IT Grado en Ingeniería Informática

Examen de Matemáticas III (Convocatoria de junio)

Ejercicio 1: (0.75 ptos) Sean A y B dos sucesos tales que $P(A) = 0.2$, $P(A|_B) = 0.25$ y $P(B|_A) = 0.5$. Calcular $P(\overline{A}|_{\overline{B}})$.

Ejercicio 2: Una empresa A fabrica chips con un porcentaje de defectuosos del 5% y los pone a la venta en cajas de 5 unidades. Otra empresa ilegal B vende imitaciones indistinguibles del mismo chip con un porcentaje de defectuosos del 50% y los comercializa en el mismo envase de 5 unidades que la empresa anterior. Teniendo en cuenta que el 10% de las cajas que hay en el mercado son ilegales,

(a) (0.5 ptos) ¿Cuál es la probabilidad de que una caja de chips de la empresa A contenga, exactamente, dos chips defectuosos?¿Y si la caja procede de la empresa B?

(b) (0.5 ptos) ¿Se compra una caja que resulta tener dos chips defectuosos. ¿Cuál es la probabilidad de que la caja proceda de la empresa B?

(c) (1 pto) ¿Tras estudiar las ventas del último año, la empresa A llegó a la conclusión de que el número de unidades diarias que vende se puede modelar de acuerdo a una variable aleatoria X que sigue una distribución normal de media 10000 y que la probabilidad de que un día venda, a lo más, 9000 unidades es de 0.1056. Si el beneficio, en euros, obtenido en un día viene dado por $B(X) = -3500 - X + 0.00015X^2$, ¿cuál es el beneficio esperado cada día?

Ejercicio 3: La duración, en horas, de los estacionamientos en un parking público, es una variable aleatoria X con función de densidad de la forma:

$$f(x) = \begin{cases} \frac{cx}{2} & \text{si} \quad 0 < x < 2 \\ c(\frac{-x}{4} + \frac{3}{2}) & \text{si} \quad 2 < x < 6 \\ 0 & \text{en otro caso} \end{cases}$$

(a) (0.5 ptos) Calcular razonadamente el valor de c.

En lo que sigue, suponer que $c = 1/3$.

(b) (0.5 ptos) Calcular la probabilidad de que un estacionamiento dure entre una y tres horas.

(c) (0.75 ptos) Cuando un conductor aparca su vehículo en el parking, la empresa propietaria del mismo le cobra dos euros si el estacionamiento dura menos de una hora, 3 si dura entre 1 y tres horas y 4 si la duración es superior a tres horas. Determinar el beneficio esperado de la empresa para cada vehículo que llega al parking.

Ejercicio 4. Como parte del estudio del comportamiento de un servidor web se desea predecir el número de errores de timeout a partir del número de peticiones simultaneas por segundo (pets/s) recibidas por el servidor. El resultado aparece en la siguiente tabla:

N^o pets/s	25	50	75	100	125	150	175	200
Errores de la conexión	47	90	103	105	122	157	194	218

(a) (0.75 ptos) Ajustar a los datos un modelo del tipo $y = \beta_0 + \beta_1 x + \varepsilon$ que permita predecir el número de errores a partir del número de peticiones simultaneas por segundo. Calcular e interpretar una medida de la bondad del ajuste realizado.

(b) (0.75 ptos) Estimar, mediante un intervalo de confianza al 90%, el número esperado de errores que se producirán cuando se reciban 90 peticiones de conexión simultáneas.

Ejercicio 5: El número de defectos en cada placa de cobre de 1m^2, resultante de un proceso de fabricación, es una variable aleatoria que sigue una distribución de Poisson de media 0.1 defectos.

(a) (0.25 ptos) ¿Cuál es la probabilidad de encontrar al menos dos defectos en una placa?

(b) (0.75 ptos) ¿Cuál es la probabilidad de encontrar, a lo más, 40 defectos en 500 placas?

