Departamento de Matemática Aplicada a las T.I.C.

ASIGNATURA: ESTADÍSTICA Y PROCESOS ESTOCÁSTICOS
EXAMEN EXTRAORDINARIO RESUELTO (CURSO 2018-19) **Duración:** 3 horas
FECHA: 27 de Junio de 2019 **Fecha publicación notas:** 8 de Julio de 2019
Fecha revisión examen: 12 de Julio de 2019

SOLUCIONES

Ejercicio 1 *(1 punto) Se forma una palabra de 7 letras ordenadas aleatoriamente y sin repetición con las letras A, B, C, D, E, F y G, por ejemplo BFACGDE o GFACBDE.*

a) Calcula la probabilidad de que aparezca la secuencia BCD.

b) Halla la probabilidad de que, en una misma palabra, estén las secuencias BA y GF.

Solución.

a) Todos los casos posibles para la formación de palabras son permutaciones de 7 elementos, 7!
Para averiguar los casos posibles consideramos las agrupaciones A, BCD, E, F y G, con lo que todas las ordenaciones posibles son las permutaciones de 5 elementos. Así, si S_1 es el suceso "aparece la secuencia BCD" se tiene

$$P(S_1) = \frac{\text{Casos favorables de que ocurra } S_1}{\text{Casos posibles}} = \frac{5!}{7!} = \frac{1}{42} \approx 0,0238$$

b) Sea el suceso S_2 de que estén las secuencias BA y GF en una misma palabra. Los casos posibles siguen siendo 7!
Ahora consideramos las agrupaciones BA, C, D, E y GF, entonces todas las ordenaciones posibles de las mismas son 5!, así que

$$P(S_2) = \frac{\text{Casos favorables de que ocurra } S_2}{\text{Casos posibles}} = \frac{5!}{7!} = \frac{1}{42} \approx 0,0238$$

Ejercicio 2 *(1 punto) Lanzamos un dado equilibrado tres veces y sean X_1, X_2 y X_3 los resultados de las tres tiradas.*

a) Sabiendo que el resultado del tercer lanzamiento ha sido mayor o igual que 5, ¿cuál es la probabilidad de que sea estrictamente mayor que los resultados de las dos primeras tiradas?

b) ¿Cuál es la probabilidad de que el resultado del primer lanzamiento sea 5 sabiendo que es estrictamente mayor que el resultado del segundo lanzamiento?

Solución.

a) Por ser el dado equilibrado todos los resultados son equiprobables y el número de casos posibles es $6 \cdot 6 \cdot 6 = 6^3$. Debemos hallar

$$P[(X_3 > X_1) \cap (X_3 > X_2)/X_3 \geq 5] = \frac{P[(X_3 > X_1) \cap (X_3 > X_2) \cap (X_3 \geq 5)]}{P(X_3 \geq 5)}$$

$P(X_3 \geq 5) = \dfrac{6 \cdot 6 \cdot 2}{6^3} = \dfrac{2}{6}$. Además, como el resultado de cada lanzamiento es independiente de los demás.

$$P[(X_3 > X_1) \cap (X_3 > X_2) \cap (X_3 \geq 5)] = P[(X_3 > X_1) \cap (X_3 > X_2) \cap (X_3 = 5) + P[(X_3 > X_1) \cap (X_3 > X_2) \cap (X_3 = 6) =$$

$$= P(X_3 = 5) \cdot P(X_1 < X_3 / X_3 = 5) \cdot P(X_2 < X_3 / (X_3 = 5 \cap X_1 < X_3) +$$

$$+ P(X_3 = 6) \cdot P(X_1 < X_3 / X_3 = 6) \cdot P(X_2 < X_3 / X_3 = 6 \cap X_1 < X_3) =$$

$$= P(X_3 = 5) \cdot P(X_1 < 5) \cdot P(X_2 < 5) + P(X_3 = 6) \cdot P(X_1 < 6) \cdot P((X_2 < 6) =$$

$$= \frac{1}{6} \cdot \frac{4}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} = \frac{4^2 + 5^2}{6^3} \approx 0,1898$$

