ASIGNATURA: ESTADÍSTICA Y PROCESOS ESTOCÁSTICOS
Examen Final (Otoño 2018) **FECHA:** 9 de Enero de 2019 **Duración:** 3 horas

SOLUCIONES

Ejercicio 1 *(1,5 puntos) Una empresa comprueba si ciertos componentes electrónicos son defectuosos (suceso D) o no defectuosos. Para verificar si un componente es defectuoso se realizan una prueba que da apto (suceso A) o no apto. Se sabe que el 90 % de los componentes son no defectuosos, que un componente no defectuoso da no apto con probabilidad 0,1 y un componente defectuoso da apto con probabilidad 0,05.*

1. *Calcula la probabilidad de que un componente de apto.*
2. *Calcula la probabilidad de que un componente que ha dado apto sea defectuoso.*
3. *Calcula la probabilidad de que un componente que ha dado no apto no sea defectuoso.*
4. *Calcula la probabilidad de que de 100 componentes que han dado apto, exactamente uno sea defectuoso.*

Solución.
Sean los sucesos:

$$\begin{array}{llll} D = & \text{Componente defectuoso}; & \overline{D} = & \text{Componente no defectuoso} \\ A = & \text{Componente apto}; & \overline{A} = & \text{Componente no apto} \end{array}$$

Según el enunciado: $P(\overline{D}) = 0,9$; $P(\overline{A}/\overline{D}) = 0,1$; $P(A/D) = 0,05$.

1. *Utilizando el Teorema de la Probabilidad Total con el sistema completo de sucesos* $\{D,\ \overline{D}\}$:

$$P(A) = P(D)P(A/D)+P(\overline{D})P(A/\overline{D}) = (1-0,9)\times 0,05+0,9\times(1-0,1) = 0,1\times 0,05+0,9\times 0,9 = 0,815$$

2.
$$P(D/A) = \frac{P(A\cap D)}{P(A)} = \frac{P(D)P(A/D)}{P(A)} = \frac{0,1\times 0,05}{0,815} \approx 0.0061$$

3.
$$P(\overline{D}/\overline{A}) = \frac{P(\overline{D}\cap\overline{A})}{P(\overline{A})} = \frac{P(\overline{D})P(\overline{A}/\overline{D})}{1-P(A)} = \frac{0,9\times 0,1}{1-0,815} \approx 0.4865$$

4. *La variable aleatoria*

$$N \equiv \text{número de componentes defectuosas de entre las 100 que han dado apto}$$

tiene distribución binomial de parámetros $n = 100$ *y*

$$p = P(D/A) \approx 0.0061$$

Entonces, la probabilidad de que exactamente 1 uno sea defectuoso, es

$$P(N=1) = \binom{100}{1} p(1-p)^{99} = 100\times p\times(1-p)^{99} \approx 0,332856$$

Ejercicio 2 *(1,5 puntos) Tenemos un dado de cuatro caras equiprobables, que corresponden a los números 1, 2, 3 y 4. La variable aleatoria A representa el resultado de una tirada del dado.*

1. *Halla la media y la varianza de A.*
2. *Si lanzamos quince veces el dado, ¿qué distribución de probabilidad sigue el número de unos obtenido? Calcula la probabilidad de sacar un uno exactamente cuatro veces.*
3. *Emplea el teorema central del límite para aproximar por una normal la suma S de los resultados de 400 tiradas. Halla la probabilidad de que S valga menos de 950 y la probabilidad de que S esté entre 1000 y 1050. ¿Cuál es la normal que mejor aproxima a la variable $\frac{1}{400}S$?*

Solución.