Ejercicio 6: Sea $X_1, \ldots, X_n$ una muestra aleatoria simple de una variable aleatoria cuya función de probabilidad viene dada por

$$P[X = k] = \begin{cases} \left(\frac{\theta}{2}\right)^{|k|} \cdot (1-\theta)^{1-|k|} & \text{si} \quad k \in \{-1, 0, 1\} \\ \\ 0 & \text{en otro caso} \end{cases}$$

(a) (0.5 ptos) Demostrar que $E[X] = 0$ y que $Var[X] = \theta$.

(b) (0.5 ptos) Es $T = \frac{\sum_{i=1}^{n} X_i^2}{5n}$ un estimador insesgado de θ? En caso negativo calcular el sesgo de T como estimador de θ.

(c) (0.75 ptos) Determinar razonadamente la función de probabilidad conjunta de la muestra. ¿Es $T_1 = \frac{\sum_{i=1}^{n} |X_i|}{n} + 1$ un estadístico suficiente para estimar θ.

Ejercicio 7: En un programa de capacitación industrial, algunos aprendices son instruidos con el método A, consistente en recibir instrucción mecanizada, y algunos son capacitados con el método B, que entraña también la atención personal de un instructor. Se toman dos muestras aleatorias de aprendices capacitados por cada uno de los métodos y se les somete a un test de aprovechamiento. La siguiente tabla resume los resultados de las puntuaciones obtenidas en ambas muestras:

	Tamaño muestral	Media	Varianza
Método A	10	70	10.2
Método B	8	74.5	13.35

Supuesto que las puntuaciones de los aprendices siguen distribuciones normales, se pide:

(a) (0.5 ptos) Plantear y resolver un contraste de hipótesis adecuado, con $\alpha = 0.05$, para estudiar la igualdad de varianzas de ambas poblaciones.

(b) (0.75 ptos) ¿Se puede afirmar que la puntuación promedio que obtienen en el test los aprendices capacitados con el método B es superior en más de tres puntos a la puntuación promedio que obtienen los aprendices capacitados con el método A? (Plantear y resolver un contraste de hipótesis adecuado con un nivel de significación del 5%).

24 de junio de 2015

DURACIÓN: HASTA LAS 13:10H.

Ejercicio 1

$$P(\bar{A}|_{\bar{B}}) = 1 - P(A|_{\bar{B}}) = 1 - \frac{P(A\cap\bar{B})}{P(\bar{B})} = 1 - \frac{P(A)-P(A\cap B)}{1-P(B)}$$

Sabemos que $P(A|_B) = 0'25$ y que $P(B|_A) = 0'5$. Entonces,

$$0'25 = P(A|_B) = \frac{P(A\cap B)}{P(B)} \Rightarrow P(A\cap B) = 0'25\cdot P(B)$$

$$0'5 = P(B|_A) = \frac{P(A\cap B)}{P(A)} \Rightarrow P(A\cap B) = 0'5\cdot P(A)$$

$\Rightarrow 0'25\cdot P(B) = 0'5\cdot P(A)$
$\Rightarrow P(B) = \frac{0'5}{0'25}\cdot P(A)$
$\Rightarrow \boxed{P(B) = 2\cdot P(A)}$

De los datos del problema, $P(A) = 0'2$, luego

$\boxed{P(B) = 2\cdot 0'2 = 0'4}$ y $\boxed{P(A\cap B) = 0'5\cdot P(A) = 0'5\cdot 0'2 = 0'1}$

Finalmente, sustituyendo, $P(\bar{A}|_{\bar{B}}) = 1 - \frac{0'2-0'1}{1-0'4} = 1 - \frac{0'1}{0'6} = \frac{5}{6}$

Ejercicio 2:

a) Consideramos las variables aleatorias X_A, X_B = número de chips defectuosos en una caja fabricada por la empresa A ó B, respectivamente. Dado que el 5% de los chips fabricados por A son defectuosos y que el 50% de los fabricados por B también lo son, resulta que $X_A \sim B(5, 0'05)$ y $X_B \sim B(5, 0'5)$. Entonces,

$$P(X_A = 2) = \binom{5}{2} 0'05^2 \cdot 0'95^3 \simeq 0'0214$$

$$P(X_B = 2) = \binom{5}{2}\cdot 0'5^2\cdot 0'5^3 = 0'3125$$

b) Consideramos ahora los sucesos:

D = una caja elegida al azar contiene exactamente dos chips defectuosos.