Por tanto, la probabilidad pedida es

$$P[(X_3 > X_1) \cap (X_3 > X_2) / X_3 = 5] = \frac{\frac{41}{6^3}}{\frac{2}{6}} = \frac{41}{2 \cdot 6^2} = \frac{41}{72} \approx 0,5694$$

b) Ahora debemos calcular $P[X_1 = 5/(X_2 < X_1)] = \dfrac{P[(X_1 = 5) \cap (X_2 < X_1)]}{P[(X_2 < X_1)]}$. Para calcular el valor del denominador, $P(X_2 < X_1)$, utilizamos el teorema de la probabilidad total

$$P(X_2 < X_1) = \sum_{i=1}^{6} P(X_2 < X_1 / X_1 = i) \cdot P(X_1 = i) = P(X_2 < X_1 / X_1 = 1) \cdot P(X_1 = 1) + P(X_2 < X_1 / X_1 = 2) \cdot P(X_1 = 2) +$$

$$+ P(X_2 < X_1 / X_1 = 3) \cdot P(X_1 = 3) + P(X_2 < X_1 / X_1 = 4) \cdot P(X_1 = 4) +$$

$$+ P(X_2 < X_1 / X_1 = 5) \cdot P(X_1 = 5) + P(X_2 < X_1 / X_1 = 6) \cdot P(X_1 = 6) =$$

$$= \frac{0}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{6} + \frac{2}{6} \cdot \frac{1}{6} + \frac{3}{6} \cdot \frac{1}{6} + \frac{4}{6} \cdot \frac{1}{6} + \frac{5}{6} \cdot \frac{1}{6} = \frac{15}{6^2}$$

Otra forma: Los casos posibles para que $P(X_1 = X_2)$ son 6. De las restantes posibilidades, $36 - 6$, la mitad (15) cumple $X_1 > X_2$ y la otra mitad (15) cumple $X_1 < X_2$. Así que $P(X_2 < X_1) = \dfrac{15}{36}$.

Utilizando el teorema de Bayes

$$P[X_1 = 5/(X_2 < X_1)] = \frac{P[(X_1 = 5) \cap (X_2 < X_1)]}{P(X_2 < X_1)} = \frac{P(X_2 < X_1 / X_1 = 5) \cdot P(X_1 = 5)}{P(X_2 < X_1)} = \frac{\frac{4}{6} \cdot \frac{1}{6}}{\frac{15}{6^2}} = \frac{4}{15} \approx 0,2667$$

Ejercicio 3 *(1,5 puntos) El ayuntamiento ha puesto detectores de vehículos en 8 calles de similares características y con tráfico independiente. El número de vehículos $N(t)$ que circulan por cada una de esas calles es un proceso de Poisson de tasa $\lambda = 3$ vehículos por minuto. Por tanto, $N(1)$ sigue una distribución de Poisson de media 3.*

a) Nos fijamos en una única calle, ¿cual es la probabilidad de que en un minuto pasen a lo sumo dos vehículos?

b) ¿Cuál es la probabilidad de que en un determinado minuto haya habido exactamente seis calles por las que han circulado más de dos vehículos (por cada una)?

c) *Se anota el número de vehículos que circulan por una determinada calle durante dos minutos. Calcula la probabilidad de que durante el primer minuto pasara como mucho un vehículo pero al final de los dos minutos se contabilizaran por lo menos seis.*

Solución.

a) Sea $N(1) \equiv$ el número de vehículos que pasa en un minuto. Como $N(1)$ sigue una distribución de Poisson de media $\lambda = 3$

$$P(N(1) \leq 2) = P(N(1) = 0) + P(N(1) = 1) + P(N(1) = 2) = e^{-3}\left(\frac{3^0}{0!} + \frac{3^1}{1!} + \frac{3^2}{2!}\right) = 0,423$$

b) La probabilidad p de que en una calle hayan circulado más de dos vehículos es la complementaria a la calculada en el apartado anterior, $p = 1 - 0,423 = 0,577$.

