Estadística I — Examen

Grado en Matemáticas / Doble Grado Matemáticas-Ing. Informática — 10 de enero de 2013

---

1. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una variable discreta X con función de probabilidad

$$\mathbb{P}_\theta(X = x) = f(x;\theta) = \frac{\theta(-\log\theta)^x}{x!}, \quad x = 0, 1, 2, \ldots, \quad \theta \in (0,1). \tag{1}$$

Indicación: $\mathbb{E}_\theta(X) = -\log\theta = V_\theta(X)$.

(a) Calcular el estimador $\hat{\theta}_n$ de θ por el método de máxima verosimilitud, probar que es asintóticamente normal y obtener la distribución asintótica.

(b) Definir la "cantidad de información de Fisher", $I(\theta)$, y explicar muy brevemente su importancia en la teoría estadística. Calcular el valor de $I(\theta)$ para el modelo (1).

(c) Probar que $\mathbb{E}_\theta[(\hat{\theta}_n - \theta)^2] \leq -\log\theta/n$. Deducir de aquí que $n^{1/3}(\hat{\theta}_n - \theta) \xrightarrow{P} 0$.

[3 p.]

2. Una marca de detergente concentrado vende su producto en paquetes cuyo contenido nominal medio es 800 gramos. Se seleccionan al azar 20 paquetes y se obtiene para ellos un contenido medio de 793 gr. con una cuasi-desviación típica de 15. ¿Hay suficiente evidencia estadística, al nivel 0.05, para afirmar que la empresa fabricante vende su producto con un peso medio menor que el valor nominal 800? Indicar si el p-valor del correspondiente contraste es mayor o menor que 0.01. Explicar claramente las suposiciones que se necesiten para garantizar la validez del procedimiento que se utilice. **[2 p.]**

3. Se considera el problema de contrastar $H_0 : \mu \leq 0$ frente a $H_1 : \mu > 0$ a partir de una muestra de tamaño 100 de una $N(\mu, \sigma)$ (con σ conocida). Se utiliza para ello el test cuya región crítica es $R = \{(x_1, \ldots, x_n) : \bar{x} > z_\alpha \frac{\sigma}{\sqrt{n}}\}$, donde $\alpha \in (0,1)$ es el nivel de significación elegido.

Demostrar que la función de potencia $\beta_\alpha(\mu)$ de este test es una función monótona creciente en $[0, \infty)$. Calcular $\beta_{0.05}(1)$ en el caso en que $\sigma = 2$. **[1.5 p.]**

4. Sea X una v.a. con distribución geométrica de parámetro θ. Esto significa que X es discreta con $\mathbb{P}_\theta(X = x) = (1-\theta)^x\theta$, $x = 0, 1, 2, \ldots$. Se desea estimar θ a partir de una muestra de tamaño n de X, usando la metodología bayesiana. Para ello se supone que la distribución a priori de θ es Beta de parámetros $\alpha > 0$ y $\beta > 0$, es decir que la función de densidad a priori es $\pi(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}\mathbb{I}_{[0,1]}(\theta)$ (la correspondiente media es $\alpha/(\alpha+\beta)$). Calcular el estimador Bayes de θ y estudiar su consistencia casi segura.
Indicación: $\mathbb{E}_\theta(X) = \frac{1-\theta}{\theta}$. **[2.5 p.]**

5. En el directorio de trabajo del R tenemos un fichero con 1000 datos (en dos columnas de 500) llamado `datos.txt`. Redactar un código que realice las siguientes o operaciones:

(a) Leer el fichero `datos.txt`.

(b) Definir un vector llamado x con los valores de la primera columna y otro llamado y con los de la segunda.

(c) Dibujar en un mismo gráfico los dos diagramas de caja de x e y.

(d) Obtener la ecuación de la recta de mínimos cuadrados de y respecto a x (es decir, y debe ser la "variable respuesta").

[1 p.]