ASIGNATURA: ESTADÍSTICA Y PROCESOS ESTOCÁSTICOS
EXAMEN EXTRAORDINARIO RESUELTO (CURSO 2017-18) **Duración:** 3 horas
FECHA: 5 de Julio de 2018

Ejercicio 1 *(1.5 puntos) Cierta convocatoria consta de dos exámenes que se realizan uno a continuación del otro. Por convocatorias anteriores se sabe que la probabilidad de aprobar el primer examen es* 0.6*, la probabilidad de aprobar el segundo si se ha suspendido el primero es* 0.4 *y la probabilidad de aprobar los dos es* 0.3*.*

a) *Calcula la probabilidad de que un alumno apruebe el segundo examen.*

Solución. Definimos los sucesos
E_1: aprobar el primer examen.
E_2: aprobar el segundo examen.

Los datos que nos dan son $P(E_1) = 0.60$, $P(E_2/\overline{E_1}) = 0.40$ y $P(E_1 \cap E_2) = 0.30$.

Vamos a usar el Teorema de la Probabilidad Total para calcular $P(E_2)$.

$$P(E_2) = P(E_1) \cdot P(E_2/E_1) + P(\overline{E_1}) \cdot P(E_2/\overline{E_1})$$

Entonces,

$$P(E_2/E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)} = \frac{0.30}{0.60} = 0.50$$

Sustituyendo, tenemos que:

$$P(E_2) = P(E_1) \cdot P(E_2/E_1) + P(\overline{E_1}) \cdot P(E_2/\overline{E_1}) = 0.60 \cdot 0.50 + 0.40 \cdot 0.40 = 0.46$$

b) *¿Cuál es la probabilidad de que un alumno apruebe alguno de los dos exámenes?*

Solución. Hay que calcular $P(E_1 \cup E_2)$. Entonces,

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2) = 0.60 + 0.46 - 0.30 = 0.76$$

c) *Si sabemos que un alumno ha aprobado al menos uno de los dos exámenes, halla la probabilidad de que no haya aprobado el primer examen.*

Solución. Nos piden

$$P(\overline{E_1}/E_1 \cup E_2) = \frac{P(\overline{E_1} \cap [E_1 \cup E_2])}{P(E_1 \cup E_2)} = \frac{P(\overline{E_1} \cap E_2)}{P(E_1 \cup E_2)} =$$
$$= \frac{0.16}{0.76} = 0.2105$$

donde hemos utilizado que

$$P(\overline{E_1} \cap E_2) = P(\overline{E_1}) \cdot P(E_2/\overline{E_1}) = 0.40 \cdot 0.40 = 0.16$$

Ejercicio 2 *(1.5 puntos) Se sabe que un lote tiene 2 artículos defectuosos y 3 no defectuosos. Si se inspeccionan al azar uno después de otro,*

a) Obtén la función de probabilidad de la variable aleatoria X que mide "el número de artículos que habrá que inspeccionar a fin de obtener todos los defectuosos".

Solución. X toma los valores 2, 3, 4 y 5.

$$P(X=2) = P(D_1 \cap D_2) = \frac{2}{5} \cdot \frac{1}{4} = \frac{1}{10}$$

$$P(X=3) = P(D_1 \cap \overline{D}_2 \cap D_3) + P(\overline{D}_1 \cap D_2 \cap D_3) = \frac{2}{5} \cdot \frac{3}{4} \cdot \frac{1}{3} + \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{1}{3} = \frac{1}{5}$$

$$P(X=4) = P(D_1 \cap \overline{D}_2 \cap \overline{D}_3 \cap D_4) + P(\overline{D}_1 \cap D_2 \cap \overline{D}_3 \cap D_4) + P(\overline{D}_1 \cap \overline{D}_2 \cap D_3 \cap D_4)$$

$$= \frac{2}{5} \cdot \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} + \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{2}{3} \cdot \frac{1}{3} + \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} = \frac{3}{10}$$

$$P(X=5) = 1 - \sum_{i=2}^{4} P(X=i) = \frac{4}{10}$$

b) Calcula la esperanza de X.

Solución.

$$E(X) = 2P(X=2) + 3P(X=3) + 4P(X=4) + 5P(X=5) = 4$$

Ejercicio 3 *(1.5 puntos) Las ventas trimestrales, V, de una empresa, en miles de euros, se distribuyen según una variable aleatoria normal de media 50 y varianza 3, y los costes trimestrales, C, también en miles de euros, siguen una distribución normal de media 45 y varianza 1. Las variables V y C son independientes.*

a) Calcula la probabilidad de que en un mismo trimestre se obtengan más de 51000 euros de ventas y menos de 44000 euros de costes.

