47'2000

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 12-9-2000 (Tarde)

APELLIDOS Y NOMBRE:________________

Nº DE EXPEDIENTE:____________ CURSO: 3º GRUPO:______

PROBLEMA 1

(2'5 puntos)

Se tiene una colección de 40 disquetes de ordenador, donde 18 están formateados y 22 sin formatear. Se utilizan dos formas de extracción de muestras de n disquetes:

a) Un disquete se elige aleatoriamente, anotamos si está formateado o no, y lo devolvemos a la colección. Después, aleatoriamente, volvemos a seleccionar un disquete al que tratamos de igual manera, y, así, sucesivamente las n veces.

b) Se toma simultáneamente la muestra aleatoria de n disquetes.

1) Si se seleccionara una muestra de 6 disquetes, de los modos indicados en a) y b), se pide ¿cuál sería la probabilidad, en cada uno de estos casos a) y b), de que la muestra tuviera, en su mayoría (más de la mitad), discos formateados? *(1'4p)*

Si, una vez concluido el proceso de las 6 extracciones indicadas en el apartado 1), hubiéramos obtenido sólo 3 discos formateados, qué probabilidad tendríamos de conseguir la ansiada mayoría (más de la mitad), en la siguiente extracción (séptima), en los casos:

2) Las 6+1 extracciones realizadas fueran del tipo a). *(0'4p)*

3) Tras realizar las 6 primeras extracciones simultáneamente, según el tipo b), esta 7ª se realizara, aleatoriamente, sobre la colección restante, en la que no se hubieran repuesto las 6 extraídas previamente. *(0'7p)*

SOLUCIÓN:

1a) CON REPOSICIÓN:

Cada extracción parte de la misma situación, al ser repuesto el disquete anterior. La probabilidad de éxito es constante en las 6 pruebas de Bernouilli sucesivas, luego se dan las condiciones para poder definir una V.A. Binomial.

Definimos el suceso éxito $S=\{\text{"extraigo un disquete formateado"}\}$; $\quad p=P(S)=\frac{18}{40}=0'45$

Definimos la v.a. binomial: $X=\{\text{nº de formateados (éxitos) extraídos en las 6 pruebas}\}$

El suceso del cual nos piden calcular la probabilidad es:

$A=\{\text{"de los 6 con reposición, más de la mitad de los disquetes son formateados"}\}$;

$$A\equiv[4\le X\le 6]$$

$$\boxed{P(A)=}\ P[4\le X\le 6]=\sum_{k=4}^{6}\binom{6}{k}\cdot 0'45^k\cdot 0'55^{6-k}=\binom{6}{4}\cdot 0'45^4\cdot 0'55^2+\binom{6}{5}\cdot 0'45^5\cdot 0'55^1+$$

$$+\binom{6}{6}\cdot 0'45^6\cdot 0'55^0\quad\cong 0'18607+0'06089+0'00830\quad \boxed{=0'255298}$$

1b) SIN REPOSICIÓN:

En este caso la v.a. Y, definida como X, pero sin reposición será hipergeométrica.

$$N=40\qquad N_F=18\qquad N_{NF}=22$$

Definimos la v.a. binomial: $Y=\{\text{nº de formateados (éxitos) extraídos en las 6 pruebas}\}$

El suceso del cual nos piden calcular la probabilidad es:

$B=\{\text{"de los 6 sin reposición, más de la mitad de los disquetes son formateados"}\}$;

$$B\equiv[4\le Y\le 6]$$

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 12-9-2000 (Tarde)

APELLIDOS Y NOMBRE:

Nº DE EXPEDIENTE: CURSO: 3º GRUPO:

PROBLEMA 1: Continuación

$$\boxed{P(B) = P[4 \leq Y \leq 6] = \sum_{k=4}^{6} \frac{\binom{18}{k} \cdot \binom{22}{6-k}}{\binom{40}{6}} = \frac{706.860 + 188.496 + 18.564}{3.838.380} \cong 0'2381}$$

2) Nos piden la probabilidad condicionada a que anteriormente se hubiesen hecho 6 con reposición y hubiesen salido 3 éxitos.
Debido a la independencia de las pruebas, es inmediato que no se condiciona:

$$\boxed{p = \frac{18}{40} = 0'45}$$

Nota: puede trabajarse en la forma de probabilidad de la intersección entre la de la condición, obteniendo el mismo resultado.

3) Ya hemos extraído 3 discos formateados y por tanto, en las 6 pruebas han salido otros 3 no formateados.
Como se ha tratado de pruebas sin reposición, al ir a extraer la 7ª, tendremos la siguiente situación:

$$\left\{ \begin{array}{l} 15 \text{ discos formateados} \\ 19 \text{ discos sin formato} \end{array} \right\} \text{Total} : 34 \text{ disquetes}$$

luego la probabilidad de obtener un disco formateado bajo la condición de haber extraído ya 3 formateados y otros 3 sin formato, sería de:

$$\boxed{p = \frac{\text{casos favorables equiprobables bajo la condición}}{\text{posibles}} = \frac{15}{34} \cong 0'441176}$$

45'2000

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 12-9-2000 (Tarde)

APELLIDOS Y NOMBRE:___________________________

Nº DE EXPEDIENTE:_______________ CURSO: 3º GRUPO:_______________

PROBLEMA 2 *(2 puntos)*

A una determinada ventanilla de un aeropuerto, llegan los pasajeros, en promedio, a razón de 192 por día. Estas llegadas son aleatorias a lo largo del tiempo, por lo que seguirán un proceso homogéneo de Poisson. Se pide:

a) Probabilidad de que en media hora lleguen 5 personas exactamente. *(0'5p)*

b) Probabilidad de que al menos lleguen 2 personas en 1 hora. *(0'5p)*

c) Probabilidad de que pasen menos de 15 minutos desde que llega un pasajero hasta que llega el próximo. *(0'5p)*

d) Si sabemos que han transcurrido dos horas sin que haya aparecido ningún pasajero, ¿qué probabilidad existe de que en las dos horas siguientes tampoco aparezca ninguno?. *(0'5p)*

SOLUCIÓN:

$$\lambda = 192 \ \frac{\text{pasaj.}}{\text{día}} = 192 \ \frac{\text{pasaj.}}{\text{día}} \cdot \frac{1 \text{ día}}{24 \text{ horas}} = \frac{192}{24} \ \frac{\text{pasajeros}}{\text{hora}} = 8 \ \frac{\text{pasajeros}}{\text{hora}}$$

a) $$P\left(X\left(\tfrac{1}{2}\right)=5\right)=(e^{-8\frac{1}{2}})\frac{\left(8\ \tfrac{1}{2}\right)^5}{5!}=e^{-4}\frac{4^5}{5!}=0'1563$$

b) $$P(X(1)\geq 2)=1-P(X(1)<2)=1-P(X(1)=0)-P(X(1)=1)=$$

$$1-(e^{-8})\frac{(8)^0}{0!}-(e^{-8})\frac{(8)^1}{1!}=1-0'0003-0'0027=0'997$$

c) $$P\left(x\left(\tfrac{1}{4}\right)>0\right)=1-P\left(x\left(\tfrac{1}{4}\right)=0\right)=1-(e^{-8\frac{1}{4}})\frac{\left(8\ \tfrac{1}{4}\right)^0}{0!}=1-(e^{-2})\frac{(2)^0}{0!}=1-0'1353=0'8647$$

d) La distribución exponencial no tiene memoria. Por tanto, la probabilidad pedida es la misma que la probabilidad de que en 2 horas no aparezca ningún pasajero:

$$P(X(2)=0)=\left(e^{-8.2}\right)\frac{(8.2)^0}{0!}=0'00000012$$

44'2000

PROBLEMA 3 *(3 puntos)*

Dada la variable aleatoria bidimensional (X,Y), se sabe que su función de densidad conjunta es nula en todo R^2 excepto en la región $R_1 = \{ (x,y) \in R^2 \ / \ 0<x<1,\ 0<y<1\}$. En esta región, su función de distribución vale:

$$F(x,y) = \frac{1}{2}\, x\, y\, (x+y)$$

Calcule:

a) Función de densidad conjunta de (X,Y) *(0'5p)*

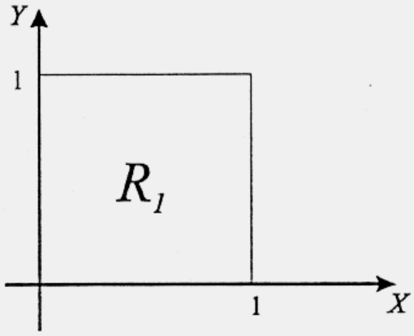
b) Funciones de densidad marginales *(0'5p)*

c) ¿Son independientes X e Y? *(0'6p)*

d) Coeficiente de correlación lineal entre X e Y *(0'9p)*

e) Recta de regresión de Y/X *(0'5p)*

SOLUCIÓN



a)

- En la región R_1 será:

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} = \frac{\partial}{\partial x}\left[\frac{\partial F(x,y)}{\partial y}\right] = \frac{\partial}{\partial x}\left[\frac{\partial}{\partial y}\left(\frac{1}{2}xy(x+y)\right)\right] = \frac{\partial}{\partial x}\left[\frac{\partial}{\partial y}\left(\frac{x^2 y}{2} + \frac{xy^2}{2}\right)\right]$$

$$= \frac{\partial}{\partial x}\left[\frac{x^2}{2} + \frac{2xy}{2}\right] \quad = x + y$$

- Fuera de R_1 sabemos por el enunciado que $f(x,y) = 0$.

Luego:

$$f(x,y) = \begin{cases} x+y & \forall (x,y) \in R_1 \\ 0 & \text{resto} \end{cases}$$

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 12-9-2000 (Tarde)

APELLIDOS Y NOMBRE:________________

Nº DE EXPEDIENTE:__________ CURSO: 3º GRUPO:__________

PROBLEMA 3: Continuación A

b) Función de densidad marginal de X:

$\forall\, x \le 0$ ó $x \ge 1$: $\quad f_1(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy = \int_{-\infty}^{+\infty} 0\,dy = 0$

$\forall\, 0 < x < 1$: $\quad f_1(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy = \int_0^1 (x+y)\,dy = \left[xy + \frac{y^2}{2}\right]_{y=0}^{y=1} = x + \frac{1}{2}$

luego

$$f_1(x) = \begin{cases} x + \frac{1}{2} & \forall\ 0 < x < 1 \\ 0 & \text{resto} \end{cases}$$

Función de densidad marginal de Y:

$\forall\, y \le 0$ ó $y \ge 1$: $\quad f_2(y) = \int_{-\infty}^{+\infty} f(x,y)\,dx = \int_{-\infty}^{+\infty} 0\,dx = 0$

$\forall\, 0 < y < 1$: $\quad f_2(y) = \int_{-\infty}^{+\infty} f(x,y)\,dx = \int_0^1 (x+y)\,dx = \left[\frac{x^2}{2} + xy\right]_{x=0}^{x=1} = y + \frac{1}{2}$

luego

$$f_2(y) = \begin{cases} y + \frac{1}{2} & \forall\ 0 < y < 1 \\ 0 & \text{resto} \end{cases}$$

c) Para que X e Y sean independientes debe verificarse: $f(x,y) = f_1(x) f_2(y) \qquad \forall (x,y) \in \Re^2$

Pero si $(x,y) \in R_1$ es:

$$f_1(x) f_2(y) = \left(x + \frac{1}{2}\right)\left(y + \frac{1}{2}\right) = xy + \frac{x}{2} + \frac{y}{2} + \frac{1}{4}$$

mientras que $f(x,y) = x + y$

Estas dos expresiones no son una identidad.

En efecto, p. ej., para $(x,y) = \left(\frac{1}{3}, \frac{1}{3}\right)$ es

$$f_1\left(\frac{1}{3}\right) f_2\left(\frac{1}{3}\right) = \left(\frac{1}{3} + \frac{1}{2}\right)\left(\frac{1}{3} + \frac{1}{2}\right) = \left(\frac{5}{6}\right)\left(\frac{5}{6}\right) = \frac{25}{36} \ne \frac{1}{3} + \frac{1}{3} = \frac{2}{3} = \frac{24}{36} = f(x,y)$$

luego X e Y no son independientes.

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID
Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 12-9-2000 (Tarde)

APELLIDOS Y NOMBRE:
Nº DE EXPEDIENTE: CURSO: 3º GRUPO:

PROBLEMA 3: Continuación B

d) Nos piden $\rho = \frac{\mu_{11}}{\sigma_x \sigma_y}$ siendo:

$$\mu_{11} = \alpha_{11} - \alpha_{10}\alpha_{01} = E[XY] - E[X]E[Y]$$
$$\sigma_X^2 = \alpha_{20} - \alpha_{10}^2 = E[X^2] - E^2[X]$$
$$\sigma_Y^2 = \alpha_{02} - \alpha_{01}^2 = E[Y^2] - E^2[Y]$$

$$E[X] = \int_{-\infty}^{+\infty} x f_1(x)dx = \int_0^1 x\left(x + \frac{1}{2}\right)dx = \int_0^1 \left(x^2 + \frac{x}{2}\right)dx = \left[\frac{x^3}{3} + \frac{x^2}{4}\right]_0^1 = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}$$

$$E[X^2] = \int_{-\infty}^{+\infty} x^2 f_1(x)dx = \int_0^1 x^2\left(x + \frac{1}{2}\right)dx = \int_0^1 \left(x^3 + \frac{x^2}{2}\right)dx = \left[\frac{x^4}{4} + \frac{x^3}{6}\right]_0^1 = \frac{1}{4} + \frac{1}{6} = \frac{5}{12}$$

Por la simetría en el recinto y en la forma de $f_1(x)$ y $f_2(y)$, se puede ver que:

$E[Y] = E[X] = \frac{7}{12}$ y que $E[Y^2] = E[X^2] = \frac{5}{12}$

$$E[XY] = \iint_{R^2} xyf(xy)dxdy = \int_0^1 dx \int_0^1 xyf(xy)dy = \int_0^1 dx \int_0^1 xy(x+y)dy = \int_0^1 dx \int_0^1 (x^2y + xy^2)dy$$

$$= \int_0^1 \left[\frac{x^2y^2}{2} + \frac{xy^3}{3}\right]_{y=0}^{y=1} dx = \int_0^1 \left(\frac{x^2}{2} + \frac{x}{3}\right)dx = \left[\frac{x^3}{6} + \frac{x^2}{6}\right]_0^1 = \frac{1}{3}$$

Entonces,

$$\mu_{11} = \frac{1}{3} - \frac{7}{12}\frac{7}{12} = \frac{48-49}{144} = -\frac{1}{144}$$
$$\sigma_X^2 = \frac{5}{12} - \frac{7}{12}\frac{7}{12} = \frac{60-49}{144} = \frac{11}{144} = \sigma_Y^2$$

quedando $\sigma_X = \sqrt{\sigma_X^2} = \sqrt{11/144} = \sigma_Y$

y:
$$\rho = \frac{\mu_{11}}{\sigma_X \sigma_Y} = \frac{\frac{-1}{144}}{\sqrt{\frac{11}{144}}\sqrt{\frac{11}{144}}} = -\frac{1}{11}$$

e) Recta de regresión de Y sobre X:

$$y - \bar{y} = \frac{\mu_{11}}{\sigma_X^2}(x - \bar{x}) \Rightarrow y - \frac{7}{12} = \frac{-\frac{1}{144}}{\frac{11}{144}}\left(x - \frac{7}{12}\right) = -\frac{1}{11}\left(x - \frac{7}{12}\right) \Rightarrow$$

$$y = -\frac{1}{11}x + \frac{7}{11 \cdot 12} + \frac{7}{12} \Rightarrow y = -\frac{1}{11}x + \frac{7}{11}$$

