Universidad de Castilla-La Mancha
Facultad Ciencias Químicas
Estadística. Primer Curso de Ingeniería Química
Convocatoria 28-06-2004

INSTRUCCIONES

- Escribe claramente tus apellidos y nombre, en todas y cada una de las hojas de examen.
- No se admitirá más de una hoja por problema. No escribas partes de problemas diferentes en una misma hoja.
- Razona todas las respuestas

EJERCICIOS

1. En un aula hay 50 alumnos. Hallar la probabilidad de que al menos dos de ellos cumplan años el mismo día del año.

2. La duración, en segundos, precisa para que se produzca una reacción entre dos compuestos sigue una distribución $N(\mu, \sigma)$. El 60% de las veces dura más de 40 segundos y el 55% dura menos de 50. Hallar μ y σ.

3. La proporción en sangre de dos compuestos, X e Y, en una especie común de ratones es variable. Su distribución conjunta en toda la población se caracteriza mediante la función de densidad

$$f(x,y) = \begin{cases} k(1-x)y^2 & \text{si } (x,y) \in (0,1) \times (0,1) \\ 0 & \text{en el resto} \end{cases}$$

Un ratón se considera sano si ambas concentraciones son inferiores a 3/4.

   (a) Hallar el valor de k. Decidir si X e Y son independientes.
   (b) Hallar la concentración media del compuesto Y en la especie.
   (c) Hallar la probabilidad de que un ratón elegido al azar esté sano.

4. Una población modelada mediante ξ, sigue un distribución de probabilidad cuya función de densidad viene dada por

$$f_\xi(x) = \begin{cases} \frac{1}{\theta} x^{\frac{1}{\theta}-1} & \text{si } x \in (0,1) \\ 0 & \text{en el resto} \end{cases}$$

siendo θ cierto parámetro positivo. Se pide

   (a) Calcular un estimador de θ mediante el método de los momentos.
   (b) Idem mediante el método de máxima verosimilitud.

5. Las tensiones de rotura (en Kp.) de 5 cables de un determinado metal fueron

$$660, 460, 540, 580, 550$$

Suponiendo normalidad para las tensiones. Estimar la tensión media de rotura mediante un intervalo de confianza del 95%.

SOLUCIONES

1. La probabilidad pedida coincide con uno menos la probabilidad de que todos cumplan años en días diferentes, por lo tanto

$$\begin{aligned} p &= 1 - P(\text{cumplan en día distintos}) \\ &= 1 - \left(\frac{1}{365}\frac{364}{365}\frac{363}{365}...\frac{365-49}{365}\right)(365) \end{aligned}$$

Se ha utilizado la noción de dependencia entre el hecho de que si el primero nace un día, el segundo ha de hacerlo el resto, lo cual supone considerar sólo 364 días posibes entre los 365 que hay en un año. Hacemos el cáclulo: si

$$l = \left(\frac{1}{365}\frac{364}{365}\frac{363}{365}...\frac{365-49}{365}\right)(365)$$

entonces

$$\begin{aligned} \log l &= \log\left(\frac{364}{365}\frac{363}{365}...\frac{365-49}{365}\right) \\ &= \log\left(\frac{364}{365}\right) + \log\left(\frac{363}{365}\right) + ... + \log\left(\frac{316}{365}\right) \\ &= \sum_{j=316}^{364} \log\left(\frac{j}{365}\right) = -3.5191 \\ l &= \exp(-3.5191) = 2.9626 \times 10^{-2} \end{aligned}$$

Así todo

$$\begin{aligned} p &= 1 - 2.9626 \times 10^{-2} \\ &= 0.97037 \end{aligned}$$

(basta con escribir $p = 1 - \exp\left(\sum_{j=316}^{364} \log\left(\frac{j}{365}\right)\right)$.

