Departamento de Matemática Aplicada a la I.T. de Telecomunicación

EXAMEN RESUELTO DE ESTADÍSTICA Y PROCESOS ESTOCÁSTICOS

CONVOCATORIA: ENERO 2011/2012 **FECHA:** 11 de Enero de 2012
Duración del examen: 3 horas
Fecha publicación notas: 24 de enero de 2012
Fecha revisión examen: 30 de enero de 2012

APELLIDOS:
NOMBRE:
DNI: **TITULACIÓN:**

---

1. *(1 punto)* En una ciudad se utilizan tres medios de transporte. Sea M el suceso "Una persona utiliza el metro", sea A el suceso "Una persona utiliza el autobús" y se sea C el suceso "Una persona utiliza el coche privado". Las probabilidades de que una persona elegida al azar utilice los distintos medios de transporte son:

$$P(M) = 0{,}3; \quad P(A) = 0{,}2; \quad P(C) = 0{,}15; \quad P(M \cap A) = 0,1$$

$$P(M \cap C) = 0,05; \quad P(A \cap C) = 0,06; \quad P(M \cap A \cap C) = 0,01$$

Calcula la probabilidad de que una persona utilice metro o coche, pero no autobús.

Solución.-

$P((M \cup C) \cap \overline{A}) = P((M \cap \overline{A}) \cup (C \cap \overline{A})) = P(M \cap \overline{A}) + P(C \cap \overline{A}) - P(M \cap \overline{A} \cap C)$

Pero

$P(M \cap \overline{A}) = P(M) - P(M \cap A) = 0,3 - 0,1 = 0,2$

$P(C \cap \overline{A}) = P(C) - P(C \cap A) = 0,15 - 0,06 = 0,09$

$P(M \cap \overline{A} \cap C) = P((M \cap C) \cap \overline{A}) = P(M \cap C) - P(M \cap C \cap A) = 0,05 - 0,01 = 0,04$

Por tanto,

$$P((M \cap C) \cap \overline{A}) = 0,2 + 0,09 - 0,04 = 0,25$$

2. *(1.2 puntos)* Tres compañías de seguros copan el mercado de una determinada ciudad. El 30 % de las pólizas suscritas corresponden a la compañía A, el 25 % a la B y el 45 % restante a la C. El porcentaje de pólizas de seguros de vida en cada una de ellas es del 15 %, 20 % y 25 %, respectivamente.

   a) Si una persona ha suscrito un seguro de vida, ¿cuál es la probabilidad de que su póliza sea de la compañía A?

   b) De 10 personas que han contratado un seguro de vida, ¿cuál es la probabilidad de que la mitad lo hayan hecho con la compañía A?

   c) Entre 1000 asegurados, ¿cuál es el número medio de personas que han contratado un seguro de vida?

Solución.- Consideremos los sucesos siguientes:

$A \equiv$ Un asegurado tiene un contrato en la compañía A.

$B \equiv$ Un asegurado tiene un contrato en la compañía B.

$C \equiv$ Un asegurado tiene un contrato en la compañía C.

$V \equiv$ Un asegurado tiene contratado un seguro de vida.

Sabemos que $P(A) = 0,3, \quad P(B) = 0,25, \quad P(C) = 0,45, \quad P(V/A) = 0,15,$
$P(V/B) = 0,2, \quad P(V/C) = 0,25$

a)

$$\begin{aligned} P(A/V) =& \frac{P(A \cap V)}{P(V)} = \frac{P(V/A)\ P(A)}{P(V/A)\ P(A) + P(V/B)\ P(B) + P(V/C)\ P(C)} \\ =& \frac{0,15 \cdot 0,3}{0,15 \cdot 0,3 + 0,2 \cdot 0,25 + 0,25 \cdot 0,45} \approx 0,217 \end{aligned}$$

b) Sea N la variable aleatoria "Número de personas, entre los 10 que tienen seguro de vida, que han contratado su póliza en A".
N sigue una distribución binomial de parámetros $n = 10$ y $p = P(A/V) \approx 0,217$, por tanto,

$$P(N = i) = \binom{10}{i}\ 0,217^i\ (1 - 0,217)^{10-i}, \quad i = 0, 1, \cdots 10$$

$$P(N = 5) = \binom{10}{5}\ 0,217^5\ (1 - 0,217)^5 \approx 0{,}036$$

c) Sea X la variable aleatoria "Número de personas, entre los 1000 asegurados, que tienen contratado un seguro de vida".
X sigue una distribución binomial de parámetros $n = 1000$ y $p = P(V)$.

