

# *Estimación puntual*

# ***El Problema de la Estimación***

Se busca precisar una característica totalmente desconocida de la población a partir de los datos obtenidos sobre una muestra.

## **Por ejemplo:**

- Estimar el porcentaje de votantes a favor de cierto candidato.
- Estimar la duración promedio de un producto.
- Estimar la variabilidad de una característica de calidad en la fabricación de un artículo.
- Estimar la rentabilidad promedio de una acción

Dada una v.a.  $X$ , de la que suponemos conocido el modelo de distribución que sigue, buscamos información sobre los parámetros ( $\theta$ ) del mismo; **buscamos un valor concreto para  $\theta$ .**

# Estimación Puntual

- *Dada una v.a.  $X$ , de la que suponemos conocido el modelo de distribución que sigue, buscamos información sobre  $\theta$*
- *Sea  $\theta$  un parámetro desconocido.*
- *Un estimador puntual  $\hat{\theta}$  es una función que depende de la muestra.*
- *Una estimación puntual es el valor de la estadística de la muestra correspondiente.*

$$\hat{\theta} = \hat{\theta}(x_1, x_2, \dots, x_n)$$

## Estimadores puntuales

Ejemplo: Sea una muestra aleatoria simple,  $X_1, X_2, \dots, X_n$  de una población con  $(\mu, \sigma^2)$ .

- **Estimador de la media poblacional:** es la media muestral

$$\hat{\mu} = \bar{x} = \sum_{i=1}^n \frac{x_i}{n}$$

- **Estimador de la varianza poblacional:** es la varianza muestral

$$\hat{\sigma}^2 = S^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{(n-1)}$$

## Ejemplo

$X$ : número de llamadas diarias por el móvil que hace un joven.

Suponemos  $X \sim P(\lambda)$ .

Antes de obtener los datos (n)	Después de obtener los datos (n=39)
$X_1, \dots, X_n$ son v.a. independientes con distribución $P(\lambda)$ .	Resultado de la muestra: $\{1, 0, 3, 2, \dots, 2\} \in \mathbf{R}$
<b>Estadísticos:</b> Media muestral: $\bar{X} = \frac{X_1 + \dots + X_n}{n}$ Varianza muestral: $V = \frac{(X_1 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{n}$ Máximo: $\text{Máx}(X_1, \dots, X_n)$	Valor del estadístico:  $\bar{X} = 1.25641 \in \mathbf{R}$ $V = 1.47368 \in \mathbf{R}$ $\text{Máx}(X_i) = 5 \in \mathbf{R}$
<b>Estimadores:</b> $\lambda^* = \bar{X} ; \lambda' = V$	Estimación de $\lambda$ : $\lambda \sim \lambda^* = 1.25641 \in \mathbf{R}$ $\lambda \sim \lambda' = 1.47368 \in \mathbf{R}$

## Ejercicio

Sea  $X_1, X_2, \dots, X_n$  una m.a.s. de una población que se distribuye como una  $N(\alpha, \sigma)$  con parámetros desconocidos ¿Son estadísticos las siguientes opciones?

$$a) \sum_i x_i - \alpha \quad i = 1 \dots n$$

$$b) \sigma x_1 + \sigma x_2$$

$$c) x_1 \quad i = 1 \dots n$$

$$d) x_1^2 + x_2^2 - e^{x_3}$$

$$e) \frac{x_i}{\sigma} \quad i = 1 \dots n$$

$$f) \sum_i (x_i - \bar{x})^2$$

# Muestra Aleatoria

Sea  $X$  una v.a. con densidad  $f_{X,\theta}(x)$ .

Si  $X_1, X_2, \dots, X_n$  es una muestra aleatoria (independiente) de la variable  $X$ , la distribución conjunta esta dada por:

$$f_{X_1, X_2, \dots, X_n, \theta}(x_1, x_2, \dots, x_n) = f_{X_1, \theta}(x_1) \cdot f_{X_2, \theta}(x_2) \cdot \dots \cdot f_{X_n, \theta}(x_n)$$

**Se cumple:**

$$f_{X_i}(x_i) = f(x)$$

$$E(X_i) = E(X)$$

$$V(X_i) = V(X)$$



# Método de Estimación Puntual

## Método de Maxima Verosimilitud

Sea  $(x, f_{x,\theta})$ ,  $\theta \in \Theta$  un modelo estadístico. La función de densidad (o probabilidad) de  $x$ , considerada como función de  $\theta$  para cada  $x$  fijo y conocido, se denomina función de verosimilitud y se denota por:

$$\mathcal{L}(\theta, x), \theta \in \Theta, x \text{ fijo}$$

Intuitivamente, el estimador de  $\theta$  que denotaremos por  $\hat{\theta}_{EMV}$  es aquel valor de  $\theta$  que maximiza la probabilidad de generar las observaciones  $x_i$ . Así, el estimador de Maxima Verosimilitud (EMV) de  $\theta$  es tal que:

$$\mathcal{L}(\hat{\theta}_{EMV}, x) = \max_{\theta \in \Theta} \mathcal{L}(\theta, x)$$

