



43'2001

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA:

13-9-2001 (Tarde)

APELLIDOS Y NOMBRE: _____

Nº DE EXPEDIENTE: _____

CURSO: 3º GRUPO: _____

PROBLEMA 1

(2'5 puntos)

Se tiene una variable X' que sigue una distribución Binomial $B(n, p)$, pero donde se cumplen los requisitos necesarios para realizar la aproximación a una distribución Normal con variable X . Por otro lado, se tiene una variable Y' que sigue una distribución de Poisson con parámetro γ , pero donde se cumplen los requisitos necesarios para realizar la aproximación a una distribución Normal con variable Y . Además, hay una variable T que se distribuye $N(d, 2d)$. Se sabe que las variables X, Y, T , son independientes. Dado que:

$$W = a + cY - bX$$

$$M = bT + aW$$

- a) Obtenga la media y la varianza de la variable M , en función de los parámetros que corresponda.
- b) Suponga una nueva situación para W , en la que ahora se conoce que se distribuye $N(5, 2)$. Teniendo en cuenta que $d = 3$, y que se mantiene la relación: $M = bT + aW$, calcule la probabilidad de que la variable M esté por debajo del valor a . Considere para este apartado que $a = b$.

SOLUCIÓN:

- a) X se distribuye $N(np, \sqrt{npq})$
 Y se distribuye $N(\gamma, \sqrt{\gamma})$

$$E(W) = E(a + bY - cX) = a + cE(Y) - bE(X) = a + c\gamma - bnp$$

$$\sigma_W^2 = \sigma_{(a+cY-bX)}^2 = c^2\sigma_Y^2 + b^2\sigma_X^2 = c^2\gamma + b^2npq$$

$$W \text{ se distribuye } N(a + c\gamma - bnp, \sqrt{c^2\gamma + b^2npq})$$

$$E(M) = E(bT + aW) = bE(T) + aE(W) = \boxed{bd + a(a + c\gamma - bnp)}$$

$$\sigma_M^2 = \sigma_{(bT+aW)}^2 = b^2\sigma_T^2 + a^2\sigma_W^2 = \boxed{4b^2d^2 + a^2(c^2\gamma + b^2npq)}$$

- b) T se distribuye $N(3, 6)$

$$E(M) = bE(T) + aE(W) = 3b + 5a = 8a$$

$$\sigma_M^2 = \sigma_{(bT+aW)}^2 = b^2\sigma_T^2 + a^2\sigma_W^2 = 36b^2 + 4a^2 = 40a^2$$

$$\sigma_M = 6'32a$$

$$P(M < a) = P\left(Z < \frac{a - 8a}{6'32a}\right) = P(Z < -1'107) = P(Z \geq 1'107) = \boxed{0'1357}$$

**PROBLEMA 2**

(2'5 puntos)

En una población de una determinada región, se quiere estudiar la edad de los individuos. Para ello, se analiza una muestra de 50 individuos en dicha región, obteniéndose que la suma de las edades de las 50 personas es de 1900, y que la varianza es 4.

- a) Obtenga un intervalo de confianza del 90% para la media de la edad de la población de dicha región.
- b) Una empresa de investigación quiere seleccionar una región sobre la que poder experimentar un determinado proyecto. La empresa ha considerado según estudios anteriores, que la condición adecuada para que dicho proyecto pueda llevarse a cabo con la fiabilidad requerida, es que la edad media de la población de la región estudiada sea de 39 años. Además, se asume una probabilidad del 10% de descartar una región válida para el proyecto. ¿Puede servir la región anterior en este proyecto?.
- c) Suponga que la edad media de la población de dicha región es de 38, ¿cuál es la probabilidad de que la consideremos válida para el proyecto? Razone sobre la conveniencia o no de aumentar la muestra.

SOLUCIÓN:

$$a) \bar{X} = \frac{1900}{50} = 38$$

$$n\sigma_M^2 = (n-1)s^2 \quad s^2 = \frac{n}{n-1}\sigma_M^2 = \frac{50}{49}4 = 4'0816 \quad s = 2'02$$

$$1 - \alpha = 0'9 \quad \alpha = 0'1 \quad \frac{\alpha}{2} = 0'05 \quad Z_{0'05} = 1'645$$

Como la desviación típica poblacional es desconocida y el tamaño de la muestra supera 30:

$$\bar{X} \pm Z_{\alpha/2} \frac{s}{\sqrt{n}} = 38 \pm 1'645 \frac{2'02}{\sqrt{50}} = (37'5301, 38'4699)$$

b)

Planteamos el contraste bilateral:

$$H_0: \mu = 39$$

$$H_a: \mu \neq 39$$



44'2001

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA:

13-9-2001 (Tarde)

APELLIDOS Y NOMBRE: _____

Nº DE EXPEDIENTE: _____

CURSO: 3º

GRUPO: _____

PROBLEMA 2: Continuación

$$Z = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{38 - 39}{2'02/\sqrt{50}} = -3'5$$

$$\alpha = 0'1$$

$$\frac{\alpha}{2} = 0'05$$

$$Z_{0'05} = 1'645$$

La región de aceptación es: $|Z| \leq Z_{\alpha/2}$

Dado que $3'5 > 1'645$, se rechaza la hipótesis, por lo que no consideramos válida esta región para el proyecto.

c)

$$H_0: \mu = 39$$

$$H_a: \mu = 38$$

$$\beta = P\left(\text{Aceptar } H_0 / H_a \text{ cierta}\right) = P\left(\frac{\mu_0 - Z_{\alpha/2} \frac{s}{\sqrt{n}} - \mu_a}{s/\sqrt{n}} \leq Z \leq \frac{\mu_0 + Z_{\alpha/2} \frac{s}{\sqrt{n}} - \mu_a}{s/\sqrt{n}}\right) =$$

$$P\left(\frac{39 - 1'645 \frac{2'02}{\sqrt{50}} - 38}{2'02/\sqrt{50}} \leq Z \leq \frac{39 + 1'645 \frac{2'02}{\sqrt{50}} - 38}{2'02/\sqrt{50}}\right) = P(1'8564 \leq Z \leq 5'1463) =$$

$$P(Z \geq 1'8564) - P(Z \geq 5'1463) = 0'0322 - 0'00000017 \approx \boxed{0'0322}$$

Dado que el Error Tipo II es aproximadamente del 3%, la potencia es del 97%, por lo que se podría considerar elevada, y no ser necesario aumentar la muestra. Por otro lado, si calculáramos la potencia con una muestra ligeramente mayor y viéramos que sobrepasa el 99%, de tal forma que ese aumento de muestra implicara un coste pequeño, entonces podría interesar dicho aumento de muestra.

**PROBLEMA 3****(2'5 puntos)**

Las calificaciones de un examen de Estadística siguen una distribución normal.

- 1) Tomados 20 exámenes al azar la suma de sus calificaciones es 106 y $\sigma_M^2 = 3.8$. En estas condiciones el profesor dice que la nota media de los alumnos está en el intervalo $[4.5, a]$. Haga las hipótesis que considere necesarias y calcule:

- Valor de a
- ¿Cuál es la confianza asociada a ese intervalo?
- ¿Cuál es la probabilidad de que realmente la media sea inferior a 4.5?. Razone cuidadosamente esta respuesta.

- 2) Si la muestra hubiera sido de 15 exámenes y con los datos de ella se hubiera obtenido $\sum_i (x_i - \bar{x})^2 = 50$, obtenga un intervalo de confianza para la desviación típica de las calificaciones de todos los alumnos, con un nivel de significación del 4%.

SOLUCIÓN:

- 1) La población que queremos estudiar está formada por todos los alumnos que se han examinado de Estadística. Debemos hacer la hipótesis de que la población es infinita o que el muestreo se ha realizado con reemplazamiento.

- a) Suponemos también que el intervalo de confianza para la nota media es simétrico (respecto de la media muestral). Como la varianza de la población es desconocida y el tamaño muestral es $n = 20 < 30$, entonces

$$I = \left[\bar{x} \pm t_{n-1, \alpha/2} \frac{s}{\sqrt{n}} \right] \quad \text{y como } \bar{x} = \frac{1}{n} \sum_{i=1}^{20} x_i = \frac{106}{20} = 5.3 \quad \text{el intervalo debe ser simétrico respecto de } 5.3. \text{ La semilongitud izquierda del intervalo es } 5.3 - 4.5 = 0.8, \text{ luego la semilongitud derecha será también } 0.8, \text{ y } \underline{a} = 5.3 + 0.8 = \underline{6.1}$$

- b) El intervalo es
$$I = \left[\bar{x} \pm t_{n-1, \alpha/2} \frac{s}{\sqrt{n}} \right] = \left[5.3 \pm t_{n-1, \alpha/2} \frac{s}{\sqrt{n}} \right] = [5.3 \pm 0.8] = [4.5, 6.1]$$

luego $t_{n-1, \alpha/2} \frac{s}{\sqrt{n}} = 0.8 \quad \text{y como } s^2 = s_M^2 = \frac{n}{n-1} \sigma_M^2 = \frac{20}{19} \cdot 3.8 = 4 \Rightarrow s = 2$

luego $t_{n-1, \alpha/2} \frac{2}{\sqrt{20}} = 0.8 \Rightarrow t_{19, \alpha/2} = \frac{0.8 \cdot \sqrt{20}}{2} = 1.7889$

interpolando en la tabla de la t de Student, para 19 g.l.:



$$\frac{0.05 - 0.025}{2.093 - 1.729} = \frac{\frac{\alpha}{2} - 0.025}{2.093 - 1.7889} \Rightarrow \frac{\alpha}{2} = 0.025 + \frac{0.3041 \cdot 0.025}{0.364} = 0.04589$$

$$\Rightarrow \alpha = 0.09177 \Rightarrow 1 - \alpha = 0.90823$$

c) La media de la población no es un suceso aleatorio. Antes de calcular el intervalo de confianza, éste es un intervalo aleatorio que tiene probabilidad $1 - \alpha$ de contener a μ :

$$P\left[\bar{x} - t_{n-1, \frac{\alpha}{2}} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{n-1, \frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right] = 1 - \alpha$$

pero una vez obtenido un intervalo de confianza particular, μ pertenecerá o no a él y no se puede hablar de probabilidad de que μ esté dentro o fuera. No obstante, tenemos una

"confianza" $1 - \alpha$ de que μ esté en él, α de que esté fuera y $\frac{\alpha}{2} = 0.04589$ de que sea menor de 4.5.

$$2) (n-1)s^2 = \sum_{i=1}^n (x_i - \bar{x})^2 = 50$$

$$\alpha = 0.04 \Rightarrow \frac{\alpha}{2} = 0.02 \Rightarrow 1 - \frac{\alpha}{2} = 0.98$$

$$I = \left[\sqrt{\frac{(n-1)s^2}{\chi_{n-1, \frac{\alpha}{2}}^2}}, \sqrt{\frac{(n-1)s^2}{\chi_{n-1, 1-\frac{\alpha}{2}}^2}} \right] = \left[\sqrt{\frac{50}{\chi_{14, 0.02}^2}}, \sqrt{\frac{50}{\chi_{14, 0.98}^2}} \right] = \left[\sqrt{\frac{50}{26.873}}, \sqrt{\frac{50}{5.368}} \right] =$$

$$= [1.364, 3.052]$$



PROBLEMA 4

(2'5 puntos)

En una fábrica de pinturas se desechan aquellas mezclas que superen la proporción del 2 por mil (2‰) de agua en volumen.

Si nombramos por X dicha proporción en volumen de agua en la mezcla, en tanto por mil, y por Y la proporción de pintura producida, como fracción respecto del máximo que la fábrica es capaz de producir, la fábrica obtiene la siguiente función de densidad conjunta:

$$f(x,y) = \begin{cases} \frac{x(1+3y^2)}{4} & \forall 0 < x < 2; \quad 0 < y < 1 \\ 0 & \text{resto} \end{cases}$$

Se pide:

- Calcule la proporción media de agua en la mezcla (en ‰), y la fracción media de pintura producida respecto al máximo producible.
- Calcule la función densidad de probabilidad para ambas proporciones por separado.
- ¿Son estadísticamente independientes ambas proporciones?. ¿Incorreladas?. Razone y demuestre matemáticamente sus afirmaciones.