A = " " " " " procede de la empresa A.

B = " " " " " " " " " B.

Tenemos que calcular $P(B|_D)$:

$$P(B|_D) = \frac{P(B\cap D)}{P(D)} = \frac{P(D|_B)\cdot P(B)}{P(D|_A)\cdot P(A) + P(D|_B)\cdot P(B)} = \frac{P(X_B=2)\cdot P(B)}{P(X_A=2)\cdot P(A)+P(X_B=2)\cdot P(B)} =$$

$$= \frac{0'3125\cdot 0'1}{0'0214\cdot 0'9 + 0'3125\cdot 0'1} \simeq \underline{0'6187}$$

c) $X \sim N(10000, \sigma^2)$ y $P(X \leq 9000) = 0'1056$. Calculamos el valor de σ^2:

$$0'1056 = P(X \leq 9000) = P\left(\frac{X-10000}{\sigma} \leq \frac{9000-10000}{\sigma}\right) = P\left(Z \leq \frac{-1000}{\sigma}\right)$$
$Z \sim N(0,1)$

Entonces $0'1056 = P\left(Z \leq \frac{-1000}{\sigma}\right) = P\left(Z \geq \frac{1000}{\sigma}\right) = 1 - P\left(Z\leq \frac{1000}{\sigma}\right)$.

Despejando, $P\left(Z \leq \frac{1000}{\sigma}\right) = 1 - 0'1056 = 0'8944$

Además, de la tabla de la distribución $N(0,1)$ $P(Z \leq 1'25) = 0'8944$

$$\Rightarrow \frac{1000}{\sigma} = 1'25 \Rightarrow \boxed{\sigma = 800}$$

Como consecuencia, $X \sim N(1000, 800^2)$ y el beneficio esperado viene dado por:

$$E(B(X)) = E(-3500 - X + 0'00015\cdot X^2) = -3500 - E(X) + 0'00015\cdot E(X^2) =$$

$$= -3500 - E(X) + 0'00015\,(Var(X) + E^2(X)) =$$

$$= -3500 - 10000 + 0'00015\,(800^2 + 10000^2) = \underline{1596\ €}$$

Ejercicio 3

a) Para que f sea función de densidad, debe cumplir:

i) $f(x) \geq 0,\ \forall x \in \mathbb{R}$, ii) $\int^{+\infty} f(x)\,dx = 1$

Si imponemos que se cumpla ii, resulta:

$$1=\int_{-\infty}^{+\infty} f(x)\,dx = \int_0^2 \frac{cx}{2}\,dx + \int_2^6 c\cdot\left(\frac{-x}{4}+\frac{3}{2}\right)dx = c\cdot\left(\frac{x^2}{4}\Big|_0^2\right) + c\cdot\left(-\frac{x^2}{8}+\frac{3}{2}x\Big|_2^6\right) =$$

$$= c\cdot\left(\frac{4}{4}-0\right) + c\cdot\left(-\frac{36}{8}+\frac{18}{2}-\left(\frac{-4}{8}+\frac{6}{2}\right)\right) = c\cdot(1+2) = 3c$$

Luego, para que se cumpla ii, debe ser $\boxed{c=\frac{1}{3}}$

Además, si $c=\frac{1}{3}$ también se cumple i, ya que $c>0$, $\frac{x}{2}>0$ si $0<x<2$

y $-\frac{x}{4}+\frac{3}{2} > \frac{-6}{4}+\frac{3}{2}=0$ si $2<x<6$. En consecuencia, $\boxed{c=\frac{1}{3}}$.

b) $$P(1\le X\le 3) = \int_1^3 f(x)\,dx = \int_1^2 \frac{1}{3}\cdot\frac{x}{2}\,dx + \int_2^3 \frac{1}{3}\left(\frac{-x}{4}+\frac{3}{2}\right)dx =$$

$$= \int_1^2 \frac{x}{6}\,dx + \int_2^3 \frac{-x}{12}+\frac{1}{2}\,dx = \frac{x^2}{12}\Big|_1^2 + \left(-\frac{x^2}{24}+\frac{1}{2}x\Big|_2^3\right) = \frac{1}{4}+\frac{7}{24} = \underline{\underline{\frac{13}{24}}}$$

c) Definimos la variable aleatoria beneficio como:

$$B(X) = \begin{cases} 2 & \text{si } X<1 \\ 3 & \text{si } 1\le X\le 3 \\ 4 & \text{si } X>3 \end{cases}$$

