ASIGNATURA: ESTADÍSTICA Y PROCESOS ESTOCÁSTICOS
EXAMEN FINAL (Otoño 2017-18) **Duración:** 3 horas
FECHA: 10 de Enero de 2018
Fecha publicación notas: 18 de Enero de 2018
Fecha revisión examen: 22 de Enero de 2018

SOLUCIONES

Ejercicio 1 *(1.5 puntos) n una juguetería hay nueve bicicletas de las que cuatro son verdes, tres azules y dos rojas. Un niño A entra en la juguetería y elige al azar (de forma equiprobable) una de las nueve bicicletas. Sus padres, la compran. Inmediatamente después, el niño B entra en la juguetería y elige aleatoriamente una bicicleta entre las ocho que quedan.*

a) ¿Cuál es la probabilidad de que B elija una bicicleta verde?

b) Si B ha elegido una bici verde, ¿cuál es la probabilidad de que A haya elegido una verde?

c) ¿Cuál es la probabilidad de que al menos uno de ellos haya elegido una verde?

d) ¿Cuál es la probabilidad de que A y B elijan bicicletas de diferente color?

a) ***Solución.***

Llamemos A_v, A_a y A_r a los sucesos "A elige una bici verde, azul y roja", respectivamente.

De igual modo, llamemos B_v, B_a y B_R a los sucesos "B elige una bici verde, azul y roja", respectivamente.

La probabilidad de que elijan una bici verde A y B es la misma, es decir, $P(B_v) = 4/9$.

También se puede obtener esa probabilidad utilizando el teorema de la Probabilidad Total,

$$P(B_v) = P(B_v/A_v)\cdot P(A_v)+P(B_v/A_a)\cdot P(A_a)+P(B_v/A_r)\cdot P(A_r) = \frac{3}{8}\cdot\frac{4}{9}+\frac{4}{8}\cdot\frac{3}{9}+\frac{4}{8}\cdot\frac{2}{9} = \frac{4}{9} = 0.\widehat{4}$$

b) ***Solución.***

$$P(A_v/B_v) = \frac{P(A_v\cap B_v)}{P(B_v)} = \frac{P(B_v/A_v)\cdot P(A_v)}{P(B_v)} = \frac{\frac{3}{8}\cdot\frac{4}{9}}{\frac{4}{9}} = \frac{3}{8} = 0.375$$

c) ***Solución.***

$$P(A_v\cup B_v) = P(A_v)+P(B_v)-P(A_v\cap B_v) = \frac{4}{9}+\frac{4}{9}-\frac{3}{8}\cdot\frac{4}{9} = \frac{13}{18} = 0.7\widehat{2}$$

d) ***Solución.***

$$P(\text{elijan las bicis de diferente color}) = 1 - P(\text{elijan las bicis del mismo color})$$

$$P(\text{elijan las bicis del mismo color}) = P[(A_v\cap B_v)\cup(A_a\cap B_a)\cup(A_r\cap B_r)] = P(A_v\cap B_v)+P(A_a\cap B_a)+P(A_r\cap B_r)$$

$$P(A_a\cap B_a) = P(B_a/A_a)\cdot P(A_a) = \frac{2}{8}\cdot\frac{3}{9}$$
$$P(A_r\cap B_r) = P(B_r/A_r)\cdot P(A_r) = \frac{1}{8}\cdot\frac{2}{9}$$

$$P(\text{elijan las bicis de diferente color}) = 1-\left(\frac{3}{8}\cdot\frac{4}{9}+\frac{2}{8}\cdot\frac{3}{9}+\frac{1}{8}\cdot\frac{2}{9}\right) = \frac{13}{18} = 0.7\widehat{2}$$

Ejercicio 2 *(1.5 puntos) El número de peticiones que tiene un servicio informático a lo largo de una hora es una variable aleatoria de Poisson de media 3.*

a) Halla la probabilidad de que no haya más de una petición en una hora.

b) Se considera el número de peticiones que llegan entre las 10 y las 11 horas en cuatro días distintos y suponemos independencia entre las peticiones en los distintos días. Halla la probabilidad de que en más de 3 de esos días no haya más de una petición.

c) Si sabemos que hay, al menos, una petición en una hora ¿cuál es la probabilidad de que sean menos de 3?

