

ESTIMACIÓN

TEMA 5: Estimación puntual I. Propiedades de los estimadores

TEMA 6: Estimación puntual II. Métodos de estimación puntual

TEMA 7: Estimación por intervalos

CONTRASTES DE HIPÓTESIS

TEMA 8: Contrastes paramétricos

TEMA 9: Contrastes no paramétricos

TEMA 5:

ESTIMACIÓN PUNTUAL I.

PROPIEDADES DE LOS ESTIMADORES

5.1. Concepto de estimador

5.2. Estimadores insesgados

5.3. Comparación de estimadores. Error cuadrático medio

5.4. Estimador insesgado de varianza mínima. Eficiencia

5.5. Estimadores consistentes

5.6. Estimadores suficientes

OBJETIVOS:

- proponer distintos estimadores para un mismo parámetro;
- evaluar las propiedades de un estimador e interpretarlas en términos estadísticos;
- comparar dos estimadores en términos de sesgo y varianza.

1. CONCEPTO DE ESTIMADOR

Modelo: $X \rightarrow f(x; \vartheta)$,

ϑ = parámetro(s) desconocido(s), $\vartheta \in \Theta$ = espacio paramétrico

Problema: Estimar ϑ a partir de los datos: (X_1, \dots, X_n) m.a.s. (i.i.d.) de X

Estimación puntual: “proponer” **un valor** plausible para el parámetro ϑ ¿cómo?

Estimador: función de la muestra (**v.a.**) que toma valores en el espacio paramétrico Θ . Se denota por $\hat{\vartheta} = \hat{\vartheta}(X_1, \dots, X_n)$

Ejemplo 1: (X_1, \dots, X_n) m.a.s. de $b(p) \Rightarrow \hat{p} = (X_1 + X_n)/2$; $\hat{p} = (X_1 + \dots + X_n)/n \Rightarrow$ SI

$\hat{p} = 2X_1$, $\hat{p} = X_1 + X_2 \Rightarrow$ NO

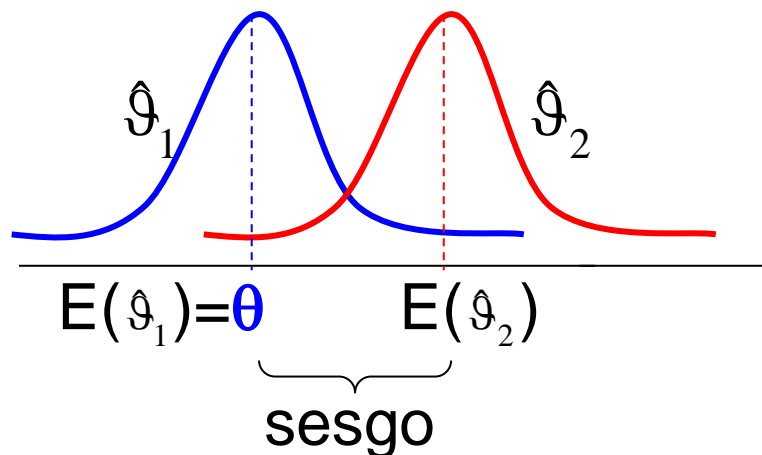
Estimación: valor numérico del estimador para una muestra concreta (**nº**)

2. ESTIMADORES INSESGADOS

La distribución de $\hat{\vartheta}$ está centrada en el parámetro que se estima

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

- Estimador **sesgado**: no incesgado $\Rightarrow E(\hat{\vartheta}) \neq \theta \Rightarrow \text{sesgo}(\hat{\vartheta}) = E(\hat{\vartheta}) - \theta$



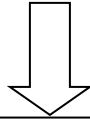
$\hat{\vartheta}_1$ es incesgado

$\hat{\vartheta}_2$ es sesgado

- Estimador **asintóticamente incesgado**: $\lim_{n \rightarrow \infty} E(\hat{\vartheta}) = \theta$.