Solución. V y C son independientes, por lo que

$$P(V > 51, C < 44) = P(V > 51) \cdot P(C < 44) = P\left(Z > \frac{1}{\sqrt{3}}\right) \cdot P(Z < -1) =$$

$$= [1 - \Phi(0.58)] \cdot [1 - \Phi(1)] = (1 - 0.71904) \cdot (1 - 0.84134) = 0.28096 \cdot 0.15866 = 0.044577$$

donde Φ es la función de distribución de la variable aleatoria $Z \sim N(0,1)$.

b) Halla la probabilidad de que los beneficios brutos, $B = V - C$, estén comprendidos entre 4000 y 6000 euros.

Solución. Por ser $B = V - C$ combinación lineal de variables aleatorias normales, B es una variable aleatoria normal. $E[B] = E[V - C] = E[V] - E[C] = 50 - 45 = 5$.
Por ser V y C independientes, $Var[V - C] = Var[V] + (-1)^2 Var[C] = 3 + 1 = 4$. Por tanto, $B \sim \mathcal{N}(5,\ 4)$.

Para hallar $P(4 \leq B \leq 6)$ tipificamos la variable aleatoria B, como $Z = \dfrac{B-5}{\sqrt{4}}$ que ya es una normal de media 0 y varianza 1. Ahora, utilizando la simetría y consultando las tablas

$$P(4 \leq B \leq 6) = P\left(\frac{4-5}{2} \leq Z \leq \frac{6-5}{2}\right) = P(-0.5 \leq Z \leq 0.5) = P(Z \leq 0.5) - P(Z \leq -0.5) =$$

$$= P(Z \leq 0.5) - (1 - P(Z \leq 0.5)) = 2 \cdot P(Z \leq 0.5) - 1 = 2 \cdot 0.69146 - 1 = 0.38292.$$

Así, la probabilidad de que los beneficios brutos estén comprendidos entre 4000 y 6000 euros es 0.38292.

c) Si los primeros 2000 euros de beneficios están exentos de impuestos y sobre los restantes se paga un 20 %, halla la probabilidad de pagar más de 500 euros de impuestos en un trimestre.

Solución. Si el 20 % de $(B-2)$ es 0.5, $\dfrac{20}{100}(B-2) = 0.5 \Rightarrow B - 2 = 5 \cdot 0.5 \Rightarrow B = 4.5$. Entonces,

$$P(B \geq 4.5) = P\left(\frac{B-5}{2} \geq \frac{4.5-5}{2}\right) = P(Z \geq -0.25) = P(Z < 0.25) = 0.59871$$

Ejercicio 4 *(1.5 puntos) Sea (X, Y) una variable aleatoria bidimensional con función de densidad conjunta*

$$f(x,y) = \begin{cases} \dfrac{2x+4y+1}{10} & si \quad 0 \leq x < 2, \quad 0 < y < 1 \\ 0 & en\ el\ resto \end{cases}$$

a) Halla las funciones de densidad marginales de X e Y.

Solución.

$$f_X(x) = \begin{cases} \displaystyle\int_0^1 \frac{2x+4y+1}{10}\,dy = \frac{1}{10}\left(2xy + 2y^2 + y\Big|_0^1\right) = \frac{2x+3}{10} & \text{si} \quad 0 \leq x < 2 \\ 0 & \text{en el resto} \end{cases}$$

$$f_Y(y) = \begin{cases} \displaystyle\int_0^2 \frac{2x+4y+1}{10}\,dx = \frac{1}{10}\left(x^2 + (4y+1)x\Big|_0^2\right) = \frac{4y+3}{5} & \text{si} \quad 0 \leq y < 1 \\ 0 & \text{en el resto} \end{cases}$$

b) Calcula $P(X > Y)$.

Solución.

$$P(X > Y) = \frac{1}{10}\int_0^1 dy \int_y^2 (2x+4y+1)dx = \frac{1}{10}\int_0^1 \left(x^2 + (4y+1)x\Big|_y^2\right) dy$$

$$= \frac{1}{10}\int_0^1 (-5y^2 + 7y + 6)dy = \frac{1}{10}\left(-\frac{5y^3}{3} + \frac{7y^2}{2} + 6y\Bigg|_0^1\right) = \frac{47}{60} = 0.7833$$

