# Ejercicios de Estimación

Bernardo D'Auria

Departamento de Estadística

Universidad Carlos III de Madrid

GRUPO MAGISTRAL

GRADO EN INGENIERÍA DE SISTEMAS AUDIOVISUALES

Otros

# Ejercicio Sep. 2001

Los siguientes datos corresponden a la longitud (en milímetros) de piezas fabricadas por una máquina

| | | | | |
|---|---|---|---|---|
| 104 | 109 | 111 | 109 | 87 |
| 86 | 80 | 119 | 88 | 122 |
| 91 | 103 | 99 | 108 | 96 |
| 104 | 98 | 98 | 83 | 107 |
| 79 | 87 | 94 | 92 | 97 |

Sabiendo que $\sum_i x_i = 2451$ y $\sum_i x_i^2 = 243505$:

a) Calcular un estimador insesgado para la media de la población.

b) Calcular un estimador insesgado para la varianza de la población.

# Ejercicio Sep. 2001

Los siguientes datos corresponden a la longitud (en milímetros) de piezas fabricadas por una máquina

| | | | | |
|---|---|---|---|---|
| 104 | 109 | 111 | 109 | 87 |
| 86 | 80 | 119 | 88 | 122 |
| 91 | 103 | 99 | 108 | 96 |
| 104 | 98 | 98 | 83 | 107 |
| 79 | 87 | 94 | 92 | 97 |

Sabiendo que $\sum_i x_i = 2451$ y $\sum_i x_i^2 = 243505$:

a) Calcular un estimador insesgado para la media de la población.

b) Calcular un estimador insesgado para la varianza de la población.

## SOLUCIÓN:

a) $\hat{\mu} = 98.04$;

b) $\hat{\sigma^2} = \frac{n}{n-1}\left(\overline{x^2} - \bar{x}^2\right) = 133.71$.

# Ejercicio

M12

Sean $X_1, \ldots, X_n, X_n + 1, \ldots, X_{n+m}$ variables aleatorias independientes e idénticamente distribuidas con densidad *Normal* de media $\mu$ y varianza $\sigma^2$. Se desea estimar $\mu$, pero los valores individuales de la variables se han extraviado y se dispone sólo de las medias

$$\bar{x}_1 = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \bar{x}_2 = \frac{1}{m}\sum_{i=n+1}^{n+m} x_i.$$

Para integrar las dos fuentes de información se utiliza un estimador de la forma

$$\hat{x} = \lambda \bar{x}_1 + (1-\lambda)\bar{x}_2,$$

donde $0 \leq \lambda \leq 1$.
Probar que un estimador de este tipo es centrado para $\mu$ indicando además qué valor define el mejor de todos ellos.

# Ejercicio

M12

Sean $X_1, \dots, X_n, X_n + 1, \dots, X_{n+m}$ variables aleatorias independientes e idénticamente distribuidas con densidad *Normal* de media $\mu$ y varianza $\sigma^2$. Se desea estimar $\mu$, pero los valores individuales de la variables se han extraviado y se dispone sólo de las medias

$$\bar{x}_1 = \frac{1}{n} \sum_{i=1}^{n} x_i, \quad \bar{x}_2 = \frac{1}{m} \sum_{i=n+1}^{n+m} x_i.$$

Para integrar las dos fuentes de información se utiliza un estimador de la forma

$$\hat{x} = \lambda \bar{x}_1 + (1 - \lambda)\bar{x}_2,$$

donde $0 \leq \lambda \leq 1$.
Probar que un estimador de este tipo es centrado para $\mu$ indicando además qué valor define el mejor de todos ellos.

## SOLUCIÓN:

$$\mathbb{E}[\hat{x}] = \mu;$$

$$\lambda = \frac{n}{n+m}.$$

# Ejercicio

I3

Se desea estimar la media $\mu$ de una variable aleatoria $X$. Para ello se toman $10$ datos y se calcula su media muestral $\bar{X}$ y la varianza de dichos datos $s_X^2$.
Comentar si son verdaderas o falsas las siguientes afirmaciones:

a) Por el teorema central del límite sabemos que $\mu$ será una variable aleatoria normal

b) Un estimador de $X$ es $\bar{X}$

c) La media muestral con un conjunto de datos es un número y no una variable aleatoria

d) La media muestral $\bar{X}$ tiene siempre una distribución en el muestreo según una normal

e) Si $X$ es normal, $\bar{X}$ es siempre normal

f) Para disminuir la varianza de $\bar{X}$ a la mitad habría que tomar al menos $100$ datos

g) Como la media muestral es un estimador insesgado, tenemos asegurado que $\bar{X} = \mu$