41'2000

APELLIDOS Y NOMBRE:__________
Nº DE EXPEDIENTE:__________ CURSO: 3º GRUPO:__________

PROBLEMA 4 *(2'5 puntos)*

Se quieren hacer inferencias sobre una población, mediante un muestreo aleatorio simple, con reemplazamiento. Se conocen los siguientes datos:

- la población tiene distribución normal
- tamaño de la muestra: 25
- media muestral: 197
- cuasivarianza muestral: 225

a) Obtenga un intervalo de confianza del 95% para la media poblacional. ***(0'5p)***

b) Obtenga un intervalo de confianza del 90% para la desviación típica de esta población. ***(0'5p)***

c) Si se quiere tener una confianza del 99% de que la estimación realizada de la media se encuentre a una distancia de menos de 3 unidades de la verdadera media de la población, ¿cuántas observaciones más deberían tomarse? ***(0'5p)***

d) Con la muestra inicial, realice un contraste unilateral, con un nivel de significación del 0'05, para ver si se puede aceptar o no que la media de la población es menor o igual que 194. ***(0'5p)***

e) Con la muestra inicial, realice un contraste bilateral para decidir aceptar o no que la varianza de la población sea 150, de manera que sea sólo del 5% la probabilidad de que si esta varianza realmente fuera 150, el contraste sin embargo nos dijera que no. ***(0'5p)***

SOLUCIÓN

a) Estimación de la media de una población normal de varianza desconocida y n=25<30.

$$I = \left[\bar{x} \pm t_{n-1,\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}\right]$$

$$1-\alpha = 0'95 \Rightarrow \alpha = 0'05 \Rightarrow \frac{\alpha}{2} = 0'025 \Rightarrow t_{24,\,0'025} = 2'064$$

$$I = \left[197 \pm 2'064 \cdot \frac{\sqrt{225}}{\sqrt{25}}\right] = [197 \pm 6'192] = [190'808,\ 203'192]$$

b) $$I = \left[\sqrt{\frac{(n-1)\cdot s^2}{\chi^2_{\frac{\alpha}{2},(n-1))}}};\ \sqrt{\frac{(n-1)\cdot s^2}{\chi^2_{1-\frac{\alpha}{2},(n-1)}}}\right]$$

$$1-\alpha = 0'9 \Rightarrow \alpha = 0'1 \Rightarrow \frac{\alpha}{2} = 0'05 \Rightarrow \begin{cases} \chi^2_{0'05,24} = 36'415 \\ \chi^2_{0'95,24} = 13'848 \end{cases}$$

$$I = \left[\sqrt{\frac{24\cdot 225}{36'415}};\sqrt{\frac{24\cdot 225}{13'848}}\right] = \left[\sqrt{148'29};\ \sqrt{389'95}\right] = [12'18;\ 19'75]$$

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 12-9-2000 (Tarde)

APELLIDOS Y NOMBRE:

Nº DE EXPEDIENTE: CURSO: 3º GRUPO:

PROBLEMA 4: Continuación

c) Del apartado a) sabemos que para n=25 y 1-α=0'95, el error en la estimación era inferior a 6.192 unidades. Si queremos un error aún más pequeño y con mayor confianza, parece claro que n deberá ser bastante más grande y seguramente será n>30. Luego sería:

$I = \left[\bar{x} \pm z_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} \right]$ cumpliéndose que $P\left(\bar{x} - z_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} \right) = 1 - \alpha$

error: $\varepsilon = z_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} \leq 3 \Rightarrow n \geq \left(z_{\frac{\alpha}{2}} \right)^2 \cdot \frac{s^2}{3^2}$

$1-\alpha = 0'99 \Rightarrow \alpha = 0'01 \Rightarrow \frac{\alpha}{2} = 0'005 \Rightarrow z_{\frac{\alpha}{2}} = 2'5757$

$n \geq 2'5757^2 \cdot \frac{225}{9} = 165'86 \Rightarrow n = 166$

luego se deberían tomar 166-25=141 observaciones más.

d) Se trata de un contraste unilateral sobre la media de una población normal con varianza desconocida y muestra pequeña, n=25<30

$H_0 : \mu \leq 194$ $\qquad$ $H_1 : \mu > 194$

estadístico de contraste: $t = \frac{\bar{x} - \mu}{s / \sqrt{n}} = \frac{197 - 194}{15 / \sqrt{25}} = 1$

región de aceptación: $t \leq t_{\alpha, n-1} = t_{0'05, 24} = 1'711$

como $t = 1 < 1'711 = t_{\alpha, n-1}$, a este nivel de significación no hay motivo para rechazar la hipótesis de que la media es inferior a 194.

e) Del enunciado, $\alpha = 0'05 \Rightarrow \frac{\alpha}{2} = 0'025$

$H_0 : \sigma^2 = 150$ $\qquad$ $H_1 : \sigma^2 \neq 150$

estadístico de contraste: $\chi^2 = \frac{(n-1) \cdot s^2}{\sigma^2} = \frac{24 \cdot 225}{150} = 36$

región de aceptación: $\left[\chi^2_{1-\frac{\alpha}{2}; n-1} ; \chi^2_{\frac{\alpha}{2}; n-1} \right] = \left[\chi^2_{0'975; 24} ; \chi^2_{0'025; 24} \right] = [12'401;\ 39'364]$

como el valor del estadístico de contraste está contenido en la región de aceptación, no hay motivo para rechazar la hipótesis de que $\sigma^2 = 150$, con ese nivel de significación.

39'2000

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID
Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 12-9-2000 (Mañana)

APELLIDOS Y NOMBRE: ________________

Nº DE EXPEDIENTE: ________ CURSO: 3º GRUPO: ________

PROBLEMA 1 *(2 puntos)*

Se tienen 2 urnas, la primera con 6 bolas blancas y 4 rojas, y la segunda con 5 blancas y 7 rojas. De la primera urna se toman 2 bolas al azar y se introducen en la segunda urna. Posteriormente se toma 1 bola al azar de la segunda urna y se introduce en la primera. A continuación se realiza la extracción de una bola de la primera urna. Calcular la probabilidad de que dicha extracción sea de una bola blanca.

SOLUCIÓN:

Sacamos 2 bolas de la urna 1y las introducimos en la 2, con lo que tenemos 3 casos:

2 B: $P_2(B_1 \cap B_2) = \left(6/10\right)\left(5/9\right) = 0'333$

1B y1 R: $P_2((B_1 \cap R_2) \cup (R_1 \cap B_2)) = \left(6/10\right)\left(4/9\right) + \left(4/10\right)\left(6/9\right) = 0'533$

2 R: $P_2(R_1 \cap R_2) = \left(4/10\right)\left(3/9\right) = 0'133$

La composición de la urna 2 puede ser una de estas 3:

$$\underbrace{7B\ y\ 7R}_{\text{tipo 1}} \quad \underbrace{6B\ y\ 8R}_{\text{tipo 2}} \quad \underbrace{5B\ y\ 9R}_{\text{tipo 3}}$$

Sacamos una bola de la urna 2 e introducimos en la urna 1, con lo que tenemos 6 casos:

B desde tipo 1: $P_1(B) = 0'333\left(7/14\right) = 0'166$

R desde tipo 1: $P_1(R) = 0'333\left(7/14\right) = 0'166$

B desde tipo 2: $P_1(B) = 0'533\left(6/14\right) = 0'228$

R desde tipo 2: $P_1(R) = 0'533\left(8/14\right) = 0'304$

B desde tipo 3: $P_1(B) = 0'133\left(5/14\right) = 0'047$

R desde tipo 3: $P_1(R) = 0'133\left(9/14\right) = 0'085$

La composición de la urna 1 puede ser una de estas 6:

$$\underbrace{5B\ y\ 4R}_{\text{tipo 1}} \quad \underbrace{4B\ y\ 5R}_{\text{tipo 2}} \quad \underbrace{6B\ y\ 3R}_{\text{tipo 3}} \quad \underbrace{5B\ y\ 4R}_{\text{tipo 4}} \quad \underbrace{7B\ y\ 2R}_{\text{tipo 5}} \quad \underbrace{6B\ y\ 3R}_{\text{tipo 6}}$$

La probabilidad de sacar una bola blanca de la urna 1 es:

$$P(B) = 0'166\left(5/9\right) + 0'166\left(4/9\right) + 0'228\left(6/9\right) + 0'304\left(5/9\right) + 0'047\left(7/9\right) + 0'085\left(6/9\right) = 0'578$$

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID
Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 12-9-2000 (Mañana)

APELLIDOS Y NOMBRE:______________________________
Nº DE EXPEDIENTE:__________________ CURSO: 3º GRUPO:______________

PROBLEMA 2 *(2'5 puntos)*

Se realiza un estudio sobre la composición de un determinado medicamento en su fabricación. Se toman 8 unidades y se observa la cantidad en miligramos del componente A, en cada unidad, con estos resultados:

30'5 30'1 30'7 29'8 29'9 30'1 28'7 29'2

Se sabe que la población tiene distribución Normal.

a) La cantidad de componente A que debe formar parte del medicamento es de 30 miligramos. Con un nivel de significación de 0.05, ¿se puede afirmar que la participación del componente A en el medicamento está siendo la adecuada? ***(1'2p)***

b) Con una prueba como la anterior, pero observando 50 unidades, calcule la probabilidad de que se detecte que un medicamento con μ=31 miligramos no tiene la cantidad adecuada del componente A. Suponga para ello que los valores muestrales media y cuasivarianza de la nueva muestra resultan los mismos que los obtenidos con la muestra anterior. ***(1'3p)***

NOTA: *No es necesario interpolar. Tómese en la tabla el valor más próximo.*

SOLUCIÓN:

a) Se trata de un contraste bilateral de la media, con σ desconocida y $n \leq 30$

$$H_0: \quad \mu = \mu_0 = 30$$
$$H_a: \quad \mu \neq \mu_0 = 30$$

para el que utilizaremos como estadístico de contraste: $t = \dfrac{\overline{X} - 30}{s_M / \sqrt{n}}$

La región de aceptación es: $|t_r| \leq t_{\alpha/2;n-1}$

- calculamos la realización del estadístico como sigue

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} = 29'875 \quad \Rightarrow \quad s_M^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2 = 0'43 \quad \Rightarrow \quad s_M = \sqrt{0'43} = 0'655$$

$$t_r = \frac{29'875 - 30}{\sqrt{0'43/8}} = -0'539 \qquad \Rightarrow \qquad |t_r| = 0'539$$

- Calculamos los puntos críticos $\quad t_{\alpha/2;n-1} = t_{0'025;7} = 2'365$

Dado que se cumple $|t_r| = 0'539 \leq 2'365 = t_{\alpha/2;n-1}$, se acepta la hipótesis nula. La participación del componente A está siendo la adecuada.

37'2000

APELLIDOS Y NOMBRE:
Nº DE EXPEDIENTE: CURSO: 3º GRUPO:

PROBLEMA 2: Continuación

b) $H_0 : \mu = \mu_0 = 30$, pero sea realmente $\mu = \mu_a = 31$

con : $\bar{x} = 29'875 \qquad s_M = 0'655 \qquad n = 50$

Ahora $n > 30$, luego se usa la distribución normal, es decir, la región de aceptación es : $|z| \leq z_{\alpha/2}$

$$P = 1 - \beta \qquad \beta = P\left(\text{Aceptar } H_0 \Big/ H_a \text{ cierta}\right)$$

El intervalo de confianza bajo H_0, nos dará la región de aceptación :

$$\mu_0 \pm z_{\alpha/2} \frac{s_M}{\sqrt{n}} = \quad 30 \pm 1'96 \frac{0'655}{\sqrt{50}} = \quad [29'818 \ , \ 30'181]$$

Siendo $\mu = \mu_a = 31$, la distribución de la $\bar{x}$ es $N\left(31, \frac{0'655}{\sqrt{50}}\right)$, luego :

$$\beta = P(29'818 \leq \bar{x} \leq 30'181) = P\left(\frac{29'818 - 31}{\frac{0'655}{\sqrt{50}}} \leq z \leq \frac{30'181 - 31}{\frac{0'655}{\sqrt{50}}}\right) = P(-1'276 \leq z \leq -0'884) =$$

$$= P(z \geq 0'884) - P(z \geq 1.276) = 0'0891$$

Potencia $= 1 - 0'0891 = 0'9109$

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID
Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 12-9-2000 (Mañana)

APELLIDOS Y NOMBRE:______________________
Nº DE EXPEDIENTE:______________ CURSO: 3º GRUPO:______________

PROBLEMA 3 *(3 puntos)*

Sea X el número de empresas en las que ha trabajado por cuenta ajena un programador en ensamblador, e Y el número de veces que el programador necesita ejecutar un programa suyo para depurarlo. La distribución de probabilidad conjunta de (X,Y) viene dada por:

P(x,y)	y = 1	2	3	4	5	
x = 0	0'02	0'03	0'1	0'08	0'02	0'25
1	0'04	0'06	0'13	0'07	0'02	0'32
2	0'05	0'1	0'07	0'05	0'02	0'29
3	0'04	0'06	0'03	0'01	0	0'14
	0'15	0'25	0'33	0'21	0'06	1

a) Calcule la probabilidad de que un programador, elegido al azar, haya trabajado a lo sumo en una empresa por cuenta ajena, y necesite no más de dos ejecuciones para depurar un programa suyo. ***(0'3p)***
b) Obtenga las distribuciones de probabilidad marginales de X e Y. ***(0'3p)***
c) Obtenga las esperanzas de ambas variables. ***(0'4p)***
d) Obtenga las varianzas de ambas variables. ***(0'4p)***
e) Calcule la probabilidad de que un programador cualquiera que haya trabajado en dos empresas por cuenta ajena, elegido al azar, necesite realizar al menos tres ejecuciones de un programa, para depurarlo. ***(0'4p)***
f) Calcule $\mathrm{Cov}(X,Y)$. ***(0'4p)***
g) Explique si son o no independientes ambas variables. ***(0'4p)***
h) Explique qué es más probable: ***(0'4p)***
I- que un programador que ha trabajado en tres empresas depure un programa en una ejecución;
II- que un programador que ha trabajado en dos empresas depure un programa en una ejecución;

SOLUCIÓN:

a) A: Un programador ha trabajado a lo sumo en 1 empresa por cuenta ajena.
B: Un programador necesita no más de 2 ejecuciones para depurar su programa.
$A \equiv X \in \{0,1\}$ $\quad B \equiv Y \in \{1,2\}$ $\quad P(A \cap B) = P(0,1) + P(0,2) + P(1,1) + P(1,2) = 0'15$

b) marginal de X

X	$P(x_i)$
0	0'25
1	0'32
2	0'29
3	0'14
	1

marginal de Y

Y	$P(y_j)$
1	0'15
2	0'25
3	0'33
4	0'21
5	0'06
	1

c)

X	$x_i \cdot P(x_i)$
0	0
1	0'32
2	0'58
3	0'42
	1'32

$= \sum x_i \cdot P(x_i) = E(X)$

Y	$y_j \cdot P(y_j)$
1	0'15
2	0'5
3	0'99
4	0'84
5	0'3
	2'78

$= \sum y_j \cdot P(y_j) = E(Y)$

35'2000

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 12-9-2000 (Mañana)

APELLIDOS Y NOMBRE:____________________

Nº DE EXPEDIENTE:____________ CURSO: 3º GRUPO:____________

PROBLEMA 3: Continuación

d)

X	$x^2_i \cdot P(x_i)$
0	0
1	0'32
2	1'16
3	1'26
	2'74

$= \sum x_i^2 \cdot P(x_i) = E(X^2)$

Y	$y^2_j \cdot P(y_j)$
1	0'15
2	1
3	2'97
4	3'36
5	1'5
	8'98

$= \sum y_j^2 \cdot P(y_j) = E(Y^2)$

$\sigma_X^2 = E(X^2) - E^2(X) = 2'74 - 1'32^2 = 0'9976$; $\sigma_Y^2 = E(Y^2) - E^2(Y) = 8'98 - 2'78^2 = 1'2516$

e) A: Un programador ha trabajado en 2 empresas por cuenta ajena.
B: Un programador necesita al menos 3 ejecuciones para depurar su programa.

$$A \equiv X = 2 \qquad B \equiv Y \in \{3,4,5\}$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(2,3) + P(2,4) + P(2,5)}{P(x=2)} = \frac{0'07 + 0'05 + 0'02}{0'29} = 0'48275862$$

f) $\sigma_{XY} = E(X \cdot Y) - E(X) \cdot E(Y)$

$$E(X \cdot Y) = \sum x_i y_j p_{ij} = 1 \cdot 0'04 + 2 \cdot 0'06 + 3 \cdot 0'13 + 4 \cdot 0'07 + 5 \cdot 0'02$$
$$+ 2 \cdot 0'05 + 4 \cdot 0'1 + 6 \cdot 0'07 + 8 \cdot 0'05 + 10 \cdot 0'02 +$$
$$+ 3 \cdot 0'04 + 6 \cdot 0'06 + 9 \cdot 0'03 + 12 \cdot 0'01 \qquad = \quad 3'32$$

$\sigma_{XY} = 3'32 - 1'32 \cdot 2'78 = -0'3496$

g) No son independientes, ya que:
la covarianza es no nula, por lo que
no son incorreladas, por lo que no pueden ser independientes.

