

## Auxiliar 6 MA34B-03

### Problema 1 (control 1)

Suponga que  $x_1, \dots, x_n$  es una m.a.s. de una distribución cuya fdp  $f(x/\theta)$  es:

$$f(x/\theta) = (1/2) \exp(-|x - \theta|) \text{ para } -\infty < x < \infty, \text{ con } \theta \text{ desconocido } (-\infty < \theta < \infty)$$

Muestre en que casos el estimador de máxima verosimilitud de  $\theta^*$  es único.  
(Indicación: Escriba la función de verosimilitud en función de los estadísticos de orden  $x_{(i)}$ . ( $x_{(1)}, x_{(2)}, \dots, x_{(n)}$  es la muestra ordenada de menor a mayor).

Calcule la pendiente de la función a maximizar en cada intervalo  $[x_{(k)}, x_{(k+1)}]$  con  $k$  en  $\{1, \dots, n-1\}$ .

**Sol:**

La función de verosimilitud es

$$f_n(\vec{x}, \theta) = \left(\frac{1}{2}\right)^n e^{-\sum_i |x_i - \theta|}$$

es equivalente a maximizar el logaritmo de  $f_n$ :

$$\ln(f_n) = K - \sum_i |x_i - \theta|$$

la función para  $\theta$  en  $[x_{(k)}, x_{(k+1)}]$  es :

$$-\sum_{i=1}^k (\theta - x_{(i)}) - \sum_{i=k+1}^n (x_{(i)} - \theta)$$

la derivada c/r a  $\theta$  :

$$-k + (n - k) = -2k + n$$

que es positiva para  $k < n/2$  y negativa para  $k > n/2$

Si  $n$  es par, la derivada es cero en el intervalo  $[x_{(\frac{n}{2})}, x_{(\frac{n}{2}+1)}]$  y como la función

es continua en todo  $\mathbb{R}$  entonces es máxima en todo punto de este intervalo. Por lo tanto el estimador de M.V. es único si en la muestra  $x_{(\frac{n}{2})} = x_{(\frac{n}{2}+1)}$ .

Si  $n$  es impar, el máximo se alcanza en  $x_{\left(\frac{n+1}{2}\right)}$ .

## Problema 2

Para las próximas elecciones de la Fech, el presidente debe decidir si repostular o no a un nuevo período. Quiere saber en particular cual es la fracción "p" de estudiantes que votarían por él. Decide entonces contratarlo a usted para que tome una muestra aleatoria de la población de los electores y consulte sus intenciones de voto.

1. Supongamos que Usted observa en un total de  $n$  encuestados obtenidos de manera independiente que un número  $Z_1$  votará por él. Encuentre la distribución de la variable aleatoria  $Z_1$ .
2. Dé el estimador  $\hat{p}$  de máxima verosimilitud para  $p$ .
3. Verifique si el estimador  $\hat{p}$  es consistente.
4. Verifique si la varianza del estimador  $\hat{p}$  alcanza la varianza mínima para los estimadores insesgados.
5. Suponga que, dada su experiencia en elecciones anteriores, Vásquez le comunica que tiene una cierta idea de cómo se distribuye a priori el parámetro  $p$ . Vásquez le comunica que cree que el parámetro  $p$  se distribuye como una  $Beta(2,3)$ . Encuentre la función de densidad a posteriori de  $p$ .
6. Tomando una función de pérdida cuadrática, determine el estimador  $p^*$  de Bayes para  $p$ .
7. Dé la solución del estimador para una pérdida de la forma:  
$$L(h,p) = \begin{cases} k_1(p-h) & \text{si } h < p \\ k_2(h-p) & \text{si } h > p \end{cases}$$
8. Compare las varianzas de  $\hat{p}$  y  $p^*$ . Haga referencia a la parte 4.
9. ¿La información del punto 8 es suficiente para concluir que un estimador es mejor que otro? ¿Existe un criterio mejor?