SOLUCIÓN:

$X = \text{‰ en volumen de agua en la mezcla (tanto por mil)}$

$Y = \frac{\text{Pintura producida}}{\text{Máxima pintura producible}} \quad (\text{tanto por uno})$

$$\begin{aligned} \text{a) } E(X) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot f(x,y) dx dy = \int_{x=0}^{x=2} \int_{y=0}^{y=1} x \cdot \frac{x(1+3y^2)}{4} dx dy = \int_{y=0}^{y=1} \left[\frac{x^3}{3} \cdot \frac{(1+3y^2)}{4} \right]_{x=0}^{x=2} dy = \\ &= \frac{2}{3} \int_{y=0}^{y=1} (1+3y^2) dy = \frac{2}{3} [y + y^3]_{y=0}^{y=1} = \frac{4}{3} \end{aligned}$$

$$\begin{aligned} E(Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y \cdot f(x,y) dx dy = \int_{x=0}^{x=2} \int_{y=0}^{y=1} y \cdot \frac{x(1+3y^2)}{4} dx dy = \int_{y=0}^{y=1} \left[\frac{x^2}{2} \cdot \frac{(y+3y^3)}{4} \right]_{x=0}^{x=2} dy = \\ &= \int_{y=0}^{y=1} \frac{y+3y^3}{2} dy = \frac{1}{2} \left[\frac{y^2}{2} + \frac{3y^4}{4} \right]_{y=0}^{y=1} = \frac{1}{2} \left(\frac{1}{2} + \frac{3}{4} \right) = \frac{5}{8} \end{aligned}$$



PROBLEMA 4: Continuación

$$b) f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy \stackrel{\forall 0 < x < 2}{=} \int_{y=0}^{y=1} \frac{x(1+3y^2)}{4} dy = \frac{x}{4} [y + y^3]_{y=0}^{y=1} = \frac{x}{2} \quad \forall 0 < x < 2$$

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

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx \stackrel{\forall 0 < y < 1}{=} \int_{x=0}^{x=2} \frac{x(1+3y^2)}{4} dx = \left[\frac{x^2(1+3y^2)}{8} \right]_{x=0}^{x=2} = \frac{1+3y^2}{2} \quad \forall 0 < y < 1$$

$$f_Y(y) = \begin{cases} \frac{1+3y^2}{2} & \forall 0 < y < 1 \\ 0 & \text{resto} \end{cases}$$

c) Sí son independientes, ya que $f(x, y) = f(x) \cdot f(y) \quad \forall (x, y)$, como a continuación se prueba:

$$\begin{aligned} \forall 0 < x < 2; 0 < y < 1 & \quad f(x) \cdot f(y) = \frac{x}{2} \cdot \frac{1+3y^2}{2} = \frac{x \cdot (1+3y^2)}{4} = f(x, y) \\ \text{resto} & \quad f(x) \cdot f(y) = 0 = f(x, y) \end{aligned}$$

Por ser independientes son incorreladas. Para demostrarlo matemáticamente, calculamos:

$$\begin{aligned} E(X \cdot Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot y \cdot f(x, y) dx dy = \int_{x=0}^{x=2} \int_{y=0}^{y=1} \frac{x^2 y (1+3y^2)}{4} dx dy = \\ &= \frac{1}{4} \int_{y=0}^{y=1} \left[\frac{x^3}{3} (y+3y^3) \right]_{x=0}^{x=2} dy = \frac{2}{3} \int_{y=0}^{y=1} (y+3y^3) dy = \frac{2}{3} \left[\frac{y^2}{2} + \frac{3y^4}{4} \right]_{y=0}^{y=1} = \frac{2}{3} \left[\frac{1}{2} + \frac{3}{4} \right] = \frac{5}{6} \end{aligned}$$

$$\left. \begin{aligned} E(X) &= \frac{4}{3} \\ E(Y) &= \frac{5}{8} \end{aligned} \right\} \quad E(X) \cdot E(Y) = \frac{4}{3} \cdot \frac{5}{8} = \frac{5}{6} = E(X \cdot Y) \quad \text{por lo que son incorreladas}$$

**PROBLEMA 1****(2'5 puntos)**

En un determinado sector, cada empresa se ha especializado en una de las 3 líneas de negocio que existen: A, B ó C. La probabilidad de que una empresa de este sector se haya especializado en las líneas A ó B, es en ambas de 0'3, mientras que en C sería de 0'4. Suponga que en cada mes, la demanda por parte de los consumidores, de los artículos de cada línea de negocio, aumenta o disminuye, pero nunca se queda fija. La probabilidad de que una empresa tenga un aumento en la demanda de sus artículos es de 0'6; pero por otro lado, se sabe que si la empresa está especializada en la línea de negocio C, la probabilidad de que la demanda disminuya, es de 0'8.

- Se ha realizado un estudio sobre la evolución de la demanda de una empresa de este sector, concluyéndose que aumenta. Calcule la probabilidad de que dicha empresa se haya especializado en la línea de negocio C.
- Calcule la probabilidad de que la demanda aumente o la empresa se haya especializado en la línea de negocio B; si se sabe que cuando la demanda disminuye, la probabilidad de que la empresa se haya especializado en la línea A es de 0'1.
- Calcule la probabilidad de que, dado que la demanda disminuye, la empresa se haya especializado en la línea de negocio B.

SOLUCIÓN:

a)

$$P(A) = 0'3$$

$$P(B) = 0'3$$

$$P(C) = 0'4$$

 D = demanda aumenta

$$P(D) = 0'6$$

 D^* = demanda disminuye

$$P(D^*) = 0'4$$

$$P(D^*/C) = 0'8$$

$$P(C/D) = \frac{P(D/C)P(C)}{P(D)} = \frac{(1 - P(D^*/C))P(C)}{P(D)} = \frac{(1 - 0'8)0'4}{0'6} = \boxed{0'133}$$

$$b) P(D \cup B) = P(D) + P(B) + P(D \cap B)$$

$$P(D \cap B) = P(B/D)P(D)$$

$$P(B/D) = 1 - P(A/D) - P(C/D)$$

$$\begin{aligned} P(A/D) &= \frac{P(D/A)P(A)}{P(D)} = \frac{(1 - P(D^*/A))P(A)}{P(D)} = \frac{\left(1 - \frac{P(A/D^*)P(D^*)}{P(A)}\right)P(A)}{P(D)} = \\ &= \frac{\left(1 - \frac{0'1 \cdot 0'4}{0'3}\right)0'3}{0'6} = 0'433 \end{aligned}$$



APELLIDOS Y NOMBRE: _____

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

PROBLEMA 1: Continuación

$$P(B/D) = 1 - 0'433 - 0'133 = 0'433$$

$$P(D \cap B) = 0'433 \cdot 0'6 = 0'26$$

$$P(D \cup B) = 0'6 + 0'3 - 0'26 = \boxed{0'64}$$

c)

$$P(C/D^*) = \frac{P(D^*/C)P(C)}{P(D^*)} = \frac{0'8 \cdot 0'4}{0'4} = 0'8$$

$$P(B/D^*) = 1 - P(A/D^*) - P(C/D^*) = 1 - 0'1 - 0'8 = \boxed{0'1}$$



34'2001

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 13-9-2001 (Mañana)

APELLIDOS Y NOMBRE: _____

Nº DE EXPEDIENTE: _____

CURSO: 3º GRUPO: _____

PROBLEMA 2

(2'5 puntos)

Dada la variable aleatoria bidimensional (X,Y) , su función de densidad de probabilidad conjunta es:

$$f(x,y) = \begin{cases} k_1 y & \text{en } R_1 \\ x/7 & \text{en } R_2 \\ k_2 x & \text{en } R_3 \\ 0 & \text{en resto} \end{cases}$$

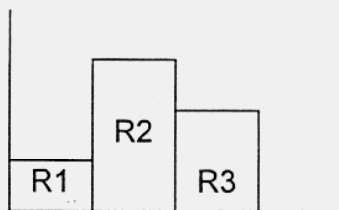
donde: R_1 es la región delimitada por: $0 \leq x < 1$; $0 \leq y \leq 1$

R_2 es la región delimitada por: $1 \leq x \leq 2$; $0 \leq y \leq 3$

R_3 es la región delimitada por $2 < x \leq 3$; $0 \leq y \leq 2$

- a) Calcule k_1 y k_2 , sabiendo que: $1 - F(2,3) = 0'3$
- b) Obtener las funciones de densidad de probabilidad marginales de las variables X e Y .
- c) Calcule la covarianza entre las variables.
- ND d) En función del resultado del apartado anterior, indique cuanto pueda sobre la existencia o no de relación entre ambas variables, y su tipo.

SOLUCIÓN:



a)

$$F(2,3) = 0'7$$

$$\int_0^1 \int_0^1 k_1 y \, dy \, dx + \int_1^2 \int_0^3 \frac{x}{7} \, dy \, dx = 0'7$$

$$\int_0^1 k_1 \left[\frac{y^2}{2} \right]_0^1 \, dx + \int_1^2 \frac{x}{7} [y]_0^3 \, dx = 0'7$$

$$\int_0^1 \frac{k_1}{2} \, dx + \int_1^2 \frac{3x}{7} \, dx = 0'7$$

$$k_1 = 0'116$$



UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA:

13-9-2001 (Mañana)

APELLIDOS Y NOMBRE: _____

Nº DE EXPEDIENTE: _____

CURSO: 3º GRUPO: _____

PROBLEMA 2: Continuación A

$$\int_2^3 \int_0^2 k_2 x \, dy \, dx = 0'3$$

$$\int_2^3 k_2 x [y]_0^2 \, dx = 0'3$$

$$\int_2^3 2k_2 x \, dx = 0'3$$

$$k_2 = 0'06$$

b)

$$\int_0^1 k_1 y \, dy = \frac{k_1}{2}$$

$$\int_0^3 \frac{x}{7} \, dy = \frac{3x}{7}$$

$$\int_0^2 k_2 x \, dy = 2xk_2$$

$$f(x) = \begin{cases} \frac{k_1}{2} & 0 \leq x \leq 1 \\ \frac{3x}{7} & 1 \leq x \leq 2 \\ 2xk_2 & 2 \leq x \leq 3 \\ 0 & \text{resto} \end{cases}$$

$$\int_0^1 k_1 y \, dx + \int_1^2 \frac{x}{7} \, dx + \int_2^3 k_2 x \, dx = k_1 y + \frac{3}{14} + \frac{5k_2}{2}$$

$$\int_1^2 \frac{x}{7} \, dx + \int_2^3 k_2 \, dx = \frac{3}{14} + \frac{5k_2}{2}$$

$$\int_1^2 \frac{x}{7} \, dx = \frac{3}{14}$$

$$f(y) = \begin{cases} k_1 y + \frac{3}{14} + \frac{5k_2}{2} & 0 \leq y \leq 1 \\ \frac{3}{14} + \frac{5k_2}{2} & 1 \leq y \leq 2 \\ \frac{3}{14} & 2 \leq y \leq 3 \\ 0 & \text{resto} \end{cases}$$



32'2001

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID
Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 13-9-2001 (Mañana)

APELLIDOS Y NOMBRE: _____

Nº DE EXPEDIENTE: _____

CURSO: 3º

GRUPO: _____

PROBLEMA 2: Continuación B

c) $Cov(x, y) = \alpha_{11} - \alpha_{10}\alpha_{01}$

$$\alpha_{11} = \int_0^1 \int_0^1 xy^2 k_1 dy dx + \int_0^2 \int_0^3 x^2 y \frac{1}{7} dy dx + \int_0^3 \int_0^2 x^2 y k_2 dy dx = k_1 + 1'5 + 12'66k_2 = 2'3756$$

$$\alpha_{10} = \int_0^1 x \frac{k_1}{2} dx + \int_0^2 \frac{3}{7} x^2 dx + \int_0^2 2k_2 x^2 dx = \frac{k_1}{4} + 1 + 1'66k_2 = 1'788$$

$$\alpha_{01} = \int_0^1 \left(k_1 y^2 + \frac{3y}{14} + \frac{5}{2} k_2 y \right) dy + \int_0^2 \left(\frac{3y}{14} + \frac{5}{2} k_2 y \right) dy + \int_0^3 \frac{3y}{14} dy = k_1 + \frac{25k_2}{4} + \frac{30}{28} = 1'562$$

$$Cov(x, y) = 2'3756 - 1'788 \cdot 1'562 = \boxed{-0'4172}$$

d) Puede resumirse en que ambas variables no son incorreladas ($\sigma_{xy} \neq 0$), y por tanto, tampoco son independientes.

Se observa cierto grado de relación lineal entre las variables X e Y , con pendiente negativa. Es decir la componente lineal de la relación indica una tendencia al aumento de una de las variables cuando disminuye la otra.



PROBLEMA 3

(2'5 puntos)

Una empresa de ámbito nacional le encarga un estudio para establecer la relación entre los gastos en publicidad y las ventas. Para ello le hace entrega de los siguientes datos, en miles de euros:

	gasto en publicidad	ventas
1	20	365
2	20	400
3	20	420
4	25	395
5	25	480
6	30	475
7	40	385
8	40	490
9	40	525
10	50	440
11	50	560
12	50	510
	$\Sigma = 410$	$\Sigma = 5445$

- 1) Obtenga la recta de regresión mínimo-cuadrática para estimar las ventas a partir de los gastos en publicidad, en euros.
- 2) Estime cuáles pueden ser las ventas si los gastos en publicidad son de 35.000 euros.

Nota: Se recomienda que haga un gráfico con la distribución de valores dato, superponiendo las respuestas, para que Vd. mismo compruebe si sus resultados son coherentes con los datos.