# Ejercicio

I3

Se desea estimar la media $\mu$ de una variable aleatoria $X$. Para ello se toman $10$ datos y se calcula su media muestral $\bar{X}$ y la varianza de dichos datos $s_X^2$.
Comentar si son verdaderas o falsas las siguientes afirmaciones:

a) Por el teorema central del límite sabemos que $\mu$ será una variable aleatoria normal
b) Un estimador de $X$ es $\bar{X}$
c) La media muestral con un conjunto de datos es un número y no una variable aleatoria
d) La media muestral $\bar{X}$ tiene siempre una distribución en el muestreo según una normal
e) Si $X$ es normal, $\bar{X}$ es siempre normal
f) Para disminuir la varianza de $\bar{X}$ a la mitad habría que tomar al menos $100$ datos
g) Como la media muestral es un estimador insesgado, tenemos asegurado que $\bar{X} = \mu$

## SOLUCIÓN:

c) Verdadera
e) Verdadera

# Ejercicio

I5

El tiempo $T$ ( en segundos) que un ordenador tarda en ejecutar una tarea sigue una variable aleatoria continua de función de densidad

$$f(t) = \frac{\alpha}{t^{\alpha+1}},$$

con $t \geq 1$ y $\alpha > 1$.

a) Utilizando el método de los momentos, propón un estimador para el parámetro $\alpha$.

b) Se ejecuta $5$ veces la tarea, y se cronometra el tiempo que ha tardado cada vez. Estos tiempos son (en segundos):

$$6, 5, 3, 7, 2.$$

Basándonos en esta muestra, y el estimador de $\alpha$ anterior, estima la probabilidad de que se tarde más de $5$ segundos en realizar la tarea.

# Ejercicio

I5

El tiempo $T$ ( en segundos) que un ordenador tarda en ejecutar una tarea sigue una variable aleatoria continua de función de densidad

$$f(t) = \frac{\alpha}{t^{\alpha+1}},$$

con $t \geq 1$ y $\alpha > 1$.

a) Utilizando el método de los momentos, propón un estimador para el parámetro $\alpha$.

b) Se ejecuta $5$ veces la tarea, y se cronometra el tiempo que ha tardado cada vez. Estos tiempos son (en segundos):

$$6, 5, 3, 7, 2.$$

Basándonos en esta muestra, y el estimador de $\alpha$ anterior, estima la probabilidad de que se tarde más de $5$ segundos en realizar la tarea.

## Solución:

a) $\hat{\alpha} = \bar{T}/(\bar{T} - 1)$;

b) $\Pr = 0.127$.

# Ejercicio

En una población hay individuos de *tres* tipos con las siguientes probabilidades:

$$\begin{aligned} \Pr(\text{Tipo1}) &= \theta^2 \\ \Pr(\text{Tipo2}) &= 2\theta(1-\theta) \quad 0 \le \theta \le 1 \\ \Pr(\text{Tipo3}) &= (1-\theta)^2 \end{aligned}$$

Obtenemos una muestra aleatoria de tamaño $n$, resultando $N_1$ individuos de Tipo $1$, $N_2$ individuos del Tipo $2$ y los restantes de Tipo $3$.

a) Estudiar si el estimador $\hat{T} = \frac{2N_1+N_2}{2n}$ es insesgado para estimar $\theta$.

b) Hallar el estimador de máxima verosimilitud de $\theta$.

# Ejercicio

En una población hay individuos de *tres* tipos con las siguientes probabilidades:

$$
\begin{aligned}
\Pr(\text{Tipo1}) &= \theta^2 \\
\Pr(\text{Tipo2}) &= 2\theta(1-\theta) \quad 0 \leq \theta \leq 1 \\
\Pr(\text{Tipo3}) &= (1-\theta)^2
\end{aligned}
$$

Obtenemos una muestra aleatoria de tamaño $n$, resultando $N_1$ individuos de Tipo $1$, $N_2$ individuos del Tipo $2$ y los restantes de Tipo $3$.

a) Estudiar si el estimador $\hat{T} = \frac{2N_1+N_2}{2n}$ es insesgado para estimar $\theta$.

b) Hallar el estimador de máxima verosimilitud de $\theta$.

## Solución:

a) $\text{Sesgo}(\hat{T}) = 0$.

b) $\hat{\theta} = \hat{T}$.

