TEMA 4

ESTIMACIÓN PUNTUAL

INDICE

- **INTRODUCCIÓN**
- **MÉTODOS DE CONSTRUCCIÓN DE ESTIMADORES**
  - **Método de los momentos.**
  - **Método de máxima verosimilitud**
- **PROPIEDADES DE UN ESTIMADOR**
  - **Insesgadez**
  - **Eficiencia**
  - **Consistencia**

INTRODUCCIÓN

La Teoría de la Estimación tiene como objetivo aproximar, lo más posible, el verdadero valor de un parámetro poblacional.

Esquema de la teoría de estimación:

- Estimación
  - **Puntual**
    - Métodos Construcción
      - **Métodos momentos**
      - **Máxima verosimilitud**
      - Mínimos cuadrados
      - Bayes
      - Bootstrap
    - Propiedades Estimadores
      - **Insesgadez**
      - **Eficiencia**
      - **Consistencia**
      - Suficiencia
  - Intervalos (Tema 5)

INTRODUCCIÓN

En lo sucesivo, sea X v.a poblacional, con función de densidad conocida salvo un parámetro θ, se denota f(x|θ) y sea X_1, X_2,....X_n una m.a.s. con reposición extraída de una población X.

ESTIMADOR: Cualquier estadístico que proporciona valores en el espacio donde está definido el parámetro desconocido θ. En general, un estimador es un estadístico muestral

$$\hat{\theta}_n = \hat{\theta}(X_1, \ldots, X_n)$$

ESTIMACIÓN PUNTUAL: Es el valor que proporciona el estimador cuando se determinan los valores concretos de la muestra.

ESTIMACIÓN POR INTERVALOS: busca evaluar el error cometido en la estimación proporcionando un intervalo entre cuyos límites se encuentra el valor del parámetro con una probabilidad de acertar fijada de antemano

Ejemplo 1

Sea $X \sim \exp\left(\frac{1}{\theta}\right)$ cuya función de densidad es: $f(x\,|\,\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}$

Tomamos una muestra de tamaño n=10 cuyos valores son:

$$(X_1, X_2, \ldots, X_{10}) = (2'5,\ 3'5,\ 2'8,\ 3,\ 2'9,\ 3'4,\ 3'4,\ 2'5,\ 3'1,\ 2'1)$$

Tabla de estimadores

Estimadores	Estimaciones puntuales
$\hat{\theta}_1 = \overline{X}_n$	$\hat{\theta}_1 = 2{,}92$
$\hat{\theta}_2 = X_1$	$\hat{\theta}_2 = 2'5$
$\hat{\theta}_3 = \frac{1}{n+1}\sum_{i=1}^{n} X_i$	$\hat{\theta}_3 = \frac{29{,}2}{11} = 2{,}65$
$\hat{\theta}_4 = 4$	$\hat{\theta}_4 = 4$

MÉTODOS CONSTRUCCIÓN

MÉTODO DE LOS MOMENTOS

a) Método intuitivo. Introducido por Pearson.
b) Consiste en igualar las características poblacionales con las muestrales y
c) Expresar los parámetros como función de las características muestrales.

Ejemplo 2:

Sea X ~ U(a,b), queremos estimar a y b .

$$\mu = E[X] = \frac{a+b}{2} \qquad \hat{\mu} = \overline{X}_n \Rightarrow b + a = 2\overline{X}_n$$

$$\sigma^2 = V[X] = \frac{(b-a)^2}{12} \qquad \hat{\sigma}^2 = S_{1,n}^2 \Rightarrow (b-a)^2 = 12S_{1,n}^2 \Rightarrow b - a = 2S_{1,n}\sqrt{3}$$

$$\boxed{\left.\begin{array}{l} \hat{b} = \overline{X}_n + S_{1,n}\sqrt{3} \\ \hat{a} = \overline{X}_n - S_{1,n}\sqrt{3} \end{array}\right\} \text{Estimadores por el método de los momentos}}$$