SOLUCIÓN:

X: gasto en publicidad

Y: ventas

(ambas en miles de euros)

- a) Nos piden la recta de regresión mínimo-cuadrática $\hat{y} - \mu_Y = \beta_{YX}(x - \mu_X)$

De la tabla, se obtiene: $\mu_Y = \frac{5.445}{12} = 453'75$ $\mu_X = \frac{410}{12} = 34'1666667$

A la hora de obtener $\beta_{YX} = \frac{\sigma_{XY}}{\sigma_X^2}$, recordaremos que:

$$\sigma_{XY} = E(X \cdot Y) - \mu_X \cdot \mu_Y \qquad \sigma_X^2 = E(X^2) - \mu_X^2$$

Dado cómo tenemos los datos, podemos bien recolectarlos en una tabla bidimensional, para obtener las frecuencias de cada par y luego utilizar las expresiones estándar

$$E(X \cdot Y) = \sum_i \sum_j x_i y_j \frac{n_{ij}}{N} \qquad E(X^2) = \sum_i x_i^2 \frac{n_{X_i}}{N}$$

o bien hacer la recolección de las frecuencias por suma de valores obtenidos:



30'2001

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID
Departamento de Electrónica y Comunicaciones
ESTADÍSTICA: 13-9-2001 (Mañana)

APELLIDOS Y NOMBRE: _____

Nº DE EXPEDIENTE: _____

CURSO: 3º

GRUPO: _____

PROBLEMA 3: Continuación

$$E(X \cdot Y) = \frac{20 \cdot 365 + 20 \cdot 400 + 20 \cdot 420 + 25 \cdot 395 + \dots + 50 \cdot 510}{12} = \frac{191.325}{12} = 15.943'75$$

$$E(X^2) = \frac{20^2 + 20^2 + 20^2 + 25^2 + \dots + 50^2}{12} = \frac{15.650}{12} = 1.304'16667$$

Y, por tanto:

$$\sigma_{XY} = 15.943'75 - 34'1666667 \cdot 453'75 = 440'625$$

$$\sigma_X^2 = 1.304'16667 - 34'1666667^2 = 136'805556$$

Así, $\beta_{YX} = \frac{\sigma_{XY}}{\sigma_X^2} = 3'220812186$

$$\hat{y} - 453'75 = 3'2208 \cdot (x - 34'1667) \Rightarrow$$

$$\hat{y} = 3'2208 \cdot x + 343'706$$

(miles de euros)

b) $\hat{y} = 3'2208 \cdot 35 + 343'706 = 456'434$

\Rightarrow

$$456.434 \text{ euros}$$



PROBLEMA 4

(2'5 puntos)

Una empresa ha desarrollado un pegamento y afirma que resiste a la tracción 5 Tm/cm^2 en término medio, cuando se utiliza entre ciertos metales. Ahora bien, se conoce el dato de que su desviación típica es de 120 Kg/cm^2 .

Para probar que la media no es menor que dichas 5 Tm/cm^2 , un comprador toma una muestra aleatoria de 50 uniones. La región crítica que obtiene es $\bar{x} < 4970$.

Se pide:

- Calcule la probabilidad de que el comprador cometa error tipo I.
- Obtenga la potencia de su contraste para $\mu = 4970$ y para $\mu = 4960$.

SOLUCIÓN:

Se ha realizado un contraste unilateral (la región crítica es $\bar{x} < \text{valor}$) para el valor medio de una población, con muestra suficientemente grande como para aproximar mediante el Teorema Central del Límite ($n=50$) y desviación típica conocida ($\sigma = 120 \text{ Kg/cm}^2$)

$$H_0: \mu = 5.000 \text{ Kg/cm}^2 \quad \text{Reg. Aceptación: } \bar{x} \geq 4.970 \text{ Kg/cm}^2$$

$$H_A: \mu < 5.000 \text{ Kg/cm}^2 \quad \text{Reg. Crítica: } \bar{x} < 4.970 \text{ Kg/cm}^2$$

$$\begin{aligned} \text{a) } \alpha &= P(\varepsilon_{\text{Tipol}}) = P(\text{rechazo } H_0 | H_0 \text{ cierto}) = P(\bar{x} < 4970 | \mu = 5000) = P\left(z < \frac{4970 - 5000}{120/\sqrt{50}}\right) = \\ &= P\left(z < -\frac{\sqrt{50}}{4}\right) = P\left(z > \frac{\sqrt{50}}{4}\right) \approx P(z > 1.7678) \approx P(z > 1.77) \underset{\text{interpolando}}{=} = 0.03858 \end{aligned}$$

$$\text{b) } 1 - \beta|_{\mu_a} = \text{POT}|_{\mu_a} = P(\text{rechazo } H_0 | \mu = \mu_a) = P(\bar{x} < 4970 | \mu = \mu_a) = P\left(z < \frac{4970 - \mu_a}{120/\sqrt{50}}\right) =$$

$$\Rightarrow 1 - \beta|_{\mu_a} = 1 - P\left(z > \frac{4970 - \mu_a}{120/\sqrt{50}}\right)$$

* $\mu_a = 4970$

$$\Rightarrow 1 - \beta|_{\mu=4970} = \text{POT}|_{\mu=4970} = 1 - P(Z > 0) = 1 - 0.5 = 0.5 \quad \Rightarrow \quad 1 - \beta|_{\mu=4970} = 0.5$$

* $\mu_a = 4960$

$$\Rightarrow 1 - \beta|_{\mu=4960} = \text{POT}|_{\mu=4960} = 1 - P\left(z > \frac{\sqrt{50}}{12}\right) = 1 - P(0.5893) = 1 - P(0.59) = 1 - 0.2776$$

$$\Rightarrow 1 - \beta|_{\mu=4960} = 0.7224$$

**PROBLEMA 1**

(2'5 puntos)

En una fábrica se produce un electrodoméstico cuyo tiempo de vida útil, X , en horas de funcionamiento, sigue una distribución normal. El 20% de ellos tiene una duración de al menos 3200 horas, pero el 30% no llega a las 2900.

- 0'8 a) ¿Cuál es la duración media, en horas de funcionamiento, de estos electrodomésticos?
- b) Si se elige un electrodoméstico al azar, calcule la probabilidad de que su duración:
- 0'1 b.1) Sea inferior a 2750 horas
- 0'1 b.2) Sea superior a 3250 horas
- 0'2 b.3) Esté comprendida entre 2800 y 3100 horas
- 0'2 b.4) Difiere de la duración media en menos de 200 horas
- 0'3 b.5) Exceda a la duración media al menos en 1'5 veces la desviación típica de X .
- 0'7 c) Se quiere dar un periodo de garantía, de forma que los electrodomésticos que se averíen en ese periodo serán sustituidos por otros nuevos. ¿Cuál debe ser ese periodo si se quiere que no haya que reemplazar más del 5% de los electrodomésticos?

Nota: Si tiene que utilizar alguna tabla, interpole la primer vez que la use; las siguientes veces aproxime al valor más cercano, o al valor medio entre dos.

SOLUCIÓN:

- a) Sabemos que la v.a. "duración del electrodoméstico" o "tiempo de vida útil", tiene distribución normal, es decir, $X: N(\mu, \sigma)$, pero no conocemos ni μ ni σ .

El 20% tiene duración de al menos 3200 horas quiere decir $P(X > 3200) = 0.2$, luego:

$$0.2 = P\left(\frac{x - \mu}{\sigma} \geq \frac{3200 - \mu}{\sigma}\right) = P\left(Z \geq \frac{3200 - \mu}{\sigma}\right) \text{ donde } Z: N(0,1) \Rightarrow \frac{3200 - \mu}{\sigma} = z_{\alpha} = z_{0.2}$$

De las propiedades de la normal, al ser $P(Z \geq z_{\alpha}) = \alpha = 0.2 < 0.5$, esto implica que $z_{\alpha} = z_{0.2} > 0$. Calculemos $z_{0.2}$, interpolando en la tabla de la $N(0,1)$: $Z_{0.1977} = 0.85$; $Z_{0.2005} = 0.84$

$$\frac{0.85 - 0.84}{0.2005 - 0.1977} = \frac{z_{0.2} - 0.84}{0.2005 - 0.20} \Rightarrow z_{0.2} = 0.84 + \frac{(0.2005 - 0.2)(0.85 - 0.84)}{0.2005 - 0.1977} = 0.841786$$

luego: $3200 - \mu = 0.841786 \sigma$ (*)

Por otra parte, que el 30% no llegue a una duración de 2900 horas quiere decir que $P(X < 2900) = 0.3$, luego:

$$P\left(\frac{x - \mu}{\sigma} \leq \frac{2900 - \mu}{\sigma}\right) = 0.3 \Rightarrow P\left(Z \leq \frac{2900 - \mu}{\sigma}\right) = 0.3, \text{ donde } Z: N(0,1).$$

Por las propiedades de la normal, al ser $P\left(Z \leq \frac{2900 - \mu}{\sigma}\right) = 0.3 < 0.5$, esto implica que

$$\frac{2900 - \mu}{\sigma} \leq 0, \text{ luego } P\left(Z \leq \frac{2900 - \mu}{\sigma}\right) = P\left(Z \geq \frac{\mu - 2900}{\sigma}\right) = 0.3.$$



PROBLEMA 3

(2'5 puntos)

Se quiere realizar un estudio sobre una variable aleatoria X , de la que se sabe que sigue una distribución normal $N(\mu, 3)$.

- Obtenga un intervalo de confianza del 95% para la media de la variable X , teniendo una muestra de tamaño 22, con valores $\bar{x} = 17$ y $s = 2'8$.
- Con el mismo nivel de confianza, qué tamaño mínimo de muestra sería necesario para que la longitud del intervalo de confianza no fuese mayor de 4.
- Estime con qué probabilidad puede asegurarse que no se cometerá un error superior a 1'5, al estimar la media de la variable aleatoria X , mediante la media muestral del apartado a), con la muestra de $n=22$.

Suponga una variable aleatoria Y que se distribuye normal, y de la que conoce los valores muestrales:

3'1 2'3 7'4 5'6 4'8

- Obtenga los valores fuera de los cuales sólo hay una probabilidad del 5% (2'5% a cada lado) de que se encuentre su varianza.

Nota: Si tiene que utilizar alguna tabla, interpole la primer vez que la use; las siguientes veces aproxime al valor más cercano, o al valor medio entre dos.

SOLUCIÓN:

- Como σ es conocida, la cuasidesviación típica muestral no se tiene en cuenta, y por tanto, la expresión adecuada para calcular el intervalo de confianza es:

$$\bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 17 \pm 1'96 \frac{3}{\sqrt{22}} = 17 \pm 1'253 = (15'747, 18'253) \quad \text{donde } Z_{0'025} = 1'96$$

- Dado que la longitud del intervalo es 4, la semiamplitud es 2, es decir, superior a 1'253 que era la semiamplitud obtenida en el apartado anterior. Por tanto, el tamaño mínimo necesario será inferior. La forma de calcularlo sería:

$$E = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \Rightarrow 2 = 1'96 \cdot \frac{3}{\sqrt{n}} \quad n = \left(\frac{1'96 \cdot 3}{2} \right)^2 = 8'64 \quad \text{Luego: } \boxed{n \geq 9}$$

- Al tener una mayor semiamplitud que en el apartado a), con un mismo tamaño de muestra, la probabilidad tiene que ser mayor. La forma de calcularlo sería:

$$E = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \Rightarrow 1'5 = Z_{\alpha/2} \frac{3}{\sqrt{22}} \quad Z_{\alpha/2} = 2'3452$$

$$\frac{\alpha}{2} = 0'009 \quad \alpha = 0'018 \quad \boxed{1 - \alpha = 0'982}$$

Es decir, se obtiene un nivel de confianza del 98'2%.

$$d) I = \left[\frac{(n-1)s^2}{\chi^2_{\alpha/2; n-1}}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2; n-1}} \right] \quad \bar{x} = \frac{3'1 + 2'3 + 7'4 + 5'6 + 4'8}{5} = 4'64$$

$$\frac{\alpha}{2} = 0'025$$

$$1 - \frac{\alpha}{2} = 0'975$$

$$n = 5$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{4} (2'37 + 5'47 + 7'61 + 0'92 + 0'025) = 4'099$$

$$I = \left[\frac{4 \cdot 4'099}{11'143}, \frac{4 \cdot 4'099}{0'484} \right] = (1'471, 33'876)$$



PROBLEMA 4

(2'5 puntos)

Una fábrica de componentes electrónicos recibe el envío de un lote de unidades de material químico con impurezas. Podemos suponer que el porcentaje de impurezas en estas unidades sigue una distribución normal. Es importante para la fábrica asegurarse de que la varianza del porcentaje de impurezas en las unidades de este lote no sea mayor de 0'2.

