Solución del Examen

Dr. Olivier Nuñez Boyer

Septiembre del 2001

Problema 1.

1. Denotamos N_{ij} el número de padres con el nivel i de colesterol y cuyo hijo tiene el nivel j. Para contrastar H_0 (independencia del nivel del padre y del hijo), se utiliza el estadístico de Pearson :

$$K_n^* = \sum_{i,j=1}^{3} \frac{\left(N_{ij} - \frac{N_{i.}N_{.j}}{n}\right)^2}{\frac{N_{i.}N_{.j}}{n}},$$

   donde $n = 3332$. Obtenemos que $K_{3531}^* \simeq 1787.25$ que comparamos con el cuantil $\chi^2_{4,5\%} = 9.48$ puesto que aquí tenemos $(3-1)(3-1) = 4$ grados de libertad. Por tanto, rechazamos la hipótesis de independencia.

2. A pesar de que la independencia haya sido rechazada, una hipótesis menos fuerte como la simetría parece acceptable aquí. Utilizamos el estadístico

$$K_n^* = \sum_{i,j=1}^{3} \frac{\left(N_{ij} - \frac{N_{ij}+N_{ji}}{2}\right)^2}{\frac{N_{ij}+N_{ji}}{2}} = \sum_{i\neq j}^{3} \frac{\left(\frac{N_{ij}-N_{ji}}{2}\right)^2}{\frac{N_{ij}+N_{ji}}{2}}$$

   Obtenemos que $K_{3531}^* \simeq 13.03$ que comparamos con el cuantil $\chi^2_{3,5\%} = 7.81$, puesto que aquí tenemos $3(3-1)/2 = 3$ grados de libertad. Rechazamos la hipótesis de simetría de la tabla.

Problema II.

1. El estimador máximo verosímil del parametro λ de una distribución exponencial basado en una muestra de N observaciones $(Z_1, \ldots, Z_N)$ es

$$\widehat{\lambda} = \frac{N}{\sum_{i=1}^{N} Z_i}.$$

Por tanto bajo cualquier hipótesis, tenemos que $\widehat{\lambda}_E = n \Big/ \sum_{i=1}^{n} X_i$ y $\widehat{\lambda}_F = m \Big/ \sum_{j=1}^{m} Y_j$. Bajo $H_0 : \lambda_E = \lambda_F = \lambda_0$, tenemos que $\widehat{\lambda}_0 = (n+m) \Big/ \left(\sum_{i=1}^{n} X_i + \sum_{j=1}^{m} Y_j \right)$.

2. La razón de verosimilitud se escribe

$$\Lambda(x, y) = \frac{\max_{H_0 \cup H_1} L(\lambda_E, \lambda_F)}{\max_{H_0} L(\lambda_E, \lambda_F)} = \frac{L(\widehat{\lambda}_E, \widehat{\lambda}_F)}{L(\widehat{\lambda}_0, \widehat{\lambda}_0)}.$$

donde

$$L(\lambda_E, \lambda_F) = \lambda_E^n \exp\left(-\lambda_E \sum_{i=1}^{n} X_i \right) \ \lambda_F^m \exp\left(-\lambda_F \sum_{i=1}^{m} Y_j \right)$$

es la verosimilitud de las observaciones. Obtenemos que

$$\ln \Lambda(x, y) = -n \ln \left(\frac{\sum_{i=1}^{n} X_i}{n} \right) - m \ln \left(\frac{\sum_{j=1}^{m} Y_j}{m} \right) + (n+m) \ln \left(\frac{\sum_{i=1}^{n} X_i + \sum_{j=1}^{m} Y_j}{n+m} \right)$$

Como bajo H_1 la dimensión del vector de parámetros es 2 y bajo H_0 es 1, tenemos que aproximadamente $2 \ln \Lambda(x, y)$ sigue un χ^2 con 1 grado de libertad. El test con nivel asintótico $\alpha = 5\%$ será :

$$\phi = \begin{cases} 1 \text{ si } 2 \ln \Lambda(x, y) > \chi^2_{1,5\%} \\ 0 \text{ en caso contrario} \end{cases}$$

3. Se obtiene que $2 \ln \Lambda(x, y) \simeq 0.25$ y $\chi^2_{1,5\%} = 3.841$ por tanto se acepta H_0.

Problema III.

1. La variable θ tiene una distribución exponencial con parametro $\lambda = 1/10 = 0.1$. Ademas, tenemos que para una función de perdidas cuadratica $d_Q = E(\theta) = 10$ y $\overline{w}_Q\left(d_Q\right) = var(\theta) = 1/\lambda^2 = 100$.

1. Tenemos que la densidad de $(X_1, ...X_n)$ condicionada por θ es (omitiendo la constante en $\sqrt{2\pi}$)

$$f(x|\theta) = \prod_{i=1}^{n} \theta^{\frac{1}{2}} e^{-\frac{\theta}{2}x_i^2} = \theta^{\frac{n}{2}} \exp\left(-\frac{\theta}{2}\sum_{i=1}^{n} x_i^2\right)$$

Por tanto la densidad de θ a posteriori será (donde C es una constante que no depende de θ y donde $q(\theta)$ la densidad de Q) :

$$\begin{aligned} f(\theta|x) &= \frac{1}{C} f(x|\theta)\ q(\theta) \\ &= \frac{1}{C}\theta^{\frac{n}{2}} \exp\left(-\frac{\theta}{2}\sum_{i=1}^{n} x_i^2\right) \lambda \exp(-\lambda\theta) \\ &= \frac{1}{C}\theta^{\frac{n}{2}} \exp\left[-\theta\left(\frac{1}{2}\sum_{i=1}^{n} x_i^2 + \lambda\right)\right]. \end{aligned}$$

Se reconoce la densidad de una Gamma $G\left(\sum_{i=1}^{n} X_i^2/\,2 + \lambda, n/2 + 1\right)$, donde $\lambda = 0.1$.

2. Deducimos de lo anterior que

$$d_Q(X) = E(\theta|X) = \frac{n+2}{\sum_{i=1}^{n} X_i^2 + 2\lambda}$$

y

$$\begin{aligned} W_Q\left(d_Q(X)\right) &= E\left[var(\theta|X)\right] \\ &= E\left[\frac{2\,(n+2)}{\left(\sum_{i=1}^{n} X_i^2 + 2\lambda\right)^2}\right]. \end{aligned}$$