Inferencia Estadística

La inferencia estadística comprende:
- Estimación de parámetros: Estimación puntual
  Estimación por intervalos
- Prueba de hipótesis.

Estimación de Parámetros.

Estimación Puntual.

Sea X una v.a con f.d f(x;θ), donde θ denota al parámetro desconocido de la población. Sea $X_1,\ldots,X_n$ una m.a extraída de esta población. Un estimador puntual del parámetro θ es cualquier función de las variables aleatorias $X_1,\ldots,X_n$ y se escribe: $\hat{\theta} = T = t(X_1,\ldots,X_n)$.

Sea $X_1,\ldots,X_n$ una m.a de f(x;θ). Sea $T = t(X_1,\ldots,X_n) = t$ un estadígrafo. Si T es usado para estimar a q(θ) , $T(X_1,\ldots,X_n) = t$ es un estimador de q(θ).

Espacio Paramétrico Θ

Sea $\Theta = (\theta_1,\ldots,\theta_k)$ un vector de k parámetros entonces al conjunto de valores posibles que puede tomar Θ se le llama espacio paramétrico.

Ejemplo 1:

1) Si X ~ Pois (λ) $\Rightarrow$ f (X; λ) es conocida además $\Theta = \{\lambda / \lambda > 0\}$
2) Si X ~ N (u, σ^2) $\Rightarrow$ f(x; u, σ^2) es conocida además $\Theta = \{ (u, \sigma^2) / -\infty < u < \infty, \sigma^2 > 0\}$.

Propiedades de los Buenos Estimadores:

Insesgabilidad – Eficiencia – Consistencia – Suficiencia.

Insesgabilidad

$\hat{\theta} = T = t(X_1,\ldots,X_n)$ es un estimador insesgado para q(θ) si:
$E(\hat{\theta}) = E(T) = q(\theta)$, $\theta \in \Theta$

Ejemplo 2: Ya se demostró que $E\left(\overline{X}\right) = \mu$ entonces $\overline{X}$ es un estimador insesgado de μ. También se demostró que $E\left(S^2\right) = \sigma^2$ entonces S^2 es un estimador insesgado de σ^2.

Ejemplo 3: Sea $X_1,\ldots,X_n$ una muestra aleatoria de tamaño n. Demuestre que $\overline{X}^2$ es un estimador sesgado de μ^2. Halle el estimador insesgado.

$$E\left(\overline{X}^2\right)=Var\left(\overline{X}\right)+\left[E\left(\overline{X}\right)\right]^2=\frac{\sigma^2}{n}+\mu^2 \rightarrow \text{Sesgo}=E\left(\overline{X}^2\right)-\mu^2=\frac{\sigma^2}{n}$$

El estimador insesgado es $T=\overline{X}^2-\frac{\sigma^2}{n} \rightarrow E\left[\overline{X}^2-\frac{\sigma^2}{n}\right]=E\left(\overline{X}^2\right)-\frac{\sigma^2}{n}=\frac{\sigma^2}{n}+\mu^2-\frac{\sigma^2}{n}=\mu^2$.

Ejemplo 4: Considere ciertos alumnos universitarios que han leído "El Aleph" de Jorge Luis Borges, y que el error respecto a la afirmación de que esos estudiantes demoran en promedio θ semanas en leer esa obra, tiene media cero y variancia conocida σ^2. Si $X_1,\ldots,X_n$ es una muestra aleatoria grande de la población que leyó "El Aleph" halle un estimador insesgado del tiempo promedio, en semanas, que les tomó a esos estudiantes para leer "La Divina Comedia" de Dante Alighieri que es $0.8\times\theta^2$. Luego suponga que en una muestra aleatoria de 100 de esos estudiantes se contabilizó un tiempo total de 350.4 semanas para leer "El Aleph" y considere $\sigma^2=1.2$ para obtener el estimador puntual del tiempo promedio que demoraron en leer la obra de Dante.
Solución

Sea e_i el error del estudiante i respecto a la afirmación de que esos estudiantes demoran en promedio θ semanas en leer "El Aleph", entonces $e_i \sim \text{Desconocida}\left(\text{Media}=0,\text{Variancia}=\sigma^2\right)$. Se puede considerar lo siguiente:

$X_i=\theta+e_i \rightarrow X_i \sim \text{Desconocida}\left(\text{Media}=\theta,\text{Variancia}=\sigma^2\right)$.

Como la muestra es grande (100) y aplicando el teorema del límite central se puede afirmar que $\bar{X}=\frac{\sum_{i=1}^{n}X_i}{n} \sim \text{Normal}\left(\text{Media}=\theta,\text{Variancia}=\frac{\sigma^2}{n}\right)$. Un candidato para estimar en forma insesgada a $0.8\times\theta^2$ es $0.8\times\bar{X}^2$, pero:

$E\left(0.8\times\bar{X}^2\right)=0.8E\left(\bar{X}^2\right)=0.8\left(\frac{\sigma^2}{n}+\theta^2\right)=0.8\theta^2+\left(0.8\frac{\sigma^2}{n}\right)$. Se aprecia que $0.8\times\bar{X}^2$ es sesgado entonces el estimador insesgado de $0.8\times\theta^2$ será $0.8\times\bar{X}^2-\left(0.8\frac{\sigma^2}{n}\right)$. El valor del estimador puntual será el siguiente:

$$0.8\times3.504^2-0.8\times\frac{1.2}{100}=9.8128128 \text{ semanas}.$$

Ejemplo 5: Sean T_1 y T_2 dos estadísticas independientes e insesgadas de θ. Si la variancia de T_1 es el doble de la de T_2. Determine los valores de las constantes

k_1 y k_2 tales que la estadística $S = k_1T_1 + k_2T_2$ sea insesgada de variancia mínima para tal combinación lineal.

$$E[S] = k_1 \underbrace{E(T_1)}_{\theta} + k_2 \underbrace{E(T_2)}_{\theta} = k_1\theta + k_2\theta = \theta \rightarrow k_1 + k_2 = 1 \rightarrow k_1 = 1 - k_2 \quad (1)$$

$$Var(S) = k_1^2 Var(T_1) + k_2^2 Var(T_2) = k_1^2\left[2Var(T_2)\right] + k_2^2 Var(T_2) \quad (2)$$

(1) en (2)

$$Var(S) = 2(1-k_2)^2\left[Var(T_2)\right] + k_2^2 Var(T_2)$$

Hallemos k_2 que minimice la $Var(S)$

$$\frac{d\left[Var(S)\right]}{dk_2} = -4(1-k_2)Var(T_2) + 2k_2Var(T_2) = 0 \rightarrow k_2 = \frac{2}{3} \text{ y } k_1 = \frac{1}{3}$$

$$\frac{d^2\left[Var(S)\right]}{dk_2^2} = 6Var(T_2) > 0$$

Ejemplo 6: Sea $X_1,\ldots,X_n$ una m.a de una distribución $\text{Uniforme}[0,\theta]$, con $Y_1,\ldots,Y_n$ las correspondientes estadísticas de orden. Demuestre que $T_1 = 2\overline{X}$ y $T_2 = \left(\frac{n+1}{n}\right)Y_n$ son estimadores insesgados de θ.

$$f(x) = \frac{1}{\theta} I_{(0,\theta)}(x)$$

$$\mu = E(X_i) = \frac{\theta}{2}, \quad Var(X_i) = \sigma^2 = \frac{\theta^2}{12}$$

$$E(T_1) = E\left(2\overline{X}\right) = 2E\left(\overline{X}\right) = 2\mu = 2\frac{\theta}{2} = \theta \rightarrow T_1 \text{ es insesgado.}$$

De otro lado:

$$E(Y_n) = \int_0^\theta y g(y)\,dy = \int_0^\theta y\left(n\frac{y^{n-1}}{\theta^n}\right)dy = \frac{n\theta}{n+1}$$

$$E(T_2) = E\left(\frac{n+1}{n}Y_n\right) = \frac{n+1}{n}E(Y_n) = \frac{n+1}{n}\left[\frac{n\theta}{n+1}\right] = \theta \rightarrow T_2 \text{ es insesgado.}$$

Ejemplo 7: Sea $X_1,\ldots,X_n$ una muestra aleatoria de tamaño n de una población con media μ y variancia σ^2. Si $S_n = \sum_{i=1}^n a_iX_i$.

a. Demuestre que S_n es un estimador insesgado para μ si $\sum_{i=1}^n a_i = 1$.

$$E(S_n) = E\left(\sum_{i=1}^n a_iX_i\right) = \sum_{i=1}^n a_iE(X_i) = \sum_{i=1}^n a_i\mu = \mu\sum_{i=1}^n a_i = \mu, \text{ sólo si } \sum_{i=1}^n a_i = 1$$

b. Considere la clase de todos los estimadores insesgados de μ, de la forma $S_n = \sum_{i=1}^{n} a_i X_i$, donde $a_i \in R$ para $i = 1, 2, \ldots, n$ tal que $\sum_{i=1}^{n} a_i = 1$. Demuestre que el estimador con la mínima variancia de esta clase de estimadores está dada para $a_i = \frac{1}{n}$ para $i = 1, 2, \ldots, n$.

$$Var(S_n) = Var\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 Var(X_i) = \sum_{i=1}^{n} a_i^2 \sigma^2 = \sigma^2 \sum_{i=1}^{n} a_i^2$$

Se tienen que hallar los parámetros a_i a fin de que la $Var(S_n)$ sea mínima con la condición $\sum_{i=1}^{n} a_i = 1$. Aplicando los multiplicadores de Lagrange a la función:

$$L(\gamma) = \sum_{i=1}^{n} a_i^2 - \gamma\left(\sum_{i=1}^{n} a_i - 1\right) = a_1^2 + a_2^2 + \ldots + a_n^2 - \gamma(a_1 + a_2 + \ldots + a_n - 1)$$

$\frac{dL(\gamma)}{a_1} = 2a_1 - \gamma, \ldots, \frac{dL(\gamma)}{a_n} = 2a_n - \gamma$. En general:

$\frac{dL(\gamma)}{a_i} = 2a_i - \gamma = 0 \rightarrow a_i = \frac{\gamma}{2}$. Con la restricción:

$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} \frac{\gamma}{2} = \frac{n\gamma}{2} = 1 \rightarrow \gamma = \frac{2}{n}$. Como:

$$\frac{dL(\gamma)}{a_i} = 2a_i - \gamma = 0 \rightarrow 2a_i - \frac{2}{n} = 0 \rightarrow a_i = \frac{1}{n}$$

$\therefore$ El estimador insesgado de mínima variancia de μ es $S_n = \sum_{i=1}^{n} a_i X_i = \frac{\sum_{i=1}^{n} X_i}{n} = \overline{X}$

Ejemplo 8: Sea $X_1, \ldots, X_n$ una muestra aleatoria de tamaño n de una población Binomial$(1, p)$. Halle un estimador insesgado de p^r (no use el estimador trivial $S = X_1 X_2 \ldots X_r \rightarrow E(S) = \prod_{i=1}^{r} E(X_1) = \prod_{i=1}^{r} p = p^r$).

