Grado en Ingeniería Informática

Examen de Matemáticas III (Convocatoria de junio)

Ejercicio 1: (1 pto) Sean A y B dos sucesos aleatorios de los que se sabe que $P(A) = 0.5$, $P(\overline{B}) = 0.6$ y $P(\overline{A} \cap \overline{B}) = 0.25$. Calcular $P(A \cap B)$ y $P(A|_{\overline{B}})$. ¿Son A y B independientes?

Ejercicio 2: (0.75 ptos) Se define una variable aleatoria X como el menor de dos números elegidos, al azar y con reemplazamiento, del conjunto $\{1, 2, 3\}$. Determinar la función de probabilidad, la esperanza y la varianza de X.

Ejercicio 3: (1.25 ptos) Dos empresas A y B fabrican un mismo tipo de componente electrónico aunque con diferentes calidades. Los fabricados por A tienen un tiempo de vida X que sigue un modelo uniforme continuo en el intervalo (0,10). Los componentes fabricados por B tienen un tiempo de vida Y, en años, con función de distribución

$$F_Y(x) = \begin{cases} 1 - e^{-x^2/30} & \text{si} \quad x > 30 \\ 0 & \text{si} \quad x \leq 30 \end{cases}$$

Se sabe que, de cada cuatro componentes que hay en el mercado, tres han sido fabricados por B. Si un componente elegido al azar dura menos de dos años, ¿cuál es la probabilidad de que fuera fabricado por A?

Ejercicio 4: Una empaquetadora automática se programa para producir paquetes de 500 gramos. Un estudio concluye que el peso real de un paquete de la producción es una variable aleatoria X que sigue una distribución $N(498, 16)$.

(a) (1.25 ptos) Se selecciona al azar una muestra de 10 paquetes. ¿Cuál es la probabilidad de que al menos tres de ellos pesen menos de 490 gramos?

(b) (0.75 pto) Se sabe que producir cada gramo de producto le cuesta a la empresa 0.05 euros y que los paquetes producidos se venden a 45 euros cada uno. Sea B la variable *beneficio de la empresa en cada paquete vendido*. Expresar la relación que existe entre la variable B y la variable X. ¿Qué distribución sigue la variable B? ¿Cuál es la probabilidad de que el beneficio obtenido por la venta de un paquete de producto supere los 20 euros?

Ejercicio 5: Sea $X_1, \ldots, X_n$ una muestra aleatoria simple de una variable aleatoria X con función de densidad:

$$f(x) = \begin{cases} \frac{x}{\beta^2} e^{-\frac{x}{\beta}} & \text{si} \quad x > 0 \\ 0 & \text{si} \quad x \leq 0 \end{cases}$$

donde $\beta > 0$ es un parámetro desconocido. Se sabe que $E[X] = 2\beta$, que $Var[X] = 2\beta^2$ y que $Var[X^2] = 84\beta^4$.

(a) (1 pto) Construir la función de densidad conjunta de la muestra y demostrar que el estadístico $T_1 = \sqrt{5\overline{X}}$ es un estadístico suficiente para estimar β.

(b) (0.75 pto) Demostrar que el estimador definido por $T_2 = \frac{1}{6n} \sum_{i=1}^{n} X_i^2$ es consistente para estimar β^2 .

Ejercicio 6: Se sospecha que la tasa de fallo de cierto componente electrónico está relacionada con la temperatura de funcionamiento del componente. Se recogen los siguientes datos relativos al número de fallos en una hora, para distintas temperaturas de funcionamiento:

Temperatura (grados Fahrenheit)	55	65	75	85	95
Número de fallos en una hora	190	193	197	200	201

a) (0.75 pto) Ajustar a los datos un modelo del tipo $y = \beta_0 + \beta_1 x + \varepsilon$ que permita predecir el número de fallos en función de la temperatura de funcionamiento. Calcular e interpretar una medida de la bondad del ajuste realizado.

b) (0.75 pto) Construir un intervalo de confianza al 90%, para la pendiente del modelo.