# Ejercicio

Sea $X$ una v.a. que sigue una distribución *Lognormal*, con función de densidad

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma\, x} \exp\left\{-\frac{1}{2\sigma^2}(\log x - \mu)^2\right\}$$

a) Calcular el estimador de máxima verosimilitud para $\mu$.

b) Se sabe que el número de descendientes de las hembras de cierta especie de insectos sigue aproximadamente una distribución Lognormal. Si $20$ hembras elegidas al azar tuvieron como descendencia

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 47 | 34 | 28 | 44 | 23 | 32 | 39 | 27 | 36 | 33 |
| 28 | 32 | 29 | 35 | 30 | 37 | 41 | 26 | 52 | 31 |

Determinar una estimación de $\mu$.

# Ejercicio

Sea $X$ una v.a. que sigue una distribución *Lognormal*, con función de densidad

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma\, x} \exp\left\{ -\frac{1}{2\sigma^2} (\log x - \mu)^2 \right\}$$

a) Calcular el estimador de máxima verosimilitud para $\mu$.

b) Se sabe que el número de descendientes de las hembras de cierta especie de insectos sigue aproximadamente una distribución Lognormal. Si $20$ hembras elegidas al azar tuvieron como descendencia

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 47 | 34 | 28 | 44 | 23 | 32 | 39 | 27 | 36 | 33 |
| 28 | 32 | 29 | 35 | 30 | 37 | 41 | 26 | 52 | 31 |

Determinar una estimación de $\mu$.

## SOLUCIÓN:

a) $\hat{\mu} = \sum \log(x_i)/n$, $\mathbb{V}\text{ar}[\hat{\mu}] = \sigma^2/n$.

b) $\hat{\mu} = 3.51$

# Ejercicio

Sea $X \sim \Gamma(2, \theta)$, donde la función de densidad de una v.a. con distribución Gamma $\Gamma(\alpha, \theta)$ y parámetros $\alpha$ y $\theta$ es

$$f(x) = \frac{\theta^{-\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-x/\theta}, \quad x > 0.$$

Calcular el estimador máximo verosímil para $\theta$ y comprobar que es insesgado.

# Ejercicio

Sea $X \sim \Gamma(2,\theta)$, donde la función de densidad de una v.a. con distribución Gamma $\Gamma(\alpha,\theta)$ y parámetros $\alpha$ y $\theta$ es

$$f(x) = \frac{\theta^{-\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-x/\theta}, \quad x > 0.$$

Calcular el estimador máximo verosímil para $\theta$ y comprobar que es insesgado.

**SOLUCIÓN:**

$$\hat{\theta} = \frac{\bar{x}}{2}$$

# Ejercicio Examen Ing. Tel. Sep'06 - P2

Se sabe que el tiempo $T$ de respuesta de un servidor web dedicado a la telemedicina se ajusta a una distribución *Rayleigh* de parámetro $\alpha > 0$ con función de densidad

$$f(t) = \begin{cases} \alpha\, t \exp\left(-\frac{\alpha}{2} t^2\right), & t > 0 \\ 0, & t \leq 0 \end{cases}$$

Se realiza un estudio en el que se recogen $n = 50$ tiempos de respuesta $t_i$ (suponer que $n = 50$ es grande) obteniéndose que

$$\sum_{i=1}^{n} t_i = 146.28 \qquad \sum_{i=1}^{n} t_i^2 = 510.58$$

Dar una estimación puntual de $\alpha$ por el método de máxima verosimilitud utilizando los datos muestrales.

# Ejercicio Examen Ing. Tel. Sep'06 - P2

Se sabe que el tiempo $T$ de respuesta de un servidor web dedicado a la telemedicina se ajusta a una distribución *Rayleigh* de parámetro $\alpha > 0$ con función de densidad

$$f(t) = \begin{cases} \alpha\, t \exp\left(-\frac{\alpha}{2} t^2\right), & t > 0 \\ 0, & t \leq 0 \end{cases}$$

Se realiza un estudio en el que se recogen $n = 50$ tiempos de respuesta $t_i$ (suponer que $n = 50$ es grande) obteniéndose que

$$\sum_{i=1}^{n} t_i = 146.28 \qquad \sum_{i=1}^{n} t_i^2 = 510.58$$

Dar una estimación puntual de $\alpha$ por el método de máxima verosimilitud utilizando los datos muestrales.

## SOLUCIÓN:

$$\hat{\alpha} = 2/\overline{t^2} = 0,1959.$$