Se utilizará el método inductivo

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{T}{n} \rightarrow T = \sum_{i=1}^{n} X_i \sim \text{Binomial}(n, p)$$

$$E(T) = \sum_{t=0}^{n} t.\binom{n}{t} p^t q^{n-t} = \sum_{t=1}^{n} t.\binom{n}{t}\left(\frac{p}{q}\right)^t q^n = \sum_{t=1}^{n} t.\binom{n}{t}\left(\frac{p}{q}\right)^t q^n = \sum_{t=1}^{n} t.\binom{n}{t}(\alpha)^t q^n =$$

$$E(T) = q^n \sum_{t=1}^{n} t.\binom{n}{t}(\alpha)^t = q^n\left[\binom{n}{1}\alpha + 2\binom{n}{2}\alpha^2 + 3\binom{n}{3}\alpha^3 + \ldots + n\binom{n}{n}\alpha^n\right]$$

Estadística Matemática.

$$(1+x)^n = 1+\binom{n}{1}x+\binom{n}{2}x^2+\binom{n}{3}x^3+\ldots+\binom{n}{n}x^n$$

derivando respecto a x

$$n(1+x)^{n-1} = \binom{n}{1}+2\binom{n}{2}x+3\binom{n}{3}x^2+\ldots+n\binom{n}{n}x^{n-1}$$

para darle la forma de $E(T)$ multiplicamos por x

$$nx(1+x)^{n-1} = \binom{n}{1}x+2\binom{n}{2}x^2+3\binom{n}{3}x^3+\ldots+n\binom{n}{n}x^n$$

reemplazamos apropiadamente en $E(T)$

$$E(T)=q^n\left[n\alpha(1+\alpha)^{n-1}\right]=q^n\left[n\frac{p}{q}\left(1+\frac{p}{q}\right)^{n-1}\right]=q^n\left[n\frac{p}{q}\left(\frac{q+p}{q}\right)^{n-1}\right]=np$$

$$E(T)=np \rightarrow E\left(\frac{T}{n}\right)=\frac{np}{n}=p\rightarrow\frac{T}{n} \text{ es insesgado de } p$$

Hallemos el insesgado de p^2

$$E\left[T(T-1)\right]=\sum_{t=0}^{n}t(t-1).\binom{n}{t}p^tq^{n-t}=\sum_{t=2}^{n}t(t-1).\binom{n}{t}\left(\frac{p}{q}\right)^t q^n=\sum_{t=2}^{n}t(t-1).\binom{n}{t}(\alpha)^t q^n=$$

$$E\left[T(T-1)\right]=q^n\sum_{t=2}^{n}t(t-1).\binom{n}{t}(\alpha)^t=q^n\left[2\times1\binom{n}{2}\alpha^2+3\times2\binom{n}{3}\alpha^3+\ldots+n(n-1)\binom{n}{n}\alpha^n\right]$$

Estadística Matemática.

$$(1+x)^n = 1+\binom{n}{1}x+\binom{n}{2}x^2+\binom{n}{3}x^3+\ldots+\binom{n}{n}x^n$$

derivando respecto a x

$$n(1+x)^{n-1} = \binom{n}{1}+2\binom{n}{2}x+3\binom{n}{3}x^2+\ldots+n\binom{n}{n}x^{n-1}$$

derivando nuevamente respecto a x

$$n(n-1)(1+x)^{n-1} = 2\times1\binom{n}{2}+3\times2\binom{n}{3}x+\ldots+n(n-1)\binom{n}{n}x^{n-2}$$

para darle la forma de $E\left[T(T-1)\right]$ multiplicamos por x^2

$$n(n-1)x^2(1+x)^{n-1} = 2\times1\binom{n}{2}x^2+3\times2\binom{n}{3}x^3+\ldots+n\times(n-1)\binom{n}{n}x^n$$

reemplazamos apropiadamente en $E\left[T(T-1)\right]$

$$E\left(T\left(T-1\right)\right)=q^{n}\left[n\left(n-1\right)\alpha^{2}\left(1+\alpha\right)^{n-2}\right]=q^{n}\left[n\left(n-1\right)\left(\frac{p}{q}\right)^{2}\left(1+\frac{p}{q}\right)^{n-2}\right]=$$

$$E\left(T\left(T-1\right)\right)=q^{n}\left[n\left(n-1\right)\frac{p^{2}}{q^{2}}\left(\frac{q+p}{q}\right)^{n-2}\right]=n\left(n-1\right)p^{2}\rightarrow\frac{T\left(T-1\right)}{n\left(n-1\right)}\text{ es insesgado de }p^{2}$$

Hallemos el estimador insesgado de p^3

$$E\left[T\left(T-1\right)\left(T-2\right)\right]=\sum_{t=0}^{n}t\left(t-1\right)\left(t-2\right).\binom{n}{t}p^{t}q^{n-t}=\sum_{t=3}^{n}t\left(t-1\right)\left(t-2\right).\binom{n}{t}\left(\frac{p}{q}\right)^{t}q^{n}=$$

$$E\left[T\left(T-1\right)\left(T-2\right)\right]=\sum_{t=3}^{n}t\left(t-1\right)\left(t-2\right).\binom{n}{t}\left(\alpha\right)^{t}q^{n}=q^{n}\sum_{t=3}^{n}t\left(t-1\right)\left(t-2\right).\binom{n}{t}\left(\alpha\right)^{t}=$$

$$E\left[T\left(T-1\right)\left(T-2\right)\right]=q^{n}\left[3\times2\times1\binom{n}{2}\alpha^{2}+4\times3\times2\binom{n}{3}\alpha^{3}+\ldots+n\left(n-1\right)\left(n-2\right)\binom{n}{n}\alpha^{n}\right]$$

Estadística Matemática.

$$\left(1+x\right)^{n}=1+\binom{n}{1}x+\binom{n}{2}x^{2}+\binom{n}{3}x^{3}+\ldots+\binom{n}{n}x^{n}$$

derivando respecto a x

$$n\left(1+x\right)^{n-1}=\binom{n}{1}+2\binom{n}{2}x+3\binom{n}{3}x^{2}+\ldots+n\binom{n}{n}x^{n-1}$$

derivando nuevamente respecto a x

$$n\left(n-1\right)\left(1+x\right)^{n-1}=2\times1\binom{n}{2}+3\times2\binom{n}{3}x+\ldots+n\times\left(n-1\right)\binom{n}{n}x^{n-2}$$

de nuevo derivando respecto a x

$$n\left(n-1\right)\left(n-2\right)\left(1+x\right)^{n-3}=3\times2\times1\binom{n}{3}+\ldots+n\left(n-1\right)\left(n-2\right)\binom{n}{n}x^{n-3}$$

para darle la forma de $E\left[T\left(T-1\right)\left(T-2\right)\right]$ multiplicamos por x^3

$$n\left(n-1\right)\left(n-2\right)x^{3}\left(1+x\right)^{n-3}=3\times2\times1\binom{n}{3}x^{3}+\ldots+n\left(n-1\right)\left(n-2\right)\binom{n}{n}x^{n}$$

reemplazamos apropiadamente en $E\left[T\left(T-1\right)\left(T-2\right)\right]$

$$E\left(T\left(T-1\right)\left(T-2\right)\right)=q^{n}\left[n\left(n-1\right)\left(n-2\right)\alpha^{3}\left(1+\alpha\right)^{n-3}\right]=q^{n}\left[n\left(n-1\right)\left(n-2\right)\left(\frac{p}{q}\right)^{3}\left(1+\frac{p}{q}\right)^{n-3}\right]=$$

$$E\left[T\left(T-1\right)\left(T-2\right)\right]=q^{n}\left[n\left(n-1\right)\left(n-2\right)\frac{p^{3}}{q^{3}}\left(\frac{q+p}{q}\right)^{n-3}\right]=n\left(n-1\right)\left(n-2\right)p^{3}$$

$\rightarrow \dfrac{T(T-1)(T-2)}{n(n-1)(n-2)}$ es el estimador insesgado de p^3

Por inducción: $\dfrac{T(T-1)(T-2)\ldots[T-(r-1)]}{n(n-1)(n-2)\ldots[n-(r-1)]}$ es un estimador insesgado de p^r .

Eficiencia.

Sean $\hat{\theta}_1$ = T_1 y $\hat{\theta}_2$ = T_2 dos estimadores insesgados para q(θ), tal que: V (T_1)< V (T_2) $\Rightarrow$ se dice que T_1 es más eficiente que T_2.

Ejemplo 9: Sea $X_1,\ldots,X_n$ una m.a de $\text{Uniforme}(0,\theta)$ con $Y_1,\ldots,Y_n$ las estadísticas de orden ¿Es T_1= 2 $\bar{X}$ más eficiente que T_2= $\left(\dfrac{n+1}{n}\right)Y_n$ para estimar a θ?

Ya se demostró que T_1 y T_2 son estimadores insesgados de θ. Ahora veamos cuál de los dos estimadores es menos variable:

$$Var(T_1) = Var(2\bar{X}) = 4Var(\bar{X}) = 4\frac{\sigma^2}{n} = 4\frac{\theta^2}{12n} = \frac{\theta^2}{3n}$$

Para hallar $Var(T_2)$ se necesita obtener $Var(Y_n)$ $\left[\text{Ya se halló } \mathrm{E}(Y_n) = \dfrac{n}{n+1}\theta\right]$

$$E(Y_n^2) = \int_0^\theta y^2\left(\frac{n}{\theta^n}y^{n-1}\right)dy = \frac{n}{\theta^n}\left(\frac{\theta^{n+2}}{n+2}\right) = \frac{n}{n+2}\theta^2$$

$$Var(Y_n) = E(Y_n^2) - [\mathrm{E}(Y_n)]^2 = \frac{n}{n+2}\theta^2 - \left[\frac{n}{n+1}\theta\right]^2 = \frac{n\theta^2}{(n+1)^2(n+2)}$$

$$Var(T_2) = Var\left(\frac{n+1}{n}Y_n\right) = \left(\frac{n+1}{n}\right)^2 Var(Y_n) = \left(\frac{n+1}{n}\right)^2\left(\frac{n\theta^2}{(n+1)^2(n+2)}\right) = \frac{\theta^2}{n(n+2)}$$

Veamos cuál de los estimadores es más eficiente:

$$\frac{Var(T_2)}{Var(T_1)} = \left(\frac{\theta^2}{n(n+2)}\right)\left(\frac{3n}{\theta^2}\right) = \frac{3}{n+2} < 1, \text{ para } n > 1 \rightarrow Var(T_2) < Var(T_1)$$

$\therefore$ el estimador T_2 es más eficiente que T_1 si $n > 1$

NOTA: Ver el Script R

Ejemplo10: Si X_1, X_2, X_3 es una muestra aleatoria de una población con media μ y variancia σ^2 y sean $T_1 = \frac{3}{4}X_1 + \frac{1}{5}X_2 + \frac{1}{20}X_3$ con $T_2 = \frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3$ dos estimadores de μ. ¿Cuál de estos estimadores es mejor? Justifique su respuesta.