De otra manera puede comprobarse la no independencia:
Las variables son independientes si y sólo si $p_{ij} = p_{x_i} \cdot p_{y_j} \quad \forall ij$
comprobando un punto, como es $P(0,1) = p_{01} = 0'02$; $p_{x=0} \cdot p_{y=1} = 0'25 \cdot 0'15 = 0'0375$
No son independientes ya que al menos en esta combinación no se cumple.

h) $[I]$ $P(1ejec|3empl) = \dfrac{P(3,1)}{P(x=3)} = \dfrac{0'04}{0'14} = 0'285714...$

$[II]$ $P(1ejec|2empl) = \dfrac{P(2,1)}{P(x=2)} = \dfrac{0'05}{0'29} = 0'17241...$

$$P[I] > P[II]$$

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: **12-9-2000 (Mañana)**

APELLIDOS Y NOMBRE:

Nº DE EXPEDIENTE: CURSO: 3º GRUPO:

PROBLEMA 4 *(2'5 puntos)*

La empresa española del sector informático donde suponga hipotéticamente que Vd. trabaja pretende lanzar al mercado un nuevo producto específico para los arquitectos. Para ello realiza un estudio de viabilidad en el mercado, efectuando una consulta, mediante muestreo aleatorio simple, a 1.750 arquitectos, entre los que se ha detectado que un 44'3% tiene la intención de realizar compra durante los próximos tres años (por el tipo de producto que es, no tiene sentido que un arquitecto realice más de una compra).

a) A partir de estos datos, la empresa realiza una presentación interna de los resultados, dando un intervalo de confianza para la proporción de intención de compra entres años de entre 0'424 y 0'462, sin informar, sin embargo, sobre el contenido probabilístico de dicho intervalo. Calcule Vd. este dato que falta. *(1'3p)*

b) Obtenga un intervalo de confianza del 95% para el porcentaje de arquitectos que tienen intención de compra en los próximos tres años. *(1'2p)*

SOLUCIÓN:

$n = 1.750 \Rightarrow \hat{p} = 0'443 \Rightarrow npq = 1750 \cdot 0'443 \cdot 0'557 > 9 \Rightarrow$ Cumple Moivre-Laplace

a) $I = \hat{p} \pm z_{\alpha/2}\sqrt{\dfrac{p\cdot(1-p)}{n}} \qquad \Rightarrow \qquad I \cong \hat{p} \pm z_{\alpha/2}\sqrt{\dfrac{\hat{p}\cdot(1-\hat{p})}{n}}$

El intervalo dado es $0'443 \pm 0'019$.

Igualando:

$$z_{\alpha/2}\sqrt{\frac{\hat{p}\cdot(1-\hat{p})}{n}} = 0'019 \qquad \Rightarrow \qquad z_{\alpha/2} = 0'019\sqrt{\frac{n}{\hat{p}\cdot(1-\hat{p})}} = 0'019\sqrt{\frac{1.750}{0'443\cdot 0'557}} = 1'6000854$$

$$\Rightarrow \quad z_{\alpha/2} \cong 1'6 \quad \Rightarrow \quad \frac{\alpha}{2} = 0'0548 \quad \Rightarrow \quad \alpha = 0'1096 \quad \Rightarrow \quad 1-\alpha = 89'04\%$$

b) $I = 0'443 \pm 1'96\sqrt{\dfrac{0'443\cdot 0'557}{1.750}} = 0'443 \pm 0'02327376 \cong [0'4197, 0'4663]$

Nota: Podría haberse resuelto por diseño en caso peor, con $p\cdot(1-p) = 0'25$. En ese caso también se cumplen las condiciones de la aproximación a Normal de Moivre-Laplace. $npq = 1750\cdot 0'5 \cdot 0'5 > 9$

Sin embargo, no cambiarían en forma visible los resultados, dado que el caso planteado es cercano al peor:

a) $z_{\alpha/2} = 0'019\sqrt{\dfrac{n}{p\cdot(1-p)}} = 0'019\sqrt{\dfrac{1.750}{0'25}} = 1'60006625$

b) $I = 0'443 \pm 1'96\sqrt{\dfrac{0'25}{1.750}} = 0'443 \pm 0'02327404 \cong [0'4197, 0'4663]$

33'2000

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 7-6-2000 (Tarde)

APELLIDOS Y NOMBRE:______________________

Nº DE EXPEDIENTE:______________ CURSO: **3º** GRUPO:______________

PROBLEMA 1 *(2'5 puntos)*

Luis tiene dos monedas diferentes en su bolsillo (M_1 y M_2). Para una de ellas, M_1, la probabilidad de obtener cara al lanzarla es 0'5, mientras que para la otra, M_2, es 0'3.

a) Si coge una de las monedas al azar y la lanza tres veces, calcule: **1p**

a1) Probabilidad de que salgan tres cruces **(NP)**

a2) Probabilidad de que salgan al menos dos caras

a3) Si ha salido una sola cruz, ¿cuál es la probabilidad de que haya cogido la moneda M_2.

b) Si saca una de las monedas de su bolsillo al azar, la lanza 40 veces, la vuelve a meter en el bolsillo, vuelve a coger otra al azar y la lanza otras 40 veces, calcule: **1'5p**

b1) Probabilidad de que con la primera moneda hayan salido menos de 30 cruces **(NP)**

b2) Probabilidad de que con la segunda moneda hayan salido entre 15 y 30 caras

b3) Probabilidad de que con cada una de las monedas hayan salido más caras que cruces

*Nota.- Si **necesitara** utilizar alguna tabla, no sería necesario interpolar, se podría coger el valor más cercano.*

SOLUCIÓN:

(a) Llamemos A_1 : se selecciona la moneda M_1

A_2 : se selecciona la moneda M_2

se verifica: $P(A_1) = P(A_2) = \frac{1}{2}$

Al lanzar la moneda escogida sea el suceso B: "sacar cara" (éxito)

Al lanzar la moneda escogida tres veces, sea la v.a. X: "nº de éxitos en los 3 lanzamientos"

como la probabilidad de éxito se mantiene de un lanzamiento a otro de la misma moneda, X tendrá distribución binomial, X : B(3, p), siendo p la probabilidad de éxito, es decir de sacar cara, con la moneda escogida

(a.1) Sea el suceso D: "sacar tres cruces en los tres lanzamientos de la moneda" = "sacar cero caras"

$P(D) = P(D/A_1)\cdot P(A_1) + P(D/A_2)\cdot P(A_2)$ (teor. probabilidad total)

$P(D/A_1) = P(X=0)$ (con la moneda M_1) $= \binom{3}{0} 0'5^0\, 0'5^3 = 0'125$

(también $P(D/A_1) = P(\text{Cruz, Cruz, Cruz})$ (con M_1) $= 0'5\cdot 0'5\cdot 0'5 = 0'125$)

$P(D/A_2) = P(X=0)$ con la moneda M_2) $= \binom{3}{0} 0'3^0\cdot 0'7^3 = 0'343$

(también $P(D/A_2) = 0'7\cdot 0'7\cdot 0'7 = 0'343$)

Estos valores se pueden leer directamente en la tabla de la distribución binomial, como $b(3,0,0'5)$ y $b(3,0,0'3)$, respectivamente

$P(D) = \frac{1}{2}(0'125 + 0'343) = \boxed{0'234}$

32'2000

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: **7-6-2000 (Tarde)**

APELLIDOS Y NOMBRE:____________________

Nº DE EXPEDIENTE:____________ CURSO: 3º GRUPO:__________

PROBLEMA 1: Continuación A

(a.2) Sea el suceso E: "sacar al menos dos caras en los tres lanzamientos de la moneda" = "sacar dos o tres caras..." = $E_2 \cup E_3$

$$P(E_2) = P(E_2/A_1)\,P(A_1) + P(E_2/A_2)\cdot P(A_2) = \frac{1}{2}\left[P(E_2/A_1) + P(E_2/A_2)\right]$$

$$P(E_2/A_1) = P(X=2)\ (\text{con la moneda } H_1) = \binom{3}{2} 0'5^2\, 0'5^1 = 0'375 = b(3,2,0'5)$$

$$P(E_2/A_2) = P(X=2)\ (\text{con la moneda } H_2) = b(3,2,0'3) = 0'189$$

luego $P(E_2) = \frac{1}{2}(0'375 + 0'189) = 0'282$

Análogamente:

$$P(E_3/A_1) = P(X=3)\ (\text{con la moneda } H_1) = \binom{3}{3} 0'5^3\, 0'5^0 = (b\ 3,3,0'5) = 0'125$$

$$P(E_3/A_2) = P(X=3)\ (\text{con la moneda } H_2) = b(3,3,0'3) = 0'027$$

luego $P(E_3) = \frac{1}{2}(0'125 + 0'027) = 0'076$

$$P(E) = P(E_2 \cup E_3) = P(E_2) + P(E_3) = 0'282 + 0'076 = \boxed{0'358}$$

(a.3) $$P(A_2/2\text{ caras}) = \frac{P(2\text{ caras}/A_2)\cdot P(A_2)}{P(2\text{ caras})} = \frac{P(E_2/A_2)\cdot P(A_2)}{P(E_2)} =$$

$$= \frac{0'189\cdot 1/2}{0'282} = \boxed{0'335}$$

(b) Cualquiera que sea la moneda seleccionada, al lanzarla 40 veces, la variable aleatoria X: "nº de éxitos (caras) en los 40 lanzamientos" tiene distribución $B(40,p)$ siendo $p = 0'5$ ó $0'3$, según la moneda elegida.

Como $np = \left\{\begin{array}{l} 40 \cdot 0'5 = 20 \\ 40\cdot 0'3 = 12\end{array}\right\} > 5$ para las dos monedas

$p = \left\{\begin{array}{l} 0'5 \\ 0'3 \end{array}\right\} \leq 0'5$ para las dos monedas

$B(n,p) \simeq N(np, \sqrt{npq})$ $B(40,p) \simeq N(40p, \sqrt{40pq})$

Nos interesa llamar "éxito" a sacar cara y no a sacar cruz, pues así mantenemos $p \leq 0'5$ en todas las monedas

(b.1) $P(\text{menos de 30 cruces}) = P(\text{al menos 10 caras}) = P_{Bin}(X \geq 10) =$
$= P_{Normal}(X \geq 9'5)$ (aplicamos la corrección de continuidad)

Sea F: "sacar menos de 30 cruces en los 40 lanzamientos de la primera moneda"

$$P(F) = P(F/A_1)\cdot P(A_1) + P(F/A_2)\cdot P(A_2) = \frac{1}{2}\left[P(F/A_1) + P(F/A_2)\right]$$

31'2000

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: **7-6-2000 (Tarde)**

APELLIDOS Y NOMBRE:______________________________

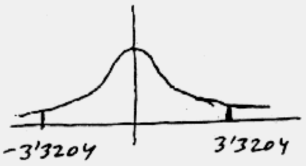
Nº DE EXPEDIENTE:_______________ CURSO: 3º GRUPO:_______________

PROBLEMA 1: Continuación B

$P(F/A_1) = P(X \geq 9'5)$, con $X : N(40 \cdot 0'5, \sqrt{40 \cdot 0'5 \cdot 0'5})$

$P(X \geq 9'5) = P\left(\frac{X-20}{\sqrt{10}} \geq \frac{9'5-20}{\sqrt{10}}\right) = P(Z \geq -3'3204) =$

$= 1 - P(Z < -3'3204) = 1 - P(Z > 3'3204) =$

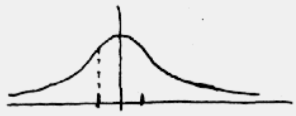


$= 1 - 0'000483 = 0'999517$

(No interpolamos)

$P(F/A_2) = P(X \geq 9'5)$, con $X : N(40 \cdot 0'3, \sqrt{40 \cdot 0'3 \cdot 0'7})$

$P(X \geq 9'5) = P\left(\frac{X-12}{\sqrt{8'4}} \geq \frac{9'5-12}{\sqrt{8'4}}\right) = P(Z \geq -0'8626) =$



$= 1 - P(Z < -0'8626) = 1 - P(Z \geq 0'8626) =$

$= 1 - 0'1949 = 0'8051$

luego, $P(F) = \frac{1}{2}(0'999517 + 0'8051) = \boxed{0'9023}$

(b.2) Los lanzamientos con la segunda moneda son independientes de los de la primera moneda.