1. *Como tenemos que* $P(A=1)=P(A=2)=P(A=3)=P(A=4)=\frac{1}{4}$ *resulta*

$$E[A]=\frac{1+2+3+4}{4}=\frac{5}{2}$$
$$E[A^2]=\frac{1+4+9+16}{4}=\frac{15}{2}$$
$$\mathrm{Var}(A)=E[A^2]-E[A]^2=\frac{5}{4}$$

2. *La variable aleatoria* $X\equiv$ *número de unos conseguidos en quince lanzamientos sigue una distribución* $X\sim\mathrm{Bin}(15,\frac{1}{4})$.

$$P(X=4)=\binom{15}{4}\left(\frac{1}{4}\right)^4\left(1-\frac{1}{4}\right)^{15-4}=0.2252$$

3. *Denotamos la suma* $S=X_1+X_2+\ldots+X_{400}$*, siendo la variable aleatoria* X_i *el resultado de la i–ésima tirada. Como las variables aleatorias* X_i *son independientes e idénticamente distribuidas el teorema central del límite concluye que* S *se distribuye aproximadamente como*

$$S\overset{\text{aprox.}}{\sim}N\left(\mu=400\cdot\frac{5}{2},\sigma^2=400\cdot\frac{5}{4}\right)=N(1000,500)$$

Si Z *es la variable aleatoria* $N(0,1)$,

$$P(S<950)=P\left(\frac{S-1000}{\sqrt{500}}<\frac{950-1000}{\sqrt{500}}\right)\approx F_Z(-2.24)=1-F_Z(2.24)=0.0125$$
$$P(1000<S<1050)=P(S<1050)-P(S<1000)\approx$$
$$\approx F_Z\left(\frac{1050-1000}{\sqrt{500}}\right)-\frac{1}{2}=F_Z(2.24)-\frac{1}{2}=0.4875.$$

La normal que mejor aproxima $\frac{1}{400}S$ *es*

$$\frac{S}{400}\sim N\left(\mu=\frac{1000}{400},\sigma^2=\frac{500}{400^2}\right)=N(2.5,0.003125)$$

Ejercicio 3 *(1.5 puntos) La variable aleatoria R, resistencia, toma los valores 1 y 2 con probabilidad 0,4 y 0,6 respectivamente. La variable aleatoria intensidad I tiene una función de densidad:*

$$f_1(x) = \begin{cases} \dfrac{x}{4} & \textit{si } 1 \leq x \leq 3 \\ \\ 0 & \textit{en el resto} \end{cases} \quad , \ \textit{si } R = 1 \qquad y \qquad f_2(x) = \begin{cases} \dfrac{x}{12} & \textit{si } 1 \leq x \leq 5 \\ \\ 0 & \textit{en el resto} \end{cases} \quad , \ \textit{si } R = 2$$

siendo r el valor que toma la variable resistencia R.

1. *¿Cuál es la probabilidad de que I valga más de 3 sabiendo que $R = 2$?*
2. *Halla $P(R = 2/I \leq 2)$.*

Solución.

1.
$$P(I > 3|R = 2) = \int_3^\infty f_2(x)\,dx = \int_3^5 \frac{x}{12}\,dx = \left[\frac{x^2}{24}\right]_3^5 = \frac{2}{3}$$

2.
$$P(R = 2\,|\,I \leq 2) = \frac{P(R = 2, I \leq 2)}{P(I \leq 2)} = \frac{P(I \leq 2|R = 2)P(R = 2)}{P(I \leq 2|R = 1)P(R = 1) + P(I \leq 2|R = 2)P(R = 2)}$$

y como

$$P(I \leq 2|R = 1) = \int_1^2 f_1(x)\,dx = \int_1^2 \frac{x}{4}\,dx = \left[\frac{x^2}{8}\right]_1^2 = \frac{3}{8}$$
$$P(I \leq 2|R = 2) = \int_1^2 f_2(x)\,dx = \int_1^2 \frac{x}{12}\,dx = \left[\frac{x^2}{24}\right]_1^2 = \frac{1}{8}$$

resulta que

$$P(R = 2\,|\,I \leq 2) = \frac{\frac{1}{8} \cdot 0.6}{\frac{3}{8} \cdot 0.4 + \frac{1}{8} \cdot 0.6} = \frac{1}{3}$$

Ejercicio 4 *(1.5 puntos) La función de densidad de la variable aleatoria bidimensional* $(X,\ Y)$ *es*

$$f(x,\ y) = \begin{cases} a(x+1)(y+1) & si\ -1 \leq x \leq 1,\ -1 \leq y \leq 1 \\ 0 & en\ el\ resto \end{cases}$$

para cierto número real $a > 0$.