Como es lo mismo maximizar  $\mathcal{L}(\theta, x)$  que su logaritmo, el valor de  $\theta_{EMV}$  puede calcularse derivando  $\mathcal{L}(\theta, x)$  con respecto a  $\theta$  e igualando a cero. Es decir:

$$\left. \frac{\partial \mathcal{L}(\theta, x)}{\partial \theta} \right|_{\theta = \hat{\theta}_{EMV}} = 0$$

o bien

$$\left. \frac{\partial \log \mathcal{L}(\theta, x)}{\partial \theta} \right|_{\theta = \hat{\theta}_{EMV}} = 0$$

## Ejercicio

Supongamos que los tiempos de fallos de ciertas componentes electrónicas,  $X$ , provienen de una distribución exponencial de parámetro  $\theta$ . Dada una muestra de  $n$  componentes, obtenga el E.M.V. de  $\theta$ .

## Solución

La función de densidad es:

$$f_{\theta}(x) = \theta e^{-\theta x} \text{ si } x > 0 \text{ con } \theta > 0$$

Y se dispone de los tiempos de fallo de  $n$  componentes elegidas al azar  $x_1, x_2, \dots, x_n$ .

La función de verosimilitud está dada por:

$$L(x_1, \dots, x_n, \theta) = \theta^n \exp\left\{-\theta \sum_{i=1}^n x_i\right\}.$$

Resolviendo la ecuación de verosimilitud

$$\frac{\partial}{\partial \theta} \ln L(x_1, \dots, x_n, \theta) = \frac{\partial}{\partial \theta} \left[ n \ln \theta - \theta \sum_{i=1}^n x_i \right] = \frac{n}{\theta} - \sum_{i=1}^n x_i = 0,$$

se concluye que el EMV para  $\theta$  por este procedimiento viene dado por

$$\hat{\theta} = \frac{1}{\overline{X}}.$$

## Ejercicio

Sean  $X_1, X_2, \dots, X_n$  v.a. iid con densidad:

$$f(x, \theta) = (\theta + 1)x^\theta \quad , 0 \leq x \leq 1$$

**Encuentre el EMV de  $\theta$ .**

## Solución

$$\begin{aligned}\mathcal{L}(x_i, \theta) &= \prod_{i=1}^n (\theta + 1) x_i^\theta \\ &= (\theta + 1)^n \prod_{i=1}^n x_i^\theta\end{aligned}$$

$$\ell(x_i, \theta) = n \log(\theta + 1) + \theta \sum_{i=1}^n \log x_i$$

$$\frac{\partial}{\partial \theta} \ell(x_i, \theta) = \frac{n}{\theta + 1} + \sum_{i=1}^n \log x_i$$

Iguando a zero

$$\longrightarrow \frac{n}{\theta + 1} + \sum_{i=1}^n \log x_i = 0$$

$$n + (\theta + 1) \sum_{i=1}^n \log x_i = 0$$

$$\theta \sum_{i=1}^n \log x_i = -n - \sum_{i=1}^n \log x_i$$

$$\hat{\theta}_{EMV} = - \frac{n + \sum_{i=1}^n \log x_i}{\sum_{i=1}^n \log x_i}$$



## Ejercicio

Sea  $X \rightarrow N(\mu, \sigma)$ , con  $\mu$  desconocido. Seleccionada una m.a.s.  $X_1, \dots, X_n$ , con realización  $x_1, \dots, x_n$ , **Encuentre el EMV de  $\theta$ .**

## Solución

Por el método de máxima verosimilitud:

$$\begin{aligned} L_{\mu}(x_1, \dots, x_n) &= \prod_{i=1}^n f_{\mu}(x_i) = \\ &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x_i - \mu)^2}{2\sigma^2}}, \end{aligned}$$

$$\ln L_{\mu}(x_1, \dots, x_n) = -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 - n \ln(\sqrt{2\pi}\sigma)$$

$$\frac{\partial}{\partial \mu} \ln L_{\mu}(x_1, \dots, x_n) = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) =$$

$$\frac{n\bar{x} - n\mu}{\sigma^2} = 0 \iff \hat{\mu} = \bar{x}$$

### ***Propiedades de los estimadores de máxima verosimilitud***

Si  $\hat{\theta}_n$  es EMV de  $\theta$  para una m.a.s. de tamaño  $n$ , entonces:

- Si  $g(x)$  es biyectiva,  $g(\hat{\theta}_n)$  es estimador de máxima verosimilitud de  $g(\theta)$ .
- Si  $n \rightarrow \infty$ ,  $E(\hat{\theta}_n) \rightarrow \theta$  (es asintóticamente centrado)
- Si  $n \rightarrow \infty$ ,  $V(\hat{\theta}_n) \rightarrow v_m$ , siendo  $v_m$  la mínima varianza posible  
(es asintóticamente eficiente)
- Si  $n \rightarrow \infty$ ,  $\frac{\hat{\theta}_n - \theta}{\sqrt{V(\hat{\theta}_n)}} \approx N(0,1)$  (es asintóticamente normal)

Es decir, asintóticamente tienen todas las propiedades “deseables” de los estimadores.

## **Realizada la estimación de un parámetro cabe preguntarse:**

- ¿Es exacta la estimación?
- ¿Es probable que la estimación sea alta o baja?
- ¿Con otra muestra se obtendría el mismo resultado, o bastante diferente?
- ¿La calidad de un procedimiento de estimación mejora bastante si la estadística de la muestra es menos variable e insesgada a la vez?

# Propiedades deseables de los estimadores

## Insesgamiento

Se dice que un estimador es insesgado (o centrado) si la esperanza del estimador coincide con el parámetro a estimar. En caso contrario se dice que es sesgado.

$$E(\hat{\theta}) = \theta$$

### Ejemplos:

- La **media muestral**  $\bar{X}$  es un estimador centrado de  $E(X)$

Por tanto,

- Si  $X \sim P(\lambda)$ ,  $\lambda^* = \bar{X}$  es un estimador centrado para  $\lambda$ .
- Si  $X \sim N(\mu, \sigma)$ ,  $\mu^* = \bar{X}$  es un estimador centrado para  $\mu$ .

- La **varianza muestral**  $\mathbf{V}$  no es un estimador centrado para  $V(X) = \sigma^2$ .

$$\text{De hecho, } E(\mathbf{V}) = \frac{n-1}{n} \sigma^2$$

Por tanto,

- Si  $X \sim P(\lambda)$ ,  $\lambda' = \mathbf{V}$  no es un estimador centrado para  $\lambda$ .

## Consistencia

Se dice que un estimador es consistente si se aproxima cada vez más al verdadero valor del parámetro a medida que se aumenta el tamaño muestral.

$$\Pr[(\hat{\theta} - \theta) > \varepsilon] \rightarrow 0 \quad n \rightarrow \infty, \varepsilon > 0$$

La distribución del estimador se concentra más alrededor del verdadero parámetro cuando el tamaño muestral aumenta.

## Eficiencia

Es claro que un estimador será tanto mejor cuanto menor sea su varianza, ya que se concentra más alrededor del verdadero valor del parámetro. Se dice que un estimador insesgado es eficiente si tiene varianza mínima.

### Ejemplo

- La varianza muestral  $V$  es más eficiente que la cuasivarianza muestral  $S^2$ :

$$\text{Como } V = \left(\frac{n-1}{n}\right) S^2 \Rightarrow V(V) = \left(\frac{n-1}{n}\right)^2 V(S^2) , \text{ por tanto } V(V) < V(S^2)$$



# Error Cuadrático Medio

- La distancia entre el estimador y el parámetro a estimar puede medirse mediante lo que se denomina el error cuadrático medio, que se define como el valor esperado del cuadrado de la diferencia entre el estimador y el verdadero parámetro.

$$ECM(\hat{\theta}) = E(\hat{\theta} - \theta)^2$$

El ECM es importante ya que puede escribirse como

$$ECM(\hat{\theta}) = VAR(\hat{\theta}) + [\theta - E(\hat{\theta})]^2$$

una es la varianza del estimador y otra el cuadrado del sesgo.

**Ejemplo:** Se puede demostrar que  $ECM(\mathbf{V}) > ECM(\mathbf{S}^2)$  para  $n > 2$ .

## Ejemplo

Sean  $Y_1, Y_2, \dots, Y_n$  variables aleatorias iid  $N(\mu, \sigma^2)$  y  $\hat{\mu}_1$  es el estimador de  $\mu$

$$\hat{\mu}_1 = \frac{2}{n(n+1)} \sum_{i=1}^n iY_i$$

¿Es este estimador insesgado?.