Se ha tomado una muestra de 25 unidades, obteniendo de ellas $\bar{x} = 2'02$ y $s^2 = 0'27$, en los porcentajes de impurezas. Se pide:

- Razone si se debe aceptar ese envío, si queremos tener una probabilidad del 1% de rechazar envíos que cumplan lo indicado.
- Si basándonos en dicha muestra tuviéramos que comprobar, con significación igual a la anterior, que el porcentaje medio de impurezas que tuviesen estas unidades no superase 1'8%, ¿qué decisión tomaríamos sobre el envío?
- Indique qué potencia tendría este último contraste, caso de que la proporción media de impurezas en el envío fuese realmente de 2'32.

SOLUCIÓN:

No se trata de evaluar ninguna proporción poblacional, sino estimaciones al rededor de la variable aleatoria normal $X = \text{"porcentaje de impurezas en cada unidad de material químico"}$

- Contraste unilateral de la varianza en población normal: $\begin{cases} H_0 : \sigma^2 \leq \sigma_0^2 = 0'2 \\ H_A : \sigma^2 > \sigma_0^2 = 0'2 \end{cases}$

$$\text{Rechazaremos si } \frac{(n-1) \cdot s_R^2}{\sigma_0^2} > \chi_{n-1, \alpha}^2, \text{ con } \alpha = 1\% = 0'01 \quad \left(\equiv s_R^2 > \chi_{n-1, \alpha}^2 \frac{\sigma_0^2}{(n-1)} \right)$$

$$\frac{(n-1) \cdot s_R^2}{\sigma_0^2} = \frac{24 \cdot 0'27}{0'2} = 32'4 < \chi_{24, 0'01}^2 = 42'98$$

Al no cumplirse la condición de rechazo, podemos concluir que no puede rechazarse la hipótesis nula con un nivel de significación del 1%. No hay suficiente evidencia en contra de la hipótesis de que la varianza de la proporción de impurezas supere el 0'2%.

Se aceptará el envío.

- Contraste unilateral de la media en población normal: $\begin{cases} H_0 : \mu \leq \mu_0 = 0'018 = 1'8\% \\ H_A : \mu > \mu_0 = 0'018 = 1'8\% \end{cases}$

Podemos trabajar todo en fracciones (tantos por uno), o todo en porcentajes.

Desconocemos σ , y la muestra es de $n = 25 < 30$, por tanto, rechazaremos si

$$\frac{\bar{x}_R - \mu_0}{s_R / \sqrt{n}} > t_{n-1, \alpha} \text{ es decir, si } \bar{x}_R \text{ supera el Punto Crítico: } \bar{x}_R > PC = \mu_0 + t_{n-1, \alpha} \frac{s_R}{\sqrt{n}}$$

$$\text{En nuestro caso, } \bar{x}_R = 2'02\% \quad \text{y} \quad PC = \mu_0 + t_{n-1, \alpha} \frac{s_R}{\sqrt{n}} = 1'8 + 2'192 \frac{\sqrt{0'27}}{\sqrt{25}} = 2'0278\%$$

Así, $\bar{x}_R = 2'02\% \leq PC = 2'0278\%$. Al no cumplirse la condición de rechazo, no podemos rechazar la hipótesis nula con un nivel de significación del 1%. No hay suficiente evidencia contra la hipótesis de que la proporción media de impurezas supere el 1'8%.

Se aceptará el envío.



APELLIDOS Y NOMBRE: _____

Nº DE EXPEDIENTE: _____

CURSO: 3º

GRUPO: _____

PROBLEMA 4: Continuación

c) La probabilidad de rechazarlo siendo $\mu = 2'32$, será:

$$1 - \beta(\mu = 2'32) = P(\text{Rechazar } H_0 | \mu = 2'32) = P(\bar{x} > PC | \mu = 2'32) \underset{t_{n-1} = \frac{\bar{x} - \mu}{s/\sqrt{n}}}{=} P\left(t_{n-1} > \frac{PC - \mu_a}{s/\sqrt{n}}\right)$$

Se trata del contraste unilateral anterior, en que $H_0: \mu \leq \mu_0 = 1'8\%$, contra la hipótesis alternativa concretada en $H_A: \mu_a = 2'32\% > \mu_0$

De b), tenemos que el Punto Crítico es $PC = \mu_0 + t_{\alpha, n-1} \frac{s_R}{\sqrt{n}}$

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

$$1 - \beta(\mu = 2'32) = P\left(t_{24} > \frac{1'8 - 2'32}{\sqrt{0'27}} \sqrt{25} + 2'192\right) = P(t_{24} > -2'8117) = 1 - P(t_{24} > 2'8117)$$

$$1 - \beta(\mu = 2'32) = 1 - \alpha(2'8117)$$

Sobre la tabla de la T-Student t_{24} , podemos apreciar que $0'005 < \alpha(2'8117) < 0'001$, en un punto muy cercano al extremo inferior. Tomándolo como aproximación: $\alpha(2'8117) \cong 0'005$

Por tanto: $1 - \beta(\mu = 2'32) \cong 99'5\%$



PROBLEMA 1

(2'5 puntos)

Un detector de intrusos por ultrasonidos decide la activación o no de una alarma, dependiendo del tipo de eco recibido. Este eco se recibe como una variable aleatoria X , con distribución normal $N(0, \sigma_0)$ cuando no hay presencia de intruso, y $N(0, \sigma_1)$ cuando la hay.

En el receptor, la variable X del eco se trabaja electrónicamente, obteniendo una nueva variable Y , que toma el valor 1 cuando X supera en valor absoluto un determinado umbral L , y -1 cuando no lo supera. Es decir:

$$Y = \begin{cases} -1 & \forall |x| \leq L \\ 1 & \forall |x| > L \end{cases}$$

a) Calcule el valor del umbral L para que Y sea de media nula cuando no haya intruso.

Para el umbral L obtenido en el apartado anterior:

b) Calcule la varianza de Y , para el caso en que no haya intruso.

c) Calcule la media y la varianza de Y , para el caso en que sí haya intruso, y $\sigma_1 = 3L$.

Dejando a un lado los resultados de los tres apartados anteriores, y continuando con el enunciado, suponga a partir de ahora que la variable Y tiene $E[Y] = 0$, $\sigma_Y^2 = 1$ cuando no hay intruso; y $E[Y] = 0.5$, $\sigma_Y^2 = 0.75$ cuando lo hay.

Se emiten 100 pulsos, obteniendo cien variables independientes $Y_1, Y_2, Y_3, \dots, Y_{100}$, todas con igual distribución a la de Y ; y se suman las 100 variables Y_i obtenidas, generando la variable suma S . El receptor genera alarma cuando dicha suma S supera el umbral de suma L_S .

d) Calcule L_S para que la probabilidad de dar alarma sin haber intruso (falsa alarma) sea 0.01.

e) Suponiendo que utilizáramos un umbral de suma $L_S = 30$, calcule la probabilidad de dar alarma caso de que hubiera intruso (alarma correcta).

Nota: Si tiene que utilizar alguna tabla, interpole la primer vez que la use; las siguientes veces aproxime al valor más cercano, o al valor medio entre dos.

SOLUCIÓN:

a) No habiendo intruso, $X_0 : N(0, \sigma_0) \Rightarrow Y|X_0 = Y_0$

$$E[Y|X_0] = E[Y_0] = (-1) \cdot P(Y_0 = -1) + (1) \cdot P(Y_0 = 1) = -P(|X_0| \leq L) + P(|X_0| > L)$$

$$\text{Siendo } P(|X_0| > L) = P(X_0 < -L) + P(X_0 > L) = 2 \cdot P(X_0 > L)$$

$$P(|X_0| \leq L) = P(X_0 > -L) + P(X_0 < L) = 1 - 2 \cdot P(X_0 > L)$$

La esperanza, que debe ser nula, queda:

$$E[Y_0] = -1 + 4 \cdot P(X_0 > L) = 4 \cdot P\left(Z > \frac{L}{\sigma_0}\right) - 1 = 0 \Rightarrow P\left(Z > \frac{L}{\sigma_0}\right) = \frac{1}{4} = 0.25$$

Entramos en la tabla de la normal, y sin más que realizar interpolación intermedia entre los valores, tendremos $\frac{L}{\sigma_0} = 0.675 \Rightarrow L = 0.675 \cdot \sigma_0$

A este mismo resultado puede llegarse razonando de la siguiente manera:

$$E[Y|X_0] = E[Y_0] = (-1) \cdot P(Y_0 = -1) + (1) \cdot P(Y_0 = 1) = 0 \Rightarrow \begin{matrix} P(Y_0 = -1) = P(Y_0 = 1) = 0.5 \\ P(Y_0 = -1) + P(Y_0 = 1) = 1 \end{matrix} \Rightarrow$$

$$\Rightarrow P(Y_0 = 1) = P(|X_0| > L) = 2 \cdot P(X_0 > L) = 0.5 \Rightarrow P\left(Z > \frac{L}{\sigma_0}\right) = \frac{1}{4} = 0.25 \dots$$



PROBLEMA 1: Continuación

b) No habiendo intruso, $X_0 : N(0, \sigma_0)$

$$\begin{aligned}\sigma_{Y_0}^2 &= E[Y_0^2] - E^2[Y_0] = E[Y_0^2] - 0 = E[Y_0^2] = (-1)^2 \cdot P(Y_0 = -1) + (1)^2 \cdot P(Y_0 = 1) = \\ &= P(Y_0 = -1) + P(Y_0 = 1) = P(X_0 \leq L) + P(X_0 > L) = 1 - 2 \cdot P(X_0 > L) + 2 \cdot P(X_0 > L) = 1\end{aligned}$$

$$\sigma_{Y_0}^2 = 1$$

c) De idéntica manera, pero habiendo intruso, $X_1 : N(0, \sigma_1)$

$$\Rightarrow Y|X_1 = Y_1$$

$$E[Y_1] = 4 \cdot P\left(Z > \frac{L}{\sigma_1}\right) - 1 = 4 \cdot P\left(Z > \frac{1}{3}\right) - 1 = 4 \cdot 0.37 - 1 = 0.48$$

$$\Rightarrow E[Y_1] = 0.48$$

$$E[Y_1^2] = 1$$

$$\sigma_{Y_1}^2 = E[Y_1^2] - E^2[Y_1] = 1 - (0.48)^2 = 0.7696$$

$$\Rightarrow \sigma_{Y_1}^2 = 0.7696$$

d) Si bien X_i son normales, Y_i no lo son; es más, son variables discretas. Pero todas las variables Y_i son independientes, tienen la misma distribución, y son 100 (más de 30), por lo que podemos aplicar el Teorema Central del Límite: $S = \sum_{i=1}^{100} Y_i$ seguirá una distribución normal. $S : N(\mu_S, \sigma_S)$

$$\text{No hay intruso} \Rightarrow \left. \begin{aligned} E[Y_i] &= 0 \\ \sigma_{Y_i}^2 &= 1 \end{aligned} \right\} \Rightarrow \left\{ \begin{aligned} \mu_S &= \sum_{i=1}^{100} \mu_{Y_i} = 0 \\ \sigma_S &= \sqrt{\sum_{i=1}^{100} \sigma_{Y_i}^2} = \sqrt{100} = 10 \end{aligned} \right.$$

$$P(S > L_S) = P\left(Z > \frac{L_S - 0}{10}\right) = 0.01 \Rightarrow \frac{L_S}{10} = Z(0.01) \cong 2.33 \Rightarrow L_S \cong 23.3$$

e) Habiendo intruso:

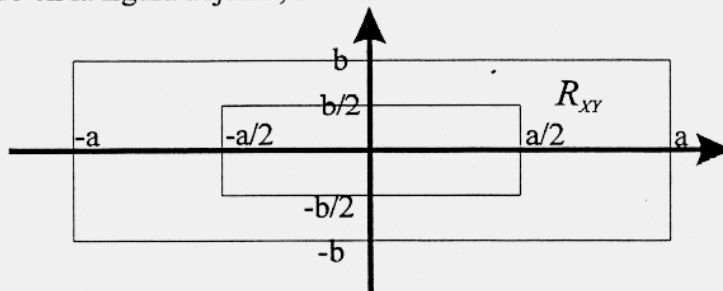
$$\left. \begin{aligned} E[Y_i] &= 0.5 \\ \sigma_{Y_i}^2 &= 0.75 \end{aligned} \right\} \Rightarrow \left\{ \begin{aligned} \mu_S &= \sum_{i=1}^{100} \mu_{Y_i} = 0.5 \cdot 100 = 50 \\ \sigma_S &= \sqrt{\sum_{i=1}^{100} \sigma_{Y_i}^2} = \sqrt{100 \cdot 0.75} = \sqrt{75} \end{aligned} \right.$$

$$P(S > L_S) = P(S > 30) = P\left(Z > \frac{30 - 50}{\sqrt{75}}\right) = P(Z > -2.3094) = 1 - P(Z > 2.3094) \cong 1 - 0.0104 = 0.9896$$

$$P(S > L_S) \cong 0.9896$$

**PROBLEMA 2****(2'5 puntos)**

Una variable aleatoria bidimensional (X,Y) tiene su f.d.p. distribuida uniformemente en el recinto R_{XY} indicado en la figura adjunta, siendo nulo en el resto.