$$P(V) = P(V/A)\ P(A) + P(V/B)\ P(B) + P(V/C)\ P(C) \approx 0,2075$$

$$E(Y) = 1000 \cdot 0,2075 = 207,5$$

3. *(1 punto)* El tiempo en años que funciona un aparato de radio está distribuido exponencialmente con una media de 8 años. Si una persona compra una radio de segunda mano que ha funcionado durante 6 años, calcula la probabilidad de que funcione, al menos, 10 años más.

Solución.- Sea X la variable aleatoria Tiempo de vida del aparato de radio.
$X \sim Exp(1/8)$

$$P(X > 16/X > 6) = \frac{P((X > 16) \cap (X > 6))}{P(X > 6)} = \frac{P(X > 16)}{P(X > 6)} = \frac{\int_{16}^{+\infty} (1/8)\ e^{-x/8} dx}{\int_{6}^{+\infty} (1/8)\ e^{-x/8} dx}$$

$$= \frac{-e^{-x/8}\Big|_{16}^{+\infty}}{-e^{-x/8}\Big|_{6}^{+\infty}} = \frac{e^{-2}}{e^{-3/4}} = e^{-5/4}$$

4. *(1 punto)* Sea X una variable aleatoria con distribución normal de media 1 y varianza σ^2. Calcula qué valor debe tener σ para que el primer cuartil de la variable aleatoria $Y = |X-1|$ sea igual a 2.

Solución.-

Debe ser $P(Y \leq 2) = 1/4$.

$$P(Y \leq 2) = P(|X-1| \leq 2) = P(-2 \leq X-1 \leq 2) = P(-1 \leq X \leq 3)$$

Pero, como X sigue una distribución normal de media 1 y desviación típica σ, se verifica que $Z = \dfrac{X-1}{\sigma} \sim N(0,1)$.

Entonces,

$$P(Y \leq 2) = P\left(\frac{-2}{\sigma} \leq Z \leq \frac{2}{\sigma}\right) = F_Z\left(\frac{2}{\sigma}\right) - F_Z\left(\frac{-2}{\sigma}\right) = 2F_Z\left(\frac{2}{\sigma}\right) - 1 = \frac{1}{4}$$

De donde, $F_Z\left(\dfrac{2}{\sigma}\right) = \dfrac{5}{8} = 0,625$

Y, a partir de las tablas de la función de distribución de Z, obtenemos que $\dfrac{2}{\sigma} \approx 0,32$.

Luego $\sigma \approx 6,25$

5. *(1,2 puntos)* Sea $f(x,\ y) = \begin{cases} kxy \text{ si } 0 < x < 1; \quad 0 < y < 2 \\ 0 \text{ en el resto} \end{cases}$ la función de densidad conjunta de la v.a. bidimensional $(X,\ Y)$.

   a) Calcula k para que efectivamente $f(x,\ y)$ sea función de densidad.

   b) ¿Cuánto vale la función de distribución conjunta en el punto $(\frac{1}{2},\ 1)$?

   c) Halla las funciones de densidad marginales. ¿Son independientes X e Y?

Solución.-

a) $\displaystyle\int_0^1 dx \int_0^2 kxy\ dy = \int_0^1 kx\frac{y^2}{2}\Big|_0^2\ dx = \int_0^1 2kx\ dx = k\ x^2\Big|_0^1 = k$

Puesto que $1 = \displaystyle\int_0^1 dx \int_0^2 kxy\ dy = k \Rightarrow k = 1$

b) Si F_{XY} es la función de distribución conjunta de $(X,\ Y)$ entonces,

$$F_{XY}(\frac{1}{2},\ 1) = \int_0^{\frac{1}{2}} dx \int_0^1 xy\ dy = \int_0^{\frac{1}{2}} (x \cdot \frac{y^2}{2})\Big|_0^1\ dx = \int_0^{\frac{1}{2}} (x \cdot \frac{1}{2})\ dx = \frac{x^2}{4}\Big|_0^{\frac{1}{2}} =$$

$$= \frac{\frac{1}{4}}{4} = \frac{1}{16}$$

c) Sean f_X y f_Y las funciones de densidad marginal de X y de Y, respectivamente, entonces $f_X(x) = 0$ si $x \notin (0,\ 1)$ y $f_Y(y) = 0$ si $y \notin (0,\ 2)$.