# Solución

Para probar insesgamiento:  $E(\hat{\mu}_1) = \mu$

$$\begin{aligned} E(\hat{\mu}_1) &= E\left(\frac{2}{n(n+1)} \sum_{i=1}^n iY_i\right) \\ &= \frac{2}{n(n+1)} \cdot E\left(\sum_{i=1}^n iY_i\right) \\ &= \frac{2}{n(n+1)} \cdot E(1 \cdot Y_1 + 2 \cdot Y_2 + \dots + n \cdot Y_n) \\ &= \frac{2}{n(n+1)} \left(E(Y_1) + 2E(Y_2) + \dots + nE(Y_n)\right) \\ &= \frac{2}{n(n+1)} \left(\mu + 2\mu + \dots + n\mu\right) \\ &= \frac{2}{n(n+1)} \left(1 + 2 + \dots + n\right) \cdot \mu \\ &= \frac{2}{n(n+1)} \cdot \mu \cdot \sum_{i=1}^n i \\ &= \frac{2}{n(n+1)} \cdot \mu \cdot \frac{n(n+1)}{2} \\ &= \mu \end{aligned}$$

## Ejercicio

Sean  $Y_1, Y_2, \dots, Y_n$  variables aleatorias iid  $N(\mu, \sigma^2)$  y  $S^2$  es el estimador de  $\sigma^2$

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

¿Es insesgado?

## Solución

$$\mathbf{E} [\mathcal{S}^2] = \mathbf{E} \left[ \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \right] = \mathbf{E} \left[ \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2 \right] = \frac{1}{n} \sum_{i=1}^n \mathbf{E} [X_i^2] - \mathbf{E} [\bar{X}^2]$$

$$\left\{ \begin{array}{l} \overbrace{\mathbf{Var} [X_i]}^{\sigma^2} = \mathbf{E} [X_i^2] - \overbrace{\mathbf{E} [X_i]^2}^{\mu^2} \Rightarrow \mathbf{E} [X_i^2] = \sigma^2 + \mu^2 \\ \underbrace{\mathbf{Var} [\bar{X}]}_{\frac{\sigma^2}{n}} = \mathbf{E} [\bar{X}^2] - \underbrace{\mathbf{E} [\bar{X}]^2}_{\mu^2} \Rightarrow \mathbf{E} [\bar{X}^2] = \frac{\sigma^2}{n} + \mu^2 \end{array} \right.$$

$$\mathbf{E} [\mathcal{S}^2] = \frac{1}{n} \sum_{i=1}^n (\sigma^2 + \mu^2) - \left( \frac{\sigma^2}{n} + \mu^2 \right) = \sigma^2 + \mu^2 - \frac{\sigma^2}{n} - \mu^2 = \frac{n-1}{n} \sigma^2.$$

## Ejercicio

Sean  $Y_1, Y_2, \dots, Y_n$  variables aleatorias iid  $N(\mu, \sigma^2)$  y  $S^2$  es el estimador de  $\sigma^2$

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

si

$$\frac{nS^2}{\sigma^2} = \sum_{i=1}^n \left( \frac{X_i - \bar{X}}{\sigma} \right)^2 \rightsquigarrow \chi^2_{n-1}$$

Obtener el valor esperado y la varianza de  $S^2$

## Solución

como

$$\frac{nS^2}{\sigma^2} = \sum_{i=1}^n \left( \frac{X_i - \bar{X}}{\sigma} \right)^2 \rightsquigarrow \chi^2_{n-1}$$

tenemos

$$\mathbf{E} \left[ \frac{nS^2}{\sigma^2} \right] = n - 1 \quad \Longleftrightarrow \quad \mathbf{E} [S^2] = \frac{n - 1}{n} \sigma^2$$

$$\mathbf{Var} \left[ \frac{nS^2}{\sigma^2} \right] = 2(n - 1) \quad \Longleftrightarrow \quad \mathbf{Var} [S^2] = \frac{2(n - 1)}{n^2} \sigma^4$$

## Ejercicio

La v. a.  $X$  sigue distribución  $U(0, \theta)$ , donde  $\theta$  es un valor positivo y desconocido. Se extrae una m. a. s. de tamaño  $n$  ( $n > 2$ ). ¿Dado los estimadores siguientes de  $\theta$ , ¿Cuáles de los siguientes errores cuadráticos medios son correctos?

$$T_1 = \bar{X} = \frac{X_1 + \dots + X_n}{n} \quad T_2 = X_1 \quad T_3 = 2\bar{X} = 2\frac{X_1 + \dots + X_n}{n}.$$

$$ECM(T_1) = \frac{\theta^2}{12n} \quad ECM(T_2) = \frac{\theta^2}{12} \quad ECM(T_3) = \frac{4\theta^2}{12n}$$