1. *Halla el valor de* a.
2. *Calcula la función de densidad marginal de* X.
3. *Calcula* $P(X > 0/Y < X)$.

Solución.

1. *Imponemos que la función de densidad tiene que integrar 1:*

$$\begin{aligned} 1 &= \iint_{\mathbb{R}^2} f(x,y)\,dxdx = \int_{-1}^{1}\int_{-1}^{1} a(x+1)(y+1)\,dxdy = \\ &= a\left(\int_{-1}^{1} x+1\,dx\right)\left(\int_{-1}^{1} y+1\,dy\right) = a\left[\frac{x^2}{2}+x\right]_{-1}^{1}\left[\frac{y^2}{2}+y\right]_{-1}^{1} = 4a \end{aligned}$$

   luego $a = \frac{1}{4}$.

2. $$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy = \frac{1}{4}(x+1)\int_{-1}^{1} y+1\,dy = \frac{x+1}{2}$$

3. $$P(X > 0/Y < X) = \frac{P\big(Y < X \cap X > 0\big)}{P(Y < X)}$$

   Como

$$\begin{aligned} P(Y < X) &= \frac{1}{4}\int_{-1}^{1}\int_{-1}^{x}(x+1)(y+1)dydx = \frac{1}{4}\int_{-1}^{1}(x+1)\left[\frac{(y+1)^2}{2}\right]_{-1}^{x} dydx \\ &= \frac{1}{8}\int_{-1}^{1}(x+1)^3dx = \frac{1}{8}\left[\frac{(x+1)^4}{4}\right]_{-1}^{1} = \frac{1}{2} \end{aligned}$$

   y

$$\begin{aligned} P\big(Y < X \cap X > 0\big) &= \frac{1}{4}\int_{0}^{1}\int_{-1}^{x}(x+1)(y+1)dydx = \frac{1}{4}\int_{0}^{1}(x+1)\left[\frac{(y+1)^2}{2}\right]_{-1}^{x} dydx = \\ &= \frac{1}{8}\int_{0}^{1}(x+1)^3dx = \frac{1}{8}\left[\frac{(x+1)^4}{4}\right]_{0}^{1} = \frac{15}{32} \end{aligned}$$

   se obtiene

$$P(X > 0/Y < X) = \frac{15/32}{1/2} = \frac{15}{16}$$

Ejercicio 5 *(2 puntos) Se desea estudiar la variable*

$$X \equiv \textit{Longitud en mm de los tornillos de cierto lote disponible}$$

1. *Si suponemos que* X *tiene distribución normal con desviación típica* $\sigma_X = 0,1$ *mm ¿Cuántos tornillos tendríamos que medir para obtener un intervalo de confianza para la longitud media,* $E[X]$*, con un nivel del 95 % y con amplitud menor de 0,02 mm (con un margen de error menor de 0,01 mm)?*

2. *Para contrastar la hipótesis nula* $H_0 \equiv$ *"La desviación típica de* X *es* $\sigma_0 = 0,1$ *mm", frente a la alternativa* $H_1 \equiv$ *"*$\sigma_0 \neq 0,1$ *mm", se miden 20 tornillos. Se obtiene que la cuasidesviación típica de las 20 longitudes resultantes es* $s_1 = 0,1315$ *mm. Utilizando el estadístico de contraste* $(n-1)S_1^2/\sigma_0^2$*, calcula el p-valor del contraste (con la aproximación que te permita la tabla de cuantiles de la chi-cuadrado).*

3. *Se contrasta la hipótesis nula* $H_0 \equiv$ *"La proporción de tornillos defectuosos es* $p = 0,01$ *(1 %)", frente a la alternativa* $H_1 \equiv$ *"*$p \neq 0,01$*", basándose en que de 300 tornillos observados ninguno es defectuoso. Calcula el p-valor del contraste. ¿Se acepta la hipótesis nula con un nivel de significación del 0,05 (5 %)?*

Solución.