$P(\text{entre } 15 \text{ y } 30 \text{ caras}) = P_{Bin}(15 \leq X \leq 30) = P_{Normal}(14'5 \leq X \leq 30'5)$

Sea G : "sacar entre 15 y 30 caras en los 40 lanzamientos de la segunda moneda"

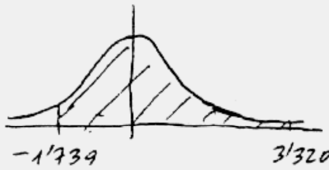
$P(G) = P(G/A_1) \cdot P(A_1) + P(G/A_2)\, P(A_2) = \frac{1}{2}\left[P(G/A_1) + P(G/A_2)\right]$

$P(G/A_1) = P(14'5 \leq X \leq 30'5)$ con $X : N(20, \sqrt{10})$

$P(14'5 \leq X \leq 30'5) = P\left(\frac{14'5-20}{\sqrt{10}} \leq Z \leq \frac{30'5-20}{\sqrt{10}}\right) =$

$= P(-1'739 \leq Z \leq 3'3204) = 1 - P(Z \geq 1'739) - P(Z \geq 3'3204) =$

$= 1 - 0'0409 - 0'000483 = 0'9586$



30'2000

APELLIDOS Y NOMBRE: —
Nº DE EXPEDIENTE: CURSO: 3º GRUPO:

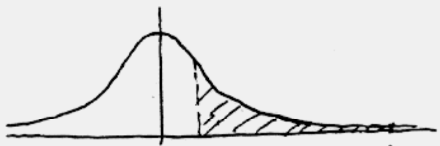
PROBLEMA 1: Continuación C

$$P(G/A_2) = P(14'5 \le X \le 30'5) \,, \quad \text{con } X: N(12, \sqrt{8'4})$$

$$\overset{\shortparallel}{} \; P\left(\frac{14'5-12}{\sqrt{8'4}} \le Z \le \frac{30'5-12}{\sqrt{8'4}}\right) = P(0'8626 \le Z \le 6'383) =$$

$$= P(Z > 0'8626) - P(Z > 6'383) =$$

$$= 0'1949 - 0 = 0'1949$$



luego $P(G) = \frac{1}{2}(0'9586 + 0'1949) = \boxed{0'5768}$

(b.3) $P(\text{más caras que cruces}) = P_{Bin}(X \geq 21) = P_{Normal}(X \geq 20'5)$

Sea H: "sacar más caras que cruces en los 40 lanzamientos de la primera moneda" = "sacar al menos 21 caras, ..."

$$P(H) = P(H/A_1)\cdot P(A_1) + P(H/A_2)\cdot P(A_2) = \frac{1}{2}\left[P(H/A_1) + P(H/A_2)\right]$$

$$P(H/A_1) = P(X \geq 20'5) \,, \quad \text{con } X: N(20, \sqrt{10})$$

$$\overset{\shortparallel}{} \; P\left(Z \geq \frac{20'5-20}{\sqrt{10}}\right) = P(Z \geq 0'1581) = 0'4364$$

$$P(H/A_2) = P(X \geq 20'5) \,, \quad \text{con } X: N(12, \sqrt{8'4})$$

$$\overset{\shortparallel}{} \; P\left(Z \geq \frac{20'5-12}{\sqrt{8'4}}\right) = P(Z \geq 2'9328) = 0'00169$$

luego $P(H) = \frac{1}{2}(0'4364 + 0'00169) = 0'219$

Con la segunda moneda la probabilidad sería la misma y es independiente de la primera moneda, luego

$P(\text{más caras que cruces con las dos monedas}) = 0'219^2 = \boxed{0'048}$

29'2000

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 7-6-2000 (Tarde)

APELLIDOS Y NOMBRE: ____________________

Nº DE EXPEDIENTE: ____________ CURSO: 3º GRUPO: ____________

PROBLEMA 2 *(2 puntos)*

Las especificaciones sobre longitud de un componente en una cadena de montaje, siguen una función de densidad de probabilidad:

$$f(x)=\frac{3x^2}{8} \qquad 0\le x\le 2$$

dada en metros.

a) Se quiere verificar que dicha cadena es de baja calidad en la fabricación del componente. Para ello se establece un control de calidad, de tal forma que si de 6 unidades de dicho componente seleccionadas al azar, al menos 5 tienen al menos 1,7 metros, se considera superado dicho control. Calcular la probabilidad de pasar dicho control de calidad. **1p**

b) Sobre los componentes de longitud superior a un metro, ¿qué probabilidad hay de seleccionar al azar uno con longitud superior a 1,8 metros? **1p**

NOTA: Si tuviera que usar tablas, coja el valor aproximado más cercano: no interpole.

SOLUCIÓN:

a) $$P(X\ge 1.7)=\int_{1'7}^{2}\frac{3x^2}{8}dx=\frac{3}{8}\left[\frac{x^3}{3}\right]_{1.7}^{2}=1-0.614=0.385$$

$$P(\text{"pasar el control"})=\binom{6}{5}(0.385)^5(0.615)^1+\binom{6}{6}(0.385)^6(0.615)^0=0.0369+0.0041=0.041$$

b) La probabilidad que se pide es:

$$P\left(X\ge 1.8 \big/ X\ge 1\right)=\frac{P[(X\ge 1.8)\cap(X\ge 1)]}{P(X\ge 1)}=\frac{P(X\ge 1.8)}{P(X\ge 1)}$$

Teniendo en cuenta que:

$$P(X\ge 1.8)=\int_{1'8}^{2}\frac{3X^2}{8}dx=\frac{3}{8}\left[\frac{x^3}{3}\right]_{1.8}^{2}=1-0.729=0.271$$

$$P(X\ge 1)=\int_{1}^{2}\frac{3x^2}{8}dx=\frac{3}{8}\left[\frac{x^3}{3}\right]_{1}^{2}=\frac{7}{8}$$

Luego:

$$P\left(X\ge 1.8 \big/ X\ge 1\right)=\frac{0.271}{\frac{7}{8}}=0.309$$

28' 2000

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 7-6-2000 (Tarde)

APELLIDOS Y NOMBRE:________________________________

Nº DE EXPEDIENTE:________________ CURSO: 3º _ GRUPO:________

PROBLEMA 3 *(3 puntos)*

Un determinado sistema consta de un componente mecánico y otro eléctrico. Antes de fallar, el componente mecánico tiene una vida útil, X, en años, caracterizada por su función de densidad:

$$f(x)=\begin{cases} 2xe^{-x^2} & \forall x>0 \\ 0 & \forall x\leq 0 \end{cases}$$

y a su vez el componente eléctrico tiene una vida útil, Y, también en años, caracterizada por su función de densidad:

$$f(y)=\begin{cases} \left(\frac{1}{3}\right)e^{-\frac{y}{3}} & \forall y>0 \\ 0 & \forall y\leq 0 \end{cases}$$

El sistema fallará tan pronto como falle uno de los dos componentes y se supone que estos pueden fallar independientemente uno del otro.

a) Compruebe si $f(x)$ y $f(y)$ son funciones de densidad o no. Si no lo fueran, hágales las correcciones necesarias para que lo sean. **0'5p**

b) Calcule la probabilidad de que el sistema falle antes de 2 años **1p**

c) Probabilidad de que el componente mecánico falle antes de 1 año, pero que el eléctrico falle entre el segundo y el tercer año. **1p**

d) ¿Cuándo es de esperar que falle el componente eléctrico? **0'5p**

SOLUCIÓN:

a) $$\int_{-\infty}^{+\infty} f(x)\,dx=\int_0^{\infty} 2xe^{-x^2}dx=\left[-e^{-x^2}\right]_0^{\infty}=1 \qquad \text{SI}$$

$$\int_{-\infty}^{+\infty} f(y)\,dy=\int_0^{\infty} \frac{1}{3}e^{-y/3}dy=\left[-e^{-y/3}\right]_0^{\infty}=1 \qquad \text{SI}$$

Los dos son funciones de densidad

27'2000

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

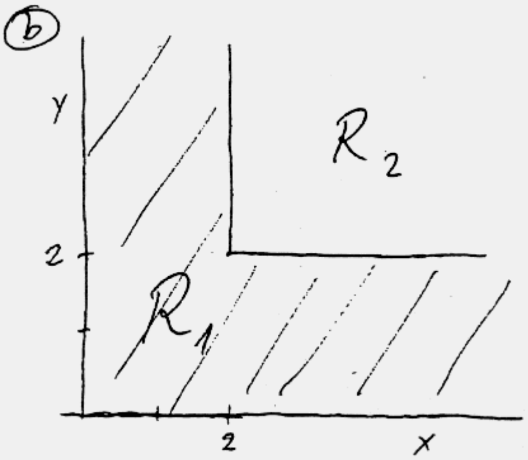
ESTADÍSTICA: **7-6-2000 (Tarde)**

APELLIDOS Y NOMBRE: ____________

Nº DE EXPEDIENTE: ____________ CURSO: 3º GRUPO: ____________

PROBLEMA 3: Continuación A

(b)



El sistema fallará antes de 2 años si la vida útil de uno al menos de los componentes es inferior a dos años la probabilidad de esto es la del area rayada, R_1

$$P(R_1) = 1 - P(R_2)$$

Dado que las vidas útiles de ambos componentes (X, Y) son independientes, podemos calcular la función de densidad conjunta como producto de las marginales

$$f(x,y) = \begin{cases} \frac{2}{3} x e^{-x^2} e^{-y/3} & x>0, \; y>0 \\ 0 & \text{resto} \end{cases}$$

$$P(R_2) = \int_2^\infty dx \int_2^\infty \frac{2}{3} x e^{-x^2} e^{-y/3} dy = \int_2^\infty 2x e^{-x^2} dx \left[-e^{-y/3}\right]_2^\infty =$$

$$= e^{-2/3} \left[-e^{-x^2}\right]_2^\infty = e^{-2/3} e^{-4} = e^{-\frac{14}{3}}$$

luego $P(\text{fallar antes de 2 años}) = 1 - e^{-\frac{14}{3}}$

También: la probabilidad de que el sistema ~~fa~~ no falle antes de 2 años es la de que no falle ninguno de los dos componentes y por la independencia de estos será el producto de las probabilidades de que no falle cada componente, con lo que podemos trabajar con las f.d. marginales, sin

26'2000

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 7-6-2000 (Tarde)

APELLIDOS Y NOMBRE:

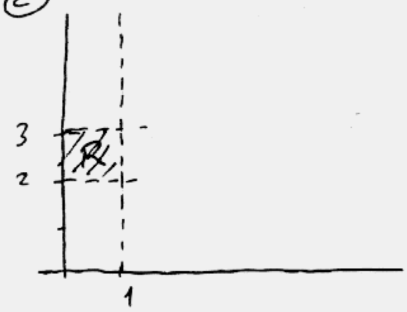
Nº DE EXPEDIENTE: CURSO: 3º GRUPO:

PROBLEMA 3: Continuación B

considerar la conjunta.

$P(\text{fallar antes de 2 años}) = 1 - P(\text{no fallar antes de 2 años}) =$
$= 1 - P(X>2, Y>2) = 1 - P(X>2) \cdot P(Y>2) \ldots$

(c)



Lo pedido es la probabilidad de la región rayada

$$P(R) = \int_0^1 2x e^{-x^2} dx \int_2^3 \frac{1}{3} e^{-y/3} dy$$

como se ve, esto equivale también a:

$$P(X<1, 2<Y<3) = P(X<1) \cdot P(2<Y<3)$$

y no necesitamos trabajar con la f.d. conjunta

$$P(X<1) = \int_0^1 2x e^{-x^2} dx = \left[-e^{-x^2}\right]_0^1 = 1 - e^{-1}$$

$$P(2<Y<3) = \int_2^3 \frac{1}{3} e^{-y/3} dy = \left[-e^{-y/3}\right]_2^3 = e^{-\frac{2}{3}} - e^{-1}$$

luego $P(X<1, 2<Y<3) = (1-e^{-1})(e^{-2/3} - e^{-1})$

(d) La f.d. de Y nos dice que Y tiene distribución exponencial de parámetro $\lambda = \frac{1}{3}$, luego

$$E[Y] = \frac{1}{\lambda} = \boxed{3 \text{ años}}$$

25 '2000

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 7-6-2000 (Tarde)

APELLIDOS Y NOMBRE:

Nº DE EXPEDIENTE: CURSO: 3º GRUPO:

PROBLEMA 4 *(2'5 puntos)*

Un cañón lleva funcionando ya muchos años. La distribución de impactos de este cañón es normal. Se quiere comprobar si su distancia máxima esperada sigue siendo la de fábrica, que era de 4.000m. Si no fuera así, habría disminuido, y se mandaría realizar reajustes en el cañón. Para ello, se disparan 10 obuses, con el cañón en posición de alcance de la máxima distancia. Se pretende que la probabilidad de enviarlo a reparar, estando bien el cañón, sea del 5%.

a) Se disparan 10 obuses. Siendo x_i la distancia alcanzada por cada uno de los diez obuses, en metros, las mediciones obtenidas son:

$$\sum_{i=1}^{10} x_i = 38.880 \qquad\qquad \sum_{i=1}^{10}(x_i - \bar{x})^2 = 32.400$$

¿Qué debe hacerse con ese cañón, en base a dichas medidas? **1p**

b) ¿Si la distancia máxima esperada ha disminuido y realmente es de 3.880m, qué probabilidad hay de concluir, sin embargo, que el cañón está perfectamente?. **1p**

c) Evalúe los resultados obtenidos. ¿Le parece correcta la prueba realizada?, ¿o le parece necesario aumentar el número de disparos? **0'5p**

SOLUCIÓN:

$$\bar{x} = \frac{\sum_{i=1}^{10} x_i}{n} = \frac{38.880}{10} = 3.888m \qquad\qquad s^2 = \frac{\sum_{i=1}^{10}(x_i - \bar{x})^2}{n-1} = \frac{32.400}{9} = 3.600m^2 \Rightarrow s = 60m$$

a)

1 - El parámetro que nos interesa es la media μ

2 - $H_0 : \mu = \mu_0 = 4000m$, $(\text{ó } \mu_0 \geq 4000m)$

3 - $H_a : \mu < 4000m$ Contraste unilateral

4 - $\alpha = 0'05$

5 - La población de impactos es normal. No conocemos la desviación típica de la distribución de los impactos, y tampoco es superior a 30 el tamaño de la muestra (n = 10) por lo que el estadístico $\dfrac{\bar{x}-\mu}{s/\sqrt{n}}$ seguirá una distribución T de Student t_{n-1} en el muestreo.

Si se cumple la hipótesis $H_0 : \mu = \mu_0 = 4000m$, tendremos $t_{n-1} = \dfrac{\bar{x}-\mu_0}{s_R/\sqrt{n}}$, con $t_{n-1} = t_9$. Este es el estadístico de contraste que vamos a utilizar.

6 - Se trata de un contraste unilateral, para detectar que la media no quede por debajo del valor de H_0: Rechazaremos que se cumpla la hipótesis ($H_0 : \mu = \mu_0 = 4.000m$), si $t_R < -t_{\alpha;9}$; aceptándola en caso contrario. $\alpha = 0'05 \quad \Rightarrow \quad t_C = -t_{\alpha,n-1} = -t_{0'05;9} = -1'833$

$$t_C = \frac{PC - \mu_0}{s_R/\sqrt{n}} \quad \Rightarrow \quad PC = \mu_0 + t_C\frac{s_R}{\sqrt{n}} = \mu_0 - t_{\alpha,n-1}\frac{s_R}{\sqrt{n}} = 4000 - 1'833\frac{60}{\sqrt{10}} = 3.965'22127$$

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 7-6-2000 (Tarde)

APELLIDOS Y NOMBRE:____________________

Nº DE EXPEDIENTE:__________ CURSO: 3º GRUPO:__________

7 - Cálculo del valor del estadístico: $t_R = \dfrac{\bar{x}_R - \mu_0}{s_R/\sqrt{n}} = \dfrac{3.888 - 4.000}{60/\sqrt{10}} = -5'9029183$

8 - Conclusiones estadísticas:

$t_R = -5'9 < t_C = -1'833$, luego cae fuera de la región de aceptación.

Así, se rechaza la hipótesis con un nivel de significación 0'05. (la probabilidad de que la estemos rechazando, siendo cierta la hipótesis, es del 5%)

9 - Conclusiones no estadísticas:

Partiendo de una muestra de 10 disparos, puede concluirse, con probabilidad de error inferior al 5%, que el alcance máximo esperado es inferior a los 4000m.

El cañón debe llevarse a reajustar.

Nota: Inferior al 5%, porque sería del 5% si resultara que H_0 fuese cierta, cosa que no sabemos si ocurre. Por el Th de Probabilidad Total: $P(E) = P(E|H_0)\cdot P(H_0) + P(E|H_1)\cdot P(H_1)$

Como rechazamos H_0, no cometemos error si es cierto H_1, luego $P(E|H_1) = 0$.

Así, rechazando H_0, tenemos $P(E) = P(E|H_0)\cdot P(H_0) \le P(E|H_0) = 5\%$ *Luego* $P(E) \le 5\%$

b) $\beta(\mu = \mu_a) = P(Aceptar H_0 | \mu = \mu_a) = P(\bar{x} \le PC | \mu = \mu_a) \underset{t_{n-1} = \frac{\bar{x}-\mu}{s/\sqrt{n}}}{=} P\left(t_{n-1} \le \dfrac{PC - \mu_a}{s/\sqrt{n}} \right)$

Es un contraste unilateral, con $H_0 : \mu = \mu_0 = 4000m$; $H_1 : \mu = 3.880$, y teníamos, de a), que

$$PC = \mu_0 - t_{\alpha,n-1}\frac{s_R}{\sqrt{n}} \Rightarrow \beta(\mu_a) \underset{t_{n-1} = \frac{\bar{x}-\mu}{s/\sqrt{n}}}{=} P\left(t_{n-1} \le \frac{\mu_0 - t_{\alpha,n-1}\frac{s_R}{\sqrt{n}} - \mu_a}{s/\sqrt{n}} \right) = P\left(t_{n-1} \le \frac{\mu_0 - \mu_a}{s/\sqrt{n}} - t_{\alpha,n-1} \right)$$

$$\beta(\mu = 3.880) = P\left(t_9 \le \frac{4.000 - 3.880}{60/\sqrt{10}} - 1'833 \right) = P(t_9 \ge 4'49155532) = \alpha(4'49155532)$$

Buscando sobre la tabla de la T-Student t_9, podemos apreciar que es inferior a 0'0005.