Solución.
Sea $N =$ el número de peticiones del servicio informático en una hora.
La media de una variable aleatoria de Poisson es su parámetro, con lo que la función de probabilidad de N es $P(N=k) = e^{-3} \cdot \dfrac{3^k}{k!}$, con $k = 0, 1, 2, \dots$

a) Que no haya más de una petición en una hora es equivalente a que en una hora tenga el servicio, a lo más, una petición

$$P(N \leq 1) = P(N=0) + P(N=1) = e^{-3} \cdot \left(\frac{3^0}{0!} + \frac{3^1}{1!}\right) = 4 \cdot e^{-3} \simeq 0.1991$$

b) Sea X la variable aleatoria: Número de días, de los 4, en que no haya más de una petición. Entonces, $X \sim Bin(4,\ 0.1991)$, por lo que

$$P(X > 3) = P(X = 4) = \binom{4}{4}(4 \cdot e^{-3})^4 = 0.0016$$

c)

$$P(N < 3/N \geq 1) = \frac{P(1 \leq N < 3)}{P(N \geq 1)} = \frac{P(N=1) + P(N=2)}{1 - P(N < 1)} = \frac{e^{-3} \cdot \left(\dfrac{3}{1!} + \dfrac{3^2}{2!}\right)}{1 - P(N=0)} = \frac{e^{-3}(6+9)}{2(1-e^{-3})} \simeq 0.39297$$

Ejercicio 3 *(1 punto) Un dispositivo electrónico sufre dilatación debido al calor. Esa dilatación, en milímetros, es una variable aleatoria X que tiene por función de densidad*

$$f(x) = \begin{cases} ax & si\ 0 \leq x \leq 3 \\ b & si\ 3 < x < 5 \\ \dfrac{b}{3}(8-x) & si\ 5 \leq x \leq 8 \\ 0 & en\ el\ resto \end{cases}$$

siendo a y b dos números reales.

a) Determina a y b sabiendo que $f(x)$ es continua en todo $\mathbb{R}$.

b) En un dispositivo se ha observado que la dilatación es superior a 3 mm. Si el dispositivo deja de funcionar cuando la dilatación es superior a 7 mm, ¿con qué probabilidad funciona el dispositivo?

Solución.

a) Imponemos que $\displaystyle\int_{\mathbb{R}} f(x)\,dx = 1$

$$\int_0^3 ax\,dx+\int_3^5 b\,dx+\int_5^8 \frac{b}{3}(8-x)\,dx = a\left.\frac{x^2}{2}\right|_0^3 + b\,x|_3^5 - \frac{b}{3}\left.\frac{(8-x)^2}{2}\right|_5^8 = \frac{9}{2}a+2b+\frac{b}{3}\cdot\frac{9}{2} = \frac{9a+7b}{2} = 1$$

Por otra parte, la función $f(x)$ es continua en todo $\mathbb{R}$ salvo en 0, 3, 5 y 8, por ser en cada intervalo una función polinómica.

$\lim\limits_{x\to 0^-} 0 = 0 = \lim\limits_{x\to 0^+} ax = 0$, con lo que no hay condiciones sobre a.

$\lim\limits_{x\to 3^-} ax = 3a = \lim\limits_{x\to 3^+} b = b$, con lo que $b = 3a$.

$\lim\limits_{x\to 5^-} b = b = \lim\limits_{x\to 5^+} \frac{b}{3}(8-x) = b$, con lo que no hay condiciones sobre b.

$\lim\limits_{x\to 8^-} \frac{b}{3}(8-x) = 0 = \lim\limits_{x\to 8^+} 0$, con lo que no hay condiciones sobre b.

Sustituyendo $b = 3a$ en $\frac{9a+7b}{2} = 1$ obtenemos $1 = \frac{9a+21a}{2} = 15a$. Así que $a = \frac{1}{15}$ y $b = \frac{3}{15} = \frac{1}{5}$.

b) Se ha observado que $X > 3$ y, para que funcione el dispositivo $X \leq 7$. Por lo que debemos calcular $P(X \leq 7/X > 3) = \dfrac{P(3 < X \leq 7)}{P(X > 3)}$.

$$P(3 < X \leq 7) = \int_3^5 \frac{1}{5}\,dx + \int_5^7 \frac{1}{5}\frac{8-x}{3}\,dx = \frac{5-3}{5} - \frac{1}{15}\left.\frac{(8-x)^2}{2}\right|_5^7 = \frac{10}{15} = \frac{2}{3}$$

$$P(X > 3) = 1 - P(X \leq 3) = 1 - \int_0^3 \frac{1}{15}x\,dx = 1 - \frac{1}{15}\left.\frac{x^2}{2}\right|_0^3 = 1 - 0.3 = 0.7$$

La probabilidad de que funcione el dispositivo si la dilatación es superior a 3 mm es $P(X \leq 7/X > 3) = \dfrac{2/3}{0.7} = 0.\widehat{952380}$.