1. *El cuantil* 0.975 *de una normal estandar, N(0,1), es aproximadamente* $z_{0,975} = 1,96$*. Entonces, para un nivel del* $0,95$*, el margen de error es menor de 0,01 mm, cuando el número de tornillos* n *de la muestra verifica*

$$\frac{\sigma}{\sqrt{n}} z_{0,975} = \frac{0,1}{\sqrt{n}} 1,96 < 0,01 \quad \textit{o equivalentemente} \quad n > \left(\frac{0.1}{0.01} 1.96\right)^2 = 384,16$$

Tendríamos pues que medir al menos 385 tornillos.

2. *Para la muestra disponible el estadístico de contraste* $(n-1)S_1^2/\sigma_0^2$ *toma el valor*

$$(n-1)\frac{s_1^2}{\sigma_0^2} = 19 \, \frac{0,1315^2}{0,1^2} \approx 32.85$$

Supuesto que la hipótesis nula es cierta, el estadístico de contraste $(n-1)S_1^2/\sigma_0^2$ *tiene aproximadamente distribución* χ^2_{19}*. Entonces*

$$\textit{p-valor} = 2P(\chi^2_{19} \geq 32,85) \approx 2 \times 0,025 = 0,05$$

El valor aproximado $P(\chi^2_{19} \geq 32,85) \approx 0,025$ *se obtiene utilizando las tablas de cuantiles de la chi-cuadrado.*

3. *Utilizamos el estadístico de contraste* $N \equiv$ *número de tornillos defectuosos entre 300. Supuesto que la hipótesis nula es cierta* N *tiene distribución binomial con parámetros* $n = 300$ *y* $p = 0,01$*. Entonces*

$$\textit{p-valor} = 2P(N = 0) = 2(1-p)^{300} = 2 \times 0,99^{300} \approx 0,098$$

Como el nivel de significación 0,05 es menor que el p-valor 0,098 se acepta la hipótesis nula.

Ejercicio 6 *(2 puntos) Sea $\{X(t)\ /t \in \mathbb{R}\}$ un proceso gaussiano, o normal, estacionario con media 0 y función de autocorrelación $R_X(\tau) = e^{-2|\tau|}$.*

1. *Obtén las distribuciones de probabilidad de $[X(t), X(t+\tau)]$, con $t, \tau \in \mathbb{R}$.*

2. *Halla la matriz de varianzas y covarianzas de la variable aleatoria tridimensional $\begin{pmatrix} X(1) \\ X(2) \\ X(5) \end{pmatrix}$.*

3. *Halla la probabilidad de que $2X(1) - 3X(2) + X(5) \geq 1$.*

4. *Sea un sistema lineal e invariante con función de transferencia $H(w) = \begin{cases} 1 & si\ |w| \leq 2 \\ 0 & en\ el\ resto \end{cases}$.*
   Si el proceso $Y(t)$ es la salida del sistema cuando la entrada es $X(t)$, halla la distribución de la variable aleatoria $Y(t)$ para t fijo.
   (**Indicación:** *La transformada de Fourier de $e^{-2|\tau|}$ es $\dfrac{4}{4+w^2}$, es decir $\mathcal{F}(e^{-2|\tau|}) = \dfrac{4}{4+w^2}$.*)

Solución.

1. *El proceso $X(t)$ gaussiano, o normal, así que la distribución conjunta de las variables aleatorias $\begin{pmatrix} X(t) \\ X(t+\tau) \end{pmatrix}$ es normal bidimensional.*

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

   $V[X(t)] = R_X(0) - \mu_X^2 = e^{-2\cdot 0} - 0^2 = 1$ *y* $Cov[X(t),\ X(t+\tau)]] = R_X(|t+\tau-t|) - \mu_X^2 = e^{-2|\tau|}$.

   Así que la matriz de varianzas y covarianzas es $\begin{pmatrix} 1 & e^{-2|\tau|} \\ e^{-2|\tau|} & 1 \end{pmatrix}$ y las distribuciones conjuntas de segundo orden son

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

2. *Por la misma razón que antes, ser el proceso $X(t)$ gaussiano, la distribución conjunta de las variables aleatorias $\begin{pmatrix} X(1) \\ X(2) \\ X(5) \end{pmatrix}$ es normal tridimensional.*

   El vector de medias es $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, por ser $\mu_X(t) = E[X(t)] = \mu_X = 0$.