Se verifica que ambos estimadores son de la forma $S_n = \sum_{i=1}^{n} a_i X_i$ y que $\sum_{i=1}^{n} a_i = 1$ por lo tanto ambos estimadores son insegados y como para T_2, $a_i = \frac{1}{n} = \frac{1}{3}$ entonces T_2 es el estimador insesgado de mínima variancia para μ.

Algunos cálculos extras:

$$E(T_1) = \frac{3}{4}E[X_1] + \frac{1}{5}E[X_2] + \frac{1}{20}E[X_3] = \frac{3}{4}\mu + \frac{1}{5}\mu + \frac{1}{20}\mu = \mu$$

$$E(T_2) = \frac{1}{3}E[X_1] + \frac{1}{3}E[X_2] + \frac{1}{3}E[X_3] = \frac{1}{3}\mu + \frac{1}{3}\mu + \frac{1}{3}\mu = \mu$$

$$Var(T_1) = \left(\frac{3}{4}\right)^2 Var[X_1] + \left(\frac{1}{5}\right)^2 Var[X_2] + \left(\frac{1}{20}\right)^2 Var[X_3] =$$

$$Var(T_1) = \left(\frac{3}{4}\right)^2 \sigma^2 + \left(\frac{1}{5}\right)^2 \sigma^2 + \left(\frac{1}{20}\right)^2 \sigma^2 = \frac{121}{200}\sigma^2$$

$$Var(T_2) = \left(\frac{1}{3}\right)^2 Var[X_1] + \left(\frac{1}{3}\right)^2 Var[X_2] + \left(\frac{1}{3}\right)^2 Var[X_3] =$$

$$Var(T_2) = \left(\frac{1}{3}\right)^2 \sigma^2 + \left(\frac{1}{3}\right)^2 \sigma^2 + \left(\frac{1}{3}\right)^2 \sigma^2 = \frac{\sigma^2}{3}$$

Se observa que ambos estimadores son insesgados y que $Var(T_2) < Var(T_1)$ por lo tanto $T_2 = \overline{X}$ es mejor estimador que T_1.

Error Cuadrático Medio

Para estimar un parámetro θ se cuenta con tres estimadores T_1, T_2 y T_3. Suponga que en 10 muestras se obtienen las siguientes estimaciones de θ.

Muestra	T_1	T_2	T_3
1	20.9	21.5	20.9
2	20.9	21.5	20.9
3	20.1	21.5	21.1
4	20.3	21.6	21.0
5	21.6	21.4	20.8
6	21.3	21.5	20.9
7	20.5	21.4	21.2
8	20.7	21.6	21.0
9	21.7	21.5	20.9
10	20.9	21.5	20.9
Promedio	20.89	21.50	20.96
Desviación	0.53	0.07	0.12

Suponga que $\theta = 21$

T_1 en promedio está cerca de $\theta = 21$ pero sus valores son muy dispersos.
T_2 sobreestima a θ aunque sus valores están concentrados.
T_3 en promedio está alrededor de θ y con pequeña dispersión. Se concluye que, de los tres estimadores, T_3 es el mejor estimador de θ.

El ECM del estimador $\theta = T$ de $q(\theta)$ se define como:

$$ECM(T) = E\left[\left(T - q(\theta)\right)^2\right]$$

donde $(T - q(\theta))$ es el error que se comete al estimar $q(\theta)$, y $E\left[\left(T - q(\theta)\right)^2\right]$ es el promedio de los errores al cuadrado.

<u>Observación</u>

Si el ECM(T) es finito entonces:

$$ECM(T) = Var(T) + b^2(T)$$

donde $b(T) = E(T) - q(\theta)$, mide el sesgo del estimador.
Prueba

$$ECM(T) = E\left\{\left[\left(T - E(T)\right) + \left(E(T) - q(\theta)\right)\right]^2\right\} =$$

$$= E\left\{\left[T - E(T)\right]^2\right\} + E\left\{\underbrace{\left[E(T) - q(\theta)\right]^2}_{Constante}\right\} + 2\left(E(T) - q(\theta)\right)\underbrace{E\left(T - E(T)\right)}_{cero} =$$

$$ECM(T) = Var(T) + \left[E(T) - q(\theta)\right]^2 = Var(T) + b^2(T)$$

Una conclusión es que: ECM(T) = Var(T) sólo si T es insesgado o sea $E(T) = q(\theta)$.

Si $b(T) = E(T) - q(T) > 0$ se dice que T sobreestima a θ.
Si $b(T) = E(T) - q(T) < 0$ se dice que T subestima a θ.
Si entre los estimadores T_1 y T_2, T_1 tiene menor sesgo y varianza se concluye que T_1 es mejor que T_2. Pero si T_1 tiene menor sesgo pero mayor varianza que T_2 el mejor estimador es el de menor error cuadrático medio.

Ejemplo 11: Sea $X_1, \ldots, X_n$ una muestra aleatoria de la distribución $N\left(\mu, \sigma^2\right)$.

a. Sea $\bar{X}$ un estimador de μ. Halle el $ECM\left(\bar{X}\right)$ y el $\text{Sesgo}\left(\bar{X}\right)$.

$$\text{Sesgo}\left(\bar{X}\right) = E\left(\bar{X}\right) - \mu = \mu - \mu = 0$$

$$ECM\left(\bar{X}\right) = Var\left(\bar{X}\right) + \left[\text{Sesgo}\left(\bar{X}\right)\right]^2 = \frac{\sigma^2}{n} + 0^2 = \frac{\sigma^2}{n}$$

b. Sea $T_1 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$ un estimador de σ^2. Halle el $ECM\left(T_1\right)$ y el $\text{Sesgo}\left(T_1\right)$.

$$\text{Sesgo}\left(T_1\right) = E\left(T_1\right) - \sigma^2 = E\left(\frac{\sigma^2}{n}\frac{nT_1}{\sigma^2}\right) - \sigma^2 = \frac{\sigma^2}{n}E\left[\frac{(n-1)S^2}{\sigma^2}\right] - \sigma^2 =$$

$$\text{Sesgo}\left(T_1\right) = \frac{\sigma^2}{n}(n-1) - \sigma^2 = -\frac{\sigma^2}{n}$$

$$E\left(T_1\right) = \frac{\sigma^2}{n}(n-1) \Rightarrow \lim_{n\to\infty}\frac{\sigma^2}{n}(n-1) = \lim_{n\to\infty}\sigma^2\left(1 - \frac{1}{n}\right) = \sigma^2 \Rightarrow T_1 \text{ es asintóticamente insesgado.}$$

.

$$ECM\left(T_1\right) = Var\left(T_1\right) + \left[\text{Sesgo}\left(T_1\right)\right]^2 = Var\left(\frac{\sigma^2}{n}\frac{nT_1}{\sigma^2}\right) + \left(-\frac{\sigma^2}{n}\right)^2 =$$

$$ECM\left(T_1\right) = \frac{\sigma^4}{n^2}Var\left[\frac{(n-1)S^2}{\sigma^2}\right] + \frac{\sigma^4}{n^2} = \frac{\sigma^4}{n^2}\left[2(n-1)\right] + \frac{\sigma^4}{n^2} = \frac{(2n-1)\sigma^4}{n^2}$$

Nota: Se conoce que si $X_1,\ldots,X_n$ es una muestra aleatoria extraída de una población $N\left(\mu,\sigma^2\right) \to \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$ y que $E\left[\chi^2_{(n-1)}\right] = n-1$, $Var\left[\chi^2_{(n-1)}\right] = 2(n-1)$

c. Sea $T = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$ un estimador de σ^2. Halle el $ECM\left(T\right)$ y el $\text{Sesgo}\left(T\right)$.

$$\text{Sesgo}\left(T\right) = E\left(T\right) - \sigma^2 = E\left(\frac{\sigma^2}{n-1}\frac{(n-1)T}{\sigma^2}\right) - \sigma^2 = \frac{\sigma^2}{n-1}E\left[\frac{(n-1)S^2}{\sigma^2}\right] - \sigma^2 =$$

.

$$\text{Sesgo}\left(T\right) = \frac{\sigma^2}{n-1}(n-1) - \sigma^2 = 0$$

$$ECM\left(T\right) = Var\left(T\right) + \left[\text{Sesgo}\left(T\right)\right]^2 = Var\left(\frac{\sigma^2}{n-1}\frac{(n-1)T}{\sigma^2}\right) + (0)^2 =$$

$$ECM\left(T\right) = \frac{\sigma^4}{(n-1)^2}Var\left[\frac{(n-1)S^2}{\sigma^2}\right] = \frac{\sigma^4}{(n-1)^2}\left[2(n-1)\right] = \frac{2\sigma^4}{n-1}$$

d. ¿Qué estimador tiene menor ECM?

La diferencia entre ambos ECMs es proporcional a.

$$\frac{2}{n-1}-\frac{2n-1}{n^2}=\frac{2n^2-(n-1)(2n-1)}{(n-1)n^2}=\frac{3n-1}{(n-1)n^2}>0$$

Se concluye que el estimador sesgado T_1 tiene menor ECM que el estimador insesgado T.

e. ¿Es T_1 mejor estimador que T ? Justifique.

Cuando se tienen estimadores insesgados o asintóticamente insesgados se prefiere el de menor varianza. En este caso se prefiere $T_1=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$. Aunque se puede discutir en forma gráfica.

Teorema: Entre todos los estimadores de la varianza normal de la forma $\hat{\sigma}_k^2=\frac{1}{n-k}\sum_{i=1}^{n}(X_i-\bar{X})^2$, el de menor ECM es $\hat{\sigma}_{-1}^2=\frac{1}{n+1}\sum_{i=1}^{n}(X_i-\bar{X})^2$.

