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

ASIGNATURA: ESTADÍSTICA Y PROCESOS ESTOCÁSTICOS

Examen Final (Otoño 2020/21) **FECHA:** 5 de Febrero de 2021 **Duración:** 3 horas

SOLUCIONES

Ejercicio 1 *(2 puntos) En un lote de diez bombillas se sabe que tres son defectuosas. Se prueba cada bombilla del lote de una en una.*

a) ¿Cuál es la probabilidad de que la última de las tres bombillas defectuosas se detecte en la quinta prueba?

b) ¿Cuál es la probabilidad de que no sea necesario probar más de seis bombillas para localizar las tres defectuosas?

c) Si dos bombillas defectuosas se han localizado antes de la quinta prueba, ¿cuál es la probabilidad de que la última bombilla defectuosa se detecte en la quinta o sexta prueba?

Solución.

a) Sea $S_5 \equiv$ la última de las tres bombillas defectuosas se detecte en la quinta prueba

Para calcular $P(S_5)$ utilizaremos la Regla de Laplace. Los casos posibles son todas las posiciones, de las diez posibles, en los que la prueba sea defectuosa, puesto que son tres, $\binom{10}{3}$.

Los casos favorables para el suceso S_5 es que en las cuatro primeras pruebas se detecten 2 de las defectuosas, $\binom{4}{2}$, y en la quinta se prueba una bombilla defectuosa. Por tanto,

$$P(S_5) = \frac{\textit{Casos favorables al suceso}}{\textit{Casos posibles}} = \frac{\binom{4}{2} \cdot 1}{\binom{10}{3}} = \frac{6}{120} = \frac{1}{20} = 0.05$$

b) Si S es el suceso "no es necesario probar más de seis bombillas para localizar las tres defectuosas", los casos favorables al suceso S son los que las tres bombillas defectuosas se localizan, a lo más, en las seis primeras pruebas, $\binom{6}{3}$. Entonces,

$$P(S) = \frac{\textit{Casos favorables al suceso } S}{\textit{Casos posibles}} = \frac{\binom{6}{3}}{\binom{10}{3}} = \frac{6 \cdot 5 \cdot 4}{10 \cdot 9 \cdot 8} = \frac{1}{6} \simeq 0.1667$$

c) Sean los sucesos $S_5 \equiv$ la última de las tres bombillas defectuosas se detecte en la quinta prueba y $S_6 \equiv$ la última de las tres bombillas defectuosas se detecte en la sexta prueba

Sea A el suceso "dos bombillas defectuosas se localizan antes de la quinta prueba" y B el suceso "la última bombilla defectuosa se detecte en la quinta o sexta prueba", por lo que $B = S_5 \cup S_6$. Entonces debemos hallar $P(B/A) = \dfrac{P(A \cap B)}{P(A)} = \dfrac{P(A \cap (S_5 \cup S_6))}{P(A)} = \dfrac{P(A \cap S_5) + P(A \cap S_6)}{P(A)}$, la última igualdad se obtiene porque $S_5 \cap S_6 = \emptyset$. Usando la Regla de Laplace y teniendo en cuenta que los casos favorables para el suceso A son las dos posiciones que deben encontrarse las *bombillas defectuosas en las cuatro primeras pruebas y la tercera defectuosa en una de las seis restantes posiciones, se tiene*

$$P(B/A) = \frac{P(A \cap S_5) + P(A \cap S_6)}{P(A)} = \frac{\dfrac{\binom{4}{2} \cdot 1 + \binom{4}{2} \cdot 1}{\binom{10}{3}}}{\dfrac{\binom{4}{2} \cdot \binom{6}{1}}{\binom{10}{3}}} = \frac{2 \cdot \binom{4}{2}}{\binom{4}{2} \cdot \binom{6}{1}} = \frac{2}{6} = \frac{1}{3} \simeq 0.3333$$

