Estadística I
Grado en Matemáticas, UAM, 2018-2019

Hoja 4. Momentos/máxima verosimilitud. Sesgo y eficiencia de estimadores

Método de momentos y máxima verosimilitud

1. Calcula los estimadores por máxima verosimilitud (para muestras de tamaño n) para
a) el parámetro λ de $X \sim \text{POISS}(\lambda)$;
b) el parámetro p de $X \sim \text{BIN}(N, p)$, con N dado.

2. La función de masa de la variable aleatoria X es

valores	0	1	2	3
probabilidades	$\theta/3$	$1/3$	$1/3$	$(1-\theta)/3$

Aquí, el parámetro $\theta \in [0, 1]$.

En una muestra de X de tamaño 100 se han obtenido los siguientes resultados:

valores	0	1	2	3
número de apariciones	10	27	23	40

Halla la *estimación* de θ por momentos. Halla la *estimación* de θ por máxima verosimilitud.

3. La función de masa de la variable X viene dada por

valores	1	2	3
probabilidades	p_1	p_2	p_3

El modelo tienen únicamente dos parámetros, por ejemplo p_1 y p_2, puesto que se debe cumplir que $p_1 + p_2 + p_3 = 1$. Se tiene que $p_1 \in (0, 1)$, $p_2 \in (0, 1)$ y además que $p_1 + p_2 < 1$.

En una muestra de X de tamaño 100 se han obtenido los siguientes resultados:

valores	1	2	3
número de apariciones	20	25	55

Halla la *estimación* de p_1 y p_2 por máxima verosimilitud.

4. Una variable aleatoria X tiene función de densidad

$$f(x; \theta) = \frac{2}{\theta^2}\, x \quad \text{para } x \in [0, \theta],$$

y $f(x; \theta) = 0$ si $x \notin [0, \theta]$. El parámetro θ es positivo. Halla el estimador por máxima verosimilitud de θ para muestras de tamaño n.

5. Una variable aleatoria X tiene función de densidad

$$f(x; \theta) = \begin{cases} \theta/x^2, & \text{si } x > \theta\,, \\ 0, & \text{si } x \leq \theta\,. \end{cases}$$

El parámetro θ es positivo. Halla los estimadores por momentos y por máxima verosimilitud de θ para muestras de tamaño n.

6. Sea $X \sim \mathcal{N}(\mu, 1)$. El espacio de parámetros (para μ) es el intervalo $[-1, 1]$. Dada una muestra $(x_1, \dots, x_n)$ de X, determina la estimación de máxima verosimilitud de μ.

7. Sea $\Theta = \{0, 1\}$ y considera las dos funciones de densidad (ambas con soporte en $(0, 1)$) alternativas dadas, para $x \in (0, 1)$, por

$$f(x; 0) = 1\,, \qquad f(x; 1) = \frac{1}{2\sqrt{x}}\,.$$

Determina el estimador de máxima verosimilitud del valor de θ para muestras de tamaño n.

Sesgo y eficiencia de estimadores

8. Comprueba que si T_1 y T_2 son estimadores insesgados de un parámetro de θ, entonces, para todo $\lambda \in (0, 1)$, $Z = \lambda T_1 + (1 - \lambda) T_2$ es estimador insesgado de θ.

9. Una variable X sigue una distribución de Rayleigh de parámetro $\sigma^2 > 0$. En un tal modelo se tiene que

$$\mathbf{E}_{\sigma^2}(X) = \sqrt{\pi\sigma^2/2} \qquad \text{y} \qquad \mathbf{E}_{\sigma^2}(X^2) = 2\sigma^2.$$

Como estimadores del parámetro σ^2 (para muestras de tamaño n), se proponen los dos siguientes:

$$T_1(X_1, \dots, X_n) = \frac{2}{\pi}\, \overline{X}^2 \qquad \text{y} \qquad T_2(X_1, \dots, X_n) = \frac{1}{2}\, \overline{X^2}.$$

Comprueba que T_2 es un estimador insesgado de σ^2, mientras que T_1 no. Calcula el sesgo de T_1.