Ejemplo 12: Sea $X_1,\ldots,X_n$ una muestra aleatoria de $f(x)=\frac{2x}{\theta^2}I_{(0,\theta)}(x)$. Considere los siguientes estimadores de θ: $S=Y_n$ y $T=\frac{3}{2}\bar{x}$. Para que valores de n el estimador S es mejor que T.

Se demuestra que:

$$E(S)=E(Y_n)=\frac{2n}{2n+1}\theta=\theta-\frac{\theta}{2n+1} \text{ y } \mathrm{Var}(S)=Var(Y_n)=\frac{2n}{(2n+1)^2(2n+2)}\theta^2$$

$$ECM(S)=\frac{2}{(2n+1)(2n+2)}\theta^2$$

$$E(T)=E\left(\frac{3}{2}\bar{x}\right)=\theta \text{ y } \mathrm{Var}(T)=Var\left(\frac{3}{2}\bar{x}\right)=\frac{1}{8n}\theta^2\rightarrow ECM(T)=\frac{1}{8n}\theta^2$$

En este caso S es mejor que T porque es asintóticamente insesgado y su varianza es menor que la de T cuando n>-0.1941457205.

Ejemplo: Sea $X_1,\ldots,X_n$ una muestra aleatoria de una distribución $\mathrm{Exponencial}(\mathrm{Tasa}=\lambda)$ esto quiere decir que $f(x)=\lambda e^{-\lambda x}I_{(0,\infty)}(x)$. Verifique los siguientes resultados:

1. $\bar{X}$ es un estimador insesgado de $q(\lambda)=\frac{1}{\lambda}$.

2. Si $Y_1 = \min\left(X_1,\ldots,X_n\right) \rightarrow Y_1 \sim \text{Exponencial}\left(\text{Tasa} = n\lambda\right) \rightarrow E\left(nY_1\right) = \dfrac{1}{\lambda}$

   $\text{Var}\left(\bar{X}\right) = \dfrac{1}{n\lambda^2}$ y $\text{Var}\left(nY_1\right) = \dfrac{1}{\lambda^2} \rightarrow \bar{X}$ es mejor que nY_1.

3. Para estimar a λ se puede usar a

   $$\frac{1}{\bar{X}} \rightarrow E\left(\frac{1}{\bar{X}}\right) = E\left(\frac{n}{\sum_{i=1}^{n} X_i}\right) = E\left(\frac{n}{W}\right) = \frac{n}{n-1}\lambda, \text{ donde } W = \sum_{i=1}^{n} X_i \sim \text{Gamma}\left(n,\lambda\right).$$

   Un estimador insesgado de λ es

   $$\hat{\lambda}_1 = \frac{n-1}{n}\left(\frac{1}{\bar{X}}\right) = \frac{n-1}{\sum_{i=1}^{n} X_i}$$

   $$\rightarrow \text{Var}\left(\hat{\lambda}_1\right) = \text{Var}\left(\frac{n-1}{\sum_{i=1}^{n} X_i}\right) = \left(n-1\right)^2 \text{Var}\left(\frac{1}{W}\right) = \left(n-1\right)^2\left[E\left(\left[\frac{1}{W}\right]^2\right) - \left(E\left[\frac{1}{W}\right]\right)^2\right] \rightarrow$$

   Verificar que: $\text{Var}\left(\hat{\lambda}_1\right) = \dfrac{\lambda^2}{n-2} \rightarrow \text{ECM}\left(\hat{\lambda}_1\right) = \text{Var}\left(\hat{\lambda}_1\right) = \dfrac{\lambda^2}{n-2}$

   .

4. Sea $\hat{\lambda}_2 = \dfrac{1}{\bar{X}} \rightarrow \hat{\lambda}_2 = \dfrac{n}{n-1}\hat{\lambda}_1$

$$\text{Sesgo}\left(\hat{\lambda}_2\right) = E\left(\hat{\lambda}_2\right) - \lambda = \frac{n}{n-1}\lambda - \lambda = \frac{1}{n-1}\lambda$$

$$\text{Var}\left(\hat{\lambda}_2\right) = \text{Var}\left(\frac{n}{n-1}\hat{\lambda}_1\right) = \left(\frac{n}{n-1}\right)^2 \text{Var}\left(\hat{\lambda}_1\right) = \frac{n^2}{\left(n-1\right)^2\left(n-2\right)}\lambda^2$$

$$\text{ECM}\left(\hat{\lambda}_2\right) = \text{Var}\left(\hat{\lambda}_2\right) + \text{Sesgo}^2\left(\hat{\lambda}_2\right) = \left[\frac{n^2}{\left(n-1\right)^2\left(n-2\right)} + \frac{1}{\left(n-1\right)^2}\right]\lambda^2$$

$\hat{\lambda}_2$ es asintóticamente insesgado $\left[\lim\limits_{n\to\infty} E\left(\hat{\lambda}_2\right) = \lim\limits_{n\to\infty}\dfrac{n}{n-1}\lambda = \lambda\right]$ y su ECM llegará a cero cuando n crezca.

$\hat{\lambda}_1$ es insesgado y su ECM llegará a cero cuando n crezca $\rightarrow$ se prefiere $\hat{\lambda}_1$.

Consistencia

La consistencia es una propiedad imaginaria de un estimador: un estimador consistente, $\hat{T}_n$ se aproxima al valor verdadero del parámetro cuando el tamaño de la muestra, n tiende al infinito (obsérvese que usamos un subíndice n para indicar que el estimador es una función del tamaño de la muestra). A veces decimos "muestra grande" para indicar que $n \to \infty$. Decimos "propiedad imaginaria" porque en realidad el tamaño de la muestra nunca puede ser infinito. La consistencia es una propiedad asintótica de un estimador; Otra propiedad asintótica es el insesgamiento asintótico. Muchos estimadores consistentes tienen una distribución normal en una muestra grande.

Tres tipos de convergencia/consistencia se usan cuando $n \to \infty$:

1. Convergencia en probabilidad (Convergencia débil).

$\lim_{n\to\infty} P\left(\left|T_n - q(\theta)\right| > \varepsilon\right) = 0$. La distancia entre T_n y $q(\theta)$ se hace más pequeña cuando n crece.

2. Convergencia en sentido cuadrático.

$\lim_{n\to\infty} E\left(T_n - q(\theta)\right)^2 = 0$

3. Convergencia con probabilidad 1 (Convergencia fuerte o casi segura).

$P\left[\lim_{n\to\infty} \hat{T}_n = q(\theta)\right] = 1$

Notar que la convergencia en sentido cuadrático implica la convergencia en probabilidad y la convergencia fuerte implica la convergencia débil.

Se dice que una sucesión $T_1,\ldots,T_n$ de estimadores de $q(\theta)$, donde: $T_i = t_i\left(X_1,\ldots,X_n\right)$ $\forall i$, es consistente si dados ∈>0, δ>0 existe N >>0 tal que ∀ n > N se tiene:

$$P\left(\left|T_n - q(\theta)\right| > \varepsilon\right) \le \delta \le \frac{E\left(T_n - q(\theta)\right)^2}{\varepsilon^2} \Leftrightarrow \lim_{n\to\infty} P\left(\left|T_n - q(\theta)\right| > \varepsilon\right) = 0 \Leftrightarrow$$

$$T_n \xrightarrow[n\to\infty]{P} q(\theta)$$

Decir que T_n es consistente para estimar a $q(\theta)$ es equivalente a decir que T_n converge en probabilidades a $q(\theta)$.

Teorema 1

Un estimador insesgado $\hat{\theta}_n$ de θ es un estimador consistente para θ sí $\lim_{n\to\infty} \mathrm{Var}\left(\hat{\theta}_n\right) = 0$, en otras palabras $\hat{\theta}_n$ converge en probabilidades a θ.

Teorema 2

Si $\hat{\theta}_n$ es un estimador asintóticamente insesgado para θ $\left[\lim_{n\to\infty} E\left(\hat{\theta}_n\right) = \theta\right]$ y $\lim_{n\to\infty} \mathrm{Var}\left(\hat{\theta}_n\right) = 0$, entonces $\hat{\theta}_n \xrightarrow{P} \theta$.

Ejemplo: Sea $X_1,\ldots,X_n$ una muestra aleatoria de una distribución $\mathrm{Uniforme}(0,\theta)$. Verificar que $Y_n = \max\left(X_1,\ldots,X_n\right)$ es un estimador consistente de θ (o converge en probabilidades a θ).

$\lim\limits_{n\to\infty} E(Y_n) = \lim\limits_{n\to\infty} \frac{n}{n+1}\theta = \theta \to Y_n$ es asintóticamente insesgado para θ.

$\lim\limits_{n\to\infty} \operatorname{Var}(Y_n) = \lim\limits_{n\to\infty} \frac{n}{(n-1)^2(n-2)}\theta^2 = 0$, entonces por el Teorema 2: $Y_n \xrightarrow{P} \theta$.

Ejemplo: Sea $X_1,\dots,X_n$ una muestra aleatoria de $f(.)$, entonces $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})$ es tal que $E(S^2) = \sigma^2$ y $\lim\limits_{n\to\infty} \operatorname{Var}(S^2) = \lim\limits_{n\to\infty} \frac{E\left[(X-\mu)^4\right]}{n} - \frac{n-3}{n(n-1)}(\sigma^2)^2 = 0$ por lo tanto, por el Teorema 1 $S^2 \xrightarrow{P} \sigma^2$.

De otro lado, $S_1^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}) = \frac{n-1}{n}S^2$ es asintóticamente insesgado para σ^2 y $\lim\limits_{n\to\infty} \operatorname{Var}(S_1^2) = \lim\limits_{n\to\infty}\left(\frac{n-1}{n}\right)^2\left[\frac{E\left[(X-\mu)^4\right]}{n} - \frac{n-3}{n(n-1)}(\sigma^2)^2\right] = 0$, entonces de acuerdo al Teorema 2: $S_1^2 \xrightarrow{P} \sigma^2$.

<u>Consistencia en error cuadrático medio</u>

Una sucesión $\{T_n\}_{n\geq 1}$ de estimadores es definida como una sucesión de estimadores consistentes en ECM de $q(\theta)$ si y sólo si:

$$\lim_{n\to\infty} E\left[(T_n - q(\theta))^2\right] = 0,\ \forall\theta\in\Theta$$

Observación La consistencia en ECM implica que ambos, el sesgo y la variancia de T_n se aproximan a cero.

Ejemplo: Sea $X_1,\dots,X_n$ una muestra aleatoria de una distribución con media μ y varianza σ^2. Verifique que $\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$ es un estimador consistente de μ.