Ejemplo 1: Sea (X_1, \dots, X_n) m.a.s. de X con $E(X) = \mu$

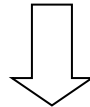
$$\bullet E(\bar{X}) = E\left(\frac{X_1 + \dots + X_n}{n}\right) = \frac{1}{n} [E(X_1) + \dots + E(X_n)] = \frac{n E(X)}{n} = E(X) = \mu$$



La **media muestral** es **siempre** estimador **insesgado** de la **media poblacional**

Ejemplo 2: Sea (X_1, \dots, X_n) m.a.s. de X con $E(X) = \mu$, $\text{Var}(X) = \sigma^2$

$$\bullet E(S^2) = E\left(\frac{\sum_{i=1}^n X_i^2}{n}\right) - E(\bar{X}^2) = \frac{1}{n} \sum_{i=1}^n E(X_i^2) - E(\bar{X}^2) = \frac{1}{n} n(\sigma^2 + \mu^2) - \left(\frac{\sigma^2}{n} + \mu^2\right) = \sigma^2 - \frac{1}{n}\sigma^2 = \frac{n-1}{n}\sigma^2 \neq \sigma^2$$

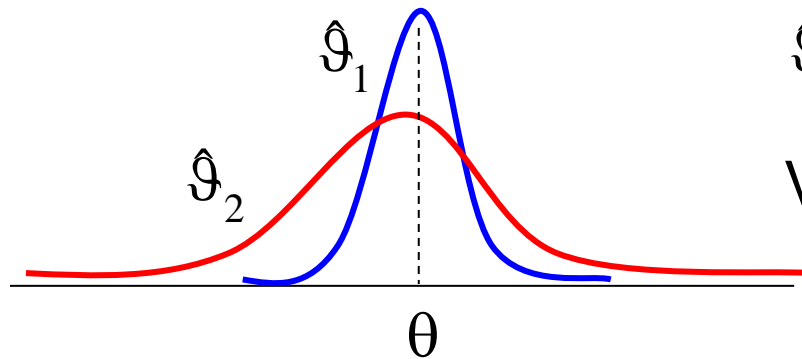


La **varianza muestral** es **siempre** estimador **sesgado** de la **varianza poblacional**

Ejercicio: construir un estimador insesgado a partir de $S^2 \Rightarrow S_c^2 = \frac{n}{n-1} S^2 \Rightarrow E(S_c^2) = \sigma^2$

3. COMPARACIÓN DE ESTIMADORES. ERROR CUADRÁTICO MEDIO

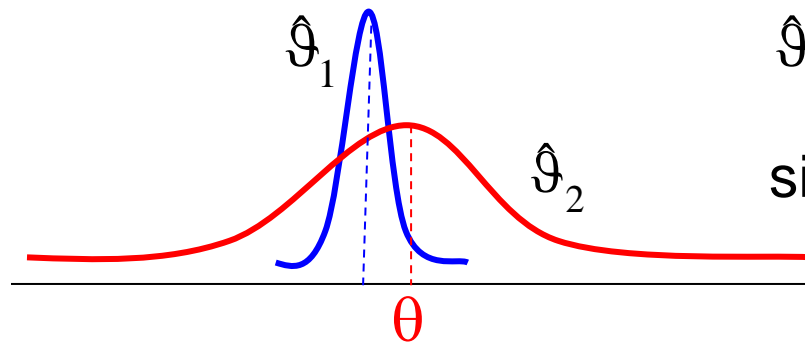
Entre dos estimadores insesgados, ¿cuál es preferible? El que tenga menos varianza porque sus valores estarán más próximos al verdadero θ .



$\hat{\vartheta}_1$ y $\hat{\vartheta}_2$ insesgados

$$\text{Var}(\hat{\vartheta}_1) < \text{Var}(\hat{\vartheta}_2)$$

Entre dos estimadores no necesariamente insesgados, ¿cómo elegir?