Se pide:

- Obtenga dicha f.d.p.
- Obtenga la esperanza de las variables X e Y .
- Compruebe si las variables X e Y están o no incorreladas.
- Calcule la f.d.p. de X y la de Y .
- Compruebe si las variables X e Y son independientes.

Nota: Puede referirse a los puntos de R_{XY} como $\forall(x,y) \in R_{XY}$

SOLUCIÓN:

a) $f(x,y) = \begin{cases} K & \forall(x,y) \in R_{XY} \\ 0 & \text{resto} \end{cases}$ para calcular K , imponemos $1 = \iint_{R_{XY}} K \cdot dx \cdot dy$

$$1 = K \cdot \left[\int_{-a}^{-a/2} dx \int_{-b}^{-b/2} dy + \int_{-a/2}^{a/2} dx \int_{-b}^{-b/2} dy + \int_{-a/2}^{a/2} dx \int_{b/2}^b dy + \int_{a/2}^a dx \int_{b/2}^b dy \right] =$$

$$= K \cdot \left[\frac{a}{2} \cdot 2b + a \cdot \frac{b}{2} + a \cdot \frac{b}{2} + \frac{a}{2} \cdot 2b \right] = K \cdot 3ab \Rightarrow$$

$$f(x,y) = \begin{cases} \frac{1}{3ab} & \forall(x,y) \in R_{XY} \\ 0 & \text{resto} \end{cases}$$

Nota: También podría realizarse el cálculo de K , basándonos en el área del recinto, ya que la función densidad es constante. $(1 = V = K \cdot A_{R_{XY}} = K \cdot 3ab)$

$$\text{Area} = 2a \cdot 2b - ab = 3ab$$

- b) Realizando los cálculos, o por simetría, observamos que ambas esperanzas son nulas

$$E(X) = \iint_{R_{XY}} Kx \, dx \, dy = K \cdot \left[\int_{-a}^{-a/2} x \, dx \int_{-b}^{-b/2} dy + \int_{-a/2}^{a/2} x \, dx \int_{-b}^{-b/2} dy + \int_{-a/2}^{a/2} x \, dx \int_{b/2}^b dy + \int_{a/2}^a x \, dx \int_{b/2}^b dy \right] = 0$$

$$E(Y) = \iint_{R_{XY}} Ky \, dx \, dy = K \cdot \left[\int_{-a}^{-a/2} dx \int_{-b}^{-b/2} y \, dy + \int_{-a/2}^{a/2} dx \int_{-b}^{-b/2} y \, dy + \int_{-a/2}^{a/2} dx \int_{b/2}^b y \, dy + \int_{a/2}^a dx \int_{b/2}^b y \, dy \right] = 0$$



PROBLEMA 2: Continuación

$$\begin{aligned} \text{c) } E(XY) &= \iint_{R_{XY}} K \cdot xy \, dx dy = \\ &= K \cdot \left[\int_{-a}^{-a/2} x \, dx \int_{-b}^b y \, dy + \int_{-a/2}^{a/2} x \, dx \int_{-b}^{-b/2} y \, dy + \int_{-a/2}^{a/2} x \, dx \int_{b/2}^b y \, dy + \int_{a/2}^a x \, dx \int_{-b}^{-b/2} y \, dy \right] = \\ &= K \cdot [0 + 0 + 0 + 0] = 0 \end{aligned}$$

Así, X e Y son incorreladas, dado que $E(X \cdot Y) = E(X) \cdot E(Y)$ ($0 = 0 \cdot 0$)

$$\text{d) } f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy = \begin{cases} \int_{-b}^b K \, dy = 2Kb = \frac{2}{3a} & \forall \frac{a}{2} \leq |x| \leq a \\ \int_{-b/2}^{b/2} K \, dy + \int_{b/2}^b K \, dy = K\left(\frac{b}{2} + \frac{b}{2}\right) = Kb = \frac{1}{3a} & \forall |x| \leq \frac{a}{2} \\ 0 & \text{resto} \end{cases}$$

Idem para Y , sustituyendo x por y , y a por b :

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) dx = \begin{cases} \frac{2}{3b} & \forall \frac{b}{2} \leq |y| \leq b \\ \frac{1}{3b} & \forall |y| \leq \frac{b}{2} \\ 0 & \text{resto} \end{cases}$$

e) A pesar de ser incorreladas, no son independientes, ya que no se cumple que

$$f_{XY}(x, y) = f_X(x) \cdot f_Y(y) \quad \forall (x, y)$$

Se puede comprobar de forma genérica, multiplicando ambas funciones, para comprobar que no se cumple; pero será suficiente con comprobar que esta condición se incumple en un punto. Por ejemplo:

$$f_{XY}(0,0) = 0 \neq f_X(0) \cdot f_Y(0) = \frac{1}{3a} \cdot \frac{1}{3b}$$

**PROBLEMA 3****(2'5 puntos)**

En una muestra de 25 individuos que tienen automóvil, se realiza una encuesta sobre lo que pagan por el seguro del coche. El valor medio obtenido es de 140.000 pts. La cantidad X que se paga por el seguro del coche, sigue una distribución normal.

- Obtenga un intervalo de confianza del 95% para el valor medio de la cantidad que se paga por el seguro del coche, si se sabe que la desviación típica poblacional de la variable aleatoria X es de 30.000 pts. Explique qué significa el resultado obtenido.
- Suponga que no conoce la desviación típica poblacional: obtenga un intervalo de confianza del 99% para el valor medio de lo que se paga por el seguro del coche, si la raíz cuadrada de la cuasivarianza de la muestra es de 12.000 pts. Dé la solución en miles de pts., con al menos 3 decimales.
- Qué tamaño mínimo muestral se necesitaría (suponiendo la situación del apartado b)), para que la diferencia entre la media muestral y la poblacional sea en valor absoluto, inferior a 8 (en miles de pts.).

SOLUCIÓN:

$$a) I = \bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}; \quad \frac{\alpha}{2} = 0'025; \quad z_{0'025} = 1'96; \quad I = 140 \pm 1'96 \frac{30}{\sqrt{25}} = [128'24, 151'76] \text{ K pts.}$$

Luego, con un nivel de confianza del 95%, podemos afirmar que la media poblacional se encuentra entre 128'24 y 151'76.

- b) No se conoce la desviación típica poblacional y la muestra es menor que 30, luego

$$I = \bar{x} \pm t_{\alpha/2, n-1} \frac{S}{\sqrt{n}}; \quad \frac{\alpha}{2} = 0'005; \quad t_{0'005, 24} = 2'797; \quad I = 140 \pm 2'797 \frac{12}{\sqrt{25}} = [133'287, 146'712] \text{ K pts.}$$

- c) Hay que ir evaluando la semiamplitud tanteando para diferentes "n", debido a que la "n" afecta a en dos sentidos. Como la situación es la del apartado b), y ahora el intervalo es algo mayor, la "n" tiene que ser menor, a la hora de ir probando.

$$\text{Semiamplitud} = t_{\alpha/2, n-1} \frac{S}{\sqrt{n}}$$

$$n = 19 \quad t_{0'005, 18} \frac{12}{\sqrt{19}} = 7'924$$

$$n = 18 \quad t_{0'005, 17} \frac{12}{\sqrt{18}} = 8'198$$

Luego el tamaño mínimo muestral es de 19, ya que con menos tamaño de muestra, la diferencia entre la media muestral y la poblacional supera a 8 en valor absoluto.



16'2001
UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA:

6-6-2001 (Mañana)

APELLIDOS Y NOMBRE: _____

Nº DE EXPEDIENTE: _____

CURSO: 3º GRUPO: _____

PROBLEMA 4

(2'5 puntos)

Una empresa fabrica un determinado tipo de piezas de maquinaria industrial. El grosor medio de estas piezas debe ser de 7 cm, pero debido a las pruebas de calidad realizadas sobre las piezas, se sospecha que dicho grosor queda reducido (en cualquier caso, el grosor de estas piezas nunca puede aumentar), por lo que no se ajustaría a las especificaciones técnicas, teniendo que tomar decisiones sobre las pruebas de calidad. Se sabe que el grosor de las piezas es normal.

- a) Realice un contraste de hipótesis e indique a qué conclusiones llegaría, asumiendo una probabilidad de Error Tipo I del 5%. Para ello dispone de una muestra de 23 piezas, sobre

las que se ha obtenido: $\sum_{i=1}^{23} X_i = 157$, $s^2 = 0'04$

- b) Se plantea la posibilidad de aumentar el tamaño de la muestra para la realización del contraste, en función de cómo responda ante piezas con grosor medio real de 6'8.

Calcule la probabilidad de concluir que el grosor medio de las piezas es 7 cm y valore si interesa o no aumentar dicha muestra.

Nota: Si tiene que utilizar alguna tabla, interpole la primer vez que la use; las siguientes veces aproxime al valor más cercano, o al valor medio entre dos.

SOLUCIÓN:

- a)

$$\bar{x} = \frac{\sum_{i=1}^{23} X_i}{23} = \frac{157}{23} = 6'82 \quad s^2 = 0'04 \Rightarrow s = 0'2$$

El contraste es unilateral: $H_0: \mu = 7 \text{ cm}$ $\alpha = 0'05$
 $H_a: \mu < 7 \text{ cm}$

Como no se conoce la desviación típica de la población y el tamaño de muestra es inferior

a 30, el estadístico $\frac{\bar{x} - \mu}{s/\sqrt{n}}$ sigue una t de Student:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{6'82 - 7}{0'2/\sqrt{23}} = -4'316$$

**PROBLEMA 4: Continuación**

Por otro lado: $t_{\alpha;n-1} = t_{0.05;22} = 1.717$

La región de aceptación corresponde a: $t > -t_{\alpha;n-1}$, pero -4.316 no es mayor que -1.717 . Por tanto, se rechaza la hipótesis nula con un nivel de significación del 5%.

Con una muestra de $n=23$, se concluye que con una probabilidad de error inferior al 5%, el grosor esperado es inferior a 7 cm, por lo que habrá que tomar alguna decisión sobre las pruebas de calidad, ya que afectan al grosor de las piezas.

(No superior al 5%, porque sería el 5% si la hipótesis nula fuera cierta, cuestión que no conocemos; luego por probabilidad total, ese 5% debe ir ponderado por la probabilidad de que realmente sea cierta la hipótesis nula, lo que reduce ese 5%).

$$b) \quad \beta = P\left(\text{Aceptar } H_0 / \mu = \mu_a\right) = P(\bar{x}_a \geq LI) = P\left(\bar{x}_a \geq \mu_0 - t_{\alpha;n-1} \frac{s}{\sqrt{n}}\right) =$$

$$P\left(\frac{\bar{x}_a - \mu_a}{s/\sqrt{n}} \geq \frac{\mu_0 - t_{\alpha;n-1} \frac{s}{\sqrt{n}} - \mu_a}{s/\sqrt{n}}\right) = P\left(t_{n-1} \geq -t_{\alpha;n-1} - \frac{(\mu_a - \mu_0)}{s/\sqrt{n}}\right) =$$

$$P\left(t_{22} \geq -1.717 - \frac{(6.8 - 7)}{0.2/\sqrt{23}}\right) = P(t_{22} \geq 3.079) = \boxed{0.005}$$

La probabilidad de error tipo II es del 0.5%, luego la potencia de contraste es del 99.5%.

Teniendo en cuenta que la potencia es muy elevada, existe poco margen de mejora; y como al aumentar la muestra se incrementa el coste del contraste, no parece recomendable utilizar una muestra mayor.



14'2001

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 29-2-2001 (Tarde)

APELLIDOS Y NOMBRE: _____

Nº DE EXPEDIENTE: _____

CURSO: 3º

GRUPO: _____

PROBLEMA 1

(2'5 puntos)

Dos equipos A y B están conectados mediante un sistema de transmisión bidireccional. Este sistema utiliza un código con tres símbolos: α , β y γ .

El sistema no está libre de ruido, que resulta independiente de los símbolos transmitidos, lo que hace que la probabilidad de que uno de estos símbolos sea interpretado correctamente en el equipo receptor es 0'6 (para cualquiera de los símbolos), siendo en caso contrario interpretado siempre erróneamente como uno de los otros dos (cada uno con probabilidad 0'2). Se pide:

- a) Si A transmite un mensaje con n símbolos, siendo equiprobables los tres tipos en cada caso, calcule la probabilidad de que sea interpretado en B con menos de 2 símbolos erróneos.
- b) Supóngase que el sistema se quiere hacer más robusto, haciendo que el equipo transmisor A envíe siempre cuatro veces cada símbolo. Si el 30% de las transmisiones realizadas son $\alpha\alpha\alpha\alpha$, el 40% son $\beta\beta\beta\beta$, y el resto $\gamma\gamma\gamma\gamma$, calcule la probabilidad de que A haya transmitido $\alpha\alpha\alpha\alpha$, si lo que se ha interpretado en B ha sido $\alpha\beta\gamma\alpha$.