   Por el apartado anterior la varianza de $X(t)$ es $V[X(t)] = R_X(0) - \mu_X^2 = e^{-2\cdot 0} = 1$, para todo t.

   Por el apartado anterior $Cov[X(1), X(2)] = R_X(|2-1|) - \mu_X^2 = e^{-2}$, $Cov[X(1), X(5)] = e^{-2\cdot 4}$, $Cov[X(2), X(5)] = e^{-2\cdot 3}$.

   Por lo que la matriz de varianzas y covarianzas de $\begin{pmatrix} X(1) \\ X(2) \\ X(5) \end{pmatrix}$ es $\quad \Sigma_{1,2,5} = \begin{pmatrix} 1 & e^{-2} & e^{-8} \\ e^{-2} & 1 & e^{-6} \\ e^{-8} & e^{-6} & 1 \end{pmatrix}$

3. *La variable aleatoria $2X(1) - 3X(2) + X(5)$ es una combinación lineal de variables aleatorias normales, por lo que es una variable aleatoria normal. Por ser el operador esperanza lineal,*

$$E[2X(1) - 3X(2) + X(5)] = 2E[X(1)] - 3E[X(2)] + E[X(5)] = 0 - 0 + 0 = 0$$

Por las propiedades del operador varianza,

$$V[2X(1)-3X(2)+X(5)] = 2^2V[X(1)] + (-3)^2V[X(2)] + V[X(5)] +$$
$$+2\cdot 2\cdot(-3)\cdot Cov[X(1),X(2)] + 2\cdot 2\cdot 1\cdot Cov[X(1),X(5)] + 2\cdot(-3)\cdot 1\cdot Cov[X(2),X(5)]$$

$$V[2X(1)-3X(2)+X(5)] = 4+9+1-12\cdot e^{-2}+4\cdot e^{-8}-6\cdot e^{-6} \approx 12,36$$

Entonces, tipificando, suponiendo que Z es la variable aleatoria $\mathcal{N}(0,\ 1)$ y consultando las tablas

$$P(2X(1)-3X(2)+X(5)) \geq 1) = P\left(\frac{2X(1)-3X(2)+X(5)-0}{\sqrt{12,36}} \geq \frac{1-0}{\sqrt{12,36}}\right) =$$
$$P(Z \geq 0,28) = 1-\Phi(0,28) = 1-0,61026 = 0,38974$$

4. *Por ser $Y(t)$ la salida de un sistema lineal e invariante del proceso estocástico $X(t)$ que es estacionario en sentido amplio, el proceso $Y(t)$ es estacionario en sentido amplio. Además, por ser $X(t)$ normal, se verifica que $Y(t)$ es, además estacionario en sentido amplio, normal. Por tanto, la variable aleatoria $Y(t)$, para t fijo, es una variable aleatoria normal.*

   $E[Y(t)] = H(0)E[X(t)] = H(0)\mu_X(t) = 1\cdot\mu_X = 1\cdot 0 = 0$

   $Var[Y(t)] = R_Y(0) - (E[Y(t)])^2 = R_Y(0) - 0^2$

$$\begin{aligned} Var[Y(t)] = R_Y(0) &= \frac{1}{2\,\pi}\int_{-\infty}^{\infty} S_Y(\omega)\, d\omega \\ &= \frac{1}{2\,\pi}\int_{-\infty}^{\infty} |H(\omega)|^2 S_X(\omega)\, d\omega = \frac{1}{2\,\pi}\int_{-2}^{2}\frac{4}{4+w^2}\, d\omega \\ &= \frac{1}{2\,\pi} 2\ arctg\left(\frac{\omega}{2}\right)\Big|_{-2}^{2} = \frac{1}{2} \end{aligned}$$

   Así que, $Y(t) \sim \mathcal{N}(0,\ \sigma^2 = \frac{1}{2})$.