Ejercicio 7: Con objeto de estudiar el contenido de folacina de dos marcas de té A y B, se ha recogido una muestra de cada marca. En la muestra de la marca A, de tamaño 7, se ha observado un valor medio de 9.27 ng/ml y una desviación típica de 1.94 ng/ml, mientras que en la muestra de la marca B, de tamaño 6, se ha observado una concentración media de 7.45 ng/ml con una desviación típica de 0.76 ng/ml.

Plantear y resolver los contrastes de hipótesis adecuados para responder a las siguientes preguntas con nivel de significación $\alpha = 0.05$:

(a) (0.75 pto) A partir de los datos observados, ¿existe evidencia significativa para rechazar la igualdad de varianzas del contenido de folacina de ambas marcas?

(b) (1 pto) ¿Podemos afirmar que el contenido promedio de folacina de la marca A supera al contenido promedio de folacina de la marca B en, al menos, 1.5 ng/ml?

22 de junio de 2016

DURACIÓN: HASTA LAS 12:50H.

Ejercicio 1: Sabemos que $P(A)=0'5$, $P(\bar{B})=0'6$ y $P(\bar{A}\cap\bar{B})=0'25$

Entonces, $0'25 = P(\bar{A}\cap\bar{B}) = P(\overline{A\cup B}) = 1-P(A\cup B) \Rightarrow P(A\cup B)=0'75$.

Además, $P(A\cup B) = P(A)+P(B)-P(A\cap B) \Rightarrow 0'75 = 0'5+(1-0'6)-P(A\cap B)$

$\Rightarrow \boxed{P(A\cap B)=0'15}$

Por otra parte, $\boxed{P(A|_{\bar{B}}) = \frac{P(A\cap\bar{B})}{P(\bar{B})} = \frac{P(A)-P(A\cap B)}{P(\bar{B})} = \frac{0'5-0'15}{0'6} = \frac{7}{12}}$

Los sucesos A y B no son independientes, pues $P(A\cap B)=0'15$ y $P(A)\cdot P(B)=0'5\cdot 0'4=0'2$, luego $P(A\cap B)\neq P(A)\cdot P(B)$.

Ejercicio 2:

El espacio muestral viene dado por el conjunto de pares ordenados:
$\Omega=\{(1,1),(1,2),(1,3),(2,1),(2,2),(2,3),(3,1),(3,2),(3,3)\}$

El primer y segundo elemento de cada par representan los elementos elegidos en primer y segundo lugar, respectivamente, del conjunto $\{1,2,3\}$.

La probabilidad de cada par viene dada por $\frac{1}{3}\cdot\frac{1}{3}=\frac{1}{9}$. Entonces,

$P(X=1) = P(\{(1,1),(1,2),(1,3),(2,1),(3,1)\}) = 5\cdot\frac{1}{9}=\frac{5}{9}$

$P(X=2) = P(\{(3,2),(2,2),(2,3)\}) = \frac{3}{9}=\frac{1}{3}$

$P(X=3) = P(\{(3,3)\}) = \frac{1}{9}$

Con lo cual, la función de probabilidad de X es:

$$P(X=k) = \begin{cases} \frac{5}{9} & \text{si } k=1 \\ \frac{1}{3} & \text{si } k=2 \\ \frac{1}{9} & \text{si } k=3 \\ 0 & \text{si } k\notin\{1,2,3\} \end{cases}$$

Finalmente, obtenemos que

$$E(X) = 1\cdot P(X=1) + 2\cdot P(X=2) + 3\cdot P(X=3) = 1\cdot\frac{5}{9} + 2\cdot\frac{1}{3} + 3\cdot\frac{1}{9} = \frac{14}{9}$$

$$Var(X) = E(X^2) - E^2(X) = \frac{26}{9} - \left(\frac{14}{9}\right)^2 = \frac{38}{81} \simeq 0'4691$$

$$E(X^2) = 1^2\cdot\frac{5}{9} + 2^2\cdot\frac{1}{3} + 3^2\cdot\frac{1}{9} = \frac{26}{9}$$

Ejercicio 3:

Sabemos que $X \sim U(0,10)$, luego su función de densidad es

$$f_X(x) = \begin{cases} \frac{1}{10} & \text{si } x \in (0,10) \\ 0 & \text{si } x \notin (0,10) \end{cases}$$

con lo cual, $P(X<2) = \int_{-\infty}^{2} f_X(x)\,dx = \int_0^2 \frac{1}{10}\,dx = \frac{2}{10}$

Por otra parte, $P(Y<2) = F_Y(2) = 1 - e^{-4/30} \simeq 0'12483$

Además, si denotamos por A y B a los sucesos que ocurren cuando el componente elegido ha sido fabricado por A o por B, respectivamente, de los datos del problema, $P(A) = \frac{1}{4}$ y $P(B) = \frac{3}{4}$.

Finalmente, obtenemos que: (denotando por D a la duración de una componente elegida al azar)

$$P(A|_{D<2}) = \frac{P(A\cap(D<2))}{P(D<2)} = \frac{P(D<2|_A)\cdot P(A)}{P(D<2|_A)\cdot P(A) + P(D<2|_B)\cdot P(B)} =$$

$$= \frac{P(X<2)\cdot P(A)}{P(X<2)\cdot P(A) + P(Y<2)\cdot P(B)} = \frac{0'2\cdot 0'25}{0'2\cdot 0'25 + 0'12483\cdot 0'75} \simeq$$

$$\simeq 0'3481$$

Ejercicio 4:

a) $X \sim N(498, 16)$. Entonces, $P(X<490) = P\left(\frac{X-498}{\sqrt{16}} < \frac{490-498}{\sqrt{16}}\right) =$

$= P(Z<-2) = P(Z>2) = 1-P(Z\leq 2) = 1-0'9772 = 0'0228.$

Consideramos la variable aleatoria Y: número de paquetes, de los 10 que forman la muestra, que pesan menos de 490 gr.

Entonces $Y \sim B(10, 0'0228)$ y $P(Y\geq 3) = 1-P(Y<3) = 1-P(Y\leq 2) =$

$= 1-\left(P(Y=0)+P(Y=1)+P(Y=2)\right) = 1-\binom{10}{0}0'0228^0 \cdot 0'9772^{10} -$

$-\binom{10}{1}0'0228^1 \cdot 0'9772^9 - \binom{10}{2}\cdot 0'0228^2 \cdot 0'9772^8 = \underline{0'001261}$

b) La relación entre B y X es: $B = 45 - 0'05\cdot X$. Además, al ser $X \sim N(498, 16)$, se tiene que $B \sim N(45-0'05\cdot 498,\ 0'05^2\cdot 16)$, esto es, $B \sim N(20'1, 0'04)$.

Entonces $P(B>20) = P\left(\frac{B-20'1}{\sqrt{0'04}} > \frac{20-20'1}{\sqrt{0'04}}\right) = P\left(Z > \frac{-0'1}{0'2}\right) =$

$= P(Z>-0'5) = P(Z<0'5) = \underline{0'6915}$

Ejercicio 5:

a) Calculamos, en primer lugar, la función de densidad conjunta:

$$f(x_1,\dots,x_n) = \prod_{i=1}^n f_{X_i}(x_i) = \prod_{i=1}^n f_X(x_i) = \prod_{i=1}^n \left(\frac{x_i}{\beta^2}\cdot e^{-\frac{x_i}{\beta}}\right) = \underline{\left(\prod_{i=1}^n x_i\right)\cdot \frac{1}{\beta^{2n}}\cdot e^{-\frac{\sum x_i}{\beta}}}$$

Por otra parte, $T_1 = \sqrt{5\cdot \bar{x}} \Rightarrow T_1^2 = 5\cdot \frac{\sum x_i}{n} \Rightarrow \sum_{i=1}^n x_i = \frac{n\cdot T_1^2}{5}$. Sustituyendo, obtenemos que

$$f(x_1,\dots,x_n) = \underbrace{\left(\prod_{i=1}^n x_i\right)}_{h(x_1,\dots,x_n)}\cdot \underbrace{\frac{1}{\beta^{2n}}\cdot e^{\frac{-nT_1^2}{5\beta}}}_{g(T_1,\beta)}$$