Ejercicio

Sea X una variable aleatoria cuya función de densidad es la siguiente:

$$f(x \mid \theta) = \begin{cases} (\theta+1)x^{\theta} & 0 < x < 1 \\ 0 & \text{resto} \end{cases}$$

Obtener, por el método de los momentos, un estimador del parámetro θ basado en una muestra de **n** observaciones independientes

FUNCIÓN VEROSIMILITUD

Dada una m.a.s. (X_1, X_2,........, X_n) extraída de una variable aleatoria X, cuya función de densidad es f(x,θ), llamaremos **función de verosimilitud de la muestra** a:

$$L(\theta\,;X)=\prod_{i=1}^{n} f\,(x_i;\theta)=f\,(x_1,x_2,...,x_n;\theta)$$

Si la v.a. X es discreta:

$$L(\theta\,;X)=\prod_{i=1}^{n} P(X=x_i;\theta)=P(x_1,x_2,...,x_n;\theta)$$

Esta función se interpreta como función del parámetro desconocido θ porque la muestra aleatoria será conocida una vez ha sido extraída y observada.

FUNCIÓN VEROSIMILITUD

Ejemplo 3:

La función de verosimilitud mide lo verosímil que es un valor concreto del parámetro θ para explicar los datos muestrales observados

Sea X~**P**(λ) y tomamos una m.a.s. con n = 2

$$L(\lambda; X) = \prod_{i=1}^{2} P(X = x_i; \lambda) = e^{-\lambda}\frac{\lambda^{x_1}}{x_1!}.e^{-\lambda}\frac{\lambda^{x_2}}{x_2!} = e^{-2\lambda}\frac{\lambda^{x_1+x_2}}{x_1!x_2!}$$

En general, para una m.a.s. de tamaño n

$$L(\lambda; X) = e^{-n\lambda}\frac{\lambda^{\sum_{i=1}^{n} x_i}}{x_1!x_2! \ \ x_n!}$$

FUNCIÓN VEROSIMILITUD

Ejemplo 3 (continuación):

Si la muestra observada es $X=(x_1, x_2) = 5, 7$

Probabilidad de que $X=(5,7)$ para cada posible valor de λ.

$$\text{Si } X = (5,7) \Rightarrow \qquad L(\lambda; (5,7)) = e^{-2\lambda} \frac{\lambda^{12}}{5!\,7!}$$

λ	L(λ,(5,7))
1	0,00000
2	0,00012
3	0,00218
4	0,00931
5	0,01833
6	**0,02211**
7	0,01903

Es más verosímil creer que $\lambda = 6$ (probabilidad mayor).

MÉTODO MÁXIMA VEROSIMILITUD

El principio de máxima verosimilitud se basa en que la muestra es representativa de la población y, por lo tanto, es más creíble el valor que maximiza la probabilidad de la muestra obtenida.

Dada la función de verosimilitud L(θ,X) el método consiste en calcular el valor de θ, que maximiza la función anterior.

$$L\left(\hat{\theta}_n, X\right) = \max_{\theta} L(\theta, X)$$

MÉTODO MÁXIMA VEROSIMILITUD

Como L(θ, X) es positiva, el máximo de L(θ, X) coincide con el máximo de ln [L(θ, X)]. El método se resume en resolver la ecuación:

$$\frac{\partial \ln L(\theta, X)}{\partial \theta} = 0 \Rightarrow \text{se obtiene } \hat{\theta}_n$$

Si la solución es no constante y depende de la muestra (X_1, X_2 ..., X_n) se llama **estimador de máxima verosimilitud.**

$\hat{\theta}_n = \hat{\theta}(X_1, \ldots, X_n)$ **M.L.E.* del parámetro θ**

*(M.L.E.) Maximun Likelihood Estimation

Ejemplo 4:

Sea X ~ N(μ,σ),¿Cuál es el estimador M.L.E. de μ?

