

**ESTADÍSTICA TEÓRICA II. Tercer curso de Economía**  
**CURSO 2008/2009. CONVOCATORIA DE SEPTIEMBRE**  
**Código de la Carrera 43 Código de la Asignatura 305.**

**PRIMERA PARTE: CUESTIONES**

1.- Estimador consistente y estimador consistente en media cuadrática. Relación entre ambos.

**Respuesta.-**

Un estimador  $\hat{\theta}_n$  de un parámetro  $\theta$ , obtenido para una muestra de tamaño  $n$ , decimos que es consistente (consistencia simple o en probabilidad) si la sucesión  $\{\hat{\theta}_n\}$  converge en probabilidad a  $\theta$ , es decir que  $\forall \varepsilon > 0$ :

$$\lim_{n \rightarrow \infty} P\left[|\hat{\theta}_n - \theta| < \varepsilon\right] = 1$$

Diremos que es consistente en media cuadrática si se verifica que

$$\lim_{n \rightarrow \infty} E\left[\hat{\theta}_n - \theta\right]^2 = 0$$

Se verifica que si un estimador es consistente en media cuadrática, también lo es en probabilidad pero el recíproco no es necesariamente cierto.

2.- Intervalo de confianza de una proporción para muestras grandes.

**Respuesta.-**

Consideremos una población con una variable que se distribuya  $B(1, p)$  (Bernoulli de parámetro  $p$ ). Sea  $\hat{p}$  la proporción muestral de éxitos para una muestra de tamaño  $n$  suficientemente grande, y  $\hat{q} = 1 - \hat{p}$ .

Sea  $Z$  normal  $N(0, 1)$  y  $z_{\frac{\alpha}{2}}$  tal que  $P\left[Z \geq z_{\frac{\alpha}{2}}\right] = \frac{\alpha}{2}$ .

Entonces un intervalo de confianza aproximado al nivel  $100(1-\alpha)\%$ , para la proporción poblacional  $p$  es:

$$\left[ \hat{p} - z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}}, \hat{p} + z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}} \right]$$

3.- Relacione las propiedades de los estimadores de máxima verosimilitud.

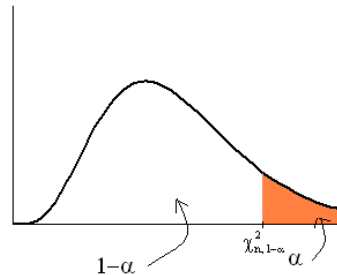
**Respuesta.-**

- 1) Son consistentes.
- 2) En general, son insesgados y en caso contrario, lo son asintóticamente
- 3) Si  $\hat{\theta}$  es eficiente, entonces es máximo verosímil y único. Pero no todo estimador máximo verosímil es eficiente, aunque sí lo es asintóticamente.
- 4) Si  $\hat{\theta}$  es máximo verosímil entonces es asintóticamente normal  $N\left(\theta, \sqrt{\text{Var}(\hat{\theta})}\right)$ , en donde  $\text{Var}(\hat{\theta})$  coincide con la cota de Frechet-Cramer-Rao.
- 5) Si  $\hat{\theta}$  es suficiente, entonces el estimador máximo verosímil de  $\theta$ , si es único, es función de  $\hat{\theta}$ .
- 6) Principio de invarianza de Zehna: Si  $\hat{\theta}$  es máximo verosímil, es invariante por una transformación biunívoca.

4.- Contraste para la varianza de una población  $N(\mu, \sigma)$  con media  $\mu$  conocida siendo la hipótesis nula  $H_0: \sigma^2 \leq \sigma_0^2$  frente a  $H_1: \sigma^2 > \sigma_0$ . Expresa analíticamente la región crítica y la regla de decisión para dicho contraste.

**Respuesta.-**

Sea  $\chi_n^2$  una variable aleatoria Chi cuadrado con n grados de libertad y sea  $\chi_{n,1-\alpha}^2$  tal que  $P[\chi_n^2 > \chi_{n,1-\alpha}^2] = \alpha$



Entonces la regla de decisión es: rechazamos  $H_0$  si  $\chi_{exp}^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{\sigma^2} > \chi_{n,1-\alpha}^2$ .

La propia regla de decisión determina la región crítica  $\chi_{exp}^2 > \chi_{n,1-\alpha}^2 \iff$

$$\iff \sum_{i=1}^n (x_i - \mu)^2 > \sigma^2 \chi_{n,1-\alpha}^2$$

**PROBLEMAS**

1.- Dada la función de densidad  $f(x; \alpha) = \frac{x}{a^2} e^{-\frac{x}{a}}$  siendo  $x \geq 0$  y  $\alpha \geq 0$ . Para muestras aleatorias simples de tamaño n:

a) Halle el estimador de máxima verosimilitud  $\hat{\alpha}$  para el parámetro  $\alpha$ .

b) Sabiendo que la varianza del estimador es  $V(\hat{\alpha}) = \frac{\alpha^2}{n}$  compruebe si el estimador  $\hat{\alpha}$  es eficiente.

