta)$ *sii* $\max_i x_i \leq \theta ; \min_i x_i \geq 0$



Como la función L(θ) es decreciente, se maximiza la seleccionar θ tan pequeño como sea posible sujeto a $x_i \epsilon (0, \theta)$.
El cálculo no funciona porque el máximo ocurre en un punto de discontinuidad.

Propiedades de los E.M.V:

- Son asintóticamente consistentes
- Son asintóticamente insesgados
- Son asintóticamente de varianza mínima

Propiedad de la invarianza:

Sea $\theta \epsilon \Theta \subseteq \mathbb{R}^k$ con $\hat{\theta}^{MV}$ sea g(θ)=a

$$\hat{a}^{MV} = g(\hat{\theta}^{MV})$$

Ejemplo: Sea X $\sim$P(λ)

Quiero encontrar el EMV de P(X=0)= $e^{-\lambda}\frac{\lambda^{o}}{o!}$= $e^{-\lambda}$ = g(λ)

$$P(\widehat{X=0}) = \widehat{g(\lambda)} \underbrace{=}_{prop\ inv.} \mathrm{g}(\hat{\lambda}^{MV})$$

Aplicamos a nuestro ejemplo:

$$P(\widehat{X=0}) = \widehat{e^{-\lambda}} \underbrace{=}_{prop\ inv.} e^{-\hat{\lambda}} \underbrace{=}_{(*)} e^{-\bar{X}}$$

(*) reemplazo por $\hat{\lambda}^{MV}$= $\bar{X}$

Por lo tanto: El EMV de P(X=0) es $e^{-\bar{X}}$