$$f\left(x_i\right)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{\left(x_i-\mu\right)^2}{2\sigma^2}}$$

$$L\left(\mu,X\right)=\prod_{i=1}^{n}f\left(x_i\right)=\frac{1}{\sigma^n\left(2\pi\right)^{n/2}}e^{-\sum_{i=1}^{n}\frac{\left(x_i-\mu\right)^2}{2\sigma^2}}$$

$$\ln L=-\frac{n}{2}\ln\left(2\pi\right)-n\ln\sigma-\frac{1}{2}\sum_{i=1}^{n}\frac{\left(x_i-\mu\right)^2}{\sigma^2}$$

$$\frac{\partial\ln L}{\partial\mu}=\sum_{i=1}^{n}\frac{\left(x_i-\mu\right)}{\sigma^2}=0\Leftrightarrow\hat{\mu}_n=\frac{1}{n}\sum_{i=1}^{n}x_i$$

PROPIEDADES DE UN ESTIMADOR

- El parámetro θ es desconocido, por lo que **nunca sabremos con exactitud si la estimación puntual** $\left(\hat{\theta}\right)$ **está cerca o lejos del verdadero valor del parámetro (θ)**

- **Utilizar una estimación puntual** en vez del verdadero valor de **θ conduce a cometer un error** más o menos importante. Un criterio de selección del estimador se debe basar en estudiar este error, ligado a la idea de precisión

- El error en la estimación es la diferencia entre el estimador y el valor del parámetro. Por tanto, como el estimador es una variable aleatoria el error también lo es

PROPIEDADES DE UN ESTIMADOR

El objetivo es encontrar el valor del parámetro θ como un valor ideal, por eso nos interesa el concepto de precisión respecto al valor ideal

"¿Cómo mediremos esa precisión?"

Mediante el grado de variabilidad o dispersión con respecto al parámetro desconocido. Para ello utilizamos:

Error cuadrático medio:

$$\mathrm{ECM}_{\theta}\left(\hat{\theta}\right) = \mathrm{E}_{\theta}\left[\left(\hat{\theta} - \theta\right)^{2}\right]$$

PROPIEDADES DE UN ESTIMADOR

Se verifica que:

$$\mathrm{ECM}_\theta\left(\hat\theta\right) = \mathrm{Var}_\theta\left(\hat\theta\right) + \left(\mathrm{E}_\theta\left[\hat\theta\right] - \theta\right)^2 = \mathrm{Var}_\theta\left(\hat\theta\right) + \left(\mathrm{Sesgo}_\theta\left[\hat\theta\right]\right)^2$$

donde

$$\mathrm{Sesgo}_\theta\left(\hat\theta\right) = \mathrm{E}_\theta\left[\hat\theta\right] - \theta$$

recibe el nombre de **sesgo** y cuantifica el error sistemático del estimador, mientras que

$$\mathrm{Var}_\theta\left(\hat\theta\right)$$

mide el error debido a la aleatoriedad del estimador que proviene del proceso de muestreo aleatorio.

.

PROPIEDADES DE UN ESTIMADOR

El mejor estimador será aquél cuyos valores estén más cerca del parámetro desconocido. Así pues, obtenemos un primer criterio de selección.

Definición:

Sean U y T, estimadores de θ, diremos que **U es más eficiente que T respecto del ECM**, si :

$$ECM_\theta(U) \leq ECM_\theta(T) \ \forall\theta \in \Theta$$

y si se cumple que :

$$ECM_\theta(U) < ECM_\theta(T) \text{ para algún } \theta$$

diremos que **U es estrictamente más eficiente que T**

PROPIEDADES DE UN ESTIMADOR

Definición:

Eficiencia relativa de U respecto de T:

$$\mathrm{Eff}_{\theta}(U;T) = \frac{\mathrm{ECM}_{\theta}(U)}{\mathrm{ECM}_{\theta}(T)}$$