Según la desigualdad de Chebyshev

$$\mathrm{P}\left[\left|\bar{X}-\mu\right|\geq\varepsilon\right]\leq\frac{E\left[\left(\bar{X}-\mu\right)^2\right]}{\varepsilon^2}=\frac{Var\left(\bar{X}\right)}{\varepsilon^2}=\frac{\sigma^2}{n\varepsilon^2}$$

$$\lim_{n\to\infty}\mathrm{P}\left[\left|\bar{X}-\mu\right|\geq\varepsilon\right]\leq\lim_{n\to\infty}\frac{\sigma^2}{n\varepsilon^2}=0\to\bar{X}\text{ es consistente.}$$

Ejemplo 13: Demuestre que los $M_r^| = \frac{1}{n}\sum_{i=1}^{n} X_i^r$; $r=1,2,3,\ldots$son estimadores consistentes de $\mu_r^| = E\left(X^r\right)$.

Según Chebyshev:
Teniendo en cuenta que $M_r^|$ es insesgado:

$$P\left[\left|M_r^|-\mu_r^|\right|\geq\varepsilon\right]\leq\frac{E\left[\left(M_r^|-\mu_r^|\right)^2\right]}{\varepsilon^2}=\frac{Var\left(M_r^|\right)}{\varepsilon^2}=\frac{\mu_{2r}^|-\left(\mu_r^|\right)^2}{n\varepsilon^2}$$

Siempre que exista $\mu_2^|$:

$$\lim_{n\to\infty}P\left[\left|M_r^|-\mu_r^|\right|\geq\varepsilon\right]\leq\lim_{n\to\infty}\frac{\mu_{2r}^|-\left(\mu_r^|\right)^2}{n\varepsilon^2}=0\to\ M_r^|\text{ converge en probabilidades a }\mu_r^|.$$

Queda demostrado.

Ejemplo 14: $X_1,\ldots,X_n$es una muestra aleatoria de $f\left(x\right)=\frac{1}{2}\left(1+\theta x\right)I_{(-1,1)}\left(x\right)$, $-1<\theta<1$. Demuestre que $\hat{\theta}=3\bar{X}$ es un estimador consistente de θ.

Se verifica que: $\mu=E\left(X\right)=\frac{\theta}{3}$, $\sigma^2=Var\left(X\right)=\frac{3-\theta^2}{9}$

$\hat{\theta}=3\bar{X}$ es un estimador insesgado de θ.

$Var\left(\hat{\theta}\right)=9Var\left(\bar{X}\right)=\frac{3-\theta^2}{n}$.

Según Chebyshev:

$$\mathrm{P}\left[\left|\hat{\theta}-\theta\right|\geq\varepsilon\right]\leq\frac{E\left[\left(\hat{\theta}-\theta\right)^2\right]}{\varepsilon^2}=\frac{Var\left(\hat{\theta}\right)}{\varepsilon^2}=\frac{3-\theta^2}{n\varepsilon^2}$$

$$\lim_{n\to\infty}\mathrm{P}\left[\left|\hat{\theta}-\theta\right|\geq\varepsilon\right]\leq\lim_{n\to\infty}\frac{3-\theta^2}{n\varepsilon^2}=0\to\hat{\theta}=3\bar{X}\text{ es consistente.}$$

Ejemplo 15: $X_1,\ldots,X_n$es una muestra aleatoria de una población con media μ y variancia σ^2. ¿Cuál de los siguientes estimadores es consistente para μ?

a. $S = 2\bar{X}$.

$E(S) = 2\mu = \mu + \mu \rightarrow S$ es sesgado con sesgo: $b(S) = \mu$

$$Var(S) = 4Var(\bar{X}) = 4\frac{\sigma^2}{n}$$

Como S es sesgado entonces $E\left[(S-\mu)^2\right]$ es su error cuadrático medio.

$$P\left[|S-\mu| \geq \varepsilon\right] \leq \frac{E\left[(S-\mu)^2\right]}{\varepsilon^2} = \frac{Var(S)+\left[b(S)\right]^2}{\varepsilon^2} = \frac{Var(S)}{\varepsilon^2} + \frac{\left[b(S)\right]^2}{\varepsilon^2} = \frac{4\sigma^2}{n\varepsilon^2} + \frac{\mu^2}{\varepsilon^2}$$

$\lim\limits_{n\to\infty}\left(\frac{4\sigma^2}{n\varepsilon^2} + \frac{\mu^2}{\varepsilon^2}\right) = \frac{\mu^2}{\varepsilon^2} \neq 0 \rightarrow S$ no es estimador consistente de μ.

b. $S = 2\sum_{i=1}^{n}\frac{iX_i}{n(n+1)}$.

$$E(S) = \frac{2}{n(n+1)}\sum_{i=1}^{n} iE(X_i) = \frac{2}{n(n+1)}\sum_{i=1}^{n} i\mu = \frac{2\mu}{n(n+1)}\sum_{i=1}^{n} i = \frac{2\mu}{n(n+1)}\frac{n(n+1)}{2} = \mu$$

$\rightarrow S$ es insesgado para μ.

$$Var(S) = \frac{4}{n^2(n+1)^2}\sum_{i=1}^{n} i^2 Var(X_i) = \frac{4}{n^2(n+1)^2}\sum_{i=1}^{n} i^2\sigma^2 = \frac{4\sigma^2}{n^2(n+1)^2}\sum_{i=1}^{n} i^2 =$$

$$Var(S) = \frac{4\sigma^2}{n^2(n+1)^2}\left(\frac{n(n+1)(2n+1)}{6}\right) = \frac{2(2n+1)\sigma^2}{3n(n+1)}$$

$$P\left[|S-\mu| \geq \varepsilon\right] \leq \frac{E\left[(S-\mu)^2\right]}{\varepsilon^2} = \frac{Var(S)}{\varepsilon^2} = \frac{Var(S)}{\varepsilon^2} = \frac{2(2n+1)\sigma^2}{3n(n+1)\varepsilon^2}$$

$\lim\limits_{n\to\infty}\left(\frac{2(2n+1)\sigma^2}{3n(n+1)\varepsilon^2}\right) = 0 \rightarrow S$ es estimador consistente de μ.

Definición. - En la práctica se usan las siguientes condiciones suficientes (a pesar de no ser necesarias) para juzgar consistencias.

Una sucesión $\{T_n\}_{n\geq 1}$ de estimadores de q(θ) es consistente si:

1) $\lim\limits_{n\to\infty} E\left[T_n\right] = q(\theta)$. Esto indica que es asintóticamente insesgado.

2) $\lim\limits_{n\to\infty} Var\left[T_n\right] = 0$.

Proposición

1. Teorema de mapeo continuo: Sea g una función continua. Si $\hat{\theta}_n \xrightarrow{P} \theta$ $\Rightarrow g\left(\hat{\theta}_n\right) \xrightarrow{P} g(\theta)$.

2. Si $\hat{\theta}_n \xrightarrow{P} \theta_1$ y $\Rightarrow$ $\hat{\theta}'_n \xrightarrow{P} \theta_2$:

i) $\hat{\theta}_n \pm \hat{\theta}'_n \xrightarrow{P} \theta_1 \pm \theta_2$

ii) $\hat{\theta}_n \hat{\theta}'_n \xrightarrow{P} \theta_1 \theta_2$

iii) $\dfrac{\hat{\theta}_n}{\hat{\theta}'_n} \xrightarrow{P} \dfrac{\theta_1}{\theta_2}$ $\quad \theta_2 \neq 0$

3. Si $\hat{\theta}_n$ tiene distribución límite F(esto es, P($\hat{\theta}_n \leq$ X) $\xrightarrow[n\to\infty]{}$ F(x)) y $\hat{\theta}'_n \xrightarrow{P} \theta$ $(\theta \neq 0)$ $\Rightarrow$ $\dfrac{\hat{\theta}_n}{\hat{\theta}'_n}$ tiene como distribución límite a F(x/θ) esto es:

P($\dfrac{\hat{\theta}_n}{\hat{\theta}'_n} \leq$ x) $\xrightarrow[n\to\infty]{}$ F(x/θ) , $\forall$x para los cuales F es continua.

Ejemplo 16: Sea $X_1,\ldots,X_n$ una m.a de UNIF (0, θ). Demuestre que $\left(\prod_{i=1}^{n} X_i\right)^{\frac{1}{n}}$ es un estimador consistente de $q(\theta) = \dfrac{\theta}{e}$. Nota: $\ln\left(\prod_{i=1}^{n} X_i\right)^{\frac{1}{n}} = \dfrac{1}{n}\sum_{i=1}^{n} \ln X_i$.

Se demuestra que:

$$Y = \frac{\ln X}{n} \to \mu = E(Y) = \frac{1}{n}\int_0^{\theta} \ln x \left(\frac{1}{\theta}\right) dx = \frac{1}{n}(\ln\theta\text{-}1) = \frac{1}{n}\ln\frac{\theta}{e} \quad \text{y} \quad \sigma^2 = Var(Y) = \frac{1}{n^2}$$

$$P\left(|Y-\mu| \geq \varepsilon\right) \leq \frac{E\left[(Y-\mu)^2\right]}{\varepsilon^2} = \frac{Var(Y)}{\varepsilon^2} = \frac{1}{n^2\varepsilon^2} \to \lim_{n\to\infty} \frac{1}{n^2\varepsilon^2} = 0 \to$$

$Y = \dfrac{\ln X}{n}$ es consistente de $\dfrac{1}{n}\ln\dfrac{\theta}{e}$ $\left(\text{o converge en probabilidades a } \dfrac{1}{n}\ln\dfrac{\theta}{e}\right)$ entonces por la proposición 1 anterior:

$\ln\left(\prod_{i=1}^{n} X_i\right)^{\frac{1}{n}} = \sum_{i=1}^{n} \dfrac{1}{n}\ln X_i$ converge en probabilidades a: $\sum_{i=1}^{n} \dfrac{1}{n}\ln\dfrac{\theta}{e} = \ln\dfrac{\theta}{e}$ en consecuencia: $\left(\prod_{i=1}^{n} X_i\right)^{\frac{1}{n}}$ es un estimador consistente de $q(\theta) = \dfrac{\theta}{e}$

Condiciones suficientes para la Consistencia

Sea $\{T_n\}$ una secuencia de estimadores tales que:

- $E(T_n)=\theta$, cuando $n\to\infty$
- $V(T_n)=0$, cuando $n\to\infty$

Entonces T_n es un estimador consistente de θ.