Ejercicio 2 *(2,5 puntos) Para la comercialización de una determinada vacuna, entre otras variables, se debe controlar el volumen de cada dosis inyectable. Esta variable sigue una distribución normal de media 3 ml. y de desviación típica 0,05 ml.*

a) Calcular la probabilidad de que una dosis inyectable tenga un volumen superior a 3,025 ml.

b) Una dosis inyectable se considera defectuosa cuando su volumen difiere de la media en más de 0,075 ml. Calcular la proporción de dosis defectuosas que se fabrican.

c) Las dosis inyectables se envasan en cajas de 10 unidades. Si una caja contiene 2 o más dosis defectuosas se elimina del mercado. Determinar el porcentaje de cajas que se retiran del mercado.

Solución. *Si X es la variable aleatoria "Volumen de cada dosis", $X \sim \mathcal{N}(\mu = 3,\ \sigma = 0,05)$.*

a) Tipificando, pasando al complementario y siendo Φ la función de distribución de la va.a estándar Z, tenemos

$$P(X > 3,025) = P\left(Z = \frac{X - \mu}{\sigma} > \frac{3,025 - \mu}{\sigma}\right) = 1 - \Phi(0,5) = 1 - 0,69146 = 0,30854$$

b) Tipificando $P(|X - 3| > 0,075) = P\left(|Z = \frac{X-3}{0,05}| > \frac{0,075}{0,05} = 1,5\right) = P(Z > 1,5) + P(Z < -1,5)$. Por simetría, $P(|X-3| > 0,075) = 2 \cdot P(Z > 1,5) = 2 \cdot (1 - \Phi(1,5)) = 2 \cdot (1 - 0,93319) = 0,1336$.

c) Sea la variable aleatoria $Y \equiv$ "número de dosis defectuosas en una caja de 10 unidades", entonces $Y \sim Bin(10,\ 0.1336)$. Si $P(Y \geq 2) = 1 - P(Y < 2) = 1 - [P(Y = 0) + P(Y = 1)] =$

$$= 1 - \left[\binom{10}{0} \cdot 0.1336^0 \cdot (1 - 0.1336)^{10} + \binom{10}{1} \cdot 0.1336^1 \cdot (1 - 0.1336)^9\right] =$$

$$= 1 - \left[(1 - 0.1336)^{10} + 10 \cdot 0.1336 \cdot (1 - 0.1336)^9\right] \backsimeq 1 - 0.2383 - 0.3675 = 0.3942$$

Ejercicio 3 *(2 puntos) Sea un dispositivo con dos componentes A y B. Consideremos las variables aleatorias $X \equiv$ número de fallos en la componente A, $\quad Y \equiv$ número de fallos en la componente B, con distribución de probabilidad conjunta:*

$X \backslash Y$	$Y = 0$	$Y = 1$	$Y = 2$
X=0	1/9	2/9	1/9
X=1	2/9	2/9	0
X=2	1/9	0	0

Sea la variable aleatoria $C = 3X + Y$ y que representa el coste de reparación de estos fallos.

a) *Calcula las probabilidades marginales de X y de Y.*

b) *Calcula $Cov[X, Y]$. ¿Son las variables aleatorias independientes?*

c) *Halla $E[C]$ y $Var[C]$*

Solución.

a) *Sumando por columnas y filas*

$X = i$	0	1	2
P(X=i)	4/9	4/9	1/9

$Y = j$	0	1	2
P(Y=j)	4/9	4/9	1/9

b) *Como $Cov(X, Y) = E[X \cdot Y] - E[X] \cdot E[Y]$ hallamos cada media por separado.*

$$E[X] = \sum_i i \cdot P(X = i) = 0 \cdot \frac{4}{9} + 1 \cdot \frac{4}{9} + 2 \cdot \frac{1}{9} = \frac{6}{9} = \frac{2}{3}$$