Si $\mathrm{Eff}_{\theta}(U;T) \leq 1 \ \forall \theta \in \Theta$ U es **más eficiente** que T

y si, para algún $\theta \in \Theta$

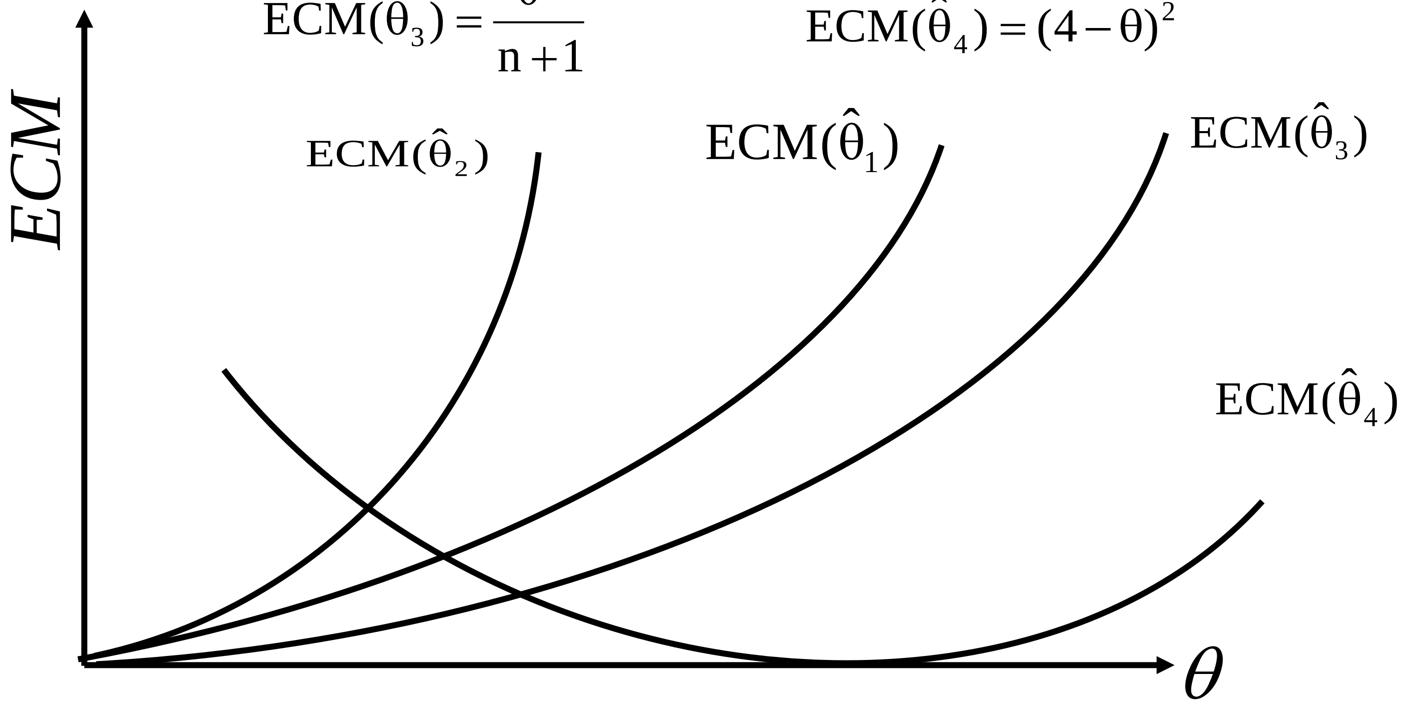
$\mathrm{Eff}_{\theta}(U;T) < 1$ U es **estrictamente más eficiente** que T

PROPIEDADES DE UN ESTIMADOR

Estudio del ECM en el ejemplo 1

$$ECM(\hat{\theta}_1) = \frac{\theta^2}{n} \qquad ECM(\hat{\theta}_2) = \theta^2$$

$$ECM(\hat{\theta}_3) = \frac{\theta^2}{n+1} \qquad ECM(\hat{\theta}_4) = (4-\theta)^2$$



No existe un estimador óptimo, respecto al ECM, para todos los valores de θ

INSESGADEZ

Dado que el ECM puede dar lugar a situaciones aberrantes provocadas por la utilización de estimadores constantes, proponemos exigir que los estimadores tengan sesgo 0 y seleccionar aquél cuya varianza sea mínima.