2. Los datos son

$$\begin{aligned} P(X &\geq 40) = 0.6 \\ P(X &\leq 50) = 0.55 \end{aligned}$$

Tipificamos estos datos y resulta que

$$\begin{aligned} P(\frac{X-\mu}{\sigma} &\geq \frac{40-\mu}{\sigma}) = 0.6 \\ P(\frac{X-\mu}{\sigma} &\leq \frac{50-\mu}{\sigma}) = 0.55 \end{aligned}$$

La primera probabilidad dice que

$$P(\frac{X-\mu}{\sigma} \leq \frac{40-\mu}{\sigma}) = 0.4,$$

luego se trata de

$$P(\frac{X-\mu}{\sigma} \leq -\frac{40-\mu}{\sigma}) = 0.6$$

Puesto que $\frac{X-\mu}{\sigma}$ se distribuye $N(0,1)$ buscamos en las tablas de la normal para llegar al sistema

$$\begin{aligned} \frac{50-\mu}{\sigma} &= 0.125 \\ -\frac{40-\mu}{\sigma} &= 0.25 \end{aligned}$$

La solución es $\mu = 46.667$ y $\sigma = 26.667$

3.

$$f(x,y) = \begin{cases} k(1-x)y^2 & \text{si } (x,y) \in (0,1)\times(0,1) \\ 0 & \text{en el resto} \end{cases}$$

Un ratón se considera sano si ambas concentraciones son inferiores a 3/4.

(a) Para que sea densidad de probabilidad es preciso que ocurrand dos cosas: que $f \geq 0$, lo ucal se consigue siempre que $k \geq 0$,y que

$$\int_{\mathbb{R}}\int_{\mathbb{R}} f(x,y)dxdy = 1;$$

pero esto último se escribe como

$$\int_0^1 \left(\int_0^1 k(1-x)y^2 dy\right) dx = k/6 = 1$$

i.e se consigue con $k = 6$

Las densidades de X y de Y se consiguen así

$$\begin{aligned} f_X(x) &= \int_{\mathbb{R}} f(x,y)dy \\ &= \begin{cases} \int_0^1 6(1-x)y^2 dy & \text{si } x \in (0,1) \\ 0 & \text{en el resto} \end{cases} \\ &= \begin{cases} 2(1-x) & \text{si } x \in (0,1) \\ 0 & \text{en el resto} \end{cases} \end{aligned}$$

Análogo con Y :

$$\begin{aligned} f_Y(y) &= \int_{\mathbb{R}} f(x,y)dx = \\ &= \begin{cases} \int_0^1 6(1-x)y^2 dx & \text{si } y \in (0,1) \\ 0 & \text{en el resto} \end{cases} \\ &= \begin{cases} 3y^2 & \text{si } x \in (0,1) \\ 0 & \text{en el resto} \end{cases} \end{aligned}$$

Como $f_X(x)f_Y(y) = f(x,y)$ entonces se trata de v.a. independientes.

(b) Nos están pidiendo la esperanza de Y :

$$E[Y] = \int_0^1 y3y^2dy = 3/4$$

(c) Se trata de calcular

$$\begin{aligned} P(X &\leq 3/4, Y \leq 3/4) \\ &= \int_0^{3/4}\left(\int_0^{3/4} 6(1-x)y^2dy\right)dx \\ &= \frac{405}{1024} \end{aligned}$$

4. Una población modelada mediante ξ, sigue un distribución de probabilidad cuya función de densidad viene dada por

$$f_\xi(x) = \begin{cases} \frac{1}{\theta}x^{\frac{1}{\theta}-1} & \text{si } x \in (0,1) \\ 0 & \text{en el resto} \end{cases}$$

siendo θ cierto parámetro positivo. Se pide

(a) El método de los momentos consiste en resolver

$$\alpha_1 = M_1$$

i.e.

$$\int_{\mathbb{R}} x f_\xi(x)dx = \overline{\xi}$$

o de otra forma más explícita:

$$\int_0^1 x\frac{1}{\theta}x^{\frac{1}{\theta}-1}dx = \frac{1}{\theta+1} = \frac{\sum_{i=1}^n \xi_i}{n}$$

con lo que

$$\theta = -\frac{-1+\frac{\sum_{i=1}^n \xi_i}{n}}{\frac{\sum_{i=1}^n \xi_i}{n}}$$

es el estimador pedido.