$$E[Y] = \sum_j j \cdot P(Y = j) = 0 \cdot \frac{4}{9} + 1 \cdot \frac{4}{9} + 2 \cdot \frac{1}{9} = \frac{6}{9} = \frac{2}{3}$$

$$E[X \cdot Y] = \sum_{i,j} i \cdot j \cdot P(X = i, Y = j) = 0 \cdot 0 \cdot \frac{1}{9} + 0 \cdot 1 \cdot \frac{2}{9} + 0 \cdot 2 \cdot \frac{1}{9} + 1 \cdot 0 \cdot \frac{2}{9} + 1 \cdot 1 \cdot \frac{2}{9} + 1 \cdot 2 \cdot 0 + 2 \cdot 0 \cdot \frac{1}{9} + 2 \cdot 1 \cdot 0 + 2 \cdot 2 \cdot 0 = \frac{2}{9}$$

Entonces, $Cov(X, Y) = E[X \cdot Y] - E[X] \cdot E[Y] = \frac{2}{9} - \frac{2}{3} \cdot \frac{2}{3} = -\frac{2}{9}$
Luego X e Y no son incorreladas y, por tanto, no son independientes.

c) $E[C] = E[3X + Y] = 3E[X] + E[Y] = 3 \cdot \frac{2}{3} + \frac{2}{3} = \frac{8}{3}$
$Var[C] = 3^2 \cdot Var[X] + Var[Y] + 2 \cdot 3 \cdot 1 \cdot Cov[x, Y]$

$$Var[X] = \left(0^2 \cdot \frac{4}{9} + 1^2 \cdot \frac{4}{9} + 2^2 \cdot \frac{1}{9}\right) - E[X]^2 = \frac{8}{9} - \frac{4}{9} = \frac{4}{9}$$

Por tener las v.a. X e Y la misma distribución $Var[Y] = \frac{4}{9}$ Así,

$$Var[C] = 9 \cdot Var[X] + Var[Y] + 6 \cdot Cov[x, Y] = 9 \cdot \frac{4}{9} + \frac{4}{9} + 6 \cdot \left(-\frac{2}{9}\right) = \frac{28}{9}$$

Ejercicio 4 *(1,5 puntos) Se desea estudiar la variable*

$$X \equiv \textit{Longitud en cm de las varillas de un lote}$$

a) *Suponemos que X sigue una distribución normal e media μ y desviación típica $\sigma_X = 1,5$ cm ¿Cuántas varillas deberemos medir para que el intervalo de confianza para la media, μ, con un nivel de confianza del 95 %, tenga una amplitud inferior a 0,5 cm?*

b) *Se desea contrastar la hipótesis nula $H_0 \equiv$ La desviación típica de X es $\sigma_0 = 1,5$ cm, frente a la alternativa $H_1 \equiv \sigma_0 \neq 1,5$ cm. Se miden 30 varillas y la cuasideviación típica de las longitudes resultantes es $s_1 = 1,8835$ cm. Calcula el p-valor del contraste (con la aproximación que te permita la tabla de cuantiles de la chi-cuadrado).*

c) Medimos 30 varillas y solo una de ellas es defectuosa. Se quiere contrastar la hipótesis nula $H_0 \equiv$ La proporción de varillas defectuosas es $p = 0,02$, frente a la alternativa $H_1 \equiv p \neq 0,02$. Calcula el p-valor del contraste. ¿Aceptarías la hipótesis nula con un nivel de significación del 0,05?