Si $0 < x < 1$ se tiene $f_X(x) = \displaystyle\int_{-\infty}^{\infty} f(x,y)\ dy = \int_0^2 xy\ dy = x \cdot \frac{y^2}{2}\Big|_0^2 = 2x.$

Si $0 < y < 2$ se tiene $f_Y(y) = \displaystyle\int_{-\infty}^{\infty} f(x,y)\ dx = \int_0^1 xy\ dx = \frac{x^2}{2} \cdot y\Big|_0^1 = \frac{y}{2}.$

Por tanto, $f_X(x) = \begin{cases} 2x \text{ si } 0 < x < 1 \\ 0 \text{ en el resto} \end{cases}$; $\quad f_Y(y) = \begin{cases} \frac{y}{2} \text{ si } 0 < y < 2 \\ 0 \text{ en el resto} \end{cases}$

$f_X(x) \cdot f_Y(y) = \begin{cases} 2x \cdot \dfrac{y}{2} \text{ si } 0 < x < 1;\ 0 < y < 2 \\ 0 \text{ en el resto} \end{cases} \Rightarrow f_X(x) \cdot f_Y(y) = f(x,\ y)$ en todo $\mathbb{R}^2$, por lo que las v.a X e Y son independientes.

6. *(1 punto)* El tiempo de vida de un bolígrafo es una variable aleatoria T de media 1 semana y desviación típica 1 semana. Utilizando el teorema central del límite calcula de forma aproximada la probabilidad de que un estudiante tenga suficiente con 25 bolígrafos para un semestre de 15 semanas.

   Solución.- Sea $ST = \sum_{i=1}^{25} T_i$, siendo T_i el tiempo de vida de cada uno de los 25 bolígrafos. Por el Teorema Central del Límite $\sum_{i=1}^{25} T_i = ST$ es aproximadamente normal de media 25 $E[T] = 25$ y desviación típica $\sqrt{25 \cdot 1} = 5$. Entonces, tipificando, utilizando la simetría y las tablas

   $$P(ST \geq 15) = P\left(\frac{ST-25}{5} \geq \frac{15-25}{5}\right) = P\left(\frac{ST-25}{5} \geq -2\right) = P\left(\frac{ST-25}{5} \leq 2\right) = 0,97725$$

7. *(1,4 puntos)* Supongamos que una señal de intensidad μ se emite desde una determinada estrella y el valor recibido en un observatorio es una variable aleatoria, X, normal con media μ y desviación típica 4. Se sospecha que la intensidad de la señal es 10. Contrasta, con un nivel de significación $\alpha = 0,06$ si esta hipótesis puede ser aceptada sabiendo que la señal se ha recibido 20 veces y la media de esos 20 valores es 11,6. ¿Se aceptaría la hipótesis con un nivel de significación $\alpha = 0,1$.

   Solución.- Debemos contrastar la hipótesis nula $H_0 : \mu = 10$, frente a la hipótesis alternativa $H_1 : \mu \neq 10$.

   Como X tiene varianza conocida, utilizamos para el contraste el estadístico $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{20}}$.

   Si H_0 es cierta, $Y = \dfrac{\overline{X}-10}{4/\sqrt{20}} \sim N(0,1)$

   El valor de Y en la muestra es $\dfrac{11,6-10}{4/\sqrt{20}} = 1,79$.

   Así pues, el p-valor del contraste es $2\ P(Z > 1,79) = 2(1 - F_Z(1,79)) = 2(1 - 0,96) = 0,08$

   Aceptaríamos que la intensidad de la señal es 10 para todo nivel de significación menor que 0,08.

   Por lo tanto, se acepta para $\alpha = 0,06$ y se rechaza para $\alpha = 0,1$

8. *(1 punto)* Sea $X(t) = \cos(t) + N(t)$, donde $N(t)$ es un proceso estocástico de media μ y función de autocorrelación $R_N(\tau)$. Obtén la media y la función de autocorrelación de $X(t)$ en términos de μ y $R_N(\tau)$. ¿Es $X(t)$ estacionario en sentido amplio?