La función de densidad conjunta se puede descomponer, por tanto, como producto de una función de la muestra, h, y una función de T_1 y de β, que hemos llamado g. En consecuencia, T_1 es suficiente de cara a estimar el parámetro β.

b) Bastará comprobar que se cumplen las condiciones:

i) $\lim_{n\to\infty} E(T_2) = \beta^2$ y ii) $\lim_{n\to\infty} Var(T_2) = 0$

$$E(T_2) = E\left(\frac{1}{6n}\cdot\sum_{i=1}^{n} X_i^2\right) = \frac{1}{6n}\cdot\sum_{i=1}^{n} E(X_i^2) \underset{E(X)=E(X_i)}{=} \frac{1}{6n}\cdot\sum_{i=1}^{n} E(X^2) = \frac{1}{6n}\cdot n\cdot E(X^2) =$$

$$= \frac{1}{6}\cdot\left(Var(X)+E^2(X)\right) = \frac{1}{6}\cdot(2\beta^2+4\beta^2) = \frac{1}{6}\cdot 6\beta^2 = \beta^2$$

$Var(X) = E(X^2) - E^2(X)$

Por otra parte,

$$Var(T_2) = Var\left(\frac{1}{6n}\cdot\sum_{i=1}^{n} X_i^2\right) = \frac{1}{36n^2}\cdot Var\left(\sum_{i=1}^{n} X_i^2\right) \underset{X_1,\ldots,X_n \text{ independientes}}{=} \frac{1}{36n^2}\cdot\sum_{i=1}^{n} Var(X_i^2) \underset{Var(X_i)=Var(X)}{=}$$

$$= \frac{1}{36n^2}\cdot\sum_{i=1}^{n} Var(X^2) = \frac{1}{36n^2}\cdot n\cdot Var(X^2) = \frac{1}{36n^2}\cdot n\cdot 84\beta^2 = \frac{7\beta^4}{3n}$$

Entonces,

$$\left.\begin{array}{l} \lim_{n\to\infty} E(T_2) = \lim_{n\to\infty} \beta^2 = \beta^2 \\ \lim_{n\to\infty} Var(T_2) = \lim_{n\to\infty} \frac{7\beta^4}{3n} = 0 \end{array}\right\} \Rightarrow T_2 \text{ es consistente para estimar } \beta^2.$$

Ejercicio 6:

a) Se desea predecir el número de fallos en función de la temperatura por lo que la variable x del modelo representará la temperatura y la variable y el número de fallos.

De los datos del problema obtenemos que:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{375}{5} = 75 \qquad S_x^2 = \frac{\sum x_i^2}{n} - \bar{x}^2 = \frac{29125}{5} - 75^2 = 200$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{981}{5} = 196'2 \qquad S_y^2 = \frac{\sum y_i^2}{n} - \bar{y}^2 = \frac{192559}{5} - 196'2^2 = 17'36$$

$$S_{xy} = \frac{\sum x_i \cdot y_i}{n} - \bar{x}\cdot\bar{y} = \frac{73865}{5} - 75 \cdot 196'2 = 58$$

La recta de regresión viene dada por

$$y - \bar{y} = \frac{S_{xy}}{S_x^2}(x - \bar{x}) \xrightarrow[\text{Sustituyendo}]{} y - 196'2 = \frac{58}{200}\cdot(x - 75) \rightarrow \boxed{y = 0'29 \cdot x + 174'45}$$

El valor del coeficiente de determinación es:

$$R^2 = \frac{S_{xy}^2}{S_x^2 \cdot S_y^2} \simeq 0'9689$$ Por tanto, podemos decir que el modelo explica el 96'89% de la variabilidad del número de fallos.

b) La expresión general del intervalo que nos piden es

$$\left(b_1 - \sqrt{\frac{S^2}{nS_x^2}} \cdot t_{n-2,1-\alpha/2},\ b_1 + \sqrt{\frac{S^2}{n\cdot S_x^2}} \cdot t_{n-2,1-\alpha/2}\right)$$

donde $S^2 = \dfrac{SS_{Tot} - SS_{Reg}}{n-2} = \dfrac{n\cdot S_y^2 - n\cdot b_1^2 \cdot S_x^2}{n-2} = \dfrac{5\cdot 17'36 - 5\cdot 0'29^2 \cdot 200}{3} = 0'9$