Solución.

a) El cuantil 0.95 *de una normal estándar,* $\mathcal{N}(0,1)$*, es aproximadamente* $z_{0,975} = 1.96$*. Entonces, para un nivel del* $0,95$ *y cuando el número de varillas de la muestra es* n*, se deberá cumplir*

$$2\frac{\sigma}{\sqrt{n}}\, z_{0,975} = 2\frac{1,5}{\sqrt{n}}\, 1.96 < 0,5 \quad \text{o equivalentemente} \quad n > \left(2\frac{1,5}{0,5}\, 1.96\right)^2 = 11,76^2 = 138,2976$$

Tendríamos pues que medir al menos 139 varillas.

b) El estadístico para realizar el contraste es $(n-1)\dfrac{S_1^2}{\sigma_0^2}$ *y en nuestro caso toma el valor*

$$(n-1)\frac{s_1^2}{\sigma_0^2} = 29\,\frac{1,8835^2}{1,5^2} \approx 45,72$$

Supuesto que la hipótesis nula es cierta, el estadístico de contraste $(n-1)S_1^2/\sigma_0^2$ *tiene aproximadamente distribución* χ^2_{29}*. Entonces,* $P(\chi^2_{29} \leq 45,72) \approx 0,975$ *y* $P(\chi^2_{29} \geq 45,72) \approx 0,025$

$$\text{p-valor} = 2P(\chi^2_{29} \geq 45,72) \simeq 2 \times 0,025 = 0,05$$

c) El estadístico de contraste es $N \equiv$ *número de varillas defectuosas entre las 30 medidas. Si llamamos éxito a que una varilla sea defectuosa* N *se distribuye como una variable aleatoria binomial,* $N \sim Bin(n, p_1)$*, con parámetros* $n = 30$ *y* $p_1 =$ *probabilidad de éxito*

Si es cierta la hipótesis nula $N \sim Bin(300,\ 0,02)$*, entonces*

$$P(N \leq 1) = P(N=0) + P(N=1) = \left[\binom{30}{0}\cdot 0.2^0 \cdot 0.8^{30} + \binom{30}{1}\cdot 0.2^1 \cdot 0.8^{29}\right] \approxeq 0,010522$$

Por tanto, $P(N \geq 1) = 1 - P(N < 1) = 1 - P(N=0) \simeq 1 - 0,001238 = 0,998762$

$$\text{p-valor} = 2\cdot \min\{P(N \leq 1),\ P(N \geq 1)\} = 2\cdot P(N \leq 1) \simeq 2\cdot 0,010522 = 0,021044$$

Puesto que el nivel de significación es mayor que el p-valor, $0,05 > 0,02$*, se rechaza la hipótesis nula con un nivel de significación del 0,05.*

Ejercicio 5 *(1 punto) El número de mensajes que llegan a un servidor en el intervalo* $[0,t)$ *es un proceso estocástico* $X(t) = 2 + N(t)$*, donde* $N(t)$ *es un proceso de Poisson de tasa 3 mensajes por segundo.*

a) Calcula $P\big(X(1) = 3, X(3) = 6\big)$*.*

b) Calcula la probabilidad de que lleguen 3 mensajes entre el segundo 2 y el segundo 4 si han llegado 5 en los primeros 4 segundos.

Solución.

a) $P\big(X(1)=3,\ X(3)=6\big)=P\big(2+N(1)=3,\ 2+N(3)=6\big)=P\big(N(1)=1,\ N(3)=4\big)$.
Por ser $N(t)$ un proceso de Poisson, para cada valor de t fijo, $N(t)\sim\mathcal{P}(3t)$.

Teniendo en cuenta que las llegadas, salvo la constante, en el intervalo $[0,1)$ es $N(1)$ y que las llegadas, salvo constante, en el intervalo $[1,3)$ es $N(3)-N(1)$, son independientes por ser intervalos disjuntos de tiempo, que $N(1)$ tiene distribución de Poisson de media $3\times 1=3$ y que $N(3)-N(1)$ tiene distribución de Poisson de media $3\times(3-1)=6$ se obtiene que

$$P\big(X(1)=3,\ X(3)=6\big)=P\big(N(1)=1\,,N(3)-N(1)=3\big)=$$

$$=P\big(N(1)=1\big)\ P\big(N(3)-N(1)=3\big)=e^{-3}\ \frac{3^1}{1!}\ e^{-6}\ \frac{6^3}{3!}=e^{-9}\cdot 3\cdot 6^2 =\simeq 0,0133$$