   Solución.-

$E[X(t)] = E[\cos(t) + N(t)] = \cos(t) + E[N(t)] = \cos(t) + \mu$

$$\begin{aligned} R_X(t, t+\tau) =& E[X(t)X(t+\tau)] = E[(\cos(t) + N(t))(\cos(t) + N(t+\tau))] \\ =& E[\cos(t)\cos(t+\tau) + \cos(t)N(t+\tau) + \cos(t+\tau)N(t) + N(t)N(t+\tau)] \\ =& \cos(t)\cos(t+\tau) + \cos(t)E[N(t+\tau)] + \cos(t+\tau)E[N(t)] + E[N(t)N(t+\tau)] \\ =& \cos(t)\cos(t+\tau) + \mu\cos(t) + \mu\cos(t+\tau) + R_N(\tau) \\ =& \cos(t)\cos(t+\tau) + \mu(\cos(t) + \cos(t+\tau)) + R_N(\tau) \end{aligned}$$

Para que el proceso fuese estacionario en sentido amplio debería ser constante la media del proceso y la autocorrelación dependiente sólo de la variable τ. No se verifica ninguna de las dos condiciones y por consiguiente el proceso no es ESA.

9. *(1,2 puntos)* Sea $\{X(t)\}_{t>0}$ un proceso Gaussiano estacionario de media cero y función de autocorrelación $R_X(\tau) = 4e^{-3|\tau|}$.

   a) Halla $P(X(1) > 1)$.
   b) Halla la distribución de la v.a. $(X(1),\ X(2),\ X(3) - X(1))$.
   c) Calcula $P(X(3) - X(1) > 1\ /\ X(2) > 1)$.

Solución.-

a) Por ser $X(t)$ un proceso Gaussiano, las distribuciones de primer orden son normales. Por tanto, $X(1) \sim N(\mu = 0, \sigma^2 = R_X(0) = 4)$

$$P(X(1) > 1) = P(Z > 1/2) = 1 - F_Z(1/2) \approx 1 - 0,69 = 0,31$$

siendo F_Z la función de distribución de la normal unitaria.

b) Por ser $X(t)$ un proceso Gaussiano, $(X(1),\ X(2),\ X(3))$ sigue una distribución normal tridimensional cuyo vector de medias es $\mathbf{m} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ y cuya matriz de covarianzas es

$$M = \begin{pmatrix} R_X(0) & R_X(1) & R_X(2) \\ R_X(1) & R_X(0) & R_X(1) \\ R_X(2) & R_X(1) & R_X(0) \end{pmatrix} = \begin{pmatrix} 4 & 4e^{-3} & 4e^{-6} \\ 4e^{-3} & 4 & 4e^{-3} \\ 4e^{-6} & 4e^{-3} & 4 \end{pmatrix}$$

Como

$$\begin{pmatrix} X(1) \\ X(2) \\ X(3) - X(1) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} X(1) \\ X(2) \\ X(3) \end{pmatrix}$$

Se verifica que $(X(1),\ X(2),\ X(3) - X(1))$ sigue una distribución normal tridimensional cuyo vector de medias es $\mathbf{m'} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$
y cuya matriz de covarianzas es

$$M' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 & 4e^{-3} & 4e^{-6} \\ 4e^{-3} & 4 & 4e^{-3} \\ 4e^{-6} & 4e^{-3} & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 4e^{-3} & -4 + 4e^{-6} \\ 4e^{-3} & 4 & 0 \\ -4 + 4e^{-6} & 0 & 8(1 - e^{-6}) \end{pmatrix}$$

c) Como podemos ver en la matriz M', $\mathrm{Cov}(X(3) - X(1),\ X(2)) = 0$.
Por tener $(X(3) - X(1),\ X(2))$ una distribución normal bidimensional, se puede afirmar que ambas variables son independientes.
Así pues,

$$P(X(3) - X(1) > 1, X(2) > 1) = P(X(3) - X(1) > 1)$$

Pero

$$X(3) - X(1) \sim N(\mu = 0, \sigma^2 = 8(1 - e^{-6}))$$

Entonces,

$$P(X(3) - X(1) > 1) = P\left(Z > \frac{1}{\sqrt{8(1 - e^{-6})}}\right) \approx P(Z > 0,35) = 1 - F_Z(0,35) = 0{,}36317$$

donde F_Z es la función de distribución de la variable aleatoria normal unitaria.