Por tanto: $\beta(H_1 : \mu = 3.880) = \alpha(4'49155532) < 0'0005 = 0'05\%$

c) La potencia del contraste es muy alta, del 99'95%, con una probabilidad de error tipo I del 5% (cosa que por ser dato, se supone acorde a las necesidades del usuario). Por tanto, no tiene sentido lanzar más obuses.

Adicionalmente puede comentarse que esta potencia es tan alta a pesar de los pocos obuses lanzados, porque se pretende detectar si la media es 4.000m, siendo realmente 3.880m. Es decir, porque el valor real está muy lejos del valor de prueba: más de seis veces la desviación estimada (4.000-3.880 > $6\cdot(\sigma/\sqrt{n})$, pues $\sigma = 60m$, y $n = 10$). Al estar tan alejados el uno del otro es fácil tener una potencia de contraste tan elevada.

23'2000

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 7-6-2000 (Mañana)

APELLIDOS Y NOMBRE:______________________

Nº DE EXPEDIENTE:______________ CURSO: 3º GRUPO:______________

PROBLEMA 1 *(2 puntos)*

Una empresa industrial tiene 3 plantas de producción de un determinado artículo, con funcionamiento independiente. La empresa tiene información previa que indica que en la primera planta (A), por cada 500 unidades de producción, 30 tienen defecto. Respecto a la segunda planta (B), se sabe que el 90% de las unidades salen sin defecto. Hay una probabilidad de 0,2 de que salga con defecto un artículo seleccionado al azar, si lo escogemos entre los de la tercera planta (C).

Se seleccionan al azar 3 unidades, siendo cada una de ellas de una planta diferente.

a) Calcule la probabilidad de que sólo una de esas 3 unidades salga defectuosa. **0'6p**

b) Calcule la probabilidad de que al menos 2 sean defectuosas. **0'6p**

c) Se ha comprobado que al menos hay una sin defecto. ¿Cuál será la probabilidad de que la de la tercera planta haya salido sin defecto? **0'8p**

SOLUCIÓN:

Probabilidades, dentro de cada planta, de que una unidad no tenga defecto :

$$P(A)=\frac{470}{500}=0.94 \qquad P(B)=0.9 \qquad P(C)=0.8$$

Probabilidades de que tenga defecto :

$$P(A^*)=0.06 \qquad P(B^*)=0.1 \qquad P(C^*)=0.2$$

a) Es la suma de 3 probabilidades :

$$P(A^*\cap B\cap C)+P(A\cap B^*\cap C)+P(A\cap B\cap C^*)=$$
$$P(A^*)P(B)P(C)+P(A)P(B^*)P(C)+P(A)P(B)P(C^*)=$$
$$0.06x0.9x0.8+0.94x0.1x0.8+0.94x0.9x0.2=0.2876$$

b) Al menos 2, es decir, los 3 casos en los que salen 2 defectuosas y el caso de las 3 defectuosas :

$$P(A^*\cap B^*\cap C)+P(A^*\cap B\cap C^*)+P(A\cap B^*\cap C^*)+P(A^*\cap B^*\cap C^*)=$$
$$P(A^*)P(B^*)P(C)+P(A^*)P(B)P(C^*)+P(A)P(B^*)P(C^*)+P(A^*)P(B^*)P(C^*)=$$
$$0.06x0.1x0.8+0.06x0.9x0.2+0.94x0.1x0.2+0.06x0.1x0.2=0.0356$$

c) $$P\left(C/\text{"al menos una sin defecto"}\right)=\frac{P\left[C\cap(\text{"al menos una sin defecto"})\right]}{P(\text{"al menos una sin defecto"})}=$$

$$\frac{P(A^*\cap B^*\cap C)+P(A^*\cap B\cap C)+P(A\cap B^*\cap C)+P(A\cap B\cap C)}{1-P(A^*\cap B^*\cap C^*)}=\frac{0.8}{0.9988}=0.8009$$

22' 2000

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 7-6-2000 (Mañana)

APELLIDOS Y NOMBRE:____________________

Nº DE EXPEDIENTE:__________ CURSO: 3º GRUPO:______

PROBLEMA 2 *(2'5 puntos)*

Se tiene la función:

$$f(x\,y)=\begin{cases} kxy & en \quad R_{xy} \\ 0 & en \quad resto \end{cases}$$

donde R_{xy} es la región delimitada por $x^2+y^2\leq 1$, y por $y\geq\dfrac{x}{2}$, con $x\geq 0$ e $y\geq 0$.

a) Calcule el valor de k para que dicha función sea de densidad de probabilidad conjunta. **0'5p**

b) Obtenga la función de densidad de probabilidad marginal de la variable Y. **0'8p**

c) Curva general de regresión de x sobre y, indicando además, con todo detalle, los valores (expresiones) que toma cada función que necesite calcular, para todo su dominio de existencia. **1'2p**

NOTA: *No es necesario sustituir el valor de k en los apartados b) y c).*

SOLUCIÓN:

a)

$$\int_0^{2/\sqrt{5}}\int_{\frac{x}{2}}^{\sqrt{1-x^2}} kxy\,dy\,dx=\int_0^{2/\sqrt{5}} kx\left[\frac{y^2}{2}\right]_{\frac{x}{2}}^{\sqrt{1-x^2}}=k\int_0^{2/\sqrt{5}} x\left[\frac{\left(\sqrt{1-x^2}\right)^2}{2}-\frac{\left(x/2\right)^2}{2}\right]dx=$$

$$=\frac{k}{2}\int_0^{2/\sqrt{5}}\left(x-\frac{5}{4}x^3\right)dx=\frac{k}{2}\left[\frac{x^2}{2}-\frac{5}{16}x^4\right]_0^{2/\sqrt{5}}=\frac{k}{10}=1 \qquad \Rightarrow \quad k=10$$

$b)$ Tiene dos partes, una en la que la "x" va de 0 a la recta, y otra en la que va de 0 a la curva. Hay que encontrar la "y" que delimita una y otra parte. El punto de corte de la recta y la curva es :

$$\sqrt{1-x^2}=\frac{x}{2} \quad \Rightarrow \quad x=\frac{2}{\sqrt{5}} \quad \Rightarrow \quad y=\frac{x}{2}=\frac{2/\sqrt{5}}{2}=\frac{1}{\sqrt{5}}$$

Así, la f.d.p marginal de Y queda:

$$0\leq y\leq\frac{1}{\sqrt{5}} \qquad f(y)=\int_0^{2y} kxy\,dx=yk\left[\frac{x^2}{2}\right]_0^{2y}=2ky^3$$

$$\frac{1}{\sqrt{5}}<y\leq 1 \qquad f(y)=\int_0^{\sqrt{1-y^2}} kxy\,dx=ky\left[\frac{x^2}{2}\right]_0^{\sqrt{1-y^2}}=\frac{ky}{2}\left(1-y^2\right)$$

$$\forall\, y>1, \forall\, y<0 \quad f(y)=0$$

21'2000

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 7-6-2000 (Mañana)

APELLIDOS Y NOMBRE:__________

Nº DE EXPEDIENTE:__________ CURSO: 3º GRUPO:__________

c) $f\left(x/y \right) = \dfrac{f(xy)}{f(y)} \qquad \forall f(y) \neq 0$

La f.d.p. condicionada queda:

En R_{xy} con $y \leq \dfrac{1}{\sqrt{5}}$ $\qquad f\left(x/y \right) = \dfrac{kxy}{2ky^3} = \dfrac{x}{2y^2}$

En R_{xy} con $y > \dfrac{1}{\sqrt{5}}$ $\qquad f\left(x/y \right) = \dfrac{kxy}{\dfrac{ky}{2}\left(1 - y^2\right)} = \dfrac{2x}{1 - y^2}$

No existe con $y < 0$ e $y > 1$; $\qquad f\left(x/y \right) = 0$ en el resto

Respecto a la curva general de regresión, queda:

$$0 \leq y \leq \frac{1}{\sqrt{5}} \qquad E\left[x/y \right] = \int_0^{2y} x \frac{x}{2y^2} dx = \frac{1}{2y^2}\left[\frac{x^3}{3} \right]_0^{2y} = \frac{4}{3} y$$

$$\frac{1}{\sqrt{5}} < y \leq 1 \qquad E\left[x/y \right] = \int_0^{\sqrt{1-y^2}} x \frac{2x}{1-y^2} dx = \frac{2}{1-y^2}\left[\frac{x^3}{3} \right]_0^{\sqrt{1-y^2}} = \frac{2}{3}\sqrt{1 - y^2}$$

No existe con $y < 0$ e $y > 1$

20'2000

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 7-6-2000 (Mañana)

APELLIDOS Y NOMBRE:____________________

Nº DE EXPEDIENTE:____________ CURSO: 3º GRUPO:__________

PROBLEMA 3 *(3 puntos)*

Se ha realizado un examen en forma de test teórico con un gran número de preguntas. Las puntuaciones finales de dicho test se dan sobre 50 puntos (enteros del 0 al 50). Las puntuaciones finales relacionadas en actas se redondean al entero más próximo a partir de las puntuaciones reales, esto es, obtienen un 12, por ejemplo, los alumnos cuyo cómputo oscile entre 11'5 y 12'5. Se exige un mínimo (redondeado) de 25 para aprobar.

Si suponemos que las puntuaciones reales (sin redondeo) siguen aproximadamente una ley normal de media 30 y varianza 25.

a) Qué máxima puntuación real (no redondeada) delimitará el 20% de notas más bajas de los alumnos presentados. **0'6p**

b) Qué mínima puntuación real (no redondeada) delimitará el 20% de notas más altas de los alumnos presentados. **0'6p**

c) Qué porcentaje de alumnos aparecerán en las actas de notas finales redondeadas, exactamente, con 25 puntos (aprobado mínimo). **0'6p**

d) Cuál será, siempre con las reglas de redondeo dadas, el porcentaje de suspensos en actas. **0'6p**

e) Se toma una muestra aleatoria de 25 alumnos de la población y se calcula su media (sin redondearla). Existe un valor, al que llamaremos N_e, tal que la probabilidad de que la media de esta muestra lo supere es de 0'2. Calcule dicho valor N_e. **0'6p**

Nota: *No es necesario que interpole en tablas.*

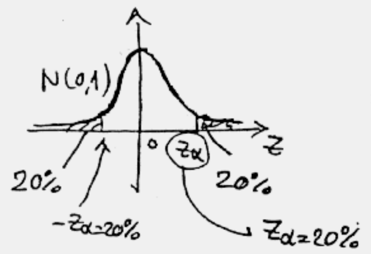
SOLUCIÓN:

a) $X : N(30, 5) \Rightarrow Z = \dfrac{X-\mu_X}{\sigma_X} : N(0,1)$

Buscamos el valor de la nota $x_{máx}$, bajo la cual haya exactamente un 20% de alumnos.

Es decir: $$0'2 = P(X \le x_{máx}) = P(X \le x_{0'8}) \underset{Z=\frac{X-\mu_X}{\sigma_X}}{=} P\left(Z \le \frac{x_{0'8}-\mu_X}{\sigma_X}\right) = P(Z \le z_{0'8})$$

De la figura: $z_{0'8} = -z_{0'2}$.



Así, $$z_{0'8} = -z_{0'2} = \frac{x_{máx}-\mu_X}{\sigma_X}$$ despejando: $$x_{máx} = x_{0'8} = \mu_X - z_{0'2}\cdot\sigma_X$$

Interpolando obtenemos: $\dfrac{z_{0'2}-0'84}{0'85-0'84}=\dfrac{0'2-0'2005}{0'1977-0'2005} \Rightarrow z_{0'2}\cong 0'841786$

La nota máxima (no redondeada) es:

$$\boxed{x_{0'8}=}\ \mu_X - z_{0'2}\cdot\sigma_X \cong 30-0'841786\cdot 5\ \boxed{=25'79107}$$

b) La nota (no redondeada) que deja por encima suyo al 20% de las notas es, $x_{0'2}$, que, por simetría de la distribución:

$$\boxed{x_{0'2}=}\ \mu_X + z_{0'2}\cdot\sigma_X \cong 30+0'841786\cdot 5\ \boxed{=34'20893}$$

c) $P(24'5\le x\le 25'5) \overset{Z=\frac{X-\mu}{\sigma}}{=} P\left(\dfrac{24'5-30}{5}\le z\le\dfrac{25'5-30}{5}\right)=P(-1'1\le z\le -0'9) \overset{\substack{\text{simetría}\\ N(0,1)}}{=} P(0'9\le z\le 1'1)=$

$$=P(z\ge 0'9)-P(z\ge 1'1)=0'1841-0'1357=0'0484\ \boxed{=4'84\%}$$

d) La nota redondeada de aprobado es 25, luego toda nota real inferior a 24'5 será suspenso.

$$P(x\le 24'5)=P(z\le -1'1)=P(z\ge 1'1)=0'1357\ \boxed{=13'57\%}$$

e) Del enunciado, $N_e=\bar{x}_{0'2}$, es decir, $P(X_{25}>N_e)=0'2=P(X_{25}>\bar{x}_{0'2})$

La distribución de la nota media de 25 alumnos será $\bar{X}: N\left(\mu_X, \dfrac{\sigma_X}{\sqrt{n}}\right)=N(30,1)$

$\Rightarrow$ Tal y como se ha trabajado en los apartados a) y b), tendremos:

Es decir: $0'2=P(X>N_e)=P(X>\bar{x}_{0'2}) \underset{Z=\frac{X-\mu_X}{\sigma_X/\sqrt{n}}}{=} P\left(Z>\dfrac{N_e-\mu_X}{\sigma_X/\sqrt{n}}\right)=P(Z>z_{0'2})$

Luego: $\boxed{N_e=}\ \bar{x}_{0'2}=\mu_X+z_{0'2}\dfrac{\sigma_X}{\sqrt{n}}=30+0'841786\ \boxed{=30'841786}$

18'2000

APELLIDOS Y NOMBRE:______________________

Nº DE EXPEDIENTE:______________ CURSO: 3º GRUPO:__________

PROBLEMA 4 *(2'5 puntos)*

Un fabricante de componentes electrónicos necesita que las resistencias que fabrica tengan, en promedio, un valor de 100Ω. Tiene dudas sobre si lo está cumpliendo actualmente o no. Caso de no cumplirlo, deberá paralizar la fábrica por un día.

Por ello, habla con el Departamento de Control de Calidad de la empresa, quienes deciden tomar una probabilidad para error de tipo I de 0'05, tomando una muestra aleatoria de 1000 resistencias, sobre la que obtienen un valor medio de 100'3Ω.

Se sabe que los valores de sus resistencias tienen una desviación típica de 3Ω.

a) ¿A qué conclusiones llegaría Vd. si estuviera encargado de dicho control de calidad?. **1p**

b) Indique entre sus conclusiones el mayor y el menor de los valores de la media de la muestra que delimiten aceptación o rechazo. **0'3p**

c) Calcule la potencia del contraste realizado si la resistencia media fuera realmente 100'3Ω. **0'7p**

d) Para una probabilidad de error tipo I del 5%, y suponiendo que en la muestra se obtenga un valor medio de 100'3Ω, calcule el valor límite del tamaño de la muestra para el cual se pasará de continuar la fabricación, a la paralización de ésta. Indique, además, si se paralizaría para mayores muestras, o para menores que el valor calculado. **0'5p**

e) Si la resistencia media fuera realmente 100'3Ω, obtenga la potencia de un contraste basado en una muestra de 200 resistencias y se mantuviera la probabilidad de error tipo I en el 5%. **NP**

f) Evalúe los resultados obtenidos, observando las mayores o menores potencias; y los rechazos/aceptaciones para mayores o menores muestras. ¿Merece la pena gastar en una prueba de 1000 resistencias, o es mejor ahorrar haciendo una de 200? Razone la respuesta. **NP**

Nota: No es necesario que realice interpolación en las tablas. Utilice el valor más aproximado.