$\hat{\vartheta}_1$ sesgado; $\hat{\vartheta}_2$ insesgado

sin embargo $\text{Var}(\hat{\vartheta}_1) < \text{Var}(\hat{\vartheta}_2)$

$$\text{E.C.M.}(\hat{\vartheta}) = E(\hat{\vartheta} - \theta)^2 = \text{Var}(\hat{\vartheta}) + \text{sesgo}^2$$

Entre dos estimadores de un mismo parámetro, es preferible aquel estimador que tenga menor error cuadrático medio:

$\hat{\vartheta}_1$ más eficiente que $\hat{\vartheta}_2$ si $\text{ECM}(\hat{\vartheta}_1) < \text{ECM}(\hat{\vartheta}_2)$

Obviamente, si $\hat{\vartheta}$ es insesgado $\Rightarrow \text{E.C.M.}(\hat{\vartheta}) = \text{Var}(\hat{\vartheta})$.

Si $\hat{\vartheta}_1$ y $\hat{\vartheta}_2$ son insesgados,

$\hat{\vartheta}_1$ es más eficiente que $\hat{\vartheta}_2$ si $\text{Var}(\hat{\vartheta}_1) < \text{Var}(\hat{\vartheta}_2)$

Ejemplo: Sea (X_1, \dots, X_n) m.a.s. de $X \rightarrow N(\mu, \sigma)$. ¿ Estimar σ^2 ?

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

$\Rightarrow S^2$ es más eficiente que s_c^2

- Distribución: $\frac{n}{\sigma^2} S^2 = \frac{n-1}{\sigma^2} s_c^2 \rightarrow \chi_{n-1}^2$
- $E(S^2) = \frac{n-1}{n} \sigma^2 \neq \sigma^2 \Rightarrow$ ¡ **sesgado** ! \Rightarrow sesgo = $\frac{n-1}{n} \sigma^2 - \sigma^2 = -\frac{1}{n} \sigma^2$
- $\text{Var}(\frac{n}{\sigma^2} S^2) = \text{Var}(\chi_{n-1}^2) = 2(n-1) \Rightarrow \text{Var}(S^2) = \frac{\sigma^4}{n^2} \text{Var}(\frac{n}{\sigma^2} S^2) = \frac{\sigma^4}{n^2} 2(n-1)$
 $\Rightarrow \text{ECM}(S^2) = \text{Var}(S^2) + \text{sesgo}^2 = \frac{\sigma^4}{n^2} (2n-1)$
- $E(s_c^2) = E(\frac{n}{n-1} S^2) = \frac{n}{n-1} \frac{n-1}{n} \sigma^2 = \sigma^2 \Rightarrow$ ¡ **insesgado** ! \Rightarrow sesgo = 0
- $\text{Var}(\frac{n-1}{\sigma^2} s_c^2) = \text{Var}(\chi_{n-1}^2) = 2(n-1) \Rightarrow \text{Var}(s_c^2) = \text{ECM}(s_c^2) = 2 \frac{\sigma^4}{(n-1)}$

4. ESTIMADOR INSESGADO DE VARIANZA MÍNIMA. EFICIENCIA

Estimador **insesgado de varianza mínima** = estimador insesgado con menor varianza dentro de la clase de los insesgados

⇒ si existe, es único ⇒ ¿cómo encontrarlo?

Teorema: Cota de Cramer-Rao

(X_1, \dots, X_n) m.a.s. de una variable aleatoria X cuya distribución verifica ciertas *condiciones de regularidad*.¹ Sea $\hat{\theta}$ un estimador **insesgado** de θ :

$$\text{Var}(\hat{\theta}) \geq \frac{1}{E \left[\left(\frac{\partial \ln f(X_1, \dots, X_n; \theta)}{\partial \theta} \right)^2 \right]}$$

- Proporciona la menor varianza posible de un estimador insesgado

Definición: *condiciones de regularidad* y $\hat{\theta}$ **insesgado**. $\hat{\theta}$ es **eficiente** si:

$$\text{Var}(\hat{\theta}) = \text{cota C-R}$$

- Si $\hat{\theta}$ es el estimador eficiente ⇒ $\hat{\theta}$ es el insesgado de mínima varianza

¹ El campo de variación de X no depende del parámetro a estimar

Teorema: bajo *condiciones de regularidad* y $\hat{\theta}$ **insesgado**

$\hat{\theta}$ estimador **eficiente** de $\theta \Leftrightarrow$

$$\frac{\partial \ln f(X_1, \dots, X_n; \theta)}{\partial \theta} = K(\theta)(\hat{\theta} - \theta)$$

