ASIGNATURA: ESTADÍSTICA Y PROCESOS ESTOCÁSTICOS
EXAMEN FINAL (Primavera 2018-19) **Duración:** 3 horas
FECHA: 31 de Mayo de 2019

SOLUCIONES

Ejercicio 1 *(1,5 puntos) Los materiales para el montaje de un ordenador personal pueden provenir de uno de los seis posibles proveedores con porcentajes 9 %, 16 %, 25 %, 25 %, 16 % y 9 %, respectivamente. La probabilidad de que un ordenador sea fiable si está hecho con los materiales que proporcionan dichos proveedores es 0.6, 0.7, 0.6, 0.8, 0.6 y 0.9, respectivamente.*

(a) Halla la probabilidad de montar un ordenador fiable.

(b) Supuesto que un ordenador no es fiable, haya la probabilidad de que se halla montado con materiales del cuarto proveedor.

(c) Se revisan sucesivamente la fiabilidad de ordenadores montados con materiales del sexto fabricante. Halla la probabilidad de que el primer ordenador que es no fiable es el 4º revisado.

Solución.

(a) Sea F el suceso "el ordenador es fiable". Sea el suceso M_i "el ordenador se monta con materiales provenientes del fabricante i-ésimo", con $i = 1, \ldots, 6$. Entonces con los datos dados tenemos

$$\begin{array}{cccccc} P(M_1) = 0,09 & P(M_2) = 0,16 & P(M_3) = 0,25 & P(M_4) = 0,25 & P(M_5) = 0,16 & P(M_6) = 0,09 \\ P(F/M_1) = 0,6 & P(F/M_2) = 0,7 & P(F/M_3) = 0,6 & P(F/M_4) = 0,8 & P(F/M_5) = 0,6 & P(F/M_6) = 0,9 \end{array}$$

Puesto que $M_i \cap M_j = \emptyset$ si $i \neq j$, $P(M_i) \neq 0$ para cualquier $i = 1, \ldots, 6$ y $\sum_{i=1}^{6} P(M_i) = 1$, los sucesos $\{M_1, \ldots M_6\}$ forman un sistema completo de sucesos. Aplicando el teorema de la probabilidad total,

$$P(F) = \sum_{i=1}^{6} P(M_i)P(F/M_i) = 0,09 \cdot 0,6 + 0,16 \cdot 0,7 + 0,25 \cdot 0,6 + 0,25 \cdot 0,8 + 0,16 \cdot 0,6 + 0,09 \cdot 0,9 = 0,693$$

(b) Aplicando el teorema de Bayes

$$P(M_4/\overline{F}) = \frac{P(\overline{F}/M_4)P(M_4)}{P(\overline{F})} = \frac{(1 - P(F/M_4))P(M_4)}{1 - P(F)} = \frac{(1 - 0,8)0,25}{1 - 0,693} \simeq 0,1629$$

(c) Sea N la variable aleatoria "número de ordenadores montados con materiales del sexto fabricante hasta obtener el primero no fiable", N es una variable aleatoria geométrica, $Geo(p = 0,1)$

$$P(X = 4) = 0,1 \cdot 0,9^3 = 0,0729$$

Ejercicio 2 *(1 punto) Sea $f_X(x) = \frac{1}{2}e^{-|x|}$, con $-\infty < x < \infty$, la función de densidad de la variable aleatoria X.*

(a) *Halla la probabilidad del suceso:* $|X| < 2$ *o* $X > 0$.

(b) *Halla la función de densidad de la variable aleatoria* $Y = 2X + 1$.

Solución.

a) $P(|X| < 2 \cup X > 0) = P(X > -2) = \int_{-2}^{\infty} f_X(x)\, dx = \int_{-2}^{0} \frac{1}{2} e^{-|x|}\, dx + \int_{0}^{\infty} \frac{1}{2} e^{-|x|}\, dx.$

$\int_{-2}^{0} \frac{1}{2} e^{-|x|}\, dx = \frac{1}{2} \int_{-2}^{0} e^{-(-x)}\, dx = \frac{1}{2}\, e^{x}\Big|_{-2}^{0} = \frac{1}{2}(1 - e^{-2}) \backsimeq 0,4323$

$\int_{0}^{\infty} \frac{1}{2} e^{-|x|}\, dx = \frac{1}{2} \int_{0}^{\infty} e^{-x}\, dx = -\frac{1}{2}\, e^{-x}\Big|_{0}^{\infty} = -\frac{1}{2}(0 - e^{0}) = \frac{1}{2}$. *Por tanto,*

$$P(|X| < 2 \cup X > 0) = P(X > -2) = \int_{-2}^{\infty} f_X(x)\, dx = 1 - \frac{e^{-2}}{2} \backsimeq 0,9323$$

b) *Calculamos la relación entre la función de distribución de* X, $F_X(x)$, *y la de* Y, $F_Y(y)$.