SOLUCIÓN:

a)

1 - El parámetro que nos interesa es la media μ

2 - $H_0 : \mu = \mu_0 = 100\Omega$

3 - $H_1 : \mu \neq 100\Omega$

4 - $\alpha = 0'05$

5 - El tamaño de la muestra es lo suficientemente grande $(1000 > 30)$ como para, por el teorema del límite central, poder aproximar la distribución del estimador $\overline{X}$ en el muestreo por una distribución normal, del tipo $\overline{X} : N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$, por lo que, $Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$ seguirá una distribución normal estándar.

Así, si se cumple la hipótesis $H_0 : \mu = \mu_0 = 100\Omega$, el estadístico $Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$ seguirá una distribución normal estándar $N(0, 1)$.

Por ello, el estadístico de contraste a utilizar será: $Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$

PROBLEMA 4: Continuación A

6 - $\alpha = 0'05 \quad \Rightarrow \quad \frac{\alpha}{2} = 0'025 \quad \Rightarrow \quad z_{CS} = z_{0'025} = 1'96$; $z_{CI} = z_{-0'025} = -1'96$

Aceptaremos que se cumple la hipótesis ($H_0 : \mu = \mu_0 = 100\Omega$), si $-1'96 \le z_R \le 1'96$; rechazaremos la hipótesis si cae fuera del intervalo.

7 - Cálculo del valor del estadístico:

$$z_R = \frac{\bar{x}_R - \mu_0}{\sigma/\sqrt{n}} = \frac{100'3 - 100}{3/\sqrt{1000}} \cong 3'162278$$

8 - Conclusiones estadísticas:

$z_R \cong 3'16 > 1'96$, luego cae fuera de la región de aceptación.

Así, **<u>se rechaza la hipótesis con un nivel de significación 0'05</u>**. (la probabilidad de que la estemos rechazando, siendo cierta la hipótesis, es del 5%)

9 - Conclusiones no estadísticas:

Partiendo de una muestra de 1000 resistencias, puede concluirse, con probabilidad de error no superior al 5%, que la resistencia media de las fabricadas no sea de 100Ω.

Nota: No superior al 5%, porque sería del 5% si resultara que H_0 fuese cierta, cosa que no sabemos si ocurre. Por el Th de Probabilidad Total: $P(E) = P(E|H_0)\cdot P(H_0) + P(E|H_1)\cdot P(H_1)$

Como rechazamos H_0, no cometemos error si es cierto H_1, luego $P(E|H_1) = 0$.

Así, $P(E) = P(E|H_0)\cdot P(H_0) \le P(E|H_0) = 5\%$ *luego* $P(E) \le 5\%$

b) Como se ha visto en a), $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$. La realización de la prueba genera un valor $\bar{x}_R = 100'3$,

con el cual se ha obtenido $z_R = \frac{\bar{x}_R - \mu_0}{\sigma/\sqrt{n}} = \frac{100'3 - 100}{3/\sqrt{1000}} \cong 3'162278$.

Siendo la región de aceptación $-1'96 \le z_R \le 1'96 \Rightarrow \quad -1'96 \le \frac{\bar{x}_R - \mu_0}{\sigma/\sqrt{n}} \le 1'96$

Luego se acepta si $\mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \bar{x}_R \le \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

Es decir: $\bar{x}_{CI} = \mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \qquad \bar{x}_{CS} = \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

quedando $100 - 1'96\frac{3}{\sqrt{1000}} \le \bar{x}_R \le 100 + 1'96\frac{3}{\sqrt{1000}}$

La región de aceptación queda:

$$99'81406 \le \bar{x}_R \le 100'18594 \quad \boxed{\begin{cases} \bar{x}_{CI} = 99'81406 \\ \bar{x}_{CS} = 100'18594 \end{cases}}$$

PROBLEMA 4: Continuación B

c) Potencia del contraste $= 1-\beta$, con $\beta(\mu_a) = P(\text{Aceptar } H_0 | \mu = \mu_a) = P(\bar{x}_{CI} \le \bar{x}_R \le \bar{x}_{CS} | \mu = \mu_a)$

Realizaremos el cálculo de β, por ser relativamente más cómodo. Para ello, podemos sustituir los valores de los puntos críticos, obtenidos en el apartado b), o directamente sus expresiones, lo que será más exacto.

$$\beta(\mu_a) = P\left(\mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \bar{x}_R \le \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \middle| \mu = \mu_a\right)_{Z=\frac{\bar{X}-\mu_a}{\sigma/\sqrt{n}}} =$$

$$= P\left(\frac{\mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} - \mu_a}{\frac{\sigma}{\sqrt{n}}} \le z \le \frac{\mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} - \mu_a}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(\frac{\mu_0-\mu_a}{\frac{\sigma}{\sqrt{n}}} - z_{\alpha/2} \le z \le \frac{\mu_0-\mu_a}{\frac{\sigma}{\sqrt{n}}} + z_{\alpha/2}\right)$$

$$\beta(100'3) = P\left(\frac{100-100'3}{3/\sqrt{1000}} - z_{0'025} \le z \le \frac{100-100'3}{3/\sqrt{1000}} + z_{0'025}\right) =$$

$$= P(-3'1623 - 1'96 \le z \le -3'1623 + 1'96) = P(-5'1223 \le z \le -1'2023) = P(1'2023 \le z \le 5'1223) =$$
$$= \alpha(1'2023) - \alpha(5'1223) \cong [0'1151 + 0'23\cdot(0'1131 - 0'1151)] - 0 \quad = 0'11464$$

Potencia del contraste $= \boxed{1-\beta \cong 1 - 0'11464 = 0'88536 = 88'54\%}$

Notas:

- A pesar de no requerirse la interpolación en tablas, ésta se ha realizado. Su cálculo ha sido de la siguiente manera: $\alpha(1'2023) = \alpha(1'20) + 0'23\cdot[\alpha(1'21) - \alpha(1'20)]$
- Dado que los valores de $\alpha(5'1223)$ son tan bajos, se ha despreciado su cálculo. Los seis primeros dígitos decimales son nulos, y el primer dígito significativo sería el 7º decimal, que despreciaremos al final, por lo que no procede calcular $\alpha(5'1223)$.

d) $z_R = \dfrac{\bar{x}_R - \mu_0}{\sigma/\sqrt{n}} = \dfrac{100'3 - 100}{3/\sqrt{n}}$

Al ser positivo z_R, rechazaríamos la hipótesis H_0, y por tanto, se paralizaría la fábrica, cuando fuera mayor que el límite superior del intervalo. Es decir, cuando $z_R > L_S = 1'96$

Es decir, funcionará si: $z_R = \dfrac{\bar{x}_R - \mu_0}{\sigma/\sqrt{n}} \le L_S \quad \Rightarrow \quad \dfrac{\bar{x}_R - \mu_0}{\sigma}\sqrt{n} \le L_S$

$$\Rightarrow \quad \text{dado que } \bar{x}_R - \mu_0 > 0 \quad \Rightarrow \quad n \le \left(\frac{\sigma\cdot L_S}{\bar{x}_R - \mu_0}\right)^2 = \left(\frac{3\cdot 1'96}{100'3-100}\right)^2 = 384'16$$

Nota: Compruebe que si fuera $\bar{x}_R - \mu_0 < 0$, también se obtendría un valor máximo de n.

PROBLEMA 4: Continuación C

Luego:

- se aceptará H_0 (funcionará la fábrica) si $n \leq 384$,
- se rechazará H_0 (se paralizará la fábrica) si $n > 385$

En el caso expuesto, con n=1000, ya hemos visto que se rechaza.

e) Repitiendo los cálculos del apartado c):

Potencia del contraste = $1-\beta$, con $\beta(\mu_a) = P\left(\frac{\mu_0-\mu_a}{\frac{\sigma}{\sqrt{n}}} - z_{\alpha/2} \leq z \leq \frac{\mu_0-\mu_a}{\frac{\sigma}{\sqrt{n}}} + z_{\alpha/2}\right)$

$$\beta(100'3) = P\left(\frac{100-100'3}{3/\sqrt{200}} - z_{0'025} \leq z \leq \frac{100-100'3}{3/\sqrt{200}} + z_{0'025}\right) = P(-1'4142-1'96 \leq z \leq -1'4142+1'96)$$

$$= P(-3'3742 \leq z \leq 0'5458) = \alpha(-3'3742) - \alpha(0'5458) = [1-\alpha(3'3742)] - \alpha(0'5458) =$$

$$\cong 1 - [0'000483 + (0'000337 - 0'000483)\cdot 0'742] - [0'2946 + (0'2912 - 0'2946)\cdot 0'58] =$$

$$= 1 - 0'0003747 - 0'292628 = 1 - 0'2930027 = 70'7\%$$

Potencia del contraste = $\boxed{1-\beta \cong 0'2930 = 29'30\%}$

f) Conclusiones:

Siempre con una probabilidad de rechazar H_0 siendo cierta, del 5% (error tipo I), y siempre hablando de contrastes para $\mu = 100'3\Omega$:

- Con n=200 se acepta H_0, (ver apdo e),
  Pero la probabilidad de cometer error tipo II (β = prob. aceptar H_0 siendo falso), es muy elevada $\beta = 70'7\%$, por lo que se trata de una mala prueba, de baja potencia (baja prob de rechazar H_0 siendo falso) Pot = 29'30%
- Con n = 1000 se rechaza H_0, (ver apdo d),
  Pero la probabilidad de cometer error tipo II (β = prob. aceptar H_0 siendo falso), es bastante más baja $\beta = 11'46\%$, por lo que se trata de una prueba más potente (más alta prob de rechazar H_0 siendo falso) Pot = 88'54%

Depende del coste, pero parece interesante realizar la prueba de 1000, ya que la de 200 no sería concluyente dada la gran probabilidad de aceptación de la hipótesis H_0, ya sea falso como no ($1-\alpha = 95\%$ si es cierta, y $\beta = 70'7\%$ si es falsa).

Es más, si el presupuesto lo permitiera, quizá fuera interesante aumentar la muestra a más de 1000, para aumentar aún más la potencia del contraste.

14'2000

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 15-2-2000 (Tarde)

APELLIDOS Y NOMBRE:____________________
Nº DE EXPEDIENTE:____________ CURSO: 3º GRUPO:________

PROBLEMA 1 *(2'5 puntos)*

Un sistema transmisor de datos envía palabras, con probabilidad p_T de que llegue con algún error al receptor.
Por su parte, el receptor tiene una probabilidad p_R de detectar un error de transmisión, caso de que éste ocurriera.
El receptor pedirá la repetición de la palabra transmitida si detecta que ha llegado errónea, volviendo a funcionar transmisor y receptor como se ha descrito.

a) Si se realiza el envío de una información con 2.000 palabras, ¿qué probabilidad hay de que tengamos que transmitir exactamente 2.050 (es decir, que se produzcan 50 repeticiones)? **1p**

b) Supóngase el caso de que sólo el 2'5% de las palabras transmitidas lleguen con defecto, pero que el 90% de las que hayan llegado con defecto se detecten en el receptor como erróneas. Si observamos que el envío de un mensaje ha finalizado en 2050 transmisiones de palabras (contando también repeticiones) ¿qué probabilidad hay de que el mensaje original no superara las 2.000 palabras? **1'5p**

SOLUCIÓN:

a) Se define como suceso la petición del enunciado:
A = "Que se finalice el envío de un mensaje de 2000 palabras, con exactamente 2050 transmisiones"

Si se envía un mensaje de 2000 palabras transmitiendo exactamente 2050 palabras, de éstas, 2000 serán recibidas como correctas, y 50 serán detectadas como no correctas (que habrán sido repetidas). Sin embargo, la última ha de ser recibida como correcta, o no habría finalizado el envío con esa transmisión.

Así, en 2049 realizaciones de una transmisión individual, en todas ellas con la misma probabilidad p de que haya error de transmisión y sea detectado en el receptor, se deben dar, en cualquier orden, 50 fallos detectados, y, por tanto, 1999 transmisiones correctas (o sin fallo detectado). Además, la 2050-ésima debe ser recibida como correcta.

El valor de la probabilidad de generar detección de error será $p = p_T \cdot p_R$

Esto nos posibilita la definición de una v.a. binomial X = "número de repeticiones en 2049 transmisiones"

Así, $P(A) = P(X = 50) \cdot (1-p)$

$$P(X=50) = \binom{2049}{50} \cdot p^{50} \cdot (1-p)^{1999} = \binom{2049}{50} \cdot (p_T \cdot p_R)^{50} \cdot (1 - p_T \cdot p_R)^{1999}$$

Luego $P(A) = P(X=50) \cdot (1-p) = \binom{2049}{50} \cdot (p_T \cdot p_R)^{50} \cdot (1 - p_T \cdot p_R)^{2000}$

PROBLEMA 1: Continuación

b) $p_T = 0'025$; $p_R = 0'9$ $\Rightarrow$ $p = 0'025 \cdot 0'9 = 0'0225$ $\Rightarrow$ $q = 1 - p = 0'9775$

Nota: El resultado del apartado anterior, con estos datos sería

$$P(A) = P(X = 50) \cdot 0'9775 = \binom{2049}{50} \cdot 0'0225^{50} \cdot 0'9775^{2000}$$

Pero en este caso debemos calcularlo para al menos 50 repeticiones (no exactamente 50).
B = "Que en 2050 transmisiones finalice el envío de a lo sumo 2000 palabras"
= "Que en 2050 transmisiones finalice un envío con al menos 50 detecciones de error"

Volviendo a tener en cuenta que un mensaje siempre acabará con una transmisión exitosa, podemos volver a mantener la definición de la variable X, quedando:
$P(B) = P(X \geq 50) \cdot q = P(X \geq 50) \cdot 0'9775$

Podría resolverse la binomial:

$$P(B) = P(X \geq 50) \cdot q = P(X \geq 50) \cdot 0'9775 = \sum_{i=50}^{2049} \binom{2049}{i} \cdot 0'0225^{i} \cdot 0'9775^{2049-i}$$

Pero dada la dificultad de dicho cálculo, lo resolveremos realizando la aproximación de la distribución binomial por la normal, según el teorema de Moivre-Laplace:
$npq = 2049 \cdot 0'0225 \cdot 0'9775 = 45'065 > 9$

Definiendo $Y : N\left(np, \sqrt{npq}\right)$, tendremos
$Y : N\left(2049 \cdot 0'0225, \sqrt{2049 \cdot 0'0225 \cdot 09775}\right) = N(46'1025, 6'713061429)$

y podremos aproximar:

$$P(X \geq 50) = P(Y \geq 49'5) = P\left(Z \geq \frac{49'5 - 46'1025}{6'713061429}\right) = P(Z \geq 0'506102921) \cong 0'306363977$$

Así,
$P(B) = P(X \geq 50) \cdot q = P(X \geq 50) \cdot 0'9775 \cong 0'306363977 \cdot 0'9775 = 0'299470787 \cong 0'2995$

Notas:
- El apartado a) podía resolverse mediante la normal si $npq = 2049 \cdot p \cdot q > 9$.
- El apartado b) puede puntualizarse como
  $P(X \geq 50) = P(50 \leq X \leq 2049) = P(49'5 \leq Y \leq 2049'5) = P(Y \geq 49'5) - P(Y \geq 2049'5)$
  Sin embargo, el segundo término resultará despreciable, al estar muy lejos de la media. (más de 298 desviaciones típicas)

12'2000

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID
Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 15-2-2000 (Tarde)