**Solución.-**

a) Función de verosimilitud:  $L = \prod_{i=1}^n \left( \frac{x_i}{a^2} e^{-\frac{x_i}{a}} \right) = \frac{1}{a^{2n}} \cdot \prod_{i=1}^n x_i \cdot \prod_{i=1}^n e^{-\frac{x_i}{a}} = \frac{1}{a^{2n}} \cdot \prod_{i=1}^n x_i \cdot e^{-\frac{\sum_{i=1}^n x_i}{a}} \rightarrow$

$\rightarrow \ln L = -2n \ln a + \ln \prod_{i=1}^n x_i - \ln \frac{\sum_{i=1}^n x_i}{a}$ . Derivando respecto de a e igualando a cero:

$$\frac{\partial \ln L}{\partial a} = -\frac{2n}{a} + \frac{\sum_{i=1}^n x_i}{a^2} = 0, \text{ de donde el estimador: } \hat{\alpha} = \frac{\sum_{i=1}^n x_i}{2n} = \frac{\bar{X}}{2}$$

b) Calculemos la cota de Frechet-Cramer-Rao:  $\frac{1}{-nE\left[\frac{\partial^2 \ln f(x;a)}{\partial a^2}\right]}$ .

Se tiene que  $\ln f(x;a) = \ln x - 2\ln a - \frac{x}{a} \rightarrow \frac{\partial \ln f(x;a)}{\partial a} = -\frac{2}{a} + \frac{x}{a^2} \rightarrow$   
 $\rightarrow \frac{\partial^2 \ln f(x;a)}{\partial a^2} = \frac{2}{a^2} - \frac{2x}{a^3}$ .

Calculemos ahora  $E(X) = \int_0^\infty \frac{x^2}{a^2} e^{-\frac{x}{a}} dx = (\text{cambio } \frac{x}{a} = t) = a \int_0^\infty t^2 e^{-t} dt = a\Gamma(3) = 2a$ .

Por tanto:  $E\left[\frac{\partial^2 \ln f(x;a)}{\partial a^2}\right] = E\left[\frac{2}{a^2} - \frac{2X}{a^3}\right] = \frac{2}{a^2} - \frac{2E(X)}{a^3} = -\frac{2}{a^2}$

Así pues, la cota de Frechet-Cramer-Rao sería  $\frac{a^2}{2n}$ .

Si, de acuerdo con el enunciado, la varianza del estimador fuese  $\frac{a^2}{n}$ , deberíamos concluir que el estimador no es eficiente.

Nota.- Si calculamos  $E(X^2) = \int_0^\infty \frac{x^3}{a^2} e^{-\frac{x}{a}} dx = (\text{cambio } \frac{x}{a} = t) = a^2 \int_0^\infty t^3 e^{-t} dt = a^2\Gamma(4) = 6a^2$

se tiene que  $\text{Var}(X) = 6a^2 - 4a^2 = 2a^2$ , de donde  $\text{Var}(\hat{a}) = \frac{\text{Var}(\bar{X})}{4} = \frac{\text{Var}(X)}{4n} = \frac{a^2}{2n}$  de donde se deduce que el estimador sí que es eficiente.

2.- Dada la función de densidad  $f(x; \lambda) = \lambda e^{-\lambda x}$  siendo  $x \geq 0$  se desea contrastar la hipótesis nula siendo  $H_0: \lambda = 1$  frente a  $H_1: \lambda = 5$  con un nivel de significación del 10% mediante una muestra aleatoria de tamaño uno.

- Construya la mejor región crítica utilizando el lema de Neyman-Pearson.
- Calcule la potencia del contraste.
- Si el valor muestral es 2,5 decida qué hipótesis se acepta.

**Solución.-**

a) La mejor región crítica será de la forma  $\frac{L(x;1)}{L(x;5)} = \frac{e^{-x}}{5e^{-5x}} = \frac{e^{4x}}{5} \leq k$ . Tomando

logaritmos y despejando  $x$  se obtiene una región crítica de la forma  $x \leq K$ . Puesto que  $\alpha = 0,1$  será  $0,1 = \int_0^K e^{-x} dx = 1 - e^{-K}$  de donde  $e^{-K} = 0,9 \rightarrow K = -\ln 0,9 \cong 0,105$ . La mejor región crítica sería el intervalo  $[0; 0,105]$ .

b) La potencia del contraste es  $P[X \leq 0,105 / \lambda = 5] = 5 \int_0^{0,105} e^{-5x} dx = -[e^{-5x}]_0^{0,105} = 1 - 0,9^5 \cong 0,4095$ .

c) De acuerdo con el apartado a, aceptaremos la hipótesis nula,  $\lambda = 1$ .