$$F_Y(y) = P(Y \leq y) = P(2X + 1 \leq y) = P\left(X \leq \frac{y-1}{2}\right) = F_X\left(\frac{y-1}{2}\right)$$

Para obtener la función de densidad, $f_Y(y)$, *derivamos la función de distribución*

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{dF_X\left(\frac{y-1}{2}\right)}{dy} = f_X\left(\frac{y-1}{2}\right) \cdot \frac{1}{2} = \frac{1}{4} e^{-\left|\frac{y-1}{2}\right|}, \quad -\infty < y < \infty$$

$$f_Y(y) = \begin{cases} \frac{1}{4} e^{\frac{y-1}{2}} & si\ -\infty < y \leq 1 \\ \\ \frac{1}{4} e^{-\frac{y-1}{2}} & si\ 1 < y \end{cases}$$

Ejercicio 3 *(1,5 puntos) Consideramos las variables aleatorias* X *e* Y *cuya función de probabilidad conjunta bidimensional viene dada por:*

X \ Y	-1	0	1
-2	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
-1	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{8}$
1	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{8}$
2	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$

(a) *Calcula las distribuciones marginales de* X *e* Y.

(b) *Calcula las probabilidades:* $P(X = i/Y = 1)$, *para* $i = -2,\ -1,\ 1,\ 2$; *y* $P(Y = j/X = 1)$, *para* $j = -1, 0, 1$.

(c) *Halla la covarianza entre* X *e* Y. *¿Son* X *e* Y *independientes?*

Solución.

(a)

X	-2	-1	1	2
P(X=i)	$\frac{3}{16}$	$\frac{5}{16}$	$\frac{5}{16}$	$\frac{3}{16}$

Y	-1	0	1
P(Y=j)	$\frac{3}{8}$	$\frac{2}{8}$	$\frac{3}{8}$

(b)

X/Y=1	*-2*	*-1*	*1*	*2*
P(X=i/Y=1)	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{2}{6}$	$\frac{1}{6}$

Y/X=1	*-1*	*0*	*1*
P(Y=j/X=1)	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{2}{5}$

(c) La covarianza entre dos variables aleatorias puede expresarse como $Cov[X,Y] = E[XY] - E[X]E[Y]$.

Utilizando las distribuciones marginales calculadas en el apartado (a) se obtiene

$$E[X] = -2 \cdot \frac{3}{16} + (-1) \cdot \frac{5}{16} + 1 \cdot \frac{5}{16} + 2 \cdot \frac{3}{16} = 0$$
$$E[Y] = -1 \cdot \frac{3}{8} + 0 \cdot \frac{2}{16} + 1 \cdot \frac{3}{8} = 0$$

Con la función de probabilidad conjunta por columnas tenemos,

$$E[XY] = [(-1)\cdot(-2)]\frac{1}{16} + [(-1)\cdot(-1)]\frac{1}{8} + [(-1)\cdot 1]\frac{1}{8} + [(-1)\cdot 2]\frac{1}{16} +$$
$$+[1\cdot(-2)]\frac{1}{16} + [1\cdot(-1)]\frac{1}{8} + [1\cdot 1]\frac{1}{8} + [1\cdot 2]\frac{1}{16} = 0$$

Por tanto, $COV(X,Y) = 0 - 0 \cdot 0$.

$P(X=-2)\cdot P(Y=-1) = \frac{3}{16}\cdot\frac{3}{8} = \frac{3\cdot 3}{16\cdot 8} = \frac{9}{128} \neq \frac{1}{16} = P(X=-2, Y=-1)$.

Por lo que X *e* Y *no son independientes aunque sean incorreladas.*

Ejercicio 4 *(1,5 puntos) La función de densidad conjunta del vector aleatorio* (X,Y) *viene dada por*

$$f(x,y) = \begin{cases} \frac{kx}{y^2} & si \quad x \in (0,1),\ y > 2 \\ 0 & en\ otro\ caso \end{cases}$$

(a) Calcula la constante k.

(b) Calcula las distribuciones marginales. ¿Son X *e* Y *independientes?*

(c) Calcula $P(Y \leq 4X)$.

Solución.