(b) En el método de máxima verosimilitud hay que maximizar la función

$$\begin{aligned} L &= \frac{1}{\theta}x_1^{\frac{1}{\theta}-1}\frac{1}{\theta}x_2^{\frac{1}{\theta}-1}...\frac{1}{\theta}x_n^{\frac{1}{\theta}-1} \\ &= \left(\frac{1}{\theta}\right)^n \Pi_{i=1}^n\left(x_i^{\frac{1}{\theta}-1}\right) \end{aligned}$$

Por comodidad maximizamos la segunda función de verosimilitud

$$\begin{aligned} l &= \log L = -n\log\theta + \sum \log x_i^{\frac{1}{\theta}-1} \\ &= -n\log\theta + \left(\frac{1}{\theta}-1\right)\sum \log x_i \end{aligned}$$

Derivamos e igualamos a cero:

$$\frac{-n}{\theta} - \frac{1}{\theta^2}\sum \log x_i = 0$$

La solución es

$$\theta = -\frac{\sum \log x_i}{n}$$

El estimador pedidido es

$$\theta^* = -\frac{\sum \log \xi_i}{n}$$

5. Las tensiones de rotura (en Kp.) de 5 cables de un determinado metal fueron

660, 460, 540, 580, 550

Suponiendo normalidad para las tensiones. Estimar de manera razonada, la tensión media de rotura mediante un intervalo de confianza del 95%. (2 puntos)Como se trata de estimar la esperanza con desviación desconocida tomamos como cantidad pivotal a

$$Q = \frac{\overline{\xi}-\mu}{\frac{S}{\sqrt{n}}} \sim \mathrm{t}_{n-1}.$$

como tenemos una muestra de tamaño 5El intervalo pedido es

$$\begin{aligned} I &= \left[\overline{\xi} - \mathrm{t}_{4,\frac{\alpha}{2}}\frac{S}{\sqrt{81}}, \overline{\xi} + \mathrm{t}_4\frac{S}{\sqrt{81}}\right] \\ &= \left[\overline{\xi} - \mathrm{t}_{4,\frac{0.05}{2}}\frac{S}{\sqrt{5}}, \overline{\xi} + \mathrm{t}_{4,\frac{0.05}{2}}\frac{S}{\sqrt{5}}\right]. \end{aligned}$$

Para la realización del intervalo necesitamos conocer la realización de $\overline{\xi}$ y de S :

$$\frac{\sum_{i=1}^5 x_i}{5} = \frac{660+460+540+580+550}{5} = 558$$

y

$$
\begin{aligned}
s^2 &= \frac{\sum_{i=1}^{5}(x_i-558)^2}{4} \\
&\quad \frac{(660-558)^2+(460-558)^2+(540-558)^2}{4} \\
&\quad +\frac{(580-558)^2+(550-558)^2}{4} \\
&= 5220
\end{aligned}
$$

Así $S=\sqrt[2]{5220}$y $\mathrm{t}_{4,0.025}=2.776$, y el intervalo pedido resulta ser

$$
\begin{aligned}
I &= \left[558-\mathrm{t}_{4,0.025}\frac{\sqrt[2]{5220}}{\sqrt{5}}, 558+\mathrm{t}_{4,0.025}\frac{\sqrt[2]{5220}}{\sqrt{5}}\right] \\
&= \left[558-2.776\frac{\sqrt[2]{5220}}{\sqrt{5}}, 558+2.774\frac{\sqrt[2]{5220}}{\sqrt{5}}\right] \\
&= [468.3, 647.63].
\end{aligned}
$$