Se dice que un estimador $\hat{\theta}$ es un estimador **insesgado** de θ si su sesgo es 0. En otro caso se dice que es **sesgado.**

INSESGADEZ

Suponemos un m.a.s. $(X_1, ..., X_n)$ extraída de una población X cuyas características son:

$$E[X] = \mu \text{ y } Var[X] = \sigma^2$$

Vamos a calcular los momentos de los estadísticos

- Media muestral $E[\overline{X}_n] = \mu \quad Var[\overline{X}_n] = \dfrac{\sigma^2}{n}$
- Varianza muestral $E[S_n^2] = \dfrac{n-1}{n}\sigma^2$
- Cuasivarianza muestral $E[S_{1,n}^2] = \sigma^2$
- CASO PARTICULAR: Proporción $X \approx Be(p)$, recordar $p=E[X]$

Proporción muestral $\hat{p}_n = \dfrac{1}{n}\sum_{i=1}^{n} X_i$

$$E[\hat{p}_n] = p \qquad Var[\hat{p}_n] = \frac{p(1-p)}{n}$$

INSESGADEZ

Ejemplos

$\overline{X}_n$ es insesgado para la media μ

$\hat{p}_n$ es insesgado para la proporción p

s_n^2 es sesgado para la varianza $\quad \text{Sesgo}\left(s_n^2\right) = \dfrac{-\sigma^2}{n}$

$s_{1,n}^2$ es insesgado para la varianza

INSESGADEZ

Ejemplo 1

$E(X)= \theta$ $\qquad$ $Var(X)= \theta^2$

Valor esperado	Sesgo	Insesgado
$E(\hat{\theta}_1) = E(\overline{X}_n) = \theta$	0	Si
$E(\hat{\theta}_2) = E(X_1) = \theta$	0	Si
$E(\hat{\theta}_3) = \frac{1}{n+1}\sum_{i=1}^{n} E(X_i) = \frac{n\theta}{n+1}$	$\frac{\theta}{n+1}$	No
$E(\hat{\theta}_4) = 4$	$\theta - 4$	No

MÍNIMA VARIANZA

Definición:

Sea T estimador insesgado de θ. Diremos que T es **insesgado uniformemente de mínima varianza** si

$$Var(T) \leq Var\,(T_1)\ \forall\theta \text{ de } \Theta\ ,\ \forall\, T_1 \text{ insesgado.}$$

¿Cómo sabremos cuál es la mínima varianza posible?

Cota de Frechet-Cramer-Rao (cota FCR)

MÍNIMA VARIANZA

ACOTACIÓN FCR (FRÉCHET-CRAMER-RAO)

La función de verosimilitud debe verificar las siguientes condiciones de regularidad:

i) El soporte de X no depende del parámetro.

ii) El parámetro θ está definido en un intervalo.

iii) $\exists \dfrac{\partial \left(L(\theta;X)\right)}{\partial\theta} \quad \forall\theta\in\Theta\ ,\forall X$

iv) $0 < E\left[\left\{\dfrac{\partial}{\partial\theta}\ln L(\theta;X)\right\}^2\right] < +\infty$

v) $\exists \dfrac{\partial^i E[1]}{\partial\theta^i} \Leftrightarrow \exists \dfrac{\partial^i}{\partial\theta^i}\int L(\theta;X)dX \quad \forall\theta\in\Theta \ \ i=1,2$

y si T es un estimador de θ con varianza finita, tal que $\exists \dfrac{\partial}{\partial\theta}E[T] \quad \forall\theta$ entonces:

$$V[T] \geq \frac{\left(\dfrac{\partial}{\partial\theta}E[T]\right)^2}{E\left[\left(\dfrac{\partial}{\partial\theta}\ln L(\theta;X)\right)^2\right]}$$