(a) Imponemos que la integral de la función de densidad en todo el plano tiene que ser 1:

$$1 = \iint_{\mathbb{R}^2} f(x,y)\,dxdx = k\int_0^1 dx \int_2^\infty \frac{x}{y^2}\,dy = k\int_0^1 x \left.\frac{y^{-1}}{-1}\right|_2^\infty dx =$$
$$= k\int_0^1 \frac{x}{2}\,dx = k\frac{1}{2}\left.\frac{x^2}{2}\right|_0^1 = \frac{k}{4} \Rightarrow k = 4$$

(b) Sean $f_X(x)$ *y* $f_Y(y)$ *las funciones de densidad marginales de* X *e* Y*, respectivamente.*

Si $x \in (0,1)$, $f_X(x) = \int_2^\infty \frac{4x}{y^2}\,dy = 4x\left.\frac{y^{-1}}{-1}\right|_2^\infty = 2x \Rightarrow f_X(x) = \begin{cases} 2x & si \quad x \in (0,1) \\ 0 & en\ otro\ caso \end{cases}$

Si $y > 2$, $f_Y(y) = \int_0^1 \frac{4x}{y^2}\,dx = \left.\frac{2x^2}{y^2}\right|_0^1 = \frac{2}{y^2} \Rightarrow f_Y(y) = \begin{cases} \frac{2}{y^2} & si \quad y > 2 \\ 0 & en\ otro\ caso \end{cases}$

Puesto que si $x \in (0,1)$ *e* $y > 2$, $f_X(x)\cdot f_Y(y) = 2x\cdot\frac{2}{y^2} = \frac{4x}{y^2} = f(x,y)$ *y en el resto de los puntos* $f(x,y) = 0$ *y* $f_X(x) = 0$*, o bien* $f_Y(y) = 0$*. Por tanto,* $f(x,y) = f_X(x)\cdot f_Y(y)$ *en* $\mathbb{R}^2$*, con lo que* X *e* Y *son independientes.*

(c) $P(Y \leq 4X) = \int_2^4 dy \int_{y/4}^1 \frac{4x}{y^2}\, dx = \int_2^4 2\left(\frac{1}{y^2} - \frac{1}{16}\right)\, dy = \frac{1}{4} = 0,25$

Ejercicio 5 *(1 punto) Se está llevando a cabo un estudio de la duración (en horas) de la batería de un móvil que se va a lanzar al mercado. Para ello, se examinan 100 baterías resultando una duración de la descarga de 83,5 h. Suponiendo que el tiempo de descarga de las baterías sigue una distribución normal con varianza* $\sigma^2 = 24,2$ *se pide:*

(a) *Calcula un intervalo con un nivel de confianza del* $92\,\%$ *para la duración media de las baterías.*

(b) *¿Cual debe ser el tamaño mínimo de la muestra para que la amplitud del intervalo de confianza al* $92\,\%$ *para la duración media de las baterías sea menor de una hora?*

Solución.

(a) *Sea* X *la variable aleatoria "duración de la batería", entonces* $X \sim \mathcal{N}(\mu,\ \sigma^2 = 24,2)$. *Sabemos que* $\dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ *se distribuye como una variable aleatoria* $\mathcal{N}(0,1)$. *Entonces, si* q_α *es el cuantil de orden* α *de la variable aleatoria* $\mathcal{N}(0,1)$

$$0,92 = P\left(q_{0,04} < \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} < q_{0.96}\right) = P\left(-q_{0,96} \leq \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq q_{0.96}\right) \Rightarrow P\left(\overline{X} - \frac{\sigma}{\sqrt{n}} q_{0,96} \leq \mu \leq< \overline{X} + \frac{\sigma}{\sqrt{n}} q_{0,96}\right)$$

El cuantil de orden $0,96$ *de la variable aleatoria* $\mathcal{N}(0,1)$ *es 1,75, entonces*

$$I.C._{92\,\%}(\mu) = \left(\overline{X} - \frac{\sigma}{\sqrt{n}}\cdot 1,75,\ \overline{X} + \frac{\sigma}{\sqrt{n}}\cdot 1,75\right) = \left(83,5 - \frac{\sqrt{24,2}}{10}\cdot 1,75,\ 83,5 + \frac{\sqrt{24,2}}{10}\cdot 1,75\right) = (82,64,\ 84,36)$$

(b) *La longitud del intervalo anterior es* $2\cdot\dfrac{\sqrt{24,2}}{10}\cdot 1,75 = 1,72$, *entonces para que esa longitud sea inferior a 1 se debe cumplir*

$$2\cdot\frac{\sqrt{24,2}}{\sqrt{n}}\cdot 1,75 < 1 \Rightarrow 2\sqrt{24,2}\cdot 1,75 < \sqrt{n} \Rightarrow 296,45 < n$$

Luego, el tamaño mínimo de la muestra para que la amplitud del intervalo de confianza al $92\,\%$ *para la duración media de las baterías sea menor de una hora es* $n = 297$.