Sea $C \equiv$ número de calles en las que han circulado más de dos vehículos. La variable aleatoria C sigue una distribución binomial de parámetros $n = 8$, $p = 0,577$. Por tanto la probabilidad pedida es:

$$P(C = 6) = \binom{8}{6} p^6 (1-p)^{8-6} = 0,185$$

c) Debemos hallar $P(N(1) \leq 1 \cap N(2) \geq 6)$. Sean $N_2(1)$ la variable aleatoria que mide el número de vehículos que pasa por una calle entre el primer y el segundo minuto. Las variables $N(1)$, $N_2(1)$ siguen una distribución de Poisson de media $\lambda = 3$ y son independientes.

$$P[N(1) \leq 1 \cap N(2) \geq 6] = P[(N(1) = 0 \cap (N(2) \geq 6)] + P[(N(1) = 1) \cap (N(2) \geq 6)] =$$

$$= P[(N(1) = 0) \cap (N_2(1) \geq 6)] + P[(N(1) = 1) \cap (N_2(1) \geq 5)] =$$

$$= P(N(1) = 0) \cdot P(N_2(1) \geq 6) + P(N(1) = 1) \cdot P(N_2(1) \geq 5) =$$

$$= e^{-3}\frac{3^0}{0!} \cdot (1 - P(N_2(1) < 6)) + e^{-3}\frac{3^1}{1!} \cdot (1 - P(N_2(1) < 5)) =$$

$$= e^{-3} \cdot \left(1 - e^{-3}\left(\frac{3^0}{0!} + \ldots + \frac{3^5}{5!}\right)\right) + 3 \cdot e^{-3}\left(1 - e^{-3}\left(\frac{3^0}{0!} + \ldots + \frac{3^4}{4!}\right)\right) \approx$$

$$\approx e^{-3} \cdot 0,0839 + 3 \cdot e^{-3} \cdot 0,1847 = 0,03176$$

Ejercicio 4 *(1,5 puntos) Una variable aleatoria X tiene función de densidad*

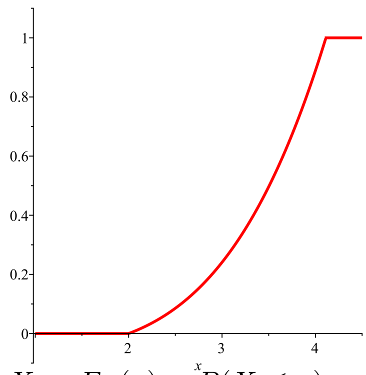
$$f(x) = \begin{cases} \dfrac{x}{6} & si \quad 2 \leq x \leq 4 \\ 0 & en\ el\ resto \end{cases}$$

a) *Halla y representa gráficamente la función de distribución de X.*

b) *Halla la media y la desviación típica de X.*

c) *Calcula el cuantil 0,4 de X.*

Solución.



a) La función de distribución de X es $F_X(x) = P(X \leq x) = \displaystyle\int_{-\infty}^{x} f(t)\ dt$ por lo que

$$F_X(x) = \begin{cases} 0 & \text{si } x < 2 \\ \displaystyle\int_2^x \frac{t}{6}\ dt = \frac{1}{12}\left(t^2\Big|_2^x\right) = \frac{x^2-4}{12} & \text{si} \quad 2 \leq x \leq 4 \\ 1 & \text{si } 4 < x \end{cases}$$

Y su gráfica es

b) Para hallar los momentos de la variable aleatoria solo consideramos el intervalo donde no se anula la función de densidad,

$$E[X] = \int_{-\infty}^{-\infty} x\ f(x)\ dx = \int_2^4 x\ \frac{x}{6}\ dx = \left(\frac{1}{18}x^3\Big|_2^4\right) = \frac{28}{9} \backsimeq 3,111$$

Para hallar la desviación típica, $\sigma_X = +\sqrt{V[X]}$, utilizamos que la varianza de X se puede expresar como $V[X] = E[X^2] - E[X]^2$.