b) $P(X(4)-X(2)=3/X(4)=5)=\dfrac{P(X(4)-X(2)=3,\ X(4)=5)}{P(X(4)=5)}=$

$$=\frac{P(2+N(4)-2-N(2)=3,\ 2+N(4)=5)}{P(2+N(4)=5)}=\frac{P(N(4)-N(2)=3,\ N(4)=3)}{P(N(4)=3)}=$$

$$=\frac{P(N(4)-N(2)=3,\ N(2)=0)}{P(N(4)=3)}$$

Las variables aleatorias $N(4)-N(2)$ y $N(2)$, que miden el número de sucesos en intervalos de tiempo disjuntos, por lo que son independientes. Por tanto, y dado que las v.a. verifican $N(4)-N(2)\sim\mathcal{P}(6)$, $\quad N(4)\sim\mathcal{P}(12)$, $\quad N(2)\sim\mathcal{P}(6)$

$$\frac{P(N(4)-N(2)=3,\ N(2)=0)}{P(N(4)=3)}=\frac{P(N(4)-N(2)=3)\ P(N(2)=0)}{e^{-12}\ \dfrac{12^3}{3!}}=\frac{e^{-6}\ \dfrac{6^3}{3!}\cdot e^{-6}\ \dfrac{6^0}{0!}}{e^{-12}\ \dfrac{12^3}{6}}$$

$$=\frac{6^3}{12^3}=\frac{1}{2^3}=0,125$$

Así, $P(X(4)-X(2)=3/X(4)=5)=0,125$.

Ejercicio 6 *(1 punto) Sea A una variable aleatoria con distribución uniforme en el intervalo (0,1) y sea el proceso estocástico $X(t)=2\cdot A^t$, con $t\geq 0$.*

a) *Calcula $E[X(t)]$ y $R_X(t,t+\tau)$, con $t,\ t+\tau\in[0,\infty)$.*

b) *Sea el proceso estocástico $Y(t)=\dfrac{X(t+1)}{X(t)}$. Estudia si $\{Y(t)\}_{t>0}$ es estacionario en sentido amplío.*

Solución.

a) *La v.a. A tiene por función de densidad* $\quad f_A(x)=\begin{cases}1 & si\ 0\leq x\leq 1\\ 0 & en\ el\ resto\end{cases}$

$$\mu_X(t)=E[X(t)]=E[2A^t]=2\int_0^1 x^t\ dx=2\ \frac{x^{t+1}}{t+1}\bigg|_0^1=\frac{2}{t+1}$$

$$R_X(t,t+\tau)=E[X(t)\cdot X(t+\tau)]=E[2\cdot A^t\cdot 2\cdot A^{t+\tau}]=4E[A^{2t+\tau}]=4\int_0^1 x^{2t+\tau}\ dx=$$

$$=4\ \frac{x^{2t+\tau+1}}{2t+\tau+1}\bigg|_0^1=\frac{4}{2t+\tau+1}$$

b) Podemos escribir el proceso $Y(t)$ *como* $Y(t) = \dfrac{X(t+1)}{X(t)} = \dfrac{2 \cdot A^{t+1}}{2 \cdot A^{t}} = A$. *Entonces,*

$\mu_Y(t) = E[A] = \displaystyle\int_0^1 x \cdot 1 \; dx = \left.\dfrac{x^2}{2}\right|_0^1 = \dfrac{1}{2}$

$R_Y(t, t+\tau) = E[Y(t) \cdot Y(t+\tau)] = E[A \cdot A] = E[A^2] = V[A] + E[A]^2 = \dfrac{1}{12} + \dfrac{1}{4} = \dfrac{1}{3}$

El proceso $Y(t)$ *es estacionario en sentido amplío porque la media del proceso y la función de autocorrelación no dependen del instante* t.