Ejemplo Sea $X_1,\dots,X_n$ una muestra aleatoria de una distribución $\text{Normal}(\mu,\sigma^2)$.
Demuestre que la $\bar{X}$ es un estimador consistente de μ.
Primera manera
Verifique que la $\bar{X}$ converge en probabilidades a μ (queda como ejercicio).
Segunda manera
Para demostrar que $\bar{X}$ converge en probabilidades a μ, utilizamos la condición suficiente:

- $E(\bar{X})=\mu$, cuando $n\to\infty$
- $V(\bar{X})=\dfrac{\sigma^2}{n}=0$, cuando $n\to\infty$

Por lo tanto, $\bar{X}$ converge en probabilidades (o es consistente) para μ.

Ejercicio: Si $X_1,\dots,X_n$ es una m.a de la distribución $f(x)=\dfrac{2}{\theta^2}(\theta-x)I_{[0,\theta]}$, verifique $3\bar{X}$ es consistente para estimar a θ. Utilice las condiciones suficientes de consistencia.

Propiedad de Invarianza de un Estimador Consistente
Si T es un estimador consistente de θ y f es una función continua entonces $f(T)$ es un estimador consistente de $f(\theta)$.

Ejercicio Sea $X_1,\dots,X_n$ una muestra aleatoria de una distribución $\text{Bernoulli}(p)$.
Demuestre que $\dfrac{\sum_{i=1}^{n}X_i}{n}\left(1-\dfrac{\sum_{i=1}^{n}X_i}{n}\right)$ es un estimador consistente de $p(1-p)$. Nota:

Para que $\hat{\theta}(1-\hat{\theta})$ sea un estimador consistente de $\theta(1-\theta)$, hay que verificar que $\hat{\theta}$ es consistente de θ y aplicar la propiedad de invarianza de los estimadores consistentes.

$$\mathrm{P}\left[\left|\bar{X}-p\right|\geq\varepsilon\right]\leq\frac{E\left[\left(\bar{X}-\mu\right)^2\right]}{\varepsilon^2}=\frac{Var\left(\bar{X}\right)}{\varepsilon^2}=\frac{p(1-p)}{n\varepsilon^2}$$

$$\lim_{n\to\infty}\mathrm{P}\left[\left|\bar{X}-\mu\right|\geq\varepsilon\right]\leq\lim_{n\to\infty}\frac{p(1-p)}{n\varepsilon^2}=0\to\bar{X}\ \text{es consistente.}$$

Convergencia en Distribución

Suponga que $\{X_n\}$ es una secuencia de variables aleatorias con FDAs $F_n(x)=P(X_n \leq x)$. Sea X otra variable aleatoria con FDA $F(x)=P(X \leq x)$. Se dice que $\{X_n\}$ converge en distribución a X si $\lim_{n\to\infty} F_n(x)=F(x)$ en todos los puntos de continuidad de F.

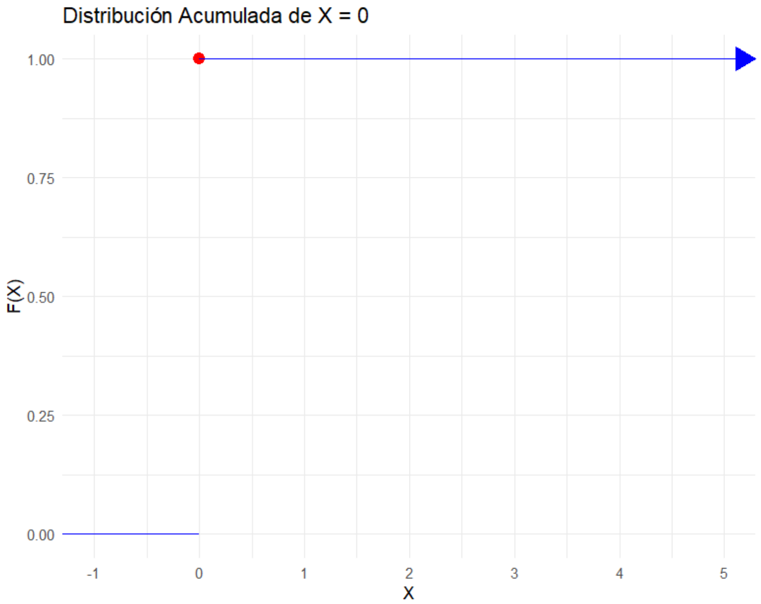
Ejemplo

Sea $a_n=\frac{1}{n}$ para $n=1,2,\ldots$ entonces:

$$F_n(x)=P(X_n \leq x)=P(a_n \leq x)=P\left(\frac{1}{n}\leq x\right)=\begin{cases}0 & \text{si } x<\frac{1}{n}\\ 1 & \text{si } x\geq\frac{1}{n}\end{cases}$$

$$\to \lim_{n\to\infty} F_n(x)=\begin{cases}0 & \text{si } x<0\\ 1 & \text{si } x\geq 0\end{cases}$$ que corresponde a la FDA de $X=0$

En conclusión: $a_n \xrightarrow{d} 0$.



Ejemplo: Si $X_1,\ldots,X_n$ es una muestra aleatoria de una distribución Pareto-II que empieza en cero. Investigue la convergencia de $Z_n=nY_1$, donde $Y_1=\min(X_1,\ldots,X_n)$. Nota: $f(x)=\frac{\beta}{(1+x)^{\beta+1}}I_{[0,\infty)}(x)$

$$F_n(z) = P(Z_n \le z) = P(nY_1 \le z) = P\left(Y_1 \le \frac{z}{n}\right) = 1 - \frac{1}{\left(1+\frac{z}{n}\right)^{n\beta}} = 1 - \left[\left(1+\frac{z}{n}\right)^n\right]^{\beta}$$

pero: $e^x = \lim\limits_{n\to\infty}\left(1+\frac{x}{n}\right)^n$

$$\therefore \lim_{n\to\infty} F_n(z) = \lim_{n\to\infty} 1 - \left[\left(1+\frac{z}{n}\right)^n\right]^{\beta} = 1-\left(e^z\right)^{-\beta} = 1-e^{-\beta z} \to \frac{d\left(1-e^{-\beta z}\right)}{dz} = \beta e^{-\beta z}$$

$\to Z_n = nY_1 \to Z_n \xrightarrow{d} Z$, donde $Z \sim \text{Exponencial}(\text{Tasa} = \beta)$

Ejemplo: Si $X_1,\ldots,X_n$ es una muestra aleatoria de una distribución Pareto que empieza en cero. Investigue la convergencia de $Z_n = Y_1$.

$$F_n(z) = P(Z_n \le z) = P(Y_1 \le z) = 1 - \frac{1}{(1+z)^{n\beta}}$$

$$F_n(z) = \begin{cases} 0 & \text{si } z<0 \\ 1-\dfrac{1}{(1+z)^{n\beta}} & \text{si } z \ge 0 \end{cases} = \begin{cases} 0 & \text{si } z<0 \\ 0 & \text{si } z=0 \\ 1 & \text{si } z>0 \end{cases} = \begin{cases} 0 & \text{si } z<0 \\ 1 & \text{si } z \ge 0 \end{cases}$$

$F_n(z)$ es la FDA de cero.

En consecuencia $Z_n = Y_1 \xrightarrow{d} 0$.

Observación: La convergencia en probabilidad es más fuerte que la convergencia en distribución.

$X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$ pero no necesariamente al revés.

Ejemplo: de X_n que converge en distribución a X pero X_n no converge en probabilidades a X.

Considerar $X_1, X_2, \ldots$ independientes con distribución común $P(X_i = \pm 1) = \frac{1}{2}$. Sea X independiente de $X_1, X_2, \ldots$ con la misma distribución $P(X = \pm 1) = \frac{1}{2}$.

$$\lim_{n\to\infty} F_n(x) = F(x) = \begin{cases} 0 & \text{si } x<-1 \\ \dfrac{1}{2} & \text{si } -1 \le x < 1 \\ 1 & \text{si } x \ge 1 \end{cases} \qquad \to X_n \xrightarrow{d} X$$

Considerando

$$P\left(\left|X_n - X\right| \geq 2\right) = P\left(\left|X_n - X\right| = 2\right) = P\left(X_n = 1, X = -1\right) + P\left(X_n = -1, X = 1\right) =$$

$$= P\left(X_n = 1\right) \times P\left(X = -1\right) + P\left(X_n = -1\right) \times P\left(X = 1\right) = \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} = \frac{1}{2}$$

$$\lim_{n \to \infty} P\left(\left|X_n - X\right| \geq 2\right) = \lim_{n \to \infty} \frac{1}{2} \neq 0 \to$$

X_n no converge en probabilidades a X .

Observación

1. $X_n \xrightarrow{P} X \;\Rightarrow\; X_n \xrightarrow{d} X$
2. $X_n \xrightarrow{d} X$ no necesariamente $X_n \xrightarrow{P} X$
3. $X_n \xrightarrow{P} c \;\Leftrightarrow\; X_n \xrightarrow{d} c$
4. $X_n \xrightarrow{P} a$ y $Y_n \xrightarrow{d} Y$ entonces $\begin{pmatrix} X_n \\ Y_n \end{pmatrix} \xrightarrow{d} \begin{pmatrix} a \\ Y \end{pmatrix}$
5. En lugar de variables se puede usar estimadores $\hat{\theta}_n \to \theta$

Teorema del mapeo continuo

Si $g : \mathbb{R} \to \mathbb{R}$ es continua:

1. $X_n \xrightarrow{P} X \;\Rightarrow\; g\left(X_n\right) \xrightarrow{P} g\left(X\right)$
2. $X_n \xrightarrow{d} X \;\Rightarrow\; \mathrm{g}\left(X_n\right) \xrightarrow{d} \mathrm{g}\left(X\right)$

Cuando se dice que $X_n \xrightarrow{P} X$ se quiere decir que cuando n crece X_n se acerca a X .

Cuando se dice que $X_n \xrightarrow{d} X$ se quiere decir que cuando n crece la distribución de X_n se acerca a la de X .

Una cosa interesante

$X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y \Rightarrow X_n + Y_n \xrightarrow{P} X + Y$

$X_n \xrightarrow{d} X$, $Y_n \xrightarrow{d} Y$ No necesariamente $X_n + Y_n \xrightarrow{d} X + Y$

Ejemplo

$X_1, X_2, \ldots$ son independientes e idénticamente distribuidas de acuerdo a $\text{Normal}\left(0,1\right)$ y $X \sim \text{Normal}\left(0,1\right)$ es independiente de las $X_1, X_2, \ldots$.

En este caso $X_n \xrightarrow{d} X$.