$$E[X^2] = \int_{-\infty}^{-\infty} x^2\ f(x)\ dx = \int_2^4 x^2\ \frac{x}{6}\ dx = \left(\frac{1}{24}x^4\Big|_2^4\right) = 10$$

Luego, $\sigma_X = +\sqrt{10 - (\frac{28}{9})^2} = \sqrt{\dfrac{26}{81}} \approx 0,56656$

c) Debemos hallar el valor de a que verifica $P(X \leq a) = 0,4$. Entonces,

$$\frac{x^2-4}{12} = 0,4 \Rightarrow x^2 = 4 + 4,8 \Rightarrow x = \pm\sqrt{9,8}$$

Luego, $a = \sqrt{8,8} \backsimeq 2,966$, ya que la v.a. X solo toma valores entre 2 y 4.

Ejercicio 5 *(1,5 puntos) Sean X e Y dos variables aleatorias uniformes en el intervalo $(0,\ 1)$ e independientes.*

a) Calcula $P(Y > 2X)$.

b) Halla la función de distribución y de densidad de $Z = \max\{X, Y\}$.

Solución.

a) Las funciones de densidad de X e Y son:

$$f_X(x) = \begin{cases} 1 & \text{si} \quad 0 < x < 1 \\ 0 & \quad \text{en el resto} \end{cases} \quad ; \quad f_Y(y) = \begin{cases} 1 & \text{si} \quad 0 < y < 1 \\ 0 & \quad \text{en el resto} \end{cases}$$

Y sus funciones de distribución $F_X(t) = F_Y(t) = \begin{cases} 0 & \text{si } t \leq 0 \\ t & \text{si} \quad 0 < t < 1 \\ 1 & \text{si } 1 \leq t \end{cases}$.

Por ser las variables aleatorias independientes la función de densidad conjunta es el producto de las marginales, $f_{XY}(x,\ y) = f_X(x) \cdot f_Y(y)$, y solo es no nula en el cuadrado $(0,\ 1) \times (0,\ 1)$. Por lo que la probabilidad pedida es la integral sobre el triángulo T de la siguiente figura

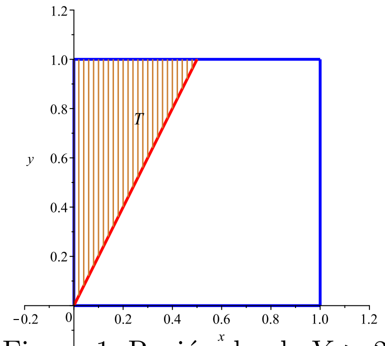


Figura 1: Región donde $Y > 2X$

Entonces, $P(Y > 2X) = 1 \cdot$ Área(T) $= \dfrac{1}{2} \cdot \dfrac{1}{2} \cdot 1 = \dfrac{1}{4}$, o bien

$$P(Y > 2X) = \iint_T f_{XY}(x,\ y)\, dx\, dy = \int_0^{0,5} \left(\int_{2x}^{1} 1\ dy \right) dx = \int_0^{0,5} (1 - 2x)\, dx = x - x^2 \Big|_0^{0.5} = \frac{1}{4} = 0,25$$

b) Sea $F_Z(z)$ la función de distribución de la variable aleatoria Z, entonces

$$F_Z(z) = P(Z \leq z) = P(\text{máx}\,\{X, Y\} \leq z) = P(X \leq z,\ Y \leq z)$$

ya que si el máximo es menor o igual que z cada variable aleatoria es menor o igual que z. Por ser las variables aleatorias X e Y independientes

$$F_Z(z) = P(Z \leq z) = P(X \leq z,\, Y \leq z) = P(X \leq z) \cdot P(Y \leq z) = F_X(z) \cdot F_Y(z) = F_X(z)^2 = \begin{cases} 0 & \text{si } z \leq 0 \\ z^2 & \text{si } 0 < z < 1 \\ 1 & \text{si } 1 \leq z \end{cases}$$