10. Sea X una variable aleatoria con función de densidad/masa $f(x; \theta)$, con $\theta \in \Theta$. Sean T_1 y T_2 dos estimadores insesgados de un parámetro θ que actúan sobre muestras de tamaño n. Formamos el estimador U (que actúa sobre muestras $(X_1, \dots, X_{2n})$ de tamaño $2n$):

$$U(X_1, \dots, X_{2n}) = \frac{1}{2}\big(T_1(X_1, \dots, X_n) + T_2(X_{n+1}, \dots, X_{2n})\big)\,.$$

Comprueba que $\mathbf{V}_\theta(U) = \frac{1}{4}\big(\mathbf{V}_\theta(T_1) + \mathbf{V}_\theta(T_2)\big)$.

11. Tenemos una variable X que toma los valores $\{-1, 0, +1\}$ con probabilidades respectivas

-1	0	$+1$
$(2+\theta)/4$	$\theta/4$	$(2-2\theta)/4$

para $\theta \in [0, 1]$.

a) Consideremos el estadístico N_1 dado por

$$N_1 = h_1(X_1, X_2, \dots, X_n)\,,$$

donde $h_1(x_1, x_2, \dots, x_n) =$ número de $\{x_j = -1\}$. Observa que $N_1 \sim \textsc{bin}(n, (2+\theta)/4)$. Comprueba que el estadístico

$$T_1 = \frac{4}{n}\, N_1 - 2$$

es un estimador insesgado de θ.

b) Consideremos el estadístico N_2 dado por

$$N_2 = h_2(X_1, X_2, \ldots, X_n)\,,$$

donde $h_2(x_1, x_2, \ldots, x_n) =$ número de $\{x_j = 0\}$. Observa que $N_2 \sim \textsc{bin}(n, \theta/4)$. Comprueba que el estadístico

$$T_2 = \frac{4}{n}\,N_2$$

es un estimador insesgado de θ para muestras aleatorias de X de tamaño n. ¿Cuál de los estimadores, T_1 o T_2, es más eficiente?

c) Supongamos que la muestra $(x_1, x_2, \ldots, x_{100})$ tiene las siguientes frecuencias:

-1	0	1
60	16	24

¿Cuál es la estimación de θ si usamos T_1? ¿Y si usamos T_2?

EJERCICIOS ADICIONALES

12. Un arquero (con escasa puntería) dispara n veces a una diana de radio θ (desconocido). En cada lanzamiento logra darle al disco, pero en lugares completamente aleatorios cuyas distancias al centro del disco son $r_1, r_2, \ldots, r_n$. Determina el estimador de máxima verosimilitud del radio del disco.

13. Sea $X \sim \textsc{unif}(a, b)$, con $a < b$. Determina los estimadores de a y b para muestras aleatorias de X de tamaño n por máxima verosimilitud y por el método de momentos.

14. Calcula el estimador por máxima verosimilitud del parámetro $\theta \in [0, 1]$ para muestras aleatorias de tamaño n de la variable X que toma tres valores $-1, 0, +1$ con probabilidades respectivas

-1	0	$+1$
$(2+\theta)/4$	$\theta/4$	$(2-2\theta)/4$

15. La variable X tiene una distribución dada por dos parámetros $\delta > 0$ y $\lambda > 0$, de manera que

$$X = \delta + Y\,, \quad \text{donde } Y \sim \textsc{exp}(\lambda)\,.$$

Los parámetros δ y λ son desconocidos.

a) Sea $m_n = \min(X_1, X_2, \ldots, X_n)$. Comprueba que

$$\mathbf{E}(m_n) = \delta + \frac{1}{n\lambda} \qquad \text{y que} \qquad \mathbf{E}(\overline{X}) = \delta + \frac{1}{\lambda}\,.$$

b) Comprueba que

$$T_1 = \frac{n}{n-1}\,(\overline{X} - m_n) \qquad \text{y} \qquad T_2 = \frac{n}{n-1}\,(m_n - \overline{X}/n)$$

son estimadores insesgado de $1/\lambda$ y δ, respectivamente.

16. Sea T un estimador de un parámetro θ. Supongamos que $\mathbf{E}_\theta(T) = \alpha\theta$ y que $\mathbf{V}_\theta(T) = \beta\theta^2$, donde α y β son dos constantes fijas.

Halla r (en función de α y β) para que $\mathrm{ECM}_\theta(rT)$ sea mínimo. Aplícalo al caso en el que $X \sim \textsc{Unif}[0, a]$ y al estimador $T = \max(X_1, \ldots, X_n)$ de a.