Ejercicio 4 *(1.5 puntos) Se preguntó a seis estudiantes cuántas veces reiniciaron el ordenador la última semana. Cuatro de ellos eran usuarios de Mac y los otros dos restantes eran usuarios de PC.*

Los dos usuarios de PC lo reiniciaron 2 y 3 veces. Los cuatro usuarios de Mac lo reiniciaron 1, 2, 2 y 8 veces.

Sea C la variable aleatoria que representa el tipo de ordenador de un estudiante elegido al azar entre los 6 (Mac=0, PC=1). Sea R el número de veces que un estudiante, elegido al azar entre los 6, reinicia el ordenador. Observamos que R toma los valores 1, 2, 3 y 8.

a) ¿Cuál es la probabilidad de que un usuario de Mac reinicie 2 veces el ordenador?, ¿Cuál es la probabilidad de que un usuario de PC reinicie 3 veces el ordenador?

b) Construye la función de probabilidad conjunta del vector aleatorio (C, R). Calcula las distribuciones marginales.

c) Calcula $E(C)$ y $E(R)$.

d) Calcula $Cov(C, R)$ e interpreta cómo es la relación entre C y R. ¿Son C y R independientes?

Solución.

a) $$P(R = 2/C = 0) = 1/2, \quad P(R = 3/C = 1) = 1/2.$$

b) $$P(C=i,R=j)=P(R=j/C=i)P(C=i), \quad i=0,1,\ j=1,2,3,8$$

Sabemos que

$$P(R=1/C=0)=1/4, \quad P(R=2/C=0)=1/2$$

$$P(R=3/C=0)=0, \quad P(R=8/C=0)=1/4$$

$$P(R=1/C=1)=0, \quad P(R=2/C=1)=1/2$$

$$P(R=3/C=1)=1/2, \quad P(R=8/C=1)=0$$

$$P(C=0)=2/3, \quad P(C=1)=1/3$$

$$P(R=j)=P(C=0,R=j)+P(C=1,R=j), \quad j=1,2,3,8$$

Por tanto, la función de probabilidad conjunta y las funciones de probabilidad marginales vendrán dadas por

R \ C	0	1	P(R=j)
1	1/6	0	1/6
2	1/3	1/6	1/2
3	0	1/6	1/6
8	1/6	0	1/6
$P(C=i)$	2/3	1/3	

c)

$$E(C)=0\cdot\frac{2}{3}+1\cdot\frac{1}{3}=\frac{1}{3}$$

$$E(R)=1\cdot\frac{1}{6}+2\cdot\frac{1}{2}+3\cdot\frac{1}{6}+8\cdot\frac{1}{6}=3$$

d)

$$Cov(C,R)=E(C\cdot R)-E(C)\cdot E(R)$$

$$E(C\cdot R)=2\cdot\frac{1}{6}+3\cdot\frac{1}{6}=\frac{5}{6}\Rightarrow Cov(C,R)=\frac{5}{6}-1=\frac{-1}{6}$$

Como la covarianza entre ambas es distinta de 0 no son independientes.

Ejercicio 5 *(1 punto) La distribución de probabilidad de la variable aleatoria X depende del parámetro θ. Sea $X_1,\dots,X_n$ es una muestra aleatoria simple de X y se sabe que el estadístico*

$$T=\frac{\text{máx}\{X_1,\dots,X_n\}-\text{mín}\{X_1,\dots,X_n\}}{n+\theta}$$

se distribuye como una chi cuadrado con $n+4$ grados de libertad ($T\sim\chi^2_{n+4}$). Calcula un intervalo de confianza para θ, con un nivel de confianza del 95 %, siendo 16, 18, 11, 23, 25, 14, 19 una muestra de X.

Solución.

$$T=\frac{\text{máx}\{X_1,\dots,X_n\}-\text{mín}\{X_1,\dots,X_n\}}{n+\theta}\sim\chi^2_{n+4}$$

Por ser $n=7$ y el nivel de confianza $0.95=1-\alpha\Rightarrow\dfrac{\alpha}{2}=0.025$. Los cuantiles 0.025 y 0.975 de una chi cuadrado con 11 grados de libertad son (en tablas) $q_{0.025}=3.8157$ y $q_{0.975}=21.92$, respectivamente.