Y la función de densidad de Z es la derivada de su función de distribución, es decir

$$f_Z(z) = \begin{cases} 2z & \text{si} \quad 0 < z < 1 \\ 0 & \quad \text{en el resto} \end{cases}$$

Ejercicio 6 *(2 puntos) Se supone que la variable aleatoria $X \equiv$ peso en gramos de una manzana (de cierta parcela), tiene distribución normal con desviación típica 20 gramos. Se desea contrastar la hipótesis nula*

$$H_0 \equiv \mu = E[X] = 200 \textit{ gramos}$$

frente a la alternativa $H_1 \equiv \mu \neq 200$ gramos.

a) Si se pesan 400 manzanas, ¿para qué valores de la media del peso de estas 400 manzanas se aceptaría H_0 con un nivel de significación del 0,04?

Solución. Sea $\overline{x}$ la media del peso de las 400 manzanas. Se aceptaría H_0 siempre que

$$\frac{\overline{x}-200}{20/\sqrt{400}} = \overline{x}-200 \in (-z_{0.98}, z_{0.98}) \approx (-2.06, 2.06)$$

o equivalentemente cuando $\overline{x} \in (197.94\,,\,202,06)$ ($z_{0.98}$ denota el cuantil 0.98 de una N(0,1)).

b) Si la media del peso de 100 manzanas es 205 gramos, ¿cuál es el p-valor del test?

Solución. Como $\frac{205-200}{20/\sqrt{100}} = 2.5$ entonces

$$\text{p-valor} = 2 \times P(Z > 2.5) = 2 \times (1 - 0.99379) = 0.01242$$

donde Z es una v.a. con distribución normal de media 0 y desviación típica 1.

c) Si con un nivel de confianza del 98 % se desea obtener una estimación para la media μ con un margen de error menor que 1 gramo (un intervalo de confianza con longitud menor que 2), ¿cuántas manzanas se tienen que pesar?

Solución. El número de manzanas n tendría que verificar

$$z_{0.99}\frac{20}{\sqrt{n}} = 2.33\frac{20}{\sqrt{n}} < 1$$

o equivalentemente $n > (2.33 \times 20)^2 = 2171.56$. Debemos pues pesar al menos 2172 manzanas.

d) Si el peso medio fuese en efecto 200 gramos, y se pesaran 25 manzanas, ¿cuál sería la probabilidad de que el peso medio de esas 25 manzanas fuese mayor que 210 gramos?

Solución. La media del peso de las 25 manzanas $\overline{X}$ tiene distribución normal con media 200 y desviación típica $\frac{20}{\sqrt{25}} = 4$. Entonces

$$P(\overline{X} > 210) = P\Big(\frac{\overline{X}-200}{4} > 2.5\Big) \approx 1 - 0.99379 = 0.00621$$

Ejercicio 7 *(1,5 puntos) Un proceso normal y estacionario $X(t)$ tiene media $E[X(t)] = 0$ y función de autocorrelación $R_X(\tau) = 6e^{-\frac{|\tau|}{2}}$.*

a) Determina las distribuciones de primer y segundo orden del proceso.

b) Calcula $P(X(1) + X(2) > 1)$.

c) Calcula la distribución de $\begin{pmatrix} X(1)+X(2) \\ X(1)-X(2) \end{pmatrix}$. ¿Qué se puede afirmar de las variables aleatorias $X(1) + X(2)$ y $X(1) - X(2)$?

Solución.

a) El proceso $X(t)$ es normal por lo que la distribución de las variables aleatorias $X(t)$, para cualquier t, es normal y queda caracterizada por su media, μ y su varianza, σ^2. Puesto que la media del proceso es $\mu_X = E[X(t)] = 0$, la media de $X(t)$ es cero. Su varianza se puede obtener como $V[X(t)] = R_X(0) - \mu_X^2 = 6\,e^{-\frac{0}{2}} - 0 = 6$. Por tanto, $\quad X(t) \sim \mathcal{N}(0,\ \sigma^2 = 6)$.