Entonces el beneficio esperado viene dado por

$$E(B(X)) = 2\cdot P(X<1) + 3\cdot P(1\le X\le 3) + 4\cdot P(X>3)$$

$$P(X<1) = \int_0^1 f(x)\,dx = \int_0^1 \frac{1}{3}\frac{x}{2}\,dx = \frac{x^2}{12}\Big|_0^1 = \frac{1}{12}$$

$$P(X>3) = 1-P(X<1)-P(1\le X\le 3) = 1-\frac{13}{24}-\frac{1}{12} = \frac{24-13-2}{24} = \frac{9}{24} = \frac{3}{8}$$

Luego $E(B(X)) = 2\cdot\frac{1}{12} + 3\cdot\frac{13}{24} + 4\cdot\frac{3}{8} = \frac{4}{24}+\frac{39}{24}+\frac{36}{24} = \frac{79}{24} \simeq 3'29$ €.

Ejercicio 4

a) Tenemos que ajustar un modelo de la forma $y = \beta_0 + \beta_1 \cdot x + \varepsilon$ en el que 'x' representa el nº de peticiones por segundo e 'y' es el número de errores. De los datos del problema se obtiene que:

$\bar{x} = 112'5 \qquad S_x^2 = 3281'25 \qquad S_{xy} = \dfrac{\sum x_i \cdot y_i}{n} - \bar{x}\cdot\bar{y} = 17531'25 - 112'5 \cdot 129'5$

$\bar{y} = 129'5 \qquad S_y^2 = 2809'25 \qquad = 2962'5$

La recta de regresión es, entonces, $y - 129'5 = \dfrac{2962'5}{3281'25}(x - 112'5)$

$$\boxed{y = 0'9029 \cdot x + 27'9286}$$

Calculamos ahora el coeficiente de determinación. Al tratarse de un ajuste de regresión lineal simple, el coeficiente de determinación se puede calcular como

$$R^2 = \left(\frac{S_{xy}}{S_x S_y}\right)^2 = \frac{S_{xy}^2}{S_x^2 \cdot S_y^2} = 0'9521$$

Podemos decir entonces que el modelo explica el 95'21% de la variabilidad de la variable número de errores.

b) El intervalo de confianza que se debe calcular viene dado por

$$\hat{y}(x_0) \pm S \cdot \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{n \cdot S_x^2}} \cdot t_{n-2,1-\alpha/2} \qquad \text{con } x_0 = 90 \text{ y } \alpha = 0'1$$

$\hat{y}(x_0) = \hat{y}(90) = 0'9029 \cdot 90 + 27'9286 = 109'1896$

$$S^2 = \frac{SS_{Res}}{n-2} = \frac{SS_{Tot} - SS_{Reg}}{n-2} = \frac{22474 - 21399'7458}{8-2} = 179'0424$$

$SS_{Tot} = n \cdot S_y^2 = 8 \cdot 2809'25 = 22474$

Finalmente, $t_{n-2,1-\alpha/2} = t_{8-2,1-0'1/2} = t_{6,0'95} = 1'943$ y el intervalo de confianza viene dado por:

$$109'1896 \pm \sqrt{179'0424} \cdot \sqrt{\frac{1}{8} + \frac{(90-112'5)^2}{8 \cdot 3281'25}} \cdot 1'943$$

esto es, $109'1896 \pm 13'3807 \cdot 0'3799 \cdot 1'943 \rightarrow \boxed{(99'3127,\ 119'0665)}$

Ejercicio 5

X = número de defectos en un metro cuadrado de placa.

