# Inferencia Estadística II
# Solución

Septiembre de 2004

**Problema 1** Denotamos $X_1, X_2, \ldots, X_n$ los gastos de la muestra de clientes de la primera operadora y $Y_1, Y_2, \ldots, Y_n$ los gastos de la muestra de clientes de la secunda operadora, donde aquí $n = 5$. Dado que se supone que estas muestras tienen una distribución exponencial, la verosimilitud de estos datos será :

$$\begin{aligned} L(\lambda_1, \lambda_2) &= \prod_{i=1}^{n} \lambda_1 \exp\left(-\lambda_1 X_i\right) \prod_{j=1}^{n} \lambda_2 \exp\left(-\lambda_2 Y_j\right) \\ &= \lambda_1^n \exp\left(-\lambda_1 \sum X_i\right) \lambda_2^n \exp\left(-\lambda_2 \sum Y_j\right) \end{aligned}$$

donde $\lambda_1$ y $\lambda_2$ son respectivamente los parámetros de la distribución exponencial de la primera y la secunda muestra. Por lo tanto, la hipótesis nula se puede también escribir $H_0 : \lambda_1 = \lambda_2$. Sabemos que bajo cualquier hipótesis, los estimadores máximo verosímiles de $\lambda_1$ y $\lambda_2$ son

$$\widehat{\lambda}_1 = \frac{1}{\overline{X}} \text{ y } \widehat{\lambda}_2 = \frac{1}{\overline{Y}}$$

Bajo $H_0$, las dos muestras siendo homogéneas, tenemos que $\widehat{\lambda}_1 = \widehat{\lambda}_2 = \widehat{\lambda} = 1/\overline{X+Y}$ donde

$$\overline{X+Y} = \frac{1}{2n}\left(\sum X_i + \sum Y_j\right).$$

La razón de verosimilitudes se define por

$$\Lambda(x, y) = \frac{L(\widehat{\lambda}_1, \widehat{\lambda}_2)}{L(\widehat{\lambda}, \widehat{\lambda})}$$

Substituyendo los estimadores de los parámetros por su expresión en $\Lambda(x, y)$ y simplificando se obtiene que

$$\Lambda(x, y) = \frac{\left(\overline{X+Y}\right)^{2n}}{\overline{X}^n \overline{Y}^n} = \frac{0.83^{10}}{0.8^5 0.86^5} = 1.007$$

Sabemos que la distribución asintótica de $2 \ln \Lambda(x, y)$ es una $\chi^2$ con $q - k$, donde $q$ y $k$ son respectivamente la dimensión del vector de parámetros bajo cualquier hipótesis y bajo $H_0$, aquí $q = 2$ y $k = 1$. Puesto, que el cuantil 5% de la $\chi^2$ con 1 grado de libertad tiene el valor $\chi^2_{1,5\%} = 3.841$, y $2 \ln \Lambda(x, y) \simeq 0$ aceptamos a la vista de estos datos la hipótesis nula.

**Problema II (3 puntos).**

1. Los datos se ajustan bastante bien a una recta.

2. El coeficiente de correlación de Pearson es aquí igual a :

$$r(x,y) = 0.953$$

Sabemos que bajo normalidad de los datos, el estadístico

$$T = \sqrt{n-2}\frac{r(x,y)}{\sqrt{1-r^2(x,y)}}$$

sigue bajo $H_0$ una Student con $n-2$ grados de libertad, $n$ siendo igual aquí a 10. Obtenemos que

$$T = \sqrt{8}\frac{0.953}{\sqrt{1-0.953^2}} = 8.89$$

Puesto que este valor es (en valor absoluto) superior al cuantil $t_{8.,2.5\%} = 2.306$, rechazamos la hipótesis nula.

3. Si no asumimos la normalidad de los datos, no conocemos en general la distribución de $T$. Por lo tanto, tenemos que recurrir a un coeficiente no paramétrico como el de Spearman $r_S$. Con estos datos obtenemos que

$$\begin{aligned} r_S &= 1 - \frac{6}{10\left(10^2-1\right)}8 \\ &= 0.951 \end{aligned}$$

Dado que el cuantil 5% de la distribución de $|r_S|$ bajo $H_0$ es 0.648, rechazamos $H_0$ también con este coeficiente.

**Problema III (3 puntos).**

1. Se trata de la practica de simulación realizada en clase titulada "test de la $\chi^2$". La hipótesis nula se traduce aquí por $H_0 : p_j = p_j^0 = \frac{1}{9}$ donde $p_j$ es la probabilidad de que una bomba caiga en la zona $Z_j$. El estadístico de test era

$$\begin{aligned} K_n &= \frac{1}{n}\sum_{j=1}^{k}\frac{N_j^2}{p_j^0} - n \\ &= \frac{1}{60}\sum_{j=1}^{9}\frac{N_j^2}{\frac{1}{9}} - 60 \end{aligned}$$

donde $N_j$ era el número observado de bombas que cayeron en la zona $Z_j$. Sabemos que este estadístico sigue asintóticamente una $\chi^2$ con $k-1 = 8$ grados de libertad. Por lo tanto rechazaremos $H_0$ si $K_n > \chi^2_{8,5\%} = 15.51$.

2. Para la configuración siguiente de los impactos

| | | |
|---|---|---|
| 8 | 7 | 3 |
| 5 | 9 | 11 |
| 6 | 4 | 7 |

obtenemos que $K_n = 7.5$, por consiguiente no podemos rechazar la hipótesis nula para el nivel de significación $\alpha = 5\%$.