EFICIENCIA

Proposición

Si el estimador T es insesgado entonces E[T]= θ y $\frac{\partial}{\partial\theta}E[T]=1$ la cota FCR es:

$$V[T] \geq \ FCR_\theta = \frac{1}{I_n(\theta)}$$

donde $I_n(\theta) = E\left[\left(\frac{\partial}{\partial\theta}\ \ln L(\theta; X)\right)^2\right] = -E\left[\frac{\partial^2}{\partial\theta^2}\ \ln L(\theta; X)\right]$

que se denomina cantidad de información de Fisher

Definición

Eficiencia absoluta de T $\frac{Var_\theta(T)}{FCR_\theta}$

T se dice **EFICIENTE** si alcanza la cota de F-C-R, es decir, si su eficiencia absoluta es 1.

CONSISTENCIA

Otra propiedad que debemos tener en cuenta al pensar en un buen estimador es la **consistencia**, un estimador de θ deberá ser mejor cuanto mayor sea la muestra (disponemos de más información).

Sea $\hat{\theta}_n$ un estimador de θ. Se dice que es un estimador **consistente** para θ si:

$$\lim_{n\to\infty} P\left\{\left|\hat{\theta}_n - \theta\right| \le \varepsilon\right\} = 1 \quad \forall\theta,\ \forall\varepsilon > 0$$

Proposición

$$\text{Si } \hat{\theta}_n \text{ es insesgado y } Var_\theta\left[\hat{\theta}_n\right] \xrightarrow[n\to\infty]{} 0 \Rightarrow \hat{\theta}_n \text{ es consistente}$$

CONSISTENCIA

Ejemplo 8:

Sea la variable X que recoge el gasto mensual de los hogares $X \sim N(\mu,\sigma)$ ¿Es $\overline{X}_n$ consistente para μ?

Sabemos que $\overline{X}_n \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$

El estimador media muestral es insesgado: $E\left[\overline{X}_n\right] = \mu$

La varianza del estimador es: $V\left[\overline{X}_n\right] = \frac{\sigma^2}{n}$

Si calculamos el límite de la varianza cuando tiende a infinito:

$$\lim_{n\to\infty} V\left[\overline{X}_n\right] = \lim_{n\to\infty} \frac{\sigma^2}{n} = 0$$

Por lo tanto, el estimador media muestral es consistente para la media poblacional μ.

PROPIEDADES DE LOS ESTIMADORES DEL MÉTODO DE LOS MOMENTOS

Los estimadores calculados con el método de los momentos:

- ✓ **Suelen ser consistentes**
- ✓ Generalmente, **no son eficientes ni insesgados**.
- ✓ **No aprovechan toda la información disponible** porque solo usa las características resumidas (media, varianza, …)

PROPIEDADES DEL M.L.E.

Si se verifican las condiciones de regularidad de FCR y existe un estimador eficiente, entonces éste es el único M.L.E.

Al revés, no es cierto (MLE no es eficiente en todos los casos, ni siquiera tiene que ser centrado)

S_n^2 no es insesgado para σ^2

Otras propiedades de los M.L.E. son las siguientes:

- ✓ Son consistentes
- ✓ Asintóticamente normales y eficientes

$$\hat{\theta}_n \cong N\left(\theta, \frac{1}{\sqrt{I(\theta)}}\right)$$

- ✓ Invariantes: si $\hat{\theta}_n$ es M.L.E. de θ entonces $g(\hat{\theta}_n)$ es el M.L.E. de $g(\theta)$, si g tiene inversa