APELLIDOS Y NOMBRE:____

Nº DE EXPEDIENTE:____________________ CURSO: 3º GRUPO:____________________

PROBLEMA 2 *(2 puntos)*

La facturación que realiza un almacén en una semana es una variable aleatoria (unidad en millón de pesetas), que tiene la función de densidad de probabilidad:

$$f(x)=\frac{1}{\sqrt{8\pi}}e^{-\frac{(x^2+25-10x)}{8}}$$

A) Calcular la probabilidad de que la facturación de una semana sea inferior a 3 millones. **1p**

B) Se ha propuesto el objetivo de facturar por encima de la media, al menos 4 semanas de 5. Cuál es la probabilidad de alcanzar dicho objetivo, suponiendo independientes unas semanas de otras. **1p**

SOLUCIÓN:

a) La función de densidad es la de una variable aleatoria con distribución normal. Para obtener la media y la desviación típica la podemos poner de esta forma:

$$f(x)=\frac{1}{2\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-5}{2}\right)^2}$$

Así podemos observar que: $\mu=5$ $\sigma=2$

Por tanto:

$$P(x\le 3)=P\left(z\le\frac{3-5}{2}\right)=P(z\le -1)=P(z\ge 1)=0{,}1587$$

b) Primero obtenemos la probabilidad de una facturación superior a 5 millones en una semana:

$$P(x\ge 5)=P\left(z\ge\frac{5-5}{2}\right)=P(z\ge 0)=0{,}5$$

Usando dicha probabilidad, como la probabilidad de éxito de una distribución binomial obtenemos la probabilidad de "al menos 4 semanas de 5":

Y: "semanas en que se supera la media", con distribución: B (5, 0´5)

$$P(Y\ge 4)=\binom{5}{4}(0'5)^4(0'5)^1+\binom{5}{5}(0'5)^5(0'5)^0=0'1562+0'0312=0'1874$$

1'2000

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 15-2-2000 (Tarde)

APELLIDOS Y NOMBRE:
Nº DE EXPEDIENTE: CURSO: 3º GRUPO:

PROBLEMA 3 *(2'5 puntos)*

Dadas las variables aleatorias X e Y, se tiene la función:

$$f(x,y)=\begin{cases} k(y-x) & \text{en } R_{xy} \\ 0 & \text{resto} \end{cases}$$

donde R_{xy} lo forman el triángulo con vértices: (-1,0); (0,0); (0,1); junto con el triángulo cuyos vértices son: (0,0); (0,2); (2,0).

Se pide:

a) Calcular k para que dicha función sea la f.d.p. conjunta. **1p**

b) Obtener la curva general de regresión de Y sobre X. **1'5p**

SOLUCIÓN:

a) Con los puntos (-1,0) y (0,1) obtenemos la recta: y = x + 1.
Con los puntos (2,0) y (0,2) obtenemos la recta: y = 2 – x.

$$1=\int_{-1}^{0}\int_{0}^{x+1}k(y-x)\,dy\,dx+\int_{0}^{2}\int_{0}^{2-x}k(y-x)\,dy\,dx=\int_{-1}^{0}k\left(\frac{y^2}{2}-xy\right)_0^{x+1}dx+\int_{0}^{2}k\left(\frac{y^2}{2}-xy\right)_0^{2-x}dx=$$

$$=\int_{-1}^{0}k\left(-\frac{x^2}{2}+\frac{1}{2}\right)dx+\int_{0}^{2}k\,(\frac{3}{2}x^2-4x+2)\,dx=k\left(-\frac{x^3}{6}+\frac{1}{2}x\right]_{-1}^{0}+k\left(\frac{x^3}{2}-2x^2+2x\right]_{0}^{2}=1$$

Luego $k\left(-\frac{1}{6}+\frac{3}{6}\right)=1$ Y, por tanto, $k=3$

b) El primer paso es obtener la función de densidad marginal de X, para posteriormente calcular la función de densidad de Y condicionada a X.

$$x<-1 \Rightarrow f(x)=0$$

$$-1\le x\le 0 \Rightarrow f(x)=\int_{0}^{x+1}k(y-x)\,dy=-\frac{3}{2}x^2+\frac{3}{2}$$

$$0\le x\le 2 \Rightarrow f(x)=\int_{0}^{2-x}k(y-x)\,dy=-\frac{9}{2}x^2-12x+6$$

$$x>2 \Rightarrow f(x)=0$$

PROBLEMA 3: Continuación 1

Ahora obtenemos las funciones de densidad condicionadas:

En el primer triángulo: $$f_1\left(y/_x\right)=\frac{3(y-x)}{-\frac{3}{2}x^2+\frac{3}{2}}=\frac{2(y-x)}{1-x^2}$$

En el segundo triángulo: $$f_2\left(y/_x\right)=\frac{3(y-x)}{\frac{9}{2}x^2-12x+6}=\frac{y-x}{\frac{3}{2}x^2-4x+2}$$

Para X <-1 y X > 2, la densidad condicionada no existe, debido a que en dichos tramos la f (x) = 0. Para X entre -1 y 2 pero fuera de las regiones delimitadas por los triángulos, la densidad condicionada es cero, al ser la f (x y) = 0.

Curva de regresión de Y sobre X:

$$-1\le x\le 0 \quad\Rightarrow\quad \hat{y}=E\left(y/_x\right)=\int y\, f_1\left(y/_x\right)dy=\int_0^{x+1} y\left(\frac{2(y-x)}{1-x^2}\right)dy=\frac{-x^3+x^2+2x+1}{3(1-x^2)}$$

$$0\le x\le 2 \quad\Rightarrow\quad \hat{y}=E\left(y/_x\right)=\int y\, f_2\left(y/_x\right)dy=\int_0^{2-x} y\left(\frac{y-x}{\frac{3}{2}x^2-4x+2}\right)dy=\frac{-5x^3+32x^2-36x+16}{9x^2-24x+12}$$

9'2000

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 15-2-2000 (Tarde)

APELLIDOS Y NOMBRE:______________________________

Nº DE EXPEDIENTE:________________ CURSO: 3º GRUPO:____________

PROBLEMA 4 *(3 puntos)*

Un cañón dispara diferentes tipos de obuses: O1; O2; O3; ... que destruyen el objetivo si hacen impacto a menos de 100m; 200m; 300m; ... del mismo, respectivamente. La distribución de impactos es normal, y la misma con todos los obuses.

El cañón puede disparar a cualquier distancia d, siendo $d_{máx}$ la distancia esperada de máximo alcance del cañón. La desviación típica es de 200m a la máxima distancia, y a cualquier otra distancia la desviación típica es proporcional a la de distancia máxima, con coeficiente de proporcionalidad $\left(\frac{d}{d_{\text{máx}}}\right)$. Es decir, $\sigma_d = \left(\frac{d}{d_{\text{máx}}}\right)\cdot\sigma_{d_{\text{máx}}}$

a) En fábrica se disparan 25 obuses, con el cañón en posición de alcance de la máxima distancia. La media de los alcances obtenidos es de 4.000m. Calcule un intervalo de confianza del 95% para el valor esperado del alcance máximo para dicho cañón. **0'8p**

b) ¿Cuántos obuses deben lanzarse en posición de máximo alcance, para calcular el valor esperado de dicho alcance máximo, con una probabilidad del 97'8% de no estar cometiendo un error mayor de 50m?. **0'7p**

c) Se dispara a un objetivo situado a una distancia mitad de la máxima. Se quiere garantizar, con una probabilidad del 99%, que el objetivo quede destruido. **1'5p**

   c1) ¿Cuántos obuses del tipo O2 han de dispararse, apuntando al objetivo?

   c2) ¿Y del tipo O3?

SOLUCIÓN:

a) La desviación típica de la distribución de impactos es conocida:

$$\sigma_{d\,\text{máx}} = 200 \Rightarrow I = \bar{x} \pm z_{\alpha/2}\frac{\sigma_{d\,\text{máx}}}{\sqrt{n}}$$

$$I = \bar{x} \pm z_{0'025}\frac{\sigma}{\sqrt{n}} = 4000 \pm 1'96\cdot\frac{200}{\sqrt{25}} = 4.000 \pm 78'4 = (3.921'6,\ 4.078'4)$$

b) $$\varepsilon \le z_{\alpha/2}\frac{\sigma_{d\,\text{máx}}}{\sqrt{n}} \Rightarrow n \ge \left(z_{\alpha/2}\frac{\sigma_{d\,\text{máx}}}{\varepsilon}\right)^2 = \left(z_{0'011}\frac{\sigma_{d\,\text{máx}}}{\varepsilon}\right)^2 = \left(2'29\frac{200}{50}\right)^2 = 83'9056 \Rightarrow n \ge 84$$

c1) Supongamos que se lanza un obús $O2$.

Definimos el suceso S_2 = "Un obús $O2$ destruye el objetivo"

El cañón disparará al objetivo, que se encuentra a distancia $d = \frac{d_{\text{máx}}}{2} \Rightarrow \sigma_d = \frac{\sigma_{d\,\text{máx}}}{2}$

Definimos la v.a. X_2 = "distancia del cañón al punto de impacto del obús", con distribución normal $X_2 : N(d,\sigma_d) = N\left(\frac{d_{\text{máx}}}{2},\frac{\sigma_{d\,\text{máx}}}{2}\right) = N\left(\frac{d_{\text{máx}}}{2},100\right)$

$$P(S_2) = P\left(-200 \le X_2 - \frac{d_{\text{máx}}}{2} \le 200\right) = P\left(-\frac{200}{100} \le z \le \frac{200}{100}\right) = P(-2 \le z \le 2) = 1 - 2\cdot\alpha(2) =$$

$$= 1 - 2\cdot 0'0228 = 0'9544$$

8'2000

PROBLEMA 4: Continuación

Definamos el suceso D_{O2n} = "El objetivo se destruye al disparar n obuses $O2$"

Definimos la v.a. Y_{O2n} = "número de destrucciones con n obuses $O2$"
Esta v.a. es binomial, con n pruebas y probabilidad de éxito $P(S) = 0'9544$

Lanzando 2 obuses (con uno, la prob de éxito ya está en el 95'44%, luego no harán falta muchos obuses)

$$P(D_{O22}) = P(Y_{O22} = 1) + P(Y_{O22} = 2)$$

$$P(Y_{O22} = 1) = \binom{2}{1} \cdot (0'9544) \cdot (1 - 0'9544) = 2 \cdot 0'9544 \cdot 0'0456 = 0'08704128$$

$$P(Y_{O22} = 2) = \binom{2}{2} \cdot (0'9544)^2 \cdot (1 - 0'9544)^0 = 1 \cdot 0'9544^2 = 0'91087936$$

$$P(D_{O22}) = P(Y_{O22} = 1) + P(Y_{O22} = 2) = 0'08704128 + 0'91087936 = 0'99792064 \cong 99'79\% > 99\%$$

Luego es suficiente con lanzar dos obuses O2

c2) Supongamos que se lanza un obús $O3$.
Definimos el suceso S_3 = "Un obús $O3$ destruye el objetivo"
El cañón disparará al objetivo, que se encuentra a distancia $d = \frac{d_{\text{máx}}}{2}$
Definimos la v.a. X_3 = "distancia del cañón al punto de impacto del obús", con distribución normal $\quad X_3 : N\left(\frac{d_{\text{máx}}}{2}, \frac{\sigma_{d\,\text{máx}}}{2}\right) = N\left(\frac{d_{\text{máx}}}{2}, 100\right)$

$$P(S_3) = P\left(-300 \le X_3 - \frac{d_{\text{máx}}}{2} \le 300\right) = P\left(-\frac{300}{100} \le z \le \frac{300}{100}\right) = P(-3 \le z \le 3) = 1 - 2 \cdot \alpha(3) =$$
$$= 1 - 2 \cdot 0'00135 = 0'9973 = 99'73\% > 99\%$$

Luego es suficiente con lanzar un obús O3

NOTA:

Lo que nos piden **no es equivalente**, aunque los resultados coincidan, al nº de disparos (tamaño de la muestra) necesario para poder garantizar, con confianza del 99%, que la **media de impactos** de dicha muestra diste del objetivo $\left(\frac{d_{\text{máx}}}{2}\right)$ no más de $\begin{cases} c1 & 200m \\ c2 & 300m \end{cases}$.

Esto se resolvería como sigue:

$$\begin{cases} c1 \Rightarrow \varepsilon_1 = 200m \Rightarrow z_{0'005}\frac{100}{\sqrt{n}} \le 200 \Rightarrow n \ge \left(2'576 \cdot \frac{100}{200}\right)^2 = 1'66 \Rightarrow n \ge 2 \\ c2 \Rightarrow \varepsilon_2 = 300m \Rightarrow z_{0'005}\frac{100}{\sqrt{n}} \le 300 \Rightarrow n \ge \left(2'576 \cdot \frac{100}{300}\right)^2 = 0'74 \Rightarrow n \ge 1 \end{cases}$$

7'2000

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: **15-2-2000 (Mañana)**

APELLIDOS Y NOMBRE:____________________________
Nº DE EXPEDIENTE:______________CURSO: **3º** GRUPO:__________

PROBLEMA 1

(2 puntos)

Hay 8000 cartas en una oficina de Correos, de las que se sabe que son de dos posibles destinos: nacional o extranjero. La probabilidad de que una carta extraída al azar tenga destino nacional es de 0'6.

a) Calcule la probabilidad de que realizando 5 extracciones sin reposición, al menos 4 tengan destino nacional. **1p**

b) Suponiendo que se realizan 45 extracciones, calcule la probabilidad de que tengan destino extranjero al menos 17. **1p**

SOLUCIÓN:

a) Al ser sin reposición, habría que usar una distribución hipergeométrica, pero dado que el número de cartas es grande, hacemos la aproximación a una binomial. La aproximación es buena con: $N > 10n$, es decir, $8000 > 10\cdot 5$.

Definimos A: "al menos 4 sean con destino nacional"

$$P(A)=\binom{5}{4}(0'6)^4(0'4)^1+\binom{5}{5}(0'6)^5(0'4)^0=\binom{5}{1}(0'4)^1(0'6)^4+\binom{5}{0}(0'4)^0(0'6)^5$$

$$P(A)=5\cdot 0'4\cdot 0'6^4+0'6^5=0'2592+0'0778=0'337$$

Como se observa, dado que la probabilidad de éxito es superior a 0'5, se ha tenido en cuenta que k éxitos implican n-k fracasos, para poder buscar en las tablas.

b) Seguiremos trabajando sobre una distribución hipergeométrica, pero la aproximación a binomial todavía sirve: $8000 > 45\cdot 5$

Pero en este caso el número de extracciones ha crecido hasta 45, de modo que:
$npq = 45\cdot 0'4\cdot 0'6 = 10'8 > 9$

Cumpliendo así las condiciones de Moivre-Laplace para aproximar la binomial a la normal.

$$\mu = n\cdot p = 45\cdot 0'4 = 18$$

$$\sigma=\sqrt{n\cdot p\cdot q}=\sqrt{45\cdot 0'4\cdot 0'6}=\sqrt{10'8}\cong 3'286335$$

Definiendo a X: "cartas con destino extranjero", y teniendo en cuenta hay que realizar la corrección de continuidad, al pasar de una variable discreta a una variable continua:

$$P(X_{BIN}\geq 17)=P(X_{NOR}\geq 16'5)=P(Z\geq \frac{16'5-18}{\sqrt{10'8}})=P(Z\geq -0'456435)=$$
$$=1-P(Z\geq 0'456435)\cong 1-0'324=0'676$$

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID
Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 15-2-2000 (Mañana)

APELLIDOS Y NOMBRE:
Nº DE EXPEDIENTE: CURSO: 3º GRUPO:

PROBLEMA 2 *(2'5 puntos)*

La vida en días de un periférico es una variable aleatoria *X*, con f.d.p.:

$$f(x)=\begin{cases}\dfrac{k}{x^2} & \forall \quad x>100\\[2ex] 0 & resto\end{cases}$$

a) Calcule *k*. **0'4p**

Si tenemos tres periféricos como éste, conectados de forma independiente, cada uno en un ordenador, y se supone que si el periférico falla no se reemplaza por uno nuevo:

b) Establezca la distribución de probabilidad de la variable aleatoria definida como número de ordenadores que han tenido un fallo en dicho periférico, en los 150 primeros días de vida. **0'8p**

c) ¿Qué probabilidad hay de que si no ha fallado ninguno en el primer año de vida (365 días), los tres sufran un fallo en el siguiente mes (30 días)? **0'8p**

d) Conectamos uno de estos periféricos con otro cuyo tiempo de vida *Y* sigue una distribución exponencial de media 200 días. Si el funcionamiento o no de uno no influye en el del otro, calcule la f.d.p. conjunta de la vida de ambos, como variable bidimensional (*X*,*Y*) **0'5p**

SOLUCIÓN:

La variable que nos dan es la vida en días, o lo que es lo mismo, el tiempo, en días, hasta que el periférico falla, como variable aleatoria.