SOLUCIÓN:

El enunciado indica que la probabilidad de interpretar correctamente un símbolo es 0'6, y por tanto, 0'4 la de interpretarlo erróneamente.

- a) Definimos los siguientes sucesos:

$S = \{\text{el mensaje tiene menos de 2 símbolos erróneos}\}$

$S_1 = \{\text{el mensaje tiene exactamente 1 símbolo erróneo}\}$

$S_0 = \{\text{el mensaje no tiene ningún símbolo erróneo}\}$

Dado que los dos últimos son disjuntos, y que su unión es S , tenemos $P(S) = P(S_0) + P(S_1)$

Estas probabilidades las obtendremos mediante la ley binomial, definiendo la v.a. binomial $X = \text{número de símbolos erróneos en } n$, con $p = 0'4$

$$P(S_0) = P(X = 0) = 0'6^n; \quad P(S_1) = P(X = 1) = \binom{n}{1} \cdot 0'4 \cdot 0'6^{n-1} = n \cdot 0'4 \cdot 0'6^{n-1}$$

Así, $P(S) = n \cdot 0'4 \cdot 0'6^{n-1} + 0'6^n$

**PROBLEMA 1: Continuación**

b) Todas las transmisiones son de uno de los tres tipos, formando la siguiente partición.:

$$P(T_{aaaa}) = 0'3$$

$$P(T_{\beta\beta\beta\beta}) = 0'4$$

$$P(T_{\gamma\gamma\gamma\gamma}) = 0'3$$

Con los datos que tenemos, es inmediato calcular la probabilidad de recibir una secuencia concreta, si se ha transmitido otra dada.

Sin embargo se pide el cálculo opuesto. Mediante el Teorema de Bayes podemos calcular la probabilidad de que A haya transmitido $\alpha\alpha\alpha\alpha$, si lo que se ha interpretado en B ha sido $\alpha\beta\gamma\alpha$:

$$P(T_{aaaa} | R_{\alpha\beta\gamma\alpha}) = \frac{P(R_{\alpha\beta\gamma\alpha} | T_{aaaa}) \cdot P(T_{aaaa})}{P(R_{\alpha\beta\gamma\alpha} | T_{aaaa}) \cdot P(T_{aaaa}) + P(R_{\alpha\beta\gamma\alpha} | T_{\beta\beta\beta\beta}) \cdot P(T_{\beta\beta\beta\beta}) + P(R_{\alpha\beta\gamma\alpha} | T_{\gamma\gamma\gamma\gamma}) \cdot P(T_{\gamma\gamma\gamma\gamma})}$$

donde, calculando las probabilidades de recepción de la secuencia dada para cada secuencia transmitida:

$$P(R_{\alpha\beta\gamma\alpha} | T_{aaaa}) = P(R_{\alpha} | T_{\alpha}) \cdot P(R_{\beta} | T_{\alpha}) \cdot P(R_{\gamma} | T_{\alpha}) \cdot P(R_{\alpha} | T_{\alpha}) = 0'6 \cdot 0'2 \cdot 0'2 \cdot 0'6 = 0'0144$$

$$P(R_{\alpha\beta\gamma\alpha} | T_{\beta\beta\beta\beta}) = P(R_{\alpha} | T_{\beta}) \cdot P(R_{\beta} | T_{\beta}) \cdot P(R_{\gamma} | T_{\beta}) \cdot P(R_{\alpha} | T_{\beta}) = 0'2 \cdot 0'6 \cdot 0'2 \cdot 0'2 = 0'0048$$

$$P(R_{\alpha\beta\gamma\alpha} | T_{\gamma\gamma\gamma\gamma}) = P(R_{\alpha} | T_{\gamma}) \cdot P(R_{\beta} | T_{\gamma}) \cdot P(R_{\gamma} | T_{\gamma}) \cdot P(R_{\alpha} | T_{\gamma}) = 0'2 \cdot 0'2 \cdot 0'6 \cdot 0'2 = 0'0048$$

Sustituyendo valores:

$$P(T_{aaaa} | R_{\alpha\beta\gamma\alpha}) = \frac{0'0144 \cdot 0'3}{0'0144 \cdot 0'3 + 0'0048 \cdot 0'4 + 0'0048 \cdot 0'3} = 0'5625$$



PROBLEMA 2

(2'5 puntos)

Juan y Felipe van a ir al mismo lugar entre las 10 y la 11 de la mañana, si bien los instantes de llegada serán escogidos al azar por cada uno de ellos, dentro del período indicado, y cada uno de ellos llegará, de forma independiente a cuando lo haga el otro.

Cada uno de ellos esperará al otro durante 10 minutos, y si no se han encontrado se irá.
¿Cuál es la probabilidad de que se encuentren?

SOLUCIÓN:

Definimos X: hora de llegada de Juan Y: hora de llegada de Felipe

Ambas son variables aleatorias independientes y uniformes entre las 10 y las 11 horas, ya que el instante de llegada de cada uno de ellos es aleatorio, escogido al azar, entre dichas horas.

Por ser uniformes:

$$f(x) = \begin{cases} 1 & \forall 10 \leq x \leq 11 \\ 0 & \text{resto} \end{cases} \quad f(y) = \begin{cases} 1 & \forall 10 \leq y \leq 11 \\ 0 & \text{resto} \end{cases}$$

Por ser independientes:

$$f(x, y) = f(x) \cdot f(y) = \begin{cases} 1 & \forall \begin{cases} 10 \leq x \leq 11 \\ 10 \leq y \leq 11 \end{cases} \\ 0 & \text{resto} \end{cases}$$

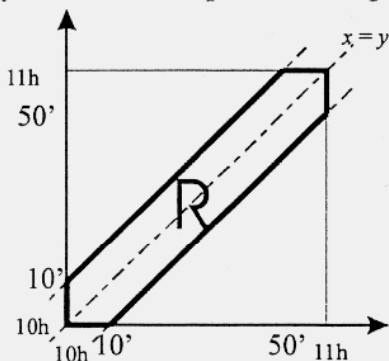
$$x - \frac{1}{6} \leq y \leq x + \frac{1}{6}$$

Se encontrarán sólo si $|X - Y| \leq 10' = \frac{1}{6}$ hora, es decir, si $y \leq x + 10' = x + \frac{1}{6}$

Definiendo S: Juan y Felipe se encuentran

$$P(S) = P\left(|X - Y| \leq \frac{1}{6}\right) = P\left(-\frac{1}{6} \leq X - Y \leq \frac{1}{6}\right) = P\left(X - \frac{1}{6} \leq Y \leq X + \frac{1}{6}\right)$$

y denominando por R a la región bidimensional así definida, tendremos $P(S) = P[(x, y) \in A]$



Podemos obtener esta probabilidad mediante una integral doble, o, al ser la función densidad constante, simplemente multiplicando su área, por la altura de la función densidad, que es 1.

$$P(S) = \Omega_R \cdot 1 = \left(1 - 2 \cdot \frac{\frac{5}{6} \cdot \frac{5}{6}}{2}\right) \cdot 1 = 1 - \frac{25}{36} = \frac{11}{36}$$

**PROBLEMA 3****(2'5 puntos)**

Un fabricante de cierto tipo de piezas de metal, quiere verificar la longitud en metros de dichas piezas. Se sabe que la longitud sigue una distribución normal, y se ha obtenido una muestra de la cadena de producción, con los datos siguientes:

Número de piezas analizadas: 20

Media muestral: 6

Varianza muestral: 0'2

- a) Obtenga un intervalo, con una probabilidad de 0'9 de que contenga a la verdadera longitud media. **(0.5p)**
- b) Cuántas piezas deberían analizarse para tener una confianza del 90% de que la media estimada a través de la muestra, no difiera de la verdadera media en más de 0'22 (considere fijo el valor de s_M^2). **(0.7p)**
- c) Cuántas piezas hay que analizar si se quiere garantizar con una probabilidad de 0'99, que la estimación de la media con la muestra, no se separa del verdadero valor medio en más de 0'12 (considere fijo el valor de s_M^2). **(0.8p)**
- d) Calcular un intervalo de confianza del 95% para la desviación típica poblacional, teniendo en cuenta el tamaño de la muestra inicial. **(0.5p)**

NOTA: No es necesario interpolar en tablas

SOLUCIÓN:

a) $n = 20$ $\bar{x} = 6$ $\sigma_M^2 = 0'2$ \Rightarrow $S_M^2 = \frac{n}{n-1} \sigma_M^2 = \frac{20}{19} 0'2 = 0'21$

Al ser la varianza de la población desconocida, y la muestra menor que 30, se obtiene el intervalo con la t-student: $1 - \alpha = 0'9$ \Rightarrow $\frac{\alpha}{2} = 0'05$

$$I = \bar{x} \pm t_{\alpha/2, n-1} \sqrt{\frac{S_M^2}{n}} = 6 \pm t_{0'05, 19} \sqrt{\frac{0'21}{20}} = 6 \pm 0'177 = [5'823, 6'177]$$

- b) Con la misma probabilidad que en el apartado anterior, y con un intervalo mayor (0'177 en el apartado a), el tamaño de muestra necesario es menor. Por tanto, seguimos con t-student, y vamos probando valores de n menores que 20 y desde 20, que también afectan a la t. Por otro lado, la cuasivarianza muestral la consideramos fija según indica el enunciado.

$$E = t_{\alpha/2, n-1} \sqrt{\frac{S_M^2}{n}} \quad \Rightarrow \quad 0'22 = t_{0'05, n-1} \sqrt{\frac{0'21}{n}}$$

Al sustituir los distintos valores, se verifica que es para $n \geq 14$.

**PROBLEMA 1: Continuación**

Por tanto, $\frac{\mu - 2900}{\sigma} = z_{\alpha} = z_{0.3}$ El valor de $z_{0.3}$, lo miramos en la tabla de la $N(0,1)$, esta

vez sin interpolar y resulta $z_{0.3} = 0.52$. Luego: $\mu - 2900 = 0.52 \sigma$ (**)

Sumando las ecuaciones (*) y (**) queda $3200 - 2900 = (0.841786 + 0.52) \sigma$, de donde

$$\sigma = \frac{300}{1.361786} = 220.3; \text{ y despejando en (**) } \mu = 2900 + 0.52 \cdot 220.3 = 3014.6$$

Es decir, la duración media de estos electrodomésticos es 3014.6 horas de funcionamiento, de forma que $X: N(3014.6, 220.3)$.

b.1) $P(X < 2750) = P\left(Z \leq \frac{2750 - 3014.6}{220.3}\right) = P(Z \leq -1.201) = P(Z \geq 1.201) = 0.1151$

donde ya no hemos interpolado

b.2) $P(X > 3250) = P\left(Z \geq \frac{3250 - 3014.6}{220.3}\right) = P(Z \geq 1.07) = 0.1423$ (sin interpolar)

b.3) $P(2800 \leq X \leq 3100) = P\left(\frac{2800 - 3014.6}{220.3} \leq Z \leq \frac{3100 - 3014.6}{220.3}\right) = P(-0.974 \leq Z \leq 0.388)$ por las propiedades de la normal, esto es igual a:

$$= 1 - P(Z \geq 0.974) - P(Z \geq 0.388) = 1 - 0.1660 - 0.3483 = 0.4857$$

b.4) $P(|X - \bar{X}| \leq 200) = P(-200 \leq X - \bar{X} \leq 200) = P\left(\frac{-200}{\sigma} \leq Z \leq \frac{200}{\sigma}\right) = P(-0.908 \leq Z \leq 0.908) =$
 $= 1 - 2 P(Z \geq 0.908) = 0.6372$

b.5) $P(X \geq \bar{X} + 1.5\sigma) = P(X - \bar{X} \geq 1.5\sigma) = P\left(\frac{X - \bar{X}}{\sigma} \geq 1.5\right) = P(Z \geq 1.5) = 0.0668$

c) Se pide el tiempo x_p , tal que $P(X \leq x_p) = 0.05$, es decir, $P\left(Z \leq \frac{x_p - 3014.6}{220.3}\right) = 0.05$

por las propiedades de la normal, al ser $0.05 < 0.5$, debe ser $\frac{x_p - 3014.6}{220.3} < 0$, luego

$$P\left(Z \geq \frac{3014.6 - x_p}{220.3}\right) = 0.05, \text{ con lo que } \frac{3014.6 - x_p}{220.3} = z_{\alpha} = z_{0.05}$$

En la tabla de la $N(0,1)$ vemos el valor de $z_{0.05}$ que es $z_{0.05} = 1.645$ (hemos interpolado directamente por ser justo el valor intermedio...).

Entonces, $3014.6 - x_p = (1.645)(220.3) = 362.4 \Rightarrow x_p = 3014.6 - 362.4 = 2652.2$ horas.