Además, como se pide un intervalo

$\alpha = 0'1 \Rightarrow t_{n-2,1-\alpha/2} = t_{3,0'95} = 2'353$ y el intervalo buscado es

$$\left(0'29 - \sqrt{\frac{0'9}{5\cdot 200}} \cdot 2'353,\ 0'29 + \sqrt{\frac{0'9}{5\cdot 200}} \cdot 2'353\right)$$

Operando $\rightarrow$ $\boxed{(0'21941, 0'36059)}$

Ejercicio 7

Datos del problema: $\bar{X}_A = 9'25$, $S_A^2 = 1'94$, $n_A = 7$, $\bar{X}_B = 7'45$, $S_B = 0'76$, $n_B = 6$.

a) Contraste a resolver: $\begin{cases} H_0 : \sigma_A^2 = \sigma_B^2 \\ H_1 : \sigma_A^2 \neq \sigma_B^2 \end{cases}$ con $\alpha = 0'05$

Rechazaremos H_0 si $\dfrac{S_{cA}^2}{S_{cB}^2} \geq f_{n_A-1, n_B-1, 1-\alpha/2}$ ó $\dfrac{S_{cA}^2}{S_{cB}^2} \leq \dfrac{1}{f_{n_B-1, n_A-1, 1-\alpha/2}}$

$f_{n_A-1, n_B-1, 1-\alpha/2} = f_{6,5,0'975} = 6'978$, $\dfrac{1}{f_{n_B-1, n_A-1, 1-\alpha/2}} = \dfrac{1}{f_{5,6,0'975}} = \dfrac{1}{5'988} \simeq 0'167$

Entonces rechazaremos H_0 si $\dfrac{S_{cA}^2}{S_{cB}^2} \geq 6'978$ ó $\dfrac{S_{cA}^2}{S_{cB}^2} \leq 0'167$

$$\left.\begin{aligned} S_{cA}^2 &= \frac{n_A \cdot S_A^2}{n_A - 1} = \frac{7 \cdot 1'94^2}{6} \simeq 4'39087 \\ S_{cB}^2 &= \frac{n_B \cdot S_B^2}{n_B - 1} = \frac{6 \cdot 0'76^2}{5} \simeq 0'69312 \end{aligned}\right\} \Rightarrow \frac{S_{cA}^2}{S_{cB}^2} = 6'33493$$ luego no

se cumple la regla de rechazo por lo que no hay evidencia para afirmar que $\sigma_A^2 \neq \sigma_B^2$

b) Tenemos que resolver el contraste $\begin{cases} H_0 : \mu_A - \mu_B = 1'5 \\ H_1 : \mu_A - \mu_B > 1'5 \end{cases}$ con $\alpha = 0'05$.

Teniendo en cuenta el resultado del apartado anterior, se trata de un contraste sobre la diferencia de medias de dos poblaciones normales con varianzas desconocidas pero iguales.

Rechazaremos H_0 si $\bar{X}_A - \bar{X}_B \geq \delta_0 + S_p \cdot \sqrt{\frac{1}{n_A} + \frac{1}{n_B}} \cdot t_{n_A+n_B-2, -\alpha}$

$S_p^2 = \dfrac{n_A \cdot S_A^2 + n_B \cdot S_B^2}{n_A + n_B - 2} = \dfrac{7 \cdot 1'94 + 6 \cdot 0'76^2}{7+6-2} = 2'7101$, luego rechazaremos

H_0 si $\bar{X}_A - \bar{X}_B \geq 1'5 + \sqrt{2'7101} \cdot \sqrt{\frac{1}{7} + \frac{1}{6}} \cdot \overbrace{t_{11,0'95}}^{=1'796} = 3'1449$

Al ser $\bar{X}_A - \bar{X}_B = 1'82$ no se cumple la regla de rechazo por lo que no tenemos evidencia para afirmar que el contenido promedio de folacina de A supere en, al menos 1'5 ng/ml, al contenido promedio de B.