Entonces,

$$0.95=P\left(3.8157\leq\frac{\text{máx}\{X_1,\dots,X_n\}-\text{mín}\{X_1,\dots,X_n\}}{n+\theta}\leq 21.92\right)$$

$$=P\left(\frac{1}{21.92}\leq\frac{n+\theta}{\text{máx}\{X_1,\dots,X_n\}-\text{mín}\{X_1,\dots,X_n\}}\leq\frac{1}{3.8157}\right)$$

$$=P\left(\frac{\text{máx}\{X_1,\dots,X_n\}-\text{mín}\{X_1,\dots,X_n\}}{21.92}-n\leq\theta\leq\frac{\text{máx}\{X_1,\dots,X_n\}-\text{mín}\{X_1,\dots,X_n\}}{3.8157}-n\right)$$

Con los datos que tenemos,

$$\text{máx}\{16,18,11,23,25,14,19\}=25,\quad\text{mín}\{16,18,11,23,25,14,19\}=11$$

El intervalo de confianza pedido es

$$I.C._{0.95}(\theta)=\left(\frac{25-11}{21.92}-7\,,\frac{25-11}{3.8157}-7\right)=\left(\frac{14}{21.92}-7,\ \frac{14}{3.8157}-7\right)=(-6.3613,\ -3.3309)$$

Ejercicio 6 *(1 punto) Se supone que la duración de un neumático, circulando a 260 km/h, sigue una distribución normal. Se observa una muestra de* 9 *neumáticos para los que se obtiene que la media es* 8.6 *minutos y la cuasivarianza es* 0.6722 *minutos.*

Plantea un contraste para decidir si la desviación típica de la duración de los neumáticos, circulando a 260 km/h, es de 2 minutos. Obtén el p-valor y toma la decisión con un nivel de significación $\alpha=0.1$ *¿para qué valores de* α *se aceptaría la hipótesis nula?*

Solución.

Planteamos el contraste

$$H_0:\sigma=2$$
$$H_1:\sigma\neq 2$$

El estadístico del contraste es $T=\dfrac{(n-1)}{\sigma^2}S_1^2$, que tiene una distribución χ^2_8 cuando H_0 es cierta.

El $p-valor$ del contraste se obtiene como

$$p-valor=2\cdot min\{P(T\leq t),P(T\geq t)\},$$

siendo t el valor muestral de T supuesto cierta H_0.

El valor del estadístico T para nuestros datos es $t=\dfrac{8}{4}\cdot 0.6722=1.3444$. Entonces,

$$p-valor=2\cdot min\{P(T\leq 1.3444),P(T\geq 1.3444)\}=2\cdot 0.005=0.01$$

Como el $p-valor<\alpha=0.1$, la decisión del contraste es rechazar H_0 y aceptar H_1 es decir, los datos indican que $\sigma\neq 2$.

Los valores de α para los que se aceptaría la hipótesis nula son aquellos para los que se verifica que $\alpha<p-valor$, es decir, para $\alpha<0.01$.

Ejercicio 7 *(2.5 puntos) Sea $X(t)$ un proceso gaussiano estacionario con media 0 y función de autocorrelación $R_X(\tau) = 4e^{-|\tau|}$. Sea A una variable aleatoria con distribución $N(0, \sigma^2 = 9)$, independiente del proceso $X(t)$. Se define el proceso estocástico*

$$Z(t) = A \cdot \cos(\pi t) + X(t)$$

a) *Obtén las distribuciones de segundo orden del proceso $X(t)$.*

b) *Obtén la distribución del vector aleatorio $[X(1) - 2X(2) + 1, 1 - X(1)]$.*
¿Cuánto vale $Cov\,[X(1) - 2X(2) + 1, 1 - X(1)]$?

c) *Estudia si $Z(t)$ es estacionario en sentido amplio.*

d) *Calcula $P(Z(1) > -3)$*

Solución.

a) Por ser el proceso $X(t)$ gaussiano la distribución conjunta de las variables aleatorias $\begin{pmatrix} X(t_1) \\ X(t_2) \end{pmatrix}$ es normal bidimensional.

Como $E[X(t)]$ es es cero para cualquier t, $\mu_X(t) = 0$, el vector de medias de $\begin{pmatrix} X(t_1) \\ X(t_2) \end{pmatrix}$ es $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

$V[X(t)] = R_X(0) - \mu_X = 4e^0 - 0 = 4$ y $Cov[X(t_1),\ X(t_2)]] = R_X(|t_1 - t_2|) - \mu_X^2 = 4e^{-|t_1 - t_2|}$.