Ejercicio 6 *(0,5 puntos) La duración de un determinado tipo de pilas sigue una distribución normal, con desviación típica de 3 h. Un comercio desea adquirir una remesa en que la duración media de dichas pilas sea exactamente de 50 h. Antes de comprar la remesa comprueba la duración de una muestra aleatoria de tamaño 9 y realiza un contraste de hipótesis con un nivel de significación del 10 %. ¿Se rechaza el lote si la media de la muestra es de 48 h. 15 min.?*

Solución. *La hipótesis nula es* $H_0 : \mu = 50$ *y la hipótesis alternativa es* $H_1 : \mu \neq 50$.

Puesto que se conoce la desviación típica el estadístico del contraste es $T = \dfrac{\overline{X}-50}{\frac{3}{\sqrt{9}}} = \overline{X} - 50$. *Si la hipótesis nula es cierta* $T = \overline{X} - 50$ *se distribuye como una variable aleatoria* $\mathcal{N}(0,1)$.

Si $T \in (q_{0.05},\ q_{0.95})$, siendo q_α es el cuantil de orden α de la variable aleatoria $\mathcal{N}(0,1)$, entonces $P(T \in (q_{0.05},\ q_{0.95}) = 0.90$ y por tanto

$$q_{0.05} \le T \le\ q_{0.95} \Rightarrow -1,64 \le \overline{X} - 50 \le 1,64 \Rightarrow 50 - 1,64 \le \overline{X} \le 50 + 1,64 \Rightarrow R.A_{0,1} = (48,36,\ 51,64)$$

Puesto que 48 h. 15 min. no pertenece a dicho intervalo rechazamos, con un nivel de significación del 10 %, que la media de la duración de las pilas es 50 h.

Ejercicio 7 *(1 punto) Sean A, B y C tres variables aleatorias incorreladas dos a dos, las tres con media 0 y varianza 1. Sea el proceso estocástico $X(t) = A + B \cdot t + C \cdot t^2$. Halla la media y la autocorrelación del proceso. ¿Es estacionario en sentido estricto?*

Solución.

$$\mu_X(t) = E[X(t)] = E[A + B \cdot t + C \cdot t^2] = E[A] + E[B] \cdot t + E[C] \cdot t^2 = 0 + 0 \cdot t + 0 \cdot t^2 = 0$$

$$R_X(t_1, t_2) = E[(A + B \cdot t_1 + C \cdot t_1^2)(A + B \cdot t_2 + C \cdot t_2^2)] =$$

$$= E[A^2 + A \cdot B \cdot t_2 + A \cdot C \cdot t_2^2 + A \cdot B \cdot t_1 + B^2 \cdot t_1 \cdot t_2 + B \cdot t_1 \cdot C \cdot t_2^2 + A \cdot C \cdot t_1^2 + B \cdot C \cdot t_1^2 \cdot t_2 + C^2 \cdot t_1^2 \cdot t_2^2] =$$

$$= E[A^2] + E[A \cdot B](t_2 + t_1) + E[A \cdot C](t_2^2 + t_1^2) + E[B^2] \cdot t_1 \cdot t_2 + E[B \cdot C] \cdot (t_1 \cdot t_2^2 + t_1^2 \cdot t_2) + E[C^2] \cdot t_1^2 \cdot t_2^2$$

$$E[A \cdot B] = Cov[A, B] + E[A]E[B] = 0 + 0 = 0, \quad E[A^2] = V[A] + E[A]^2 = 1$$
$$E[A \cdot C] = Cov[A, C] + E[A]E[C] = 0 + 0 = 0, \quad E[B^2] = V[B] + E[B]^2 = 1$$
$$E[B \cdot C] = Cov[B, C] + E[B]E[C] = 0 + 0 = 0, \quad E[C^2] = V[C] + E[C]^2 = 1$$

Por tanto, $R_X(t_1, t_2) = 1 + t_1 \cdot t_2 + t_1^2 \cdot t_2^2$. Puesto que depende de t_1 y t_2 el proceso no es estacionario en sentido amplio y, por tanto, tampoco es estacionario en sentido estricto.