Si para $n = 1, 2, \ldots$ se tiene que $Y_n = -X_n$ y se cumple que las $Y_1, Y_2 \ldots$ son independientes e idénticamente distribuidas de acuerdo a $\text{Normal}(0,1)$ en consecuencia, $Y_n \xrightarrow{d} Y$ donde $Y \sim \text{Normal}(0,1)$ no se necesita que $Y = -X$.

Como $Y_n = -X_n$ se cumple que $X_n + Y_n = 0$ para $n = 1, 2, \ldots$ por lo tanto, $X_n + Y_n \xrightarrow{d} 0$ mientras $X + Y \sim \text{Normal}(0, 2)$

$$X_n \xrightarrow{P} a \text{ y } Y_n \xrightarrow{d} Y \Rightarrow \begin{pmatrix} X_n \\ Y_n \end{pmatrix} \xrightarrow{d} \begin{pmatrix} a \\ Y \end{pmatrix}$$

Teorema de Slutsky

Si $X_n \xrightarrow{P} a$ y $Y_n \xrightarrow{d} Y$ entonces:

1. $X_n + Y_n \xrightarrow{d} a+Y$
2. $X_n Y_n \xrightarrow{d} aY$
3. $\dfrac{Y_n}{X_n} \xrightarrow{d} \dfrac{Y}{a}$ $(\text{si } a \neq 0)$

Ejemplo

Sean las v.as $Z \sim N(0,1)$ y $V \sim \chi^2_{(m)}$ son independientes, entonces la v.a $X = \dfrac{Z}{\sqrt{\dfrac{V}{m}}} \sim t_{(m)}$. Verifique que $X_m \xrightarrow{d} Z$ donde $Z \sim \text{Normal}(0,1)$.

Recordar que $E(V) = m$ y $\text{Var}(V) = 2m$ entonces $\dfrac{V}{m}$ es tal que $E\left(\dfrac{V}{m}\right) = 1$ y $\text{Var}\left(\dfrac{V}{m}\right) = \dfrac{1}{m^2}\text{Var}(V) = \dfrac{2}{m}$ entonces:

$$\lim_{m\to\infty} P\left(\left|\frac{V}{m} - 1\right| > \varepsilon\right) \leq \lim_{m\to\infty} \frac{\text{Var}\left(\frac{V}{m}\right)}{\varepsilon^2} = \lim_{m\to\infty} \frac{2}{m\varepsilon^2} = 0$$, en consecuencia $\dfrac{V}{m} \xrightarrow{P} 1$ y por el Teorema de mapeo continuo $\sqrt{\dfrac{V}{m}} \xrightarrow{P} \sqrt{1} = 1$.

Como Z no depende de n entonces $Z = Z_n \xrightarrow{d} Z \sim \text{Normal}(0,1)$.

En conclusión y de acuerdo al Teorema de Slutsky:

$$X = \frac{Z}{\sqrt{\frac{V}{m}}} \xrightarrow{d} \frac{Z}{1} = Z \sim \text{Normal}(0,1)$$

Ejemplo:

Anteriormente se demostró que Si $X_1,\ldots,X_n$ es una muestra aleatoria de una distribución Pareto-II que empieza en cero entonces $Z_n = nY_1 \to Z_n \xrightarrow{d} Z$, donde $Z \sim \text{Exponencial}(\text{Tasa} = \beta)$.

Considerando que $Z_n = Y_1 = \frac{1}{n} \times nY_1$ y teniendo en cuenta que $\frac{1}{n} \xrightarrow{P} 0$ entonces $Z_n = Y_1 = \frac{1}{n} \times nY_1 = 0 \times nY_1$. Se tiene que $\frac{1}{n} \xrightarrow{P} 0$ y $nY_1 \xrightarrow{d} Z$, donde $Z \sim \text{Exponencial}(\text{Tasa} = \beta)$ entonces por la parte 2 del Teorema de Slutsky $Y_1 \xrightarrow{d} 0$. Notar que esta es otra manera de llegar a esta conclusión.

Teorema

Sea $X_1, X_2, \ldots$ es una secuencia de variables aleatorias con $F_n(x) = P(X_n \le x)$ y sea $\Psi_n(t)$ la fgm para X_n. Considere que X es una v.a con $F(x) = P(X \le x)$ y fgm $\Psi(t)$ entonces $\underbrace{\lim_{n\to\infty} \Psi_n(t) = \Psi(t)}_{\forall t \text{ en alguna vecindad de cero}} \Rightarrow \underbrace{\lim_{n\to\infty} F_n(x) = F(x)}_{\text{en todos los puntos de continuidad}}$.

Ejemplo

Suponga que $X_1, X_2, \ldots$ son independientes con $X_n \sim \text{Binomial}\left(n, \frac{\lambda}{n}\right)$ para algún $\lambda > 0$. Investigar la convergencia en distribución de X_n.

$$\Psi_n(t) = \left(1 - \frac{\lambda}{n} + \frac{\lambda}{n} e^t\right)^n = \left[1 - \frac{\lambda\left(1 - e^t\right)}{n}\right]^n$$

$$\lim_{n\to\infty} \Psi_n(t) = \lim_{n\to\infty} \left[1 - \frac{\lambda\left(1 - e^t\right)}{n}\right]^n = e^{\lambda\left(e^t - 1\right)}$$

En conclusión $X_n \to X$ donde $X \sim \text{Poisson}(\lambda)$

Normalidad Asintótica

Se dice que $X_1, X_2, \ldots$ son tales que $X_n \overset{\text{Asint}}{\sim} \text{Normal}\left(\mu_n, \sigma_n^2\right)$, esto significa que $\frac{X_n - \mu_n}{\sigma_n} \overset{d}{\sim} Z \sim \text{Normal}(0,1)$.

Un ejemplo de normalidad asintótica se presenta cuando se aplica el Teorema del Límite Central con una muestra de una distribución que tiene varianza finita σ^2. En este caso $\bar{X}_1, \bar{X}_2, \ldots$ son tales que $\bar{X}_n \overset{\text{Asint}}{\sim} \text{Normal}\left(\mu, \dfrac{\sigma^2}{n}\right)$ lo cual significa que $\dfrac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \overset{d}{\sim} Z \sim \text{Normal}(0,1)$.

Otro ejemplo es el siguiente: $S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})$ tiene distribución normal asintótica con $\mu_n = \sigma^2$ y $\sigma_n^2 = \dfrac{1}{n}\left[\mu_4 - \dfrac{n-3}{n-1}\sigma^4\right] = \dfrac{E\left[(X-\mu)^4\right]}{n} - \dfrac{n-3}{n(n-1)}\left(\sigma^2\right)^2$.

Ejercicios

1. Sea X ~ Poisson (θ) y q(θ) = $e^{-3\theta}$. Demuestre que T(X) = $(-2)^X$ es insesgado pero absurdo para q(θ).
2. Si $X_1, \ldots, X_4$ es una muestra aleatoria de una distribución $\text{Binomial}(1, p)$ ¿Es $\theta = (-2)^T$, donde $T = \sum_{i=1}^{4} X_i$ un estimador insesgado de $\theta = (1-3p)^4$? Justifique su respuesta.
3. Sea $X_1, \ldots, X_n$ una m.a de una distribución $Unif(0, \theta)$ y $Y_1, \ldots, Y_n$ las correspondientes estadísticas de orden. Si n = 6 halle el esperado de Y_5 y determine un estimador insesgado de $\dfrac{1}{\theta^2}$.
4. Sean $X_1, \ldots, X_{n1}$ y $Y_1, \ldots, Y_{n2}$ dos m.a independientes extraídas de una población N(u, σ^2) con medias muestrales $\bar{X}$ y $\bar{Y}$ respectivamente. Un investigador pretende estimar la media poblacional u y propone como estimadores alternativos: T_1 = $\hat{u}_1$ = $\dfrac{1}{2}(\bar{X} + \bar{Y})$ y T_2 = $\hat{u}_2$ = $\dfrac{n_1\bar{X} + n_2\bar{Y}}{n_1 + n_2}$. Comparar las propiedades de insesgamiento y eficiencia de estimadores.
   a. ¿Es $T = \sum_{i=1}^{5} X_i$ más eficiente que $T = \dfrac{5}{4}\sum_{i=1}^{4} X_i$? Justifique.
5. Sea $\bar{X}_1$ la media de una muestra aleatoria de tamaño n de una distribución $N(\mu, \sigma_1^2)$, y $\bar{X}_2$ la media de una muestra aleatoria de tamaño n de una distribución $N(\mu, \sigma_2^2)$.¿Para qué valor de α la variancia de $T = \alpha\bar{X}_1 + (1-\alpha)\bar{X}_2$ es mínima? Justifique su respuesta.

6. Sea $X_1,\ldots,X_n$ una m.a de una distribución $Unif\left(0,\theta\right)$ y $Y_1,\ldots,Y_n$ las correspondientes estadísticas de orden. ¿Es $T_1 = \dfrac{n+1}{n-1}R$, donde $R = Y_n - Y_1$, un estimador, de θ, más eficiente que $T_2 = \dfrac{n+1}{n}Y_n$.

7. Sea $X_1,\ldots,X_n$ una muestra aleatoria de la distribución $f\left(x\right) = \dfrac{2x}{\theta^2} I_{(0,\theta)}\left(x\right)$ y $Y_1,\ldots,Y_n$ las respectivas estadísticas de orden. ¿Es $T_1 = \dfrac{3}{2}\overline{X}$ más eficiente que $T_2 = \dfrac{2n+1}{2n}Y_n$?

8. Si X_1, X_2, X_3 es una m.a de una distribución $Unif\left(\theta - \dfrac{1}{2}, \theta + \dfrac{1}{2}\right)$.¿Es $T_1 = \overline{X}$ más eficiente que la mediana para estimar a θ ? Justifique.

9. En la reserva nacional A la longitud máxima en cm (incluida la cola) de las ardillas grandes, en 20 mediciones, tiene distribución del valor máximo tipo II de Fréchet y su densidad es $f(x) = \dfrac{k}{\mu}\left(\dfrac{\mu}{x}\right)^{k+1} e^{-\left(\frac{\mu}{x}\right)^k} I_{[0,\infty[}(x)$. Suponga que la distribución es $\left[\text{FréchetII}\left(\mu, k=10\right)\right]$. Qué estimador de μ es mejor $T_1 = 2X_1$ o $T_2 = \overline{X}$.