O bien,

$$
\begin{aligned}
P(X>Y) &= \frac{1}{10}\int_0^1 dx\int_0^x (2x+4y+1)dy+\frac{1}{10}\int_1^2 dx\int_0^1 (2x+4y+1)dy=\\
&= \frac{1}{10}\int_0^1 \left((2xy+2y^2+y)\Big|_0^x\right)dx+\frac{1}{10}\int_1^2 \left((2xy+2y^2+y)\Big|_0^1\right)dx=\\
&= \frac{1}{10}\int_0^1 (4x^2+x)\,dx+\frac{1}{10}\int_1^2 (2x+3))\,dx=\frac{1}{10}\left(\frac{4x^3}{3}+\frac{x^2}{2}\Big|_0^1\right)+\frac{1}{10}\left(x^2+3x\Big|_1^2\right)=\\
&= \frac{1}{10}\left(\frac{4}{3}+\frac{1}{2}\right)+\frac{1}{10}(4+6-1-3)=\frac{1}{10}\frac{11}{6}+\frac{6}{10}=\frac{47}{60}=0.7833
\end{aligned}
$$

c) *Calcula $E(X\cdot Y)$.*

Solución.

$$
\begin{aligned}
E(X\cdot Y) &= \frac{1}{10}\int_0^1 dy\int_0^2 xy(2x+4y+1)dx=\frac{1}{10}\int_0^1\left(\frac{2x^3}{3}y+2y^2x^2+y\frac{x^2}{2}\Big|_0^2\right)dy\\
&= \frac{1}{10}\int_0^1(\frac{16}{3}y+8y^2+2y)dy=\frac{1}{10}\left(\frac{16}{3\cdot 2}4y^2+\frac{8y^3}{3}+y^2\Big|_0^1\right)=\frac{19}{30}=0.6333
\end{aligned}
$$

Ejercicio 5 *(1 punto) De los 10 árboles de cierta especie a los que tenemos acceso, 9 son hembras y 1 macho. Se quiere contrastar si la proporción de árboles hembra p de esta especie es $1/2$. A partir de estos datos realizamos el test con hipótesis nula*

$$H_0 \equiv p=\tfrac{1}{2}$$

y alternativa $H_1 \equiv p\neq \frac{1}{2}$. Calcula el p-valor del test. (calcula el valor exacto del p-valor, no el que se obtiene mediante una aproximación con una distribución normal).
¿Aceptarías con un nivel de significación de 0.01 que $p=1/2$?

Solución. Si la hipótesis nula es cierta, la variable aleatoria $N\equiv$ nº de árboles hembra que hay en los 10 árboles, tiene distribución

$$\text{Binomial}\,(n=10,p=1/2)$$

El p-valor es entonces

$$\text{p-valor}=2\left[P(N=9)+P(N=10)\right]=2\left[10\frac{1}{2^{10}}+\frac{1}{2^{10}}\right]=\frac{11}{2^9}\approx 0.0215$$

Como $0.01<0.0215$ se acepta la hipótesis nula $p=1/2$.

Ejercicio 6 *(1 punto) Se miden 25 tornillos, obteniendose una media de 30 mm y una cuasidesviación típica de $s_1=0.06724$ mm. Utilizando estos datos, los cuantiles*

$$
\begin{aligned}
q_{0.005}&=9.88,\ q_{0.01}=10.85,\ q_{0.025}=12.40,\ q_{0.05}=13.84,\\
q_{0.95}&=36.41,\ q_{0.975}=39.36, q_{0.99}=42.97,\ q_{0.995}=45.55
\end{aligned}
$$

de una χ^2_{24} y el estadístico $\dfrac{(n-1)S_1^2}{\sigma^2}$, que tiene distribución χ^2_{n-1}, calcula:

(a) *Un intervalo de confianza con nivel del* 90 % *para la desviación típica* σ *de la variable aleatoria*

$$X \equiv \textit{longitud en milímetros de un tornillo}$$

Solución. Los cuantiles de ordenes 0.05 y 0.95 de una chi-cuadrado con 24 grados de libertad, son

$$q_{0.05} = 13.84, \quad q_{0.95} = 36.41$$

y se verifica que

$$0.9 = P\Big(q_{0.05} \leq \frac{(n-1)S_1^2}{\sigma^2} \leq q_{0.95}\Big) = P\Big(\frac{(n-1)S_1^2}{q_{0.95}} \leq \sigma^2 \leq \frac{(n-1)S_1^2}{q_{0.05}}\Big)$$

Por tanto, el intervalo de confianza de la varianza con nivel del 95 % es

$$\Big(\frac{24 \cdot 0.06724^2}{36.41}, \frac{24 \cdot 0.06724^2}{13.84}\Big) \approx (0.00298, 0.00784)$$

El correspondiente intervalo para la desviación típica es

$$(0.0546\,,\,0.0885)\ \text{mm}$$

(b) *El p-valor del test con hipótesis nula*

$$H_0 \equiv \textit{la desviación típica de } X \textit{ es } \sigma = 0.1$$

y alternativa $H_1 \equiv \sigma \neq 0.1$ *¿Para qué valores de* α *se rechazaría* H_0*?*

Solución. Se tiene que

$$\frac{(n-1)s_1^2}{\sigma^2} = \frac{24 \cdot 0.067242}{0,01} \approx 10.85$$

y entonces

$$\text{p-valor} = 2P(\chi^2_{24} < 10.85) = 2 \cdot 0.01 = 0.02$$

Se rechazaría H_0 para valores de α mayores de 0.02.