Además: donde $K(\theta) = \frac{1}{\text{Var}(\hat{\theta})}$

Ejemplo:

(X_1, \dots, X_n) muestra aleatoria simple de una v.a. Bernoulli $b(p)$

$$f(x;p) = p(X=x) = p^x(1-p)^{1-x}$$

▪ $f(x_1, \dots, x_n; p) = f(x_1; p) \dots f(x_n; p) = p(X=x_1) \dots p(X=x_n) = p^{\sum x_i} (1-p)^{n-\sum x_i}$

▪ $\ln(f(x_1, \dots, x_n; p)) = \sum x_i \ln(p) + (n - \sum x_i) \ln(1-p)$

▪ $\frac{\partial \ln f(x_1, \dots, x_n; p)}{\partial p} = \frac{\sum x_i}{p} - \frac{n - \sum x_i}{(1-p)} = n \left[\frac{\bar{X}}{p} - \frac{1 - \bar{X}}{(1-p)} \right] = \frac{n}{p(1-p)} (\bar{X} - p) \Rightarrow \bar{X} \text{ eficiente de } p$

$K(p) = 1 / \text{Var}(\bar{X}) \Rightarrow \text{Var}(\bar{X}) = p(1-p)/n$

5. ESTIMADORES CONSISTENTES

Un estimador se define para un cierto tamaño muestral n

\Rightarrow Un estimador, calculado para distintos tamaños muestrales, puede verse como una sucesión de variables aleatorias $\{\hat{\vartheta}_n\}_n = \{\hat{\vartheta}_1, \hat{\vartheta}_2, \dots, \hat{\vartheta}_n, \dots\}$. \Rightarrow ¿Qué pasa si $n \rightarrow \infty$?

Ejemplo:

(X_1, \dots, X_n) muestra aleatoria simple de una Bernoulli $b(p) \Rightarrow \hat{p} = \bar{X}$

$$\hat{p}_1 = \bar{X}_1 = X_1; \hat{p}_2 = \bar{X}_2 = \frac{X_1 + X_2}{2}; \hat{p}_3 = \bar{X}_3 = \frac{X_1 + X_2 + X_3}{3}; \dots; \hat{p}_n = \bar{X}_n = \frac{X_1 + \dots + X_n}{n}; \dots$$

Tiramos n veces una moneda y observamos la proporción de caras:

Nº ensayos (n)	Nº caras ($\sum X_i$)	Frecuencia = $\frac{\text{Nº caras}}{\text{Nº ensayos}} = \bar{X}_n$
1000	501	0.501
2000	986	0.493
3000	1495	0.49833
4000	2031	0.50775
....
7000	3504	0.50057
8000	4001	0.500125

Notación:

$$\bar{X}_n \xrightarrow{c.p.} 0.5 \Leftrightarrow p \lim_{n \rightarrow \infty} \bar{X}_n = 0.5 \Leftrightarrow \lim_{n \rightarrow \infty} p(|\bar{X}_n - 0.5| \geq \varepsilon) = 0$$

DEFINICION:

$\hat{\vartheta}_n$ es un estimador **consistente** de θ si $\boxed{\hat{\vartheta}_n \xrightarrow{\text{c.p.}} \theta}$:

$$\text{si } \forall \varepsilon > 0, \lim_{n \rightarrow \infty} p(|\hat{\vartheta}_n - \theta| < \varepsilon) = 1 \Leftrightarrow \lim_{n \rightarrow \infty} p(|\hat{\vartheta}_n - \theta| \geq \varepsilon) = 0$$

- A partir de un determinado tamaño muestral, suficientemente grande, es muy probable que el valor de $\hat{\vartheta}_n$ difiera muy poco de θ (menos de ε).

• Notación: $\boxed{p \lim_{n \rightarrow \infty} \hat{\vartheta}_n = \theta}$

¿CÓMO PROBAR LA CONSISTENCIA DE UN ESTIMADOR?