10. Sea $X_1,\ldots,X_n$ una m.a de $N\left(0,\sigma^2\right)$ y se definen los siguientes estimadores de σ^2: $A = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$, $B = \left(\dfrac{n-2}{n}\right)S^2 + 2\overline{X}^2$ y $C = a\,A + \left(1-a\right)B$. Donde a es una constante.
a. Encuentre el valor de a que minimice la Var (C).
b. ¿Cuál de los tres estimadores es más eficiente? **Nota**: Para C considere el valor de a hallado en la subpregunta a.

11. Sea $X_1,\ldots,X_n$ una m.a de N(u, σ^2). Demuestre que $S^2 = \dfrac{\sum(X_i - \overline{X})^2}{n-1}$ y $S_1^2 = \dfrac{\sum(X_i - \overline{X})^2}{n}$ son estimadores consistentes de σ^2.

12. Si $X_1,\ldots,X_n$ es una m.a de una distribución cualquiera con variancia σ^2 y cuarto momento poblacional, en torno a la media , finito. Diga si $S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ o $S_1^2 = \dfrac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$ son estimadores consistentes de

σ^2. Justifique su respuesta con la definición de consistencia. NOTA: $\frac{1}{n}\left(\mu_4 - \frac{n-3}{n-1}\sigma^4\right)$

13. Suponga que θ_1 y θ_2 son estimadores del parámetro $\theta > 0$. Se sabe que $E(\theta_1) = \frac{\theta}{4}$, $E(\theta_2) = \frac{\theta}{3}$, $Var(\theta_1) = 5$ y $Var(\theta_2) = 6$. ¿Qué estimador es mejor? Justifique su respuesta.

14. El tiempo que dura una pareja de enamorados universitarios tiene distribución Exponencial(con media θ). Si $X_1,\ldots,X_n$ es una muestra aleatoria de esa distribución ¿Es $T = \frac{\sum_{i=1}^{n} X_i}{n+4}$ un estimador consistente de θ? Justifique.

15. Obtenga un estimador consistente de $P(X=0) = e^{-\lambda}$ en un muestreo de una distribución Poisson con parámetro λ.

16. Si $X_1,\ldots,X_n$ es una muestra aleatoria de una distribución con media μ y variancia σ^2. ¿Es $T = \frac{\sum_{i=1}^{n} i\,X_i}{2n(n+1)}$ un estimador consistente de μ? Justifique su respuesta.

17. Suponga que $E(T_1) = E(T_2) = \theta$, $V(T_1) = \sigma_1^2$ y $V(T_2) = \sigma_2^2$. Se define un nuevo estimador: $T_3 = aT_1 + (1-a)T_2$ ¿Cómo debe ser la constante a para minimizar el error cuadrático medio de T_3 si T_1 y T_2 no son independientes y son tales que $Cov(T_1, T_2) = c \neq 0$?

18. Sea X~Binomial(n,p). Si $T = \frac{x+1}{n+2}$ estima a p, obtenga el sesgo y el error cuadrático medio de T.

19. Suponga que las v.as $W_1,\ldots,W_n$ satisfacen: $W_i = \theta d_i + \varepsilon_i$, $i = 1,2,\ldots,n$. Donde las d_i son constantes fijadas y las ε_i son v.as i.i.d con distribución $N(0,\sigma^2)$, con σ^2 desconocida. Si se tienen los siguientes estimadores de θ: $A = \frac{\sum_{i=1}^{n} d_i W_i}{\sum_{i=1}^{n} d_i^2}$, $B = \frac{\sum_{i=1}^{n} W_i}{\sum_{i=1}^{n} d_i}$ y $C = \frac{1}{n}\sum_{i=1}^{n}\frac{W_i}{d_i}$.

a) Determine la distribución de C.
b) ¿Son A, B y C estimadores insesgados de θ? Justifique.
c) ¿Son A, B y C estimadores consistentes de θ? Justifique.
d) ¿Es más eficiente A que B? Justifique.

20. El tiempo de falla en años de una computadora de marca C tiene distribución Weibull$\left(r,\theta=2,\mu=2\right)$. Sea 5.8, 6.2, 6.4, 5.4, 5.2 una m.a de esa distribución. Si $T_1=2\bar{X}$ y $T_2=\bar{X}-\sqrt{\pi}$ son estimadores de r. Diga que estimador es mejor.
NOTA: Si $X\sim$ Weibull$\left(r,\theta,\mu\right)\rightarrow$
$f\left(x\right)=\dfrac{\mu}{\theta}\left(\dfrac{x-r}{\theta}\right)^{\mu-1}\exp\left[-\left(\dfrac{x-r}{\theta}\right)^{\mu}\right]I_{[r,\infty)}\left(x\right)$, $E\left(X\right)=r+\theta\,\Gamma\left(1+\dfrac{1}{\mu}\right)$ y
$Var\left(X\right)=\theta^2\left\{\Gamma\left(1+\dfrac{2}{\mu}\right)-\left[\Gamma\left(1+\dfrac{1}{\mu}\right)\right]^2\right\}$.

21. Según los Psicólogos de la escuela A el tiempo en segundos entre una expresión grotesca y otra en cierto programa "cómico" tiene densidad $f\left(x\right)=\dfrac{6x^5}{\delta^6}I_{(0,\delta)}\left(x\right)$. Si $X_1,\ldots,X_n$ es una muestra aleatoria de esa densidad y $M=Y_n$ y $P=\dfrac{7\bar{X}}{6}$ son dos estimadores de δ ; ¿para qué valores de n M es mejor estimador que P?

22. Sea $X_1,\ldots,X_n$ una m.a de una distribución $Unif\left(0,\theta\right)$ y $Y_1,\ldots,Y_n$ las correspondientes estadísticas de orden. ¿Es $T_1=\dfrac{n+1}{n-1}R$, donde $R=Y_n-Y_1$, un estimador, de θ, más eficiente que $T_2=\dfrac{n+1}{n}Y_n$?. Justifique

23. La longitud total del lagarto enano juvenil en cm tiene densidad $f\left(x\right)=\dfrac{1}{\lambda}e^{\frac{-1}{\lambda}(x-\theta)}I_{(\theta,\infty)}\left(x\right)$. Suponga que $\lambda=75$. Si Ud. tiene que elegir entre la media muestral y el mínimo de la muestra para estimar a θ ¿Cuál escogería? Justifique su respuesta.

24. El ingreso mensual de un padre de familia que vive en cierto distrito tiene densidad $f\left(x\right)=\dfrac{4x^3}{\theta^4}I_{(0,\theta)}\left(x\right)$. Sea $X_1,\ldots,X_n$ una muestra aleatoria de esa distribución y $Y_1,\ldots,Y_n$ las respectivas estadísticas de orden. Se tiene que elegir entre la media muestral y el máximo de la muestra para estimar a θ ¿Qué estimador recomendaría Ud.? Justifique su respuesta.

25. Haga la deducción del valor por el que hay que multiplicar a la $\bar{X}$ para que estime a μ con error cuadrático medio mínimo. En particular si se conoce que $\sigma^2=k\mu^2$.

26. Sea $X_1,\dots,X_n$ una m.a de $N\left(\mu,\sigma^2\right)$ y se definen los siguientes estimadores de σ^2: $A=\frac{1}{n+1}\sum_{i=1}^{n}(X_i-\bar{X})^2$, $B=\frac{1}{n+2}\sum_{i=1}^{n}(X_i-\bar{X})^2$. ¿Cuál de estos estimadores es mejor? Presente la justificación de su respuesta.
27. ¿Es el tercer momento muestral alrededor de cero un estimador consistente del tercer momento poblacional alrededor de cero?

28. $X_1,\dots,X_n$ es una m.a de BIN (1, p). Considere los siguientes estimadores de p:
$\hat{p}_1 = X_1$, $\hat{p}_2 = \frac{1}{n}\sum_{i=1}^{n-1} X_i$, $\hat{p}_3 = \frac{1}{n}\sum_{i=4}^{n} X_i$

a. ¿Cuál (es) de estos estimadores es insesgado?
b. ¿Cuál (es) de estos estimadores es consistente?
c. Usando los resultados de a) y b), modificar el estimador sesgado pero consistente para hacerlos insesgados.

29. Si $X_1,\dots,X_n$ es una muestra aleatoria de una distribución exponencial con media θ. ¿Es $T=\frac{1}{n+2}\sum_{i=1}^{n}X_i$ un estimador consistente de θ? Justifique.
30. La longitud total en cm del lagarto blanco de cierta clase III tiene distribución Normal$\left(\mu_1=150,\sigma_1^2\right)$, y la longitud total en cm del lagarto negro de cierta clase III tiene distribución Normal$\left(\mu_2=175,\sigma_2^2\right)$. Sea $X_1,\dots,X_n$ una muestra aleatoria de la longitud total en cm del lagarto blanco y $Y_1,\dots,Y_n$ una muestra aleatoria de la longitud total en cm del lagarto negro. Suponga que ambas longitudes son independientes.

a. Demuestre que $T_1=\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2$ es un estimador consistente de σ_1^2.
b. ¿Cuál de los siguientes estimadores de σ_1^2: $\frac{\sum_{i=1}^{n}(X_{i1}-\bar{X}_1)^2}{n-1}$ y $\frac{\sum_{i=1}^{n}(X_{i1}-\bar{X}_1)^2}{n}$ es mejor. Justifique su respuesta.

c. Halle el estimador consistente de $\frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}$.

31. Sean $X_1,\dots,X_n$ y $Y_1,\dots,Y_n$ v.as independientes de poblaciones con variancias σ_1^2 y σ_2^2 respectivamente. Halle el estimador consistente de $\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}$.
(Justifique su respuesta)

Teniendo en cuenta que S^2 es insesgada:

$$P\left[\left|S^2-\sigma^2\right|\geq\varepsilon\right]\leq\frac{E\left[\left(S^2-\sigma^2\right)^2\right]}{\varepsilon^2}=\frac{Var\left(S^2\right)}{\varepsilon^2}=\frac{\left[\mu_4-\frac{n-3}{n-1}\sigma^4\right]}{n\varepsilon^2}$$

Siempre que exista $\mu_2^{|}$:

$$\lim_{n\to\infty}P\left[\left|S^2-\sigma^2\right|\geq\varepsilon\right]\leq\lim_{n\to\infty}\frac{\left[\mu_4-\frac{n-3}{n-1}\sigma^4\right]}{n\varepsilon^2}=0\rightarrow$$ queda demostrado.

32. Sean $X_1,\ldots,X_n$ y $Y_1,\ldots,Y_n$ v.as independientes de poblaciones con variancias σ_1^2 *y* σ_2^2 respectivamente. Halle los estimadores consistentes de:

   a. $\dfrac{\sigma_1^2}{\sigma_2^2}$
   b. $\dfrac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}$
   c. $\sigma_1^2\sigma_2^2$