Ejercicio 7 *(2 puntos) Sea* $Z(t)$ *un proceso estocástico definido por* $Z(t) = 2\,t\,X - Y$ *donde* X, Y *son variables aleatorias independientes con distribuciones* $X \sim N(\mu = 0,\ \sigma = 1)$*,* $Y \sim N(\mu = 2,\ \sigma = 2)$*.*

a) *Calcula la media, la función de autocorrelación y la varianza de* $Z(t)$ *¿Es* $Z(t)$ *un proceso estacionario en sentido amplio?*

Solución. Calculamos $E[Z(t)]$

Calculamos $E[Z(t)]$

$$E[Z(t)] = E[2tX - Y] = 2tE[X] - E[Y] = -2$$

A continuación, calculamos la función de autocorrelación del proceso

$$R_Z(t, t+\tau) = E[Z(t) \cdot Z(t+\tau)] = E[(2tX - Y) \cdot (2(t+\tau)X - Y)] =$$

$$= E[4t(t+\tau)X^2] - 2tE[XY] - 2(t+\tau)E[XY] + E[Y^2] =$$

$$= 4t(t+\tau)E[X^2] + E[Y^2] = 4t(t+\tau) + 8$$

donde hemos utilizado que

$$E[XY] = E[X] \cdot E[Y] = 0 \text{ al ser } X \text{ e } Y \text{ independientes}$$
$$\text{VAR}(X) = 1 = E[X^2] - (E[X])^2) \Leftrightarrow E[X^2] = 1$$
$$\text{VAR}(Y) = 4 = E[Y^2] - (E[Y])^2) \Leftrightarrow E[Y^2] = 4 + (E[Y])^2) = 8$$

Podemos decir que el proceso $Z(t)$ no es estacionario en sentido amplio puesto que su función de autocorrelación $R_Z(t, t+\tau)$ depende de t.

Podemos calcular la varianza mediante

$$\text{VAR}[Z(t)] = R_Z(t,t) - E[Z(t)]^2 = 4t^2 + 8 - 4 = 4t^2 + 4$$

o bien mediante

$$\text{VAR}[Z(t)] = \text{VAR}[2tX - Y] = (2t)^2\text{VAR}[X] + \text{VAR}[Y] = 4t^2 + 4$$

b) Calcula $E[Z(1) \cdot Z(2)]$

Solución.

$$E[Z(1) \cdot Z(2)] = R_Z(t = 1, t+\tau = 2) = 4 \cdot 1 \cdot 2 + 8 = 16$$

c) Obtén la distribución de la variable aleatoria bidimensional $[Z(1), Z(2)]$.

Solución. Como X e Y son variables aleatorias normales independientes, la distribución conjunta es normal bidimensional con vector de medias

$m = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$ y matriz de varianzas $M = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$.

El vector del que nos pieden calcular la distribución, $(Z(1)Z(2))$, es una transformación lineal del vector (X, Y). En concreto:

$$\begin{pmatrix} Z(1) \\ Z(2) \end{pmatrix} = \begin{pmatrix} 2X - Y \\ 4X - Y \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 4 & -1 \end{pmatrix} \cdot \begin{pmatrix} X \\ Y \end{pmatrix} = A \cdot \begin{pmatrix} X \\ Y \end{pmatrix}$$

Por tanto, la distribución del vector $(Z(1)Z(2))$ es normal bidimensional con vector de medias $A \cdot m = \begin{pmatrix} -2 \\ -2 \end{pmatrix}$ y matriz de varianzas $A \cdot M \cdot A^t = \begin{pmatrix} 8 & 12 \\ 12 & 20 \end{pmatrix}$.

d) Calcula $P[Z(1) > -3]$

Solución. Las variables componentes de una distribución normal bidimensional son normales. Por tanto, $Z(1) \sim N(-2, \sigma^2 = 8)$. Entonces,

$$P(Z(1) > -3) = P(Z > \frac{-3+2}{\sqrt{8}}) = P(Z > -0.35) = P(Z < 0.35) = 0.6368$$

donde $Z \sim N(0, 1)$.