También por ser el proceso $X(t)$ normal las distribución conjunta de las variables aleatorias $\begin{pmatrix} X(t) \\ X(t+\tau) \end{pmatrix}$ es normal bidimensional.

Como $\mu_X(t) = E[X(t)] = \mu_X$, por ser el proceso estacionario, y $\mu_X = 0$, por ser de media cero, el vector de medias de $\begin{pmatrix} X(t) \\ X(t+\tau) \end{pmatrix}$ es $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Ya hemos probado que $\sigma_X^2 = V[X(t)] = 6$ y $Cov[X(t),\ X(t+\tau)] = R_X(|t+\tau-t|) - \mu_X^2 = 6\ e^{-\frac{|\tau|}{2}}$.

Así que la matriz de varianzas y covarianzas es $\begin{pmatrix} 6 & 6\ e^{-\frac{|\tau|}{2}} \\ 6\ e^{-\frac{|\tau|}{2}} & 6 \end{pmatrix}$ y las distribuciones conjuntas de segundo orden son normales bidimensionales

$$\begin{pmatrix} X(t) \\ X(t+\tau) \end{pmatrix} \sim \mathcal{N}_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, 6 \begin{pmatrix} 1 & e^{-\frac{|\tau|}{2}} \\ e^{-\frac{|\tau|}{2}} & 1 \end{pmatrix} \right)$$

b) La variable aleatoria $X(1) + X(2)$ es una combinación lineal de variables aleatorias normales, por lo que es una variable aleatoria normal. Por ser el operador esperanza lineal,

$$E[X(1) + X(2)] = E[X(1)] + E[X(2)] = 0 + 0 = 0$$

Por las propiedades del operador varianza,

$$V[X(1)+X(2)] = V[X(1)]+V[X(2)]+2\cdot 1\cdot 1\cdot Cov[X(1), X(2)] = 6+6+2\ 6\ e^{-\frac{|2-1|}{2}} \approx 19,28$$

Tipificando, considerando que Z es la variable aleatoria $\mathcal{N}(0,\ 1)$, con función de distribución $\Phi(x)$, y consultando las tablas

$$P(X(1)+X(2) > 1) = P\left(\frac{X(1) + X(2) - 0}{\sqrt{19,28}} > \frac{1-0}{\sqrt{19,28}} \right) = P(Z > 0,23) = 1-\Phi(0,23) = 1-0,59095 = 0,40905$$

c) La variable bidimensional dada la podemos expresar como $\begin{pmatrix} X(1) + X(2) \\ X(1) - X(2) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} X(1) \\ X(2) \end{pmatrix}$

Puesto que $X(t)$ es un proceso normal y las variables $X(1) + X(2)$ y $X(1) - X(2)$ son combinaciones lineales de variables aleatorias normales tienen, a su vez, distribución normal con vector de medias $\begin{pmatrix} E[X(1) + X(2)] \\ E[X(1) - X(2)] \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

Y la matriz de varianzas y covarianzas de es

$$\Sigma = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 6 & 6\ e^{-\frac{1}{2}} \\ 6\ e^{-\frac{1}{2}} & 6 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = 6 \begin{pmatrix} 2 + 2\ e^{-\frac{1}{2}} & 0 \\ 0 & 2 - 2\ e^{-\frac{1}{2}} \end{pmatrix}$$

Por tanto,

$$\begin{pmatrix} X(1) + X(2) \\ X(1) - X(2) \end{pmatrix} \sim \mathcal{N}_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, 6 \begin{pmatrix} 2 + 2\ e^{-\frac{1}{2}} & 0 \\ 0 & 2 - 2\ e^{-\frac{1}{2}} \end{pmatrix} = \begin{pmatrix} 19,28 & 0 \\ 0 & 4,72 \end{pmatrix} \right)$$

Por ser la covarianza de $X(1) + X(2)$ y $X(1) - X(2)$ cero, estas variables son independientes.