Ejercicio 8 *(1,5 puntos) Sea $X(t)$ un proceso estocástico que cumple que $(X(t), X(t+s) - X(t))$ sigue una distribución normal con vector de medias $(2t, 2s)$ y matriz de varianzas y covarianzas $\begin{pmatrix} t & 0 \\ 0 & s \end{pmatrix}$, para $t > 0$ y $s > 0$. Calcula*

(a) $P(X(5) > 10\ ,\ X(6) - X(5) > 2)$.

(b) La distribución de $(X(t)\ ,\ X(t+s))$ para $s > 0$.

(c) La función de autocovarianza, $C_X(t, t+s)$, del proceso $X(t)$.

Solución.

(a)

$$\begin{pmatrix} X(5) \\ X(6) - X(5) \end{pmatrix} \sim \mathcal{N}_2\left(\begin{pmatrix} 2 \cdot 5 \\ 2 \cdot (6-5) \end{pmatrix}, \begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix}\right)$$

Entonces $X(5) \sim \mathcal{N}(10, \sigma^2 = 5)$ y $X(6) - X(5) \sim \mathcal{N}(2, \sigma^2 = 1)$. Además, $X(5)$ y $X(6) - X(5)$ son independientes ya que $Cov(X(5), X(6) - X(5)) = 0$. Denotando por Z a la variable aleatoria $\mathcal{N}(0,1)$

$$P(X(5) > 10\ ,\ X(6) - X(5) > 2) = P(X(5) > 10)\ P(X(6) - X(5) > 2) = P(Z > 0)P(Z > 0) = 1/4$$

(b)

$$\begin{pmatrix} X(t) \\ X(t+s)-X(t) \end{pmatrix} \sim \mathcal{N}_2\left(\begin{pmatrix} 2t \\ 2s \end{pmatrix}, \begin{pmatrix} t & 0 \\ 0 & s \end{pmatrix}\right)$$

Entonces, como

$$\begin{pmatrix} X(t) \\ X(t+s) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} X(t) \\ X(t+s)-X(t) \end{pmatrix}$$

Será

$$\begin{pmatrix} X(t) \\ X(t+s) \end{pmatrix} \sim \mathcal{N}_2\left(\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 2t \\ 2s \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} t & 0 \\ 0 & s \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\right)$$

$$\begin{pmatrix} X(t) \\ X(t+s) \end{pmatrix} \sim N_2\left(\begin{pmatrix} 2t \\ 2(t+s) \end{pmatrix}, \begin{pmatrix} t & t \\ t & t+s \end{pmatrix}\right)$$

c) Del apartado anterior sabemos que

$$C_X(t,t+s) = Cov[X(t),X(t+s)] = t$$

Ejercicio 9 *(0,5 puntos) Contesta razonadamente cuales de las siguientes funciones pueden ser funciones de autocorrelación de procesos estacionarios en sentido amplío:*

(a) $R_1(\tau) = e^{-|\tau|}$

(b) $R_2(\tau) = e^{-\tau} \cdot \cos(\tau)$

(c) $R_3(\tau) = e^{-\tau^2}$

(d) $R_4(\tau) = e^{-\tau^2} \cdot \operatorname{sen}(\tau)$

Solución.

(a) La función $R_1(\tau)$ *sí puede ser función de autocorrelación de procesos estacionarios en sentido amplio por ser par y* $1 = |R_1(0)| \geq e^{-|\tau|} = R_1(\tau), \forall \tau \in \mathbb{R}$.

(b) $R_2(\tau)$ *no puede ser función de autocorrelación de procesos ESA porque* $R_2(0) = 1 < R_2(-2\pi) = e^{2\pi}$ *y no se cumple* $|R_2(-2\pi)| \leq R_2(0)$. *Además,*

$$R_2(\tau) = e^{-\tau} \cdot \cos(\tau) = e^{-\tau} \cdot \cos(-\tau) \neq e^{\tau} \cdot \cos(-\tau) = R_2(-\tau),$$

con lo que $R_2(\tau)$ *no es par.*

(c) $R_3(\tau)$ *si puede ser función de autocorrelación de procesos ESA por ser par y* $1 = |R_3(0)| \geq R_3(\tau) = e^{-\tau^2}, \forall \tau \in \mathbb{R}$.

(d) $R_4(\tau)$ *no puede ser función de autocorrelación de procesos ESA porque* $R_4(0) = 0$ *y no puede ser cero la varianza del proceso ya que no se verifica* $|R_4(\tau)| = e^{-\tau^2} \cdot |\operatorname{sen}(\tau)| \leq 0,\ \forall \tau \in \mathbb{R}$.