$X \sim \mathcal{P}(0'1)$

a) $P(X \geq 2) = 1 - P(X < 2) = 1 - \left[P(X=0) + P(X=1)\right] = 1 - \frac{e^{-0'1} \cdot 0'1^0}{0!} - \frac{e^{-0'1} \cdot 0'1^1}{1!}\Big] =$

$= 1 - e^{-0'1} - e^{-0'1} \cdot 0'1 = 0'0047$

b) Sea X_i = nº de defectos en el i-ésimo metro cuadrado.
$i = 1, \ldots, 500$

Entonces cada $X_i \sim \mathcal{P}(0'1)$ y las variables $X_1, \ldots, X_{500}$ son independientes, con lo cual si Y = número de defectos en 500 metros cuadrados,

$Y = \sum_{i=1}^{500} X_i$ y $Y \sim \mathcal{P}(500 \cdot 0'1) = \mathcal{P}(50)$.

Calculamos ahora $P(Y \leq 40)$. Aproximaremos esta probabilidad utilizando la distribución normal. Sea entonces $W \sim N(50, 50)$.

$$P(Y \leq 40) \simeq P(W \leq 40) = P\left(\frac{W-50}{\sqrt{50}} \leq \frac{40-50}{\sqrt{50}}\right) = P(\underset{Z \sim N(0,1)}{Z} \leq -1'41) =$$

$$= P(Z \geq 1'41) = 1 - P(Z \leq 1'41) = 1 - 0'9207 = \boxed{0'0793}$$

Ejercicio 6

a). $\cdot\ E(X) = -1 \cdot P(X=-1) + 0 \cdot P(X=0) + 1 \cdot P(X=1) = -1 \cdot P(X=-1) + 1 \cdot P(X=1) =$

$$= -1 \cdot \left(\frac{\theta}{2}\right)^1 \cdot (1-\theta)^{1-1} + 1 \cdot \left(\frac{\theta}{2}\right)^1 \cdot (1-\theta)^{1-1} = -\frac{\theta}{2} + \frac{\theta}{2} = \underline{0}$$

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

$E(X^2) = (-1)^2 \cdot P(X=-1) + 0^2 \cdot P(X=0) + 1^2 \cdot P(X=1) = P(X=-1) + P(X=1) =$

$$= \left(\frac{\theta}{2}\right) \cdot (1-\theta)^{1-1} + \left(\frac{\theta}{2}\right) \cdot (1-\theta)^{1-1} = \frac{\theta}{2} + \frac{\theta}{2} = \underline{\theta}$$

Entonces $Var(X) = \theta - 0^2 = \underline{\theta}$

b) $E(T) = E\left(\dfrac{\sum_{i=1}^n X_i^2}{5n}\right) = \dfrac{1}{5n} \cdot \sum_{i=1}^n E(X_i^2) = \dfrac{1}{5n} \cdot \sum_{i=1}^n E(X^2) =$

$$= \frac{1}{5n} \cdot \sum_{i=1}^n \left(Var(X) + E^2(X)\right) = \frac{1}{5n} \cdot \sum_{i=1}^n (\theta + 0^2) = \frac{1}{5n} \cdot n\theta = \frac{\theta}{5}$$

Entonces $E(T) = \frac{\theta}{5} \neq \theta$, luego T no es un estimador insesgado de θ.

El sesgo de T es: $|E(T) - \theta| = \left|\frac{\theta}{5} - \theta\right| = \left|\frac{\theta - 5\theta}{5}\right| = \frac{4\theta}{5}$

c) Función de probabilidad conjunta de la muestra:

$$P(X_1 = x_1; \dots; X_n = x_n) = \prod_{i=1}^n P(X_i = x_i) = \prod_{i=1}^n P(X = x_i) = \prod_{i=1}^n \left(\frac{\theta}{2}\right)^{|x_i|} \cdot (1-\theta)^{1-|x_i|} =$$

$$= \left(\frac{\theta}{2}\right)^{\sum_{i=1}^n |x_i|} \cdot (1-\theta)^{n - \sum_{i=1}^n |x_i|}$$

Por otra parte, $T_1 = \dfrac{\sum_{i=1}^n |x_i|}{n} + 1 \Rightarrow \sum_{i=1}^n |x_i| = n \cdot (T_1 - 1)$. Entonces

$$P(X_1 = x_1; \dots; X_n = x_n) = \left(\frac{\theta}{2}\right)^{\sum |x_i|} \cdot (1-\theta)^{n - \sum |x_i|} = \left(\frac{\theta}{2}\right)^{n \cdot (T_1 - 1)} \cdot (1-\theta)^{n - n(T_1 - 1)} =$$

$$= \left(\frac{\theta}{2}\right)^{n \cdot (T_1 - 1)} \cdot (1-\theta)^{2n - n \cdot T_1} \cdot 1$$

Luego T_1 es suficiente

Ejercicio 7:

X_A = puntuación de un aprendiz capacitado con el método A.
X_B = " " " " " " " " B.