Así,
$X \le t$ significa que el periférico falla entre 0 y t
$X \ge t$ significa que el periférico funciona entre 0 y t (falla después de t)
$t_1 \le X \le t_2$ significa que el periférico falla entre t_1 y t_2

a) Para resolverlo aplicamos que debe cumplir $\int_{-\infty}^{\infty} f(x)dx = 1$, lo que en este caso resulta:

$$\int_{-\infty}^{\infty} f(x)dx = \int_{100}^{\infty}\frac{k}{x^2}dx = -\frac{k}{x}\bigg|_{100}^{\infty} = 0+\frac{k}{100} = 1 \quad\Rightarrow\quad k=100$$

b) Primero calculemos la probabilidad de que uno falle en ese período, y luego se establece la sucesión de pruebas de Bernoulli, es decir, se modelarán como V.A. binomial:

S = "Un periférico falla en los primeros 150 días"

$$P(S) = P(X\le 150) = \int_{-\infty}^{150} f(x)dx = \int_{100}^{150}\frac{k}{x^2}dx = -\frac{k}{x}\bigg|_{100}^{150} = \frac{k}{100}-\frac{k}{150} = \frac{3k-2k}{300} = \frac{k}{300}$$

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID
Departamento de Electrónica y Comunicaciones

ESTADÍSTICA:	15-2-2000 (Mañana)

APELLIDOS Y NOMBRE:
Nº DE EXPEDIENTE: CURSO: 3º GRUPO:

PROBLEMA 2: Continuación A

Para el valor de k obtenido: $P(S)=\frac{1}{3}$

Definiendo la variable:
W ="número de ordenadores que han tenido un fallo en dicho periférico, en los 150 primeros días de vida"
será una v.a. binomial, con tres realizaciones, y probabilidad de éxito $P(S)$, en cada una:

$$\boxed{P(W=w_i) = \binom{3}{w_i}\cdot P_S^{w_i}\cdot P_{\overline{S}}^{3-w_i} = \binom{3}{w_i}\cdot\left(\frac{1}{3^{w_i}}\right)\cdot\left(1-\frac{1}{3}\right)^{3-w_i}}$$

$$P(W=0) = P(\overline{S},\overline{S},\overline{S}) = \binom{3}{0}\cdot P_S^0\cdot P_{\overline{S}}^3 = P_{\overline{S}}^3 = (1-P_S)^3 = \left(1-\frac{k}{300}\right)^3 = \left(\frac{2}{3}\right)^3 = \frac{8}{27}$$

$$P(W=1) = \binom{3}{1}\cdot P_S^1\cdot P_{\overline{S}}^2 = 3\cdot P_S\cdot P_{\overline{S}}^2 = 3\cdot P_S\cdot(1-P_S)^2 = 3\frac{k}{300}\left(1-\frac{k}{300}\right)^2 = 3\frac{1}{3}\left(\frac{2}{3}\right)^2 = \frac{4}{9}$$

$$P(W=2) = \binom{3}{2}\cdot P_S^2\cdot P_{\overline{S}}^1 = 3\cdot P_S{}^2\cdot(1-P_S) = 3\left(\frac{k}{300}\right)^2\left(1-\frac{k}{300}\right) = 3\left(\frac{1}{3}\right)^2\frac{2}{3} = \frac{2}{9}$$

$$P(W=3) = P(S,S,S) = \binom{3}{3}\cdot P_S^3\cdot P_{\overline{S}}^0 = P_S^3 = P_S^3 = \left(\frac{k}{300}\right)^3 = \left(\frac{1}{3}\right)^3 = \frac{1}{27}$$

Así,

W	0	1	2	3	
$P(w_i)$	8/27	12/27	6/27	1/27	1

c) A = "No hay fallos en los primeros 365 días" = "no falla entre 1 y 365 inclusive"
B = "Los tres fallan entre los días 366 y 395, ambos inclusive"

Nos piden $P(B\mid A)=\frac{P(A\cap B)}{P(A)}$

Dado que el enunciado indica que cada periférico puede fallar una sola vez (no se reemplazan por nuevos), la intersección de A y B coincide con B.

Así: $P(B\mid A)=\frac{P(B)}{P(A)}$

PROBLEMA 2: Continuación B

Calculemos ambas probabilidades. Para ello definimos
A_i = "el periférico i-ésimo no falla en los primeros 365 días"
B_i = " el periférico i-ésimo falla entre los días 366 y 395, ambos inclusive"

$P(B) = P(B_1, B_2, B_3) = P(B_i)^3$ siendo

$$P(B_i) = P(365 \le X \le 395) = \int_{365}^{395} \frac{k}{x^2}dx = k\left(\frac{1}{365} - \frac{1}{395}\right) = \left(\frac{100}{365} - \frac{100}{395}\right) \cong 0'020808045$$

Nótese que X=1 indica el final del primer día, por lo que X=365 indica el final del día 365, y por tanto principio del 366, así como X=395 indica el final del día 395.

Por otro lado:
$P(A) = P(A_1, A_2, A_3) = P(A_i)^3$ siendo

$$P(A_i) = P(X > 365) = 1 - P(X \le 365) = 1 - \int_{100}^{365} \frac{k}{x^2}dx = 1 - k\left(\frac{1}{100} - \frac{1}{365}\right) = 1 - \left(\frac{100}{100} - \frac{100}{365}\right)$$

$$P(A_i) \cong 1 - 0'726027397 = 0'273972603$$

Por tanto,

$$P(B \mid A) = \frac{P(B)}{P(A)} = \frac{P(B_i)^3}{P(A_i)^3} = \left(\frac{P(B_i)}{P(A_i)}\right)^3 \cong (0'075943364)^3 \cong 4'381 \cdot 10^{-4}$$

d) $$f(y) = \begin{cases} \dfrac{e^{-\frac{y}{200}}}{200} & y \ge 0 \\ 0 & resto \end{cases}$$

Dado que el enunciado expone la independencia entre las variables X e Y, su función densidad de probabilidad conjunta será el producto de sus marginales, para todo par (x, y)

$$f(x,y) = \begin{cases} \dfrac{k \cdot e^{-\frac{y}{200}}}{200 \cdot x^2} & \forall \begin{cases} x \ge 100 \\ y \ge 0 \end{cases} \\ 0 & resto \end{cases}$$ luego $$f(x,y) = \begin{cases} \dfrac{e^{-\frac{y}{200}}}{2x^2} & \forall \begin{cases} x \ge 100 \\ y \ge 0 \end{cases} \\ 0 & resto \end{cases}$$

3'9.000

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 15-2-2000 (Mañana)

APELLIDOS Y NOMBRE:____________________________________

Nº DE EXPEDIENTE:__________________CURSO: 3º _ GRUPO:________________

PROBLEMA 3 *(3 puntos)*

Dadas las variables aleatorias X e Y, se tiene la función:

$$f(x,y)=\begin{cases}(x-y)/2 & \text{en } R_{1\text{-}xy}\\ k & \text{en } R_{2\text{-}xy}\\ 0 & \text{resto}\end{cases}$$

siendo $R_{1\text{-}xy}$ el polígono delimitado por los puntos: (0,0); (1,0); (1,2); (0,1); y
$R_{2\text{-}xy}$ el triángulo delimitado por los puntos: (1,0); (3,0); (1,2).

Se pide:

a) Calcular el valor que debe tener *k* para que dicha función sea la f.d.p. conjunta. **0'5p**

b) Calcular las funciones de densidad marginales de X y de Y. **1p**

c) Obtener la curva general de regresión de Y sobre X. **1'5p**

SOLUCIÓN:

Los puntos (0,1) y (1,2) determinan la recta y = x + 1 ; y los puntos (1,2) y (3,0) determinan la recta y = 3 – x

a) $$\int_0^1\int_0^{x+1}\frac{x-y}{2}\,dy\,dx+\int_1^3\int_0^{3-x}k\,dy\,dx=1 \Rightarrow \int_0^1 \left.\frac{x.y-y^2/2}{2}\right]_0^{x+1}dx+\int_1^3 \left.k.y\right]_0^{3-x}dx=1$$

$$\Rightarrow \int_0^1\frac{x^2-1}{4}\,dx+\int_1^3(3k-kx)\,dx=1 \qquad \Rightarrow \qquad (\frac{1}{12}-\frac{1}{4})+\left[(9k-\frac{9k}{2})-(3k-\frac{k}{2})\right]=1$$

Luego $\boxed{k=7/12}$

b) $$y<0 \quad\Rightarrow\quad f(y)=0$$

$$0\le y\le 1 \quad\Rightarrow\quad f(y)=\int_0^1\frac{x-y}{2}\,dx+\int_1^{3-y}\frac{7}{12}\,dx=\frac{17}{12}-\frac{13y}{12}$$

$$1\le y\le 2 \quad\Rightarrow\quad f(y)=\int_{y-1}^1\frac{x-y}{2}\,dx+\int_1^{3-y}\frac{7}{12}\,dx=\frac{7}{6}-\frac{13y}{12}+\frac{y^2}{4}$$

$$y>2 \quad\Rightarrow\quad f(y)=0$$

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID
Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 15-2-2000 (Mañana)

APELLIDOS Y NOMBRE: ______
Nº DE EXPEDIENTE: ______ CURSO: 3º GRUPO: ______

PROBLEMA 3: Continuación

c) Para obtener la curva de regresión de Y sobre X, necesitamos la densidad condicionada, y para ello hay que calcular previamente la función de densidad marginal de X, ya que:

$$f\left(y/x\right)=\frac{f(x,y)}{f(x)} \quad \forall f(x)\neq 0$$

Función de densidad marginal de X:

$$x<0 \;\Rightarrow\; f(x)=0$$

$$0\le x\le 1 \;\Rightarrow\; f(x)=\int_0^{x+1}\frac{x-y}{2}\,dy=\frac{x^2-1}{4}$$

$$1\le x\le 3 \;\Rightarrow\; f(x)=\int_0^{3-x}\frac{7}{12}\,dy=\frac{21}{12}-\frac{7}{12}x$$

$$x>3 \;\Rightarrow\; f(x)=0$$

Función de densidad de Y condicionada al valor de X:

$$f_{R1-xy}\left(y/x\right)=\frac{\frac{x-y}{2}}{\frac{x^2-1}{4}}=\frac{2x-2y}{x^2-1} \qquad f_{R2-xy}\left(y/x\right)=\frac{7/12}{\frac{21}{12}-\frac{7}{12}x}=\frac{1}{3-x}$$

Para X < 0 y X > 3, dicha densidad condicionada no existe, debido a que en dichos tramos la f (x) = 0. Para X entre 0 y 3 pero fuera de R_{1-xy} y de R_{2-xy}, la densidad condicionada es cero, al ser la f (x y) = 0.

Curva de regresión de Y sobre X:

$$0\le x\le 1 \;\Rightarrow\; E\left(y/x\right)=\int_{R1} y\, f_{R1}\left(y/x\right)dy=\int_0^{x+1} y\left(\frac{2x-2y}{x^2-1}\right)dy=\frac{-\frac{2}{3}x^3-x^2-x-\frac{2}{3}}{x^2-1}$$

$$1\le x\le 3 \;\Rightarrow\; E\left(y/x\right)=\int_{R2} y\, f_{R2}\left(y/x\right)dy=\int_0^{3-x} y\left(\frac{1}{3-x}\right)dy=\frac{3-x}{2}$$

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID
Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 15-2-2000 (Mañana)

APELLIDOS Y NOMBRE:______________________

Nº DE EXPEDIENTE:____________ CURSO: 3º GRUPO:____________

PROBLEMA 4 *(2'5 puntos)*

a) Una empresa que fabrica discos de ordenador ha tomado una muestra aleatoria de 200 discos, de los cuales 12 son defectuosos. Calcule un intervalo de confianza bilateral del que podamos asegurar que existe una probabilidad del 95% de que contenga a la proporción de defectuosos poblacional. Expréselo en porcentajes, y no en tanto por uno. **1p**

b) Qué tamaño debe tener una muestra de discos si se pretende poder garantizar con una probabilidad del 95% que podemos utilizar la estimación puntual del porcentaje de defectuosos como aproximación del porcentaje real, sin cometer más error que un 2%. Utilice para este cálculo los datos obtenidos en muestras anteriores. **0'8p**

c) Realice de nuevo el apartado b), pero, en este caso, no utilice para realizar este apartado ningún resultado obtenido en muestras anteriores. Suponga que desconoce cualquier dato de la población. **0'7p**

SOLUCIÓN:

a) Si la muestra es lo suficientemente grande para que se cumplan las condiciones de Moivre-Laplace, podremos usar $I = \hat{p} \pm z_{\alpha/2}\sqrt{\dfrac{p\cdot(1-p)}{n}}$

Como valor de p, tomamos el de estimación mediante la muestra de tamaño 85 anterior.

$$p \cong \hat{p} = \frac{12}{200} = 0'06 = 6\% \quad \Rightarrow \quad npq = 200\cdot 0'06\cdot 0'94 = 11'28 > 9 \quad \Rightarrow \quad \text{Cumple Moivre}$$

$$NC = 1-\alpha = 95\% = 0'95 \quad \Rightarrow \quad z_{\alpha/2} = z_{0,025} = 1'96$$

$$I = \frac{12}{200} \pm 1'96\sqrt{\frac{\frac{12}{200}\cdot\frac{188}{200}}{200}} = 0'06 \pm 0'032914 = (0'027086,\ 0'092914) \quad \Rightarrow \quad I \cong (2'71\%,\ 9'29\%)$$

b) $$\varepsilon \geq z_{\alpha/2}\sqrt{\frac{p\cdot(1-p)}{n}} \quad \Rightarrow \quad n \geq \left(\frac{z_{\alpha/2}}{\varepsilon}\right)^2 p\cdot(1-p)$$

Como valor de p, tomamos el de estimación mediante la muestra de tamaño 85 anterior.

$$p \cong \hat{p} = \frac{12}{200} = 0'06 = 6\% \qquad \text{Así,}\quad n \geq \left(\frac{z_{\alpha/2}}{\varepsilon}\right)^2 p\cdot(1-p) = \left(\frac{1'96}{0'02}\right)^2 \frac{12}{200}\cdot\frac{188}{200} = 541'6656$$

Luego $n \geq 542$ [con este mayor tamaño de muestra sigue cumpliendo Moivre-Laplace]

c) Idem b), salvo que el peor caso posible es $p\cdot(1-p) = 0'25$. Así, para poder garantizar el 95% de confianza, desconociendo estimaciones previas de la proporción, podemos hacer:

$$n \geq \left(\frac{z_{\alpha/2}}{\varepsilon}\right)^2 p\cdot(1-p) = \left(\frac{1'96}{0'02}\right)^2 0'25 = 2401$$

[$npq = 2401\cdot 0'5\cdot 0'5 > 9$, luego cumple Moivre-Laplace y la distribución tomada es válida]