Así que la matriz de varianzas/covarianzas es $\begin{pmatrix} 4 & 4e^{-|t_1-t_2|} \\ 4e^{-|t_1-t_2|} & 4 \end{pmatrix}$ y las distribuciones conjuntas de segundo orden son

$$\begin{pmatrix} X(t_1) \\ X(t_2) \end{pmatrix} \sim \mathcal{N}_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, 4\begin{pmatrix} 1 & e^{-|t_1-t_2|} \\ e^{-|t_1-t_2|} & 1 \end{pmatrix}\right)$$

b) El vector aleatorio dado se puede expresar como

$$\begin{pmatrix} C_1 \\ C_2 \end{pmatrix} = \begin{pmatrix} X(1) - 2X(2) + 1 \\ 1 - X(1) \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} X(1) \\ X(2) \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Por ser las variables aleatorias C_1 y C_2 combinaciones lineales de variables aleatorias normales su distribución es normal y queda caracterizada por su vector de medias y su matriz de varianzas/covarianzas. El vector de medias es

$$\begin{pmatrix} E[C_1] \\ E[C_2] \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

La matriz de varianzas/covarianzas de C_1 y C_2 es Σ_C que se obtiene como

$$\Sigma_C = \begin{pmatrix} 1 & -2 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 4 & 4e^{-|2-1|} \\ 4e^{-|2-1|} & 4 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -2 & 0 \end{pmatrix} = 4\begin{pmatrix} 1 - 2e^{-1} & e^{-1} - 2 \\ -1 & -e^{-1} \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -2 & 0 \end{pmatrix} =$$

$$= 4\begin{pmatrix} 5 - 4e^{-1} & 2e^{-1} - 1 \\ -1 + 2e^{-1} & 1 \end{pmatrix}$$

Por tanto,

$$\begin{pmatrix} C_1 \\ C_2 \end{pmatrix} \sim \mathcal{N}_2\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 20 - 16e^{-1} & 8e^{-1} - 4 \\ -4 + 8e^{-1} & 4 \end{pmatrix}\right)$$

La $Cov[X(1) - 2X(2) + 1, 1 - X(1)] = -4 + 8e^{-1} = -1.05696$

c) Aplicando propiedades del operador esperanza la media del proceso $Z(t)$ es

$$\mu_Z(t) = E[Z(t)] = E[A \cdot \cos(\pi t) + X(t)] = E[A] \cdot \cos(\pi t) + E[X(t)] = 0 \cdot \cos(\pi t) + 0 = 0$$

Por ser las variables aleatorias A y $X(t)$ independientes la función de autocorrelación del proceso es

$$R_Z(t,\ t+\tau) = E[Z(t)Z(t+\tau)] = E[(A \cdot \cos(\pi t) + X(t)) \cdot (A \cdot \cos(\pi(t+\tau)) + X(t+\tau))] =$$

$$= E[A^2 \cos(\pi t)\cos(\pi(t+\tau)) + X(t)A\cos(\pi(t+\tau)) + A\cos(\pi t)X(t+\tau) + X(t)X(t+\tau)] =$$

$$= \cos(\pi t)\cos(\pi(t+\tau))E[A^2] + \cos(\pi(t+\tau))E[X(t)]E[A] + \cos(\pi t)E[A]E[X(t+\tau)] + E[X(t)X(t+\tau)] =$$

$$= \cos(\pi t)\cos(\pi(t+\tau))V[A] + \cos(\pi(t+\tau)) \cdot 0 + \cos(\pi t) \cdot 0 + R_X(\tau) = 9\cos(\pi t)\cos(\pi(t+\tau)) + 4e^{-|\tau|}$$

Por depender de t la función de autocorrelación de $Z(t)$ se concluye que $Z(t)$ no es estacionario en sentido amplio.

d) $Z(1) = A \cdot \cos(\pi) + X(1) = X(1) - A$, es una combinación lineal de variables aleatorias normales, por lo que la distribución $Z(1)$ es normal, de media $E[Z(1)] = 0$ y varianza $V[Z(1)] = R_Z(1,\ 1) = 9\cos(\pi)\cos(\pi) + 4 = 13$.

Tipificando la v.a. $Z(1)$ como Y, utilizando la simetría de la variable Y y consultando la tablas obtenemos

$$P(Z(1) > -3) = P\left(Y = \frac{Z(1) - 0}{\sqrt{13}} > \frac{-3 - 0}{\sqrt{13}}\right) = P(Y > -0.83) = P((Y < 0.83) = 0.79673$$