Inferencia Estadística II

Dr. Olivier Nuñez Boyer

19 de Septiembre del 2000

Problema 1 (5 puntos). Se trata de un problema de decisión con dos hipótesis simples:

$$H_0 : \lambda = 10 \text{ frente a } H_1 : \lambda = 2,$$

donde λ es el parámetro de una distribución exponencial.

1. Intuitivamente o utilizando el lema de Neyman-Pearson se obtiene un test cuya forma es :

   $$\phi = \begin{cases} 1 \text{ si } X > c_\alpha \\ 0 \text{ en caso contrario} \end{cases}.$$

   O sea, rechazamos H_0 si la concentración de glóbulos blancos es superior a un cierto umbral c_α que depende del nivel α del test.
   La distribución de X bajo H_0 es conocida: es una distribución exponencial con parámetro $\lambda = 10$. Por tanto, el umbral c_α verifica

   $$\begin{aligned} \alpha &= P(X > c_\alpha \,|H_0) \\ &= \int_{c_\alpha}^{\infty} 10 \exp(-10x)\, dx \\ &= [-\exp(-10x)]_{c_\alpha}^{\infty} = \exp(-10c_\alpha). \end{aligned}$$

   Deducimos que $c_\alpha = -\frac{1}{10} \ln(\alpha)$.

2. la potencia del test es la probabilidad (aquí H_1 es simple) de rechazar H_0 cuando H_0 es falso, o sea :

   $$\begin{aligned} \beta_\phi(2) &= P(X > c_\alpha \,|H_1) \\ &= \int_{c_\alpha}^{\infty} 2 \exp(-2x)\, dx \\ &= \exp(-2c_\alpha). \end{aligned}$$

Puesto que $c_\alpha = -\frac{1}{10}\ln(\alpha)$, obtenemos que $\beta_\phi(2) = \exp\left(-\frac{1}{5}\ln(\alpha)\right) = \alpha^{\frac{1}{5}}$.

3. A continuación, representamos mediante áreas el nivel y la potencia en el caso (por ejemplo) en el que $\alpha = 5\%$ y por consiguiente $c_\alpha = -\frac{1}{10}\ln(0.05) = 0.2995$.

Grafico 1: Curvas de las dos funciones de densidad exponenciales.

4. Como lo acabamos de ver para un nivel $\alpha = 5\%$, el umbral será $c_\alpha = 0.2995$. Entonces si se observa $X = 0.3$, rechazamos H_0.
   Para este nivel, deducimos que la potencia será $\beta_\phi(2) = (0.05)^{\frac{1}{5}} = 0.54928 \simeq 54\%$. Eso significa que con este test y este nivel, en el 54% de los casos concluiremos con razón que el paciente tiene el SIDA. Por otra parte, el riesgo (o error) de tipo II, siendo entonces igual a $\beta \simeq 46\%$, en el 46% de los casos no detectaremos con este test que un paciente tiene el SIDA.

5. Si bajo H_1, la media de X es superior a 0.5 mg/l, H_1 "se aleja" de H_0, y por tanto la capacidad de distinguir las dos hipótesis crece. Eso se traduce por un aumento de la potencia del test (ver gráfico).

Problema 2 (5 puntos).

1. $E(X) = \int_0^1 \theta x^\theta dx = \frac{\theta}{1+\theta}$. Cuando θ tiende hacia 0, $E(X)$ es "pequeño" (tiende también hacia 0), y cuando θ tiende hacia el infinito, $E(X)$ es "grande" (tiende hacia 1).

2. El logaritmo de la verosimilitud de la muestra es
$$l(\theta) = (\theta - 1)\sum_{i=1}^{n} \ln(X_i) + n\ln(\theta).$$
Su derivada es
$$l'(\theta) = \sum_{i=1}^{n} \ln(X_i) + \frac{n}{\theta}.$$
Por tanto, el estimador de máxima verosimilitud de θ es
$$\widehat{\theta} = -\frac{n}{\sum_{i=1}^{n} \ln(X_i)}.$$

3. Para construir el test de Wald, necesitamos la varianza asintótico $\widehat{\theta}$ (o inversa de la información de Fisher). Tenemos que
$$l''(\theta) = -\frac{n}{\theta^2}.$$
Por tanto la información de Fisher es $E_{H_0}(-l''(\theta_0)) = n$, donde $\theta_0 = 1$, valor de θ bajo H_0. En este problema el test de Wald con nivel asintótico α tiene la forma : Rechazamos H_0 si
$$\left(\widehat{\theta} - \theta_0\right)^2 E_{H_0}(-l''(\theta_0)) > \chi^2_{1;\alpha}.$$
O sea, Rechazamos H_0 si
$$\left(\frac{-n}{\sum_{i=1}^{n} \ln(X_i)} - 1\right)^2 n > \chi^2_{1;\alpha}.$$

4. Para un nivel $\alpha = 5\%$, si se ha observado para $n = 100$, $\sum_{i=1}^{100} \ln(X_i) = -150$, obtenemos la desigualdad
$$\left(\frac{100}{150} - 1\right)^2 100 > 3.841,$$
equivalente a $11.111 > 3.841$. Puesto que esta desigualdad es cierta, rechazamos H_0 con este test.

5. Sean $\theta' > \theta$, el logaritmo de la razón de verosimilitud es:

$$l(\theta') - l(\theta) = (\theta' - \theta) \sum_{i=1}^{n} \ln(X_i) + n [\ln(\theta') - \ln(\theta)].$$

Es una función creciente del estadístico $T = \sum_{i=1}^{n} \ln(X_i)$. Por tanto, para el contraste $H_0 : \theta \leq 1$ frente a $H_1 : \theta > 1$, el test UMP con nivel α será

$$\phi = \begin{cases} 1 \text{ si } T > c_\alpha \\ 0 \text{ en caso contrario} \end{cases}.$$

6. Si $n = 1$ (se observó sólo una familia), el logaritmo siendo una función creciente, el test se puede escribir de manera equivalente

$$\phi = \begin{cases} 1 \text{ si } X > u \\ 0 \text{ en caso contrario} \end{cases}.$$

Para un nivel $\alpha = 5\%$, el umbral u (que debe pertenecer al intervalo $[0, 1]$ puesto que X toma también sus valores en este intervalo) verifica

$$\begin{aligned} 5\% &= \sup_{\theta \leq 1} P_\theta (X > u) \\ &= \sup_{\theta \leq 1} \int_u^1 \theta x^{\theta - 1} dx \\ &= \sup_{\theta \leq 1} \left(1 - u^\theta\right) \\ &= 1 - u. \end{aligned}$$

Por tanto $u = 1 - 5\% = 0.95$. Por consiguiente, si se observa $X = 0.5$ aceptamos H_0.