$X_A \sim N(\mu_A, \sigma_A^2)$, $X_B \sim N(\mu_B, \sigma_B^2)$, X_A, X_B independientes.

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

Rechazaremos H_0 si $\dfrac{S_{C_A}^2}{S_{C_B}^2} \geq f_{n_A-1, n_B-1, 1-\alpha/2}$ o bien $\dfrac{S_{C_A}^2}{S_{C_B}^2} \leq \dfrac{1}{f_{n_B-1, n_A-1, 1-\alpha/2}}$

De los datos del problema, $n_A = 10$, $n_B = 8$, $S_{CA}^2 = 10'2$ y $S_{C_B}^2 = 13'3$, luego

$$S_{C_A}^2 = \frac{n_A \cdot S_A^2}{n_A - 1} = \frac{10 \cdot 10'2}{9} \simeq 11'3333 \; ; \; S_{C_B}^2 = \frac{n_B \cdot S_B^2}{n_B - 1} = \frac{8 \cdot 13'35}{7} \simeq 15'2571$$

$$f_{n_A-1, n_B-1, 1-\alpha/2} = f_{9,7,0'975} = 4'823 \; ; \; f_{n_B-1, n_A-1, 1-\alpha/2} = f_{7,9,0'975} = 4'197$$

Entonces se rechazará H_0 si $\dfrac{S_{C_A}^2}{S_{C_B}^2} \geq 4'823$ o bien $\dfrac{S_{C_A}^2}{S_{C_B}^2} \leq \dfrac{1}{4'197} \simeq 0'2383$.

Puesto que $\dfrac{S_{C_A}^2}{S_{C_B}^2} = \dfrac{11'3333}{15'2571} = 0'7428$ no se cumple la regla de rechazo, por lo que no tenemos evidencia significativa para afirmar que las varianzas de ambas poblaciones son distintas.

b) Tenemos que resolver el contraste $\begin{cases} H_0: \mu_B = \mu_A + 3 \\ H_1: \mu_B > \mu_A + 3 \end{cases}$, o lo que es igual, $\begin{cases} H_0: \mu_A - \mu_B = -3 \\ H_1: \mu_A - \mu_B < -3 \end{cases}$ con $\alpha = 0'05$. Teniendo en cuenta el resultado obtenido en el apartado anterior, se trata de un contraste sobre la diferencia de medias de poblaciones normales con varianzas desconocidas pero iguales.

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

Siendo $\delta_0 = -3$, $t_{n_A+n_B-2,1-\alpha} = t_{10+8-2,1-0'05} = t_{16,0'95} = 1'746$

y $S_p^2 = \frac{(n_A-1) S_{cA}^2 + (n_B-1) \cdot S_{cB}^2}{n_A+n_B-2} = \frac{n_A \cdot S_A^2 + n_B \cdot S_B^2}{n_A+n_B-2} = \frac{10 \cdot 10'2 + 8 \cdot 13'35}{10+8-2} \simeq 13'05$

Entonces se rechazará H_0 si

$$\bar{X}_A - \bar{X}_B \leq -3 - \sqrt{13'05} \cdot \sqrt{\frac{1}{10}+\frac{1}{8}} \cdot 1'746 = -5'9919$$

Al ser $\bar{X}_A - \bar{X}_B = 70 - 74'5 = -4'5$ no se cumple la regla de rechazo por lo que no tenemos evidencia para rechazar H_0 y no podemos afirmar que la puntuación promedio que se obtiene con el método B supere en más de tres puntos a la que se obtiene con el método A.