Proposición:

Condiciones suficientes para que $\hat{\theta}_n$ sea consistente:

- **Insesgado** ($E\hat{\theta}_n = \theta$) o **asintóticamente insesgado** ($\lim_{n \rightarrow \infty} E\hat{\theta}_n = \theta$)
- Al aumentar n , su **varianza tiende a cero** ($\lim_{n \rightarrow \infty} \text{Var}\hat{\theta}_n = 0$)

⇒ Entonces, $\hat{\theta}_n$ es consistente.

■ **Ejemplo:** (X_1, \dots, X_n) m.a.s. de X con $E(X) = \mu$, $\text{Var}(X) = \sigma^2 < \infty$

➤ $E(\bar{X}) = E(X) = \mu \Rightarrow$ estimador insesgado de μ

➤ $\text{Var}(\bar{X}) = \frac{\text{Var}(X)}{n} \xrightarrow{n \rightarrow \infty} 0$

La **media muestral** es estimador **consistente** de la **media poblacional**

Proposición: Ley débil de los grandes números

En general, los momentos muestrales a_k convergen en probabilidad a los correspondientes momentos poblacionales α_k .

- Ejemplo: $p \lim_{n \rightarrow \infty} \bar{X}_n = E(X) = \mu$

La **media muestral** es estimador **consistente** de la **media poblacional**

Proposición: Teorema de Slutsky

$$p \lim_{n \rightarrow \infty} g(\hat{\theta}_n) = g(p \lim_{n \rightarrow \infty} \hat{\theta}_n), \text{ si } g \text{ es continua}$$

Operaciones básicas con “plim” igual que con límites de n^{os} (“lim”)

- Ejemplo: Sea (X_1, \dots, X_n) m.a.s. de X con $E(X) = \mu$, $\text{Var}(X) = \sigma^2 < \infty$

$$p \lim_{n \rightarrow \infty} S^2 = p \lim_{n \rightarrow \infty} \left(\frac{\sum_{i=1}^n X_i^2}{n} - \bar{X}^2 \right) = p \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n X_i^2}{n} - \left(p \lim_{n \rightarrow \infty} \bar{X} \right)^2 = E(X^2) - (EX)^2 = \sigma^2$$

La **varianza muestral** es estimador **consistente** de la **varianza poblacional**

6. ESTIMADORES SUFICIENTES

Recoge **toda la información** contenida en la muestra acerca de ϑ .

Formalmente:

$T=T(X_1, \dots, X_n)$ es **estadístico suficiente** para el parámetro ϑ si la distribución condicionada de la muestra, (X_1, \dots, X_n) , dado el valor de T , no depende de ϑ :

$$p(X_1=x_1, \dots, X_n=x_n / T=t) \text{ no depende de } \vartheta$$

\Rightarrow conocido T , la muestra (X_1, \dots, X_n) ya no tiene nada que decir sobre ϑ

Ejemplo: (X_1, X_2) m.a.s. de Bernoulli $\Rightarrow T=X_1+X_2$ es suficiente

Teorema de factorización de Neyman:

Sea (X_1, \dots, X_n) m.a.s. de una variable aleatoria cuya distribución depende de un parámetro ϑ . Sea $T=T(X_1, \dots, X_n)$ un estadístico.

$$\mathbf{T} \text{ es suficiente} \Leftrightarrow p(X_1=x_1, \dots, X_n=x_n) = \underbrace{g(\mathbf{T}(x_1, \dots, x_n); \vartheta)}_{\text{depende de } \vartheta \text{ y de la muestra sólo a través del estadístico } \mathbf{T}} \underbrace{h(x_1, \dots, x_n)}_{\text{no depende de } \vartheta}$$

depende de ϑ y de la muestra sólo a través del estadístico \mathbf{T}

no depende de ϑ

Ejemplo: (X_1, \dots, X_n) m.a.s. de Bernoulli $\Rightarrow T=\sum X_i$ es suficiente

- Un **estimador suficiente** es un estimador que como estadístico es un estadístico suficiente.

Propiedad:

Una función 1:1 de un estadístico suficiente es un estadístico suficiente.

Ejemplo

⇒ Si la suma, $T = \sum X_i$, es suficiente,

⇒ También lo es la media muestral: $\bar{X} = \varphi(T) = T/n = \sum X_i/n$