PROBLEMA 2

(2'5 puntos)

Se tienen dos variables aleatorias independientes X e Y . La variable X es continua y sigue una distribución uniforme en el intervalo $[1,4]$. Por otro lado, la variable Y tiene como función de densidad:

$$f(y) = \begin{cases} K \cdot y^3 & 0 < y < 2 \\ 0 & \text{resto} \end{cases}$$

- Calcule K .
- Calcule la función de densidad conjunta.
- Calcule la probabilidad de que Y sea mayor que X .
- Calcule el momento respecto al origen, cuyo orden para cada variable es 1.
- Calcule el coeficiente de correlación lineal entre las dos variables utilizando la expresión matemática correspondiente. Dé una explicación razonada sobre el resultado obtenido.

SOLUCIÓN:

$$a) \int_{-\infty}^{\infty} f(y) dy = \int_0^2 K \cdot y^3 dy = K \left[\frac{y^4}{4} \right]_0^2 = 4K = 1 \Rightarrow \boxed{K = \frac{1}{4}}$$

- b) Dado que la variable X es uniforme con $a=1$ y $b=4$, en el tramo distinto de cero hay que evaluar la expresión: $\frac{1}{b-a}$, quedando la función de densidad de la variable X :

$$f(x) = \begin{cases} \frac{1}{3} & 1 \leq x \leq 4 \\ 0 & \text{resto} \end{cases}$$

Como las dos variables son independientes, la función de densidad conjunta es el producto de las funciones de densidad marginales:

$$f(x,y) = \begin{cases} \frac{1}{12} y^3 & 1 \leq x \leq 4, \quad 0 < y < 2 \\ 0 & \text{resto} \end{cases}$$

$$c) P(y > x) = \int_1^2 \int_x^2 \frac{1}{12} y^3 dy dx = \int_1^2 \frac{1}{12} \left(4 - \frac{x^4}{4} \right) dx = \int_1^2 \frac{1}{3} dx - \int_1^2 \frac{x^4}{48} dx = \frac{1}{3} [x]_1^2 - \frac{1}{48} \left[\frac{x^5}{5} \right]_1^2 = \boxed{0'204}$$

$$d) \alpha_{11} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x y f(x,y) dx dy = \int_1^4 \int_0^2 x y \frac{y^3}{12} dx dy = \int_1^4 \frac{x}{12} \left[\frac{y^5}{5} \right]_0^2 dx = \int_1^4 \frac{x}{12} \cdot \frac{32}{5} dx = \frac{32}{60} \left[\frac{x^2}{2} \right]_1^4 = \boxed{4}$$



UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA:

6-6-2001 (Tarde)

APELLIDOS Y NOMBRE: _____

Nº DE EXPEDIENTE: _____

CURSO: 3º

GRUPO: _____

PROBLEMA 2: Continuación

e) $\rho = \frac{\mu_{11}}{\sigma_x \cdot \sigma_y}$ $\mu_{11} = \alpha_{11} - \alpha_{10} \cdot \alpha_{01}$

$$\alpha_{10} = \frac{a+b}{2} = \frac{1+4}{2} = 2.5$$

$$\alpha_{01} = \int_0^2 y f(y) dy = \int_0^2 y \frac{y^3}{4} dy = \frac{1}{4} \left[\frac{y^5}{5} \right]_0^2 = 1.6$$

$$\mu_{11} = 4 - 2.5 \cdot 1.6 = 0 \quad \Rightarrow \quad \boxed{\rho = 0}$$

Teniendo en cuenta que las variables son independientes y que la independencia implica ausencia de correlación, el coeficiente de correlación lineal debía ser cero.

**PROBLEMA 3: Continuación**

- c) Al ser un intervalo menor que en el apartado a), y con una mayor probabilidad, necesitará un tamaño de muestra mayor, posiblemente mayor que 30.

$$1 - \alpha = 0.99 \Rightarrow \frac{\alpha}{2} = 0.005$$

$$\text{Probamos con 30: } t_{0.005;29} \sqrt{\frac{0.21}{30}} = 0.23 > 0.12$$

Luego, el tamaño de la muestra será mayor que 30, y utilizamos la normal:

$$E = Z_{\alpha/2} \sqrt{\frac{S_M^2}{n}} \Rightarrow 0.12 = Z_{0.005} \sqrt{\frac{0.21}{n}}$$

Aproximando sin interpolar: $Z_{0.005} = 2.58$, y sustituyendo y despejando n , se obtiene:
 $n \geq 105$.

$$\text{d) } 1 - \alpha = 0.95 \Rightarrow \frac{\alpha}{2} = 0.025 \quad ; \quad 1 - \frac{\alpha}{2} = 0.975$$

$$I = \left[\sqrt{\frac{(n-1)S_M^2}{\chi_{\alpha/2}^2; n-1}}, \sqrt{\frac{(n-1)S_M^2}{\chi_{1-\alpha/2}^2; n-1}} \right] = \left[\sqrt{\frac{19 \cdot 0.21}{\chi_{0.025;19}^2}}, \sqrt{\frac{19 \cdot 0.21}{\chi_{0.975;19}^2}} \right] = (0.348, 0.669)$$



PROBLEMA 4

(2'5 puntos)

Una cadena de televisión realiza un estudio sobre el tiempo de emisión de anuncios.

Si emite demasiados anuncios pierde audiencia, lo que derivará en menores ingresos futuros.

Por otro lado, si emite pocos anuncios, tendrá menores ingresos obtenidos de las empresas anunciantes.

De estudios anteriores, se conoce que el objetivo que permitiría una situación óptima para la cadena de televisión, es un promedio de 15 minutos de emisión de anuncios por hora.

Se analiza una muestra de 72 horas, generando una media de 17 minutos por hora. Suponga que la desviación típica poblacional es de 6 minutos por hora. El análisis se quiere realizar con una probabilidad de error de tipo I de 0'05.

- a) Mediante un contraste de hipótesis, razone a qué conclusiones se llega sobre la emisión de anuncios de esta cadena de televisión. (0.7p)
- b) Si partimos de que el valor real del promedio de minutos de emisión de anuncios por hora fuese 17 minutos por hora, obtenga la potencia del contraste. (0.9p)
- c) Razone sobre la conveniencia o no de que el estudio se realizara sobre una muestra de 200 horas, respecto a la muestra anterior de 72 horas. Para ello, haga uso de la nueva región de aceptación y de la potencia del contraste. Suponga que la media de la nueva muestra sigue siendo 17 minutos por hora. (0.9p)

SOLUCIÓN:

- a) El parámetro a estudiar es μ . Dado que el objetivo es un promedio de 15 minutos, y que las repercusiones para la cadena son negativas, tanto por exceso como por defecto, el contraste debe ser bilateral.

$$H_0: \mu = \mu_0 = 15$$

$$H_1: \mu \neq 15$$

$$\alpha = 0'05 \quad \Rightarrow \quad \frac{\alpha}{2} = 0'025$$

Como el tamaño de la muestra es suficientemente grande ($72 > 30$), por el teorema central de límite, aproximamos la distribución de la media muestral \bar{x} , por $N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$. Además, teniendo en cuenta que la varianza poblacional es conocida, el estadístico de contraste será:

$$Z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}, \text{ que seguirá una } N(0,1).$$

La región de aceptación viene dada por: $|Z| \leq Z_{\alpha/2} \Rightarrow \frac{|\bar{x} - \mu_0|}{\frac{\sigma}{\sqrt{n}}} \leq Z_{\alpha/2} = Z_{0'025} = 1'96$

$$Z = \frac{17 - 15}{\frac{6}{\sqrt{72}}} = 2'828$$

Cae fuera de dicha región: $2'828 > 1'96$, y por tanto, se rechaza, con un nivel de significación de 0'05, que el promedio por hora de emisión de anuncios sea de 15 minutos.



8'2001

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID
Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 29-2-2001 (Tarde)

APELLIDOS Y NOMBRE: _____

Nº DE EXPEDIENTE: _____

CURSO: 3º GRUPO: _____

PROBLEMA 4: Continuación A

b) Potencia del contraste $= 1 - \beta$; donde: $\beta = P\left(\frac{\text{Aceptar } H_0}{H_1 \text{ cierta}}\right)$

Ahora tenemos: $H_0: \mu_0 = 15$
 $H_1: \mu_1 = 17$

En primer lugar calculamos el intervalo bajo el que aceptaremos la hipótesis nula:

$$\mu_0 \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 15 \pm 1.96 \frac{6}{\sqrt{72}} = (13.614, 16.386)$$

La condición de la probabilidad β , es que la hipótesis alternativa es cierta, por lo que es necesaria la distribución de \bar{x} bajo dicha hipótesis, para poder tipificar: $N\left(17, \frac{6}{\sqrt{72}}\right)$.

$$\begin{aligned}\beta &= P(13.614 \leq \bar{x} \leq 16.386) = P\left(\frac{13.614 - 17}{6/\sqrt{72}} \leq Z \leq \frac{16.386 - 17}{6/\sqrt{72}}\right) = \\ &= P(-4.789 \leq Z \leq -0.868) = P(Z \geq 0.868) - P(Z \geq 4.789) = 0.19219\end{aligned}$$

Potencia del contraste $= 1 - 0.19219 = 0.807 \Rightarrow 80.7\%$

c) $n = 200$ $Z = \frac{17 - 15}{6/\sqrt{200}} = 4.717$ es decir, también rechaza la hipótesis nula como en a).

De forma análoga al apartado b): $\mu_0 \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 15 \pm 1.96 \frac{6}{\sqrt{200}} = (14.169, 15.831)$

Bajo la hipótesis alternativa: $\bar{X}: N\left(17, \frac{6}{\sqrt{200}}\right)$

$$\begin{aligned}\beta &= P(14.169 \leq \bar{x} \leq 15.831) = P\left(\frac{14.169 - 17}{6/\sqrt{200}} \leq Z \leq \frac{15.831 - 17}{6/\sqrt{200}}\right) = \\ &= P(-6.676 \leq Z \leq -2.757) = P(Z \geq 2.757) - P(Z \geq 6.676) = 0.00297\end{aligned}$$

Potencia del contraste $= 1 - 0.00297 = 0.997 \Rightarrow 99.7\%$

Con ambas muestras se rechaza la hipótesis nula, pero con $n=200$, la potencia del contraste es muy alta. Por tanto, si el coste de aumentar la muestra a 200 horas no fuera muy alto, sería de interés realizar dicha muestra para obtener una mayor fiabilidad en las conclusiones del estudio.

**PROBLEMA 1****(2'5 puntos)**

- a) Una compañía de autocares interurbanos, tiene negocio en 2 países: Francia y Alemania. Por cada 2 autocares que tiene en Francia, tiene uno en Alemania.

Se dedica en Alemania a la explotación de 2 líneas interurbanas, de tal forma que la probabilidad diaria de que un autocar se averíe si es de la primera línea es de 0'02, y de 0'01 si es de la segunda línea. Por otro lado, se sabe que el 80% de los autocares con los que cuenta en Alemania son de la primera de estas líneas.

En Francia trabaja en 3 líneas interurbanas con probabilidad diaria 0'04 de que un autocar elegido de la primera de estas líneas se averíe, mientras que para los de la segunda línea dicha probabilidad es de 0'01, y para los de la tercera es de 0'02. Estas dos últimas líneas disponen de un mismo número de autocares, sumando entre ambas el 30% de autocares que la compañía tiene en Francia.

Elegido un autocar, calcule la probabilidad de que se averíe a lo largo de un día. **(0.8p)**

- b) Dada la siguiente función: $f(x) = \frac{k}{\sqrt{2\pi}} e^{-\frac{(x-5)^2}{18}}$ obtenga el valor de k y calcule: $P(4 \leq x \leq 7)$ **(0.8p)**

- c) Si la variable X del apartado anterior mide salario anual (en millones) en un cierto colectivo de personas, qué salario anual mínimo obtiene el 70% de dicho colectivo. **(0.9p)**

NOTA: No es necesario interpolar.