**Sol.:**

1.  $Z_1 \sim \text{Binomial}(n, p)$

2.  $Z_1 = \sum x_i$  con  $x_i \sim \text{Bernoulli}(p) : f(x_i; p) = p^{x_i} (1-p)^{1-x_i}$

$\Rightarrow$  la función de verosimilitud es:  $f_n(x_1, \dots, x_n; p) = p^{\sum x_i} (1-p)^{n-\sum x_i}$

Derivando  $\log(f_n(x_1, \dots, x_n; p)) = (\sum x_i) \log(p) + (n - \sum x_i) \log(1-p)$  y anulando la derivada, se obtiene:  $\hat{p} = \frac{1}{n} \sum x_i = \bar{x}$ .

3. Hay distinta manera de probar la consistencia:  $\lim_{n \rightarrow \infty} P(|\hat{p} - p| < \varepsilon) = 1 \quad \forall \varepsilon > 0$ .

Una condición suficiente para que  $\hat{p}$  sea consistente es que converge en media cuadrática hacia el parámetro, es decir que  $E(\hat{p}) \rightarrow p$  y  $\text{Var}(\hat{p}) \rightarrow 0$ .

Aquí  $E(\hat{p}) = E(\bar{x}) = p$  y  $\text{Var}(\hat{p}) = \frac{p(1-p)}{n}$ .

4. Hay que calcular la cota inferior de la desigualdad de Cramer Rao:

$$\text{Var}(\tilde{p}) \geq \frac{1}{nI(p)} \quad \forall b \text{ con } E(\tilde{p}) = p$$

en donde es la cantidad de información de Fisher:  $I(p) = -E\left(\frac{\partial^2 \log(f)}{\partial p^2}\right)$ .

$$\frac{\partial \log(f)}{\partial p} = \frac{x}{p} - \frac{1-x}{1-p}; \quad \frac{\partial^2 \log(f)}{\partial p^2} = -\frac{x}{p^2} - \frac{1-x}{(1-p)^2};$$

$$-E\left(\frac{\partial^2 \log(f)}{\partial p^2}\right) = \frac{E(x)}{p^2} + \frac{1-E(x)}{(1-p)^2}. \text{ Con } E(x) = p \text{ se deduce que } I(p) = \frac{1}{p(1-p)} \text{ y}$$

$$\text{Var}(\tilde{p}) \geq \frac{1}{nI(p)} \quad \forall b \text{ con } E(\tilde{p}) = p. \text{ Como } \text{Var}(\hat{p}) = \frac{p(1-p)}{n}, \text{ se deduce que } \hat{p}$$

es de mínima varianza entre los estimadores insesgados de p.

5.  $\xi(p | x_1, \dots, x_n) = \frac{p^{\sum x_i} (1-p)^{n-\sum x_i} p(1-p)^2}{K} = \frac{p^{1+\sum x_i} (1-p)^{2+n-\sum x_i}}{K}$  que es una  $\text{Beta}(2 + \sum x_i, 3 + n - \sum x_i)$ .

6.  $p^*$  es la esperanza de la  $\text{Beta}(2 + \sum x_i, 3 + n - \sum x_i)$ :  $p^* = \frac{2 + \sum x_i}{5 + n}$ .

7. El estimador de Bayes minimiza la perdida media donde la función de perdida define la penalización de equivocarse en la estimación de  $p$  por  $h$ :  $L(h,p)$ .

Si  $L(h,p)=(h-p)^2$ , el estimador de Bayes es la esperanza de la función de densidad  $\xi(p | x_1, \dots, x_n)$  a posteriori.

Si  $L(h,p) = \begin{cases} k_1(p-h) & \text{si } h < p \\ k_2(h-p) & \text{si } h > p \end{cases}$ , el estimador de Bayes es el fractil  $\frac{k_1}{k_1 + k_2}$

8.  $Var(\hat{p}) = \frac{p(1-p)}{n}$ ,  $Var(p^*) = \frac{np(1-p)}{(5+n)^2}$ .  $\frac{Var(p^*)}{Var(\hat{p})} = \frac{n^2}{(5+n)^2} < 1$ . Si bien  $\hat{p}$  es de

mínima varianza entre los estimadores insesgados de  $p$ ,  $p^*$ , que es sesgado, tiene menor varianza.

9. Como  $p^*$  que es sesgado, hay que considerar como criterio el error cuadrático medio: ¿  $E((p^* - p)^2)$  es menor o mayor que  $E((\hat{p} - p)^2)$  ?