SOLUCIÓN:

- a) Se pide la probabilidad total de que en un día se averíe un autocar concreto:

B = que en un día se averíe un autocar

A_1 = autocar sea de Alemania

A_2 = autocar sea de Francia

Además, en Alemania dados los porcentajes de participación en las 2 líneas, las ponderaciones son de 0'8 y 0'2, ya que cada línea tiene distinta probabilidad de avería. Por otro lado, respecto a Francia del enunciado se deduce que las ponderaciones de las 3 líneas son 0'7, 0'15 y 0'15, respectivamente. Por tanto, la probabilidad pedida es:

$$P(B) = P(A_1)P\left(\frac{B}{A_1}\right) + P(A_2)P\left(\frac{B}{A_2}\right) =$$

$$\frac{1}{3}(0'8 \cdot 0'02 + 0'2 \cdot 0'01) + \frac{2}{3}(0'7 \cdot 0'04 + 0'15 \cdot 0'01 + 0'15 \cdot 0'02) = 0'0276$$



6'2001

UNIVERSIDAD PONTIFICIA DE SALAMANCA EN MADRID

Departamento de Electrónica y Comunicaciones

ESTADÍSTICA: 29-2-2001 (Mañana)

APELLIDOS Y NOMBRE: _____

Nº DE EXPEDIENTE: _____

CURSO: 3º GRUPO: _____

PROBLEMA 1: Continuación

b) $f(x) = \frac{k}{\sqrt{2\pi}} e^{-\frac{(x-5)^2}{18}} = \frac{k}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-5}{3}\right)^2}$

Por tanto, la esperanza matemática y la desviación típica son: $\mu = 5$
 $\sigma = 3$

Luego: $k = \frac{1}{\sigma} = \frac{1}{3}$

$$P(4 \leq x \leq 7) = P\left(\frac{4-5}{3} \leq z \leq \frac{7-5}{3}\right) = P(-0'33 \leq z \leq 0'66) = 1 - P(z \geq 0'66) - P(z \geq 0'33) = 0'3747$$

c) $P(x \geq x_{min}) = 0'7 \Rightarrow P\left(z \geq \frac{x_{min}-5}{3}\right) = 0'7$

$$P\left(z \leq \frac{x_{min}-5}{3}\right) = 0'3 \Rightarrow P\left(z \geq -\frac{x_{min}-5}{3}\right) = 0'3$$

Dado que: $z_{0'3} = 0'52$ se tiene que: $0'52 = -\frac{x_{min}-5}{3}$

Por lo tanto, al despejar obtenemos: $x_{min} = 3'44$

**PROBLEMA 2****(2'5 puntos)**

Dada la variable aleatoria bidimensional (x,y) , su función de densidad de probabilidad conjunta es:

$$f(x,y) = \begin{cases} xy/4 & \text{en } R_1 \\ k & \text{en } R_2 \\ 0 & \text{en resto} \end{cases}$$

donde: R_1 es la región delimitada por: $0 \leq x \leq 1$; $0 \leq y \leq 2$

R_2 es la región delimitada por: $1 \leq x \leq 2$; $0 \leq y \leq 1$

a) Calcule el valor de k .

(0.5p)

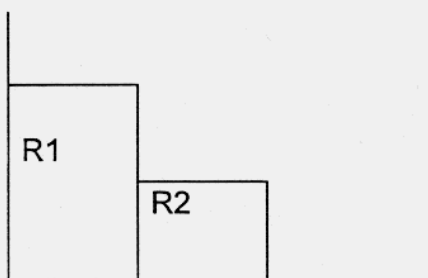
b) Obtenga las funciones de densidad marginales de cada una de las variables.

(0.5p)

c) Compruebe si son o no independientes las variables x e y .

(0.5p)

d) Calcule la recta de regresión de y sobre x .

(1.0p)**SOLUCIÓN:**

a) $\int_0^1 \int_0^2 \frac{xy}{4} dy dx + \int_1^2 \int_0^1 k dy dx = 1$

$$\int_0^1 \frac{x}{4} \left[\frac{y^2}{2} \right]_0^2 dx + \int_1^2 k[y]_0^1 dx = 1$$

$$\int_0^1 \frac{x}{2} dx + \int_1^2 k dx = 1$$

$$k = \frac{3}{4}$$

b) $\int_0^2 \frac{xy}{4} dy = \frac{x}{2} \int_0^1 3 dy = \frac{3}{4}$

$$\text{luego: } f(x) = \begin{cases} \frac{x}{2} & 0 \leq x \leq 1 \\ \frac{3}{4} & 1 \leq x \leq 2 \\ 0 & \text{resto} \end{cases}$$



PROBLEMA 2: Continuación

$$\int_0^1 \frac{xy}{4} dx + \int_1^2 \frac{3}{4} dx = \frac{y}{8} + \frac{3}{4}$$

$$\int_0^1 \frac{xy}{4} dx = \frac{y}{8}$$

$$\text{luego: } f(y) = \begin{cases} \frac{y}{8} + \frac{3}{4} & 0 \leq y \leq 1 \\ \frac{y}{8} & 1 \leq y \leq 2 \\ 0 & \text{resto} \end{cases}$$

- c) Verificamos la condición de independencia para el punto (2,2), ya que es suficiente que incumpla en algún punto para que no sean independientes:

$$f(x,y) = f(x)f(y) \Rightarrow 0 \neq \frac{3}{4} \cdot \frac{2}{8}$$

Por tanto, las variables no son independientes.

d) $\hat{y} - \mu_y = \frac{\mu_{11}}{\sigma_x^2} (x - \mu_x)$

$$\alpha_{10} = \int_0^1 x \frac{x}{2} dx + \int_1^2 x \frac{3}{4} dx = \frac{31}{24}$$

$$\alpha_{01} = \int_0^1 y \left(\frac{y}{8} + \frac{3}{4} \right) dy + \int_1^2 y \frac{y}{8} dy = \frac{17}{24}$$

$$\alpha_{11} = \int_0^1 \int_0^2 xy \frac{xy}{4} dy dx + \int_1^2 \int_0^1 xy \frac{3}{4} dy dx = \int_0^1 \frac{2}{3} x^2 dx + \int_1^2 \frac{3x}{4} dx = \frac{43}{72}$$

$$\mu_{11} = \alpha_{11} - \alpha_{10}\alpha_{01} = \frac{43}{72} - \frac{31}{24} \cdot \frac{17}{24} = -0'31771$$

$$\alpha_{20} = \int_0^1 x^2 \frac{x}{2} dx + \int_1^2 x^2 \frac{3}{4} dx = \frac{15}{8}$$

$$\sigma_x^2 = \alpha_{20} - \alpha_{10}^2 = \frac{15}{8} - \left(\frac{31}{24} \right)^2 = 0'2066$$

Luego, la recta de regresión es: $\hat{y} - \frac{17}{24} = \frac{-0'31771}{0'2066} \left(x - \frac{31}{24} \right)$

**PROBLEMA 3****(2'5 puntos)**

Conocidas las funciones de distribución $F(x)$ y de densidad $f(x)$ de una variable aleatoria continua, X , determine la función de distribución condicional $F(x|S)$, y la función densidad condicional $f(x|S)$, siendo S el suceso $S = \{X \leq a\}$, dejándolas expresadas en función de a , $F(x)$ y/o $f(x)$.

SOLUCIÓN

$$F(x|S) = P(X \leq x | S) = \frac{P((X \leq x) \cap S)}{P(S)} = \frac{P[(X \leq x) \cap (X \leq a)]}{P(X \leq a)}$$

$$P[(X \leq x) \cap (X \leq a)] = \begin{cases} P(X \leq x) & \forall x < a \\ P(X \leq a) & \forall x \geq a \end{cases}$$

Por tanto,

$$F(x|S) = \begin{cases} \frac{P(X \leq x)}{P(X \leq a)} = \frac{F(x)}{F(a)} & \forall x < a \\ \frac{P(X \leq a)}{P(X \leq a)} = 1 & \forall x \geq a \end{cases}$$

$$F(x|S) = \begin{cases} \frac{F(x)}{F(a)} & \forall x < a \\ 1 & \forall x \geq a \end{cases}$$

Para obtener la función densidad de probabilidad, derivamos:

$$f(x|S) = \frac{dF(x|S)}{dx} = \begin{cases} \frac{f(x)}{F(a)} & \forall x < a \\ 0 & \forall x \geq a \end{cases}$$



PROBLEMA 4

(2'5 puntos)

Vd. está encargado del mantenimiento informático en una empresa y sospecha del mal funcionamiento de las fuentes de alimentación de unos ordenadores que están dando problemas. Las especificaciones del fabricante dicen que el voltaje de salida sigue una distribución normal, de media 5V, con desviación típica 0'25V.

Se desea comprobar que efectivamente tienen ese voltaje promedio, y para ello se realizan medidas sobre 8 equipos.

Un compañero suyo le echa una mano, y tras unos cálculos (no erróneos), le sugiere que considere correctas las fuentes si la suma de los 8 voltajes medidos se encuentra entre 38'8V y 41'2V.

- ¿Con qué probabilidad la suma de los voltajes de unas fuentes que funcionaran como asegura el fabricante se encontrarían en dicho intervalo de valores? (1'25p)
- Si realmente el verdadero voltaje promedio de las fuentes de alimentación de este fabricante no fuera 5V, sino 5'1V, ¿qué probabilidad hay de que, haciendo caso a su compañero, la suma de las 8 quedase fuera de dicho intervalo de valores? (1'25p)

SOLUCIÓN

Se trata de un contraste bilateral de la media de una población normal, con varianza conocida.

$$H_0 \equiv \mu = \mu_0 \Rightarrow \mu = 5v$$

$$H_a \equiv \mu \neq \mu_0 \Rightarrow \mu \neq 5v$$

Región de aceptación: $\bar{x} \in \mu_0 \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$

No nos ofrecen un intervalo de aceptación para $\bar{x} = \frac{\sum x_i}{n} = \frac{T}{n}$, sino para T, luego:

$$T \in (38'8, 41'2) \Leftrightarrow \bar{x} \in \left(\frac{38'8}{8}, \frac{41'2}{8} \right) = (4'85, 5'15)$$

$$a) \bar{x}_{CI} = 4'85 = \mu_0 - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \Rightarrow z_{\frac{\alpha}{2}} = (\mu_0 - 4'85) \cdot \frac{\sqrt{n}}{\sigma} = (5 - 4'85) \cdot \frac{\sqrt{8}}{0'25} \cong 1'697056$$

Interpolando valores en la tabla,

$$\begin{array}{ccc} \frac{z}{1'69} & \rightarrow & \frac{\alpha/2}{0'0455} \\ z = 1'697 & \Rightarrow & \Rightarrow \frac{\alpha}{2} = 0'04487 \\ 1'70 & \rightarrow & 0'0446 \end{array}$$

$$\Rightarrow \frac{\alpha}{2} = 0'04487 \Rightarrow \alpha = 0'08974 \Rightarrow \boxed{1 - \alpha = 0'91026} \cong 91\%$$

Hubiese sido el mismo resultado y de cálculo equivalente de haberse realizado a partir de \bar{x}_{CS} , ya que μ_0 es el punto medio del intervalo de aceptación.



PROBLEMA 4: Continuación

De forma similar, podría haberse trabajado con el intervalo completo:

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \quad N(0, 1)$$

$$\begin{aligned} 1 - \alpha &= P\left(\frac{4'85 - 5}{0'25 / \sqrt{8}} \leq z \leq \frac{5'15 - 5}{0'25 / \sqrt{8}}\right) = P(-0'6 \cdot \sqrt{8} \leq z \leq 0'6 \cdot \sqrt{8}) = P(-1'697056 \leq z \leq 1'697056) = \\ &= 1 - 2 \cdot P(z \geq 1'697056) = 1 - 2 \cdot 0'04487 = 0'91026 \cong 91\% \end{aligned}$$

Y, en cualquier caso, podría haberse trabajado directamente mediante la distribución de la suma T , en lugar de la de la media muestral \bar{x} .

$$\begin{aligned} X: N(\mu, \sigma) &\Rightarrow \bar{X}: N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) \Rightarrow T = n\bar{X}: N(n\mu, \sigma\sqrt{n}) \\ (5, 0'25) &\quad \quad \quad \left(5, \frac{0'25}{\sqrt{8}}\right) \quad \quad \quad (40, 0'25\sqrt{8}) \end{aligned}$$

b) Siendo $\mu = 5'1$, debemos calcular la probabilidad de caer fuera del intervalo de aceptación. Es decir, la potencia del contraste $1 - \beta(\mu = 5'1)$.

Calcularemos, sin embargo β , que es la probabilidad de caer en dicho intervalo; es decir, la probabilidad de error tipo II, o de aceptar, si realmente la media es 5'1.

$$\begin{aligned} \beta &= P((4'85 \leq \bar{x} \leq 5'15) | \mu = 5'1) = P\left(\frac{4'85 - 5'1}{0'25 / \sqrt{8}} \leq z \leq \frac{5'15 - 5'1}{0'25 / \sqrt{8}}\right) = P(-2'8284 \leq z \leq 0'5657) = \\ &= P(z \geq -2'8284) - P(z \geq 0'5657) = 1 - P(z \geq 2'8284) - P(z \geq 0'5657) = 1 - 2'34 \cdot 10^{-3} - 0'2858 \end{aligned}$$

$$\beta = 0'71186$$

Así, $\boxed{1 - \beta = 1 - 0'71186 = 0'28814} \cong 28'81\%$

Igualmente se podría haber realizado el cálculo a partir de las expresiones de \bar{x}_{CI} y \bar{x}_{CS} , en

lugar de sus valores: $\bar{x}_{CI} = 4'85 = \mu_0 - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ $\bar{x}_{CS} = 5'15 = \mu_0 + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$