

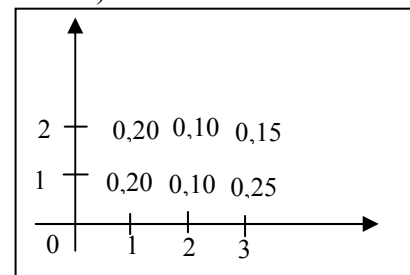
EJERCICIOS DE ESTADÍSTICA EMPRESARIAL**VARIABLE ALEATORIA BIDIMENSIONAL (TEMA 1)**

1º) Se mide la polución en gramos para cierto volumen de aire en los alrededores de una fábrica de cemento. Designamos por X la cantidad de polución recogida cuando no se utiliza un filtro y por Y la recogida cuando se utiliza. Si $f(x, y) = k$, $0 \leq x \leq 3$; $0 \leq y \leq 1$, $3y \leq x$, siendo cero en el resto del plano, calcular: **1)** el valor de k ; **2)** $P(X \geq 4Y)$

$$\text{Sol.: } k = \frac{2}{3}; \text{ b) } P(X \geq 4Y) = \frac{3}{4}$$

2º) Una variable aleatoria bidimensional discreta tiene la función de probabilidad que aparece en la figura adjunta. Calcular la probabilidad condicionada: $P[Y \leq 1 / 2 \leq X \leq 3]$.

$$\text{Sol.: } P[Y \leq 1 / 2 \leq X \leq 3] = \frac{7}{12}$$



3º) Las variables aleatorias X e Y tienen la función de densidad condicional $f(y/x) = 2yx^{-2}$, para $0 \leq y \leq x$, siendo cero en el resto. Además $f(x) = 4x^3$ para $0 \leq x \leq 1$, siendo cero en el resto. Hallar, indicando los intervalos de variación: **1)** la función de densidad conjunta; **2)** la función de densidad marginal de Y ; **3)** $f(x/y)$.

$$\text{Sol.: } \text{1) } f(x, y) = 8xy, 0 \leq y \leq x \leq 1; \text{ 2) } f(y) = 4y(1 - y^2), 0 \leq y \leq 1; \text{ 3) } f(x/y) = \frac{2x}{1 - y^2},$$

$y \leq x \leq 1$.

4º) Se supone que los salarios X_1 y X_2 superiores a 35 unidades monetarias en dos actividades económicas diferentes, tienen una función de densidad conjunta $f(x_1, x_2) = A(x_1 x_2)^{-2}$, $x_1 \geq 35$, $x_2 \geq 35$. Determinar: **1)** la constante A ; **2)** la función de distribución conjunta de las variables aleatorias X_1 y X_2 y **3)** la probabilidad de que dos trabajadores elegidos al azar, uno de cada actividad, tengan cada uno salarios superiores a 100 u.m..

$$\text{Sol.: } \text{1) } A = 35^2; \text{ 2) } F(x_1, x_2) = \left(\frac{35}{x_1} - 1\right)\left(\frac{35}{x_2} - 1\right); \text{ 3) } P[X_1 > 100, X_2 > 100] = 0,35^2 =$$

$= 0,1225$.

5º) La variable aleatoria bivalente (X, Y) tiene la función de densidad $f(x, y) = K(x^2 + y^2)$, $0 \leq y \leq x \leq 1$, siendo cero en el resto del plano. Hallar el valor de k , la función de densidad marginal $f(x)$ y $f(y/x)$.

$$\text{Sol.: } k = 3; f(x) = 4x^3, 0 \leq x \leq 1; f(y/x) = \frac{3(x^2 + y^2)}{4x^3}, 0 \leq y \leq x.$$

6º) Dada la función de densidad $f(x, y) = ke^{-x}$, $0 < \frac{x}{2} < y < x$, hallar k y las funciones de distribución condicionadas.

$$\text{Sol.: } k = 2; f(y/x) = \frac{2}{x}, \frac{x}{2} < y < x; f(x/y) = \frac{e^{-x}}{e^{-y} - e^{-2y}}, y < x < 2y. \text{ De aquí}$$

obtenemos las funciones de distribución condicionadas.

$$F(x/y) = \begin{cases} 0, & x < y \\ \int_y^x \frac{e^{-x}}{e^{-y} - e^{-2y}} dx = \frac{e^{-y} - e^{-x}}{e^{-y} - e^{-2y}}, & y \leq x < 2y \\ 1, & x \geq 2y \end{cases}; F(y/x) = \begin{cases} 0, & y < \frac{x}{2} \\ \int_{\frac{x}{2}}^y \frac{2}{x} dy = \frac{2}{y} \left(y - \frac{x}{2} \right) = 2 - \frac{x}{y}, & \frac{x}{2} \leq y < x \\ 1, & y \geq x \end{cases}$$

7º) La función de probabilidad conjunta de la variable aleatoria bidimensional (X, Y) es:

$$P(x, y) = \begin{cases} k, & 0 \leq x \leq 3, \quad 0 \leq y \leq 5, \quad x \text{ e } y \text{ enteros} \\ 0, & \text{en cualquier otro caso} \end{cases}$$

Calcular las probabilidades: $P(X = 1, Y = 4)$ y $P(X + Y < 3)$.

$$\text{Sol.: } P(X = 1, Y = 4) = \frac{1}{24} \text{ y } P(X + Y < 3) = \frac{6}{24}.$$

8º) La función de probabilidad conjunta de la variable aleatoria bidimensional discreta (X, Y) es:

$$P(x, y) = \begin{cases} k(x + 2y), & 0 \leq x \leq 3, \quad 0 \leq y \leq 2, \quad x \text{ e } y \text{ enteros} \\ 0, & \text{en cualquier otro caso} \end{cases}$$

Calcular: 1) el valor de la constante k; 2) $P(X \leq 1, Y \geq 1)$; 3) $P(X = 1, Y = 1)$; 4) $P(X \leq 1/Y \geq 1)$

$$\text{Sol.: 1) } k = \frac{1}{42}; \text{ 2) } P(X \leq 1, Y \geq 1) = \frac{1}{3}; \text{ 3) } P(X = 1, Y = 1) = \frac{1}{14}; \text{ 4) } P(X \leq 1/Y \geq 1) = \frac{7}{18}$$

9º) Dada la variable aleatoria bidimensional continua (X, Y), con función de densidad conjunta:

$$f(x, y) = \begin{cases} K(3x^2 + 2y), & 0 \leq x \leq 1, \quad 0 \leq y \leq 1 \\ 0, & \text{en el resto} \end{cases}$$

hallar: 1) la constante K; 2) la función de distribución conjunta; 3) $P(0,3 < X \leq 0,8; 0,1 < Y \leq 0,5)$.

$$\text{Sol.: 1) } K = \frac{1}{2}; \text{ 2) } F(x, y) = \frac{x^3 y + y^2 x}{2}, \quad 0 \leq x \leq 1; \quad 0 \leq y \leq 1; \text{ 3) } P(0,3 < X \leq 0,8; 0,1 < Y \leq 0,5) = 0,157.$$

10º) Una variable aleatoria bidimensional discreta tiene la función de probabilidad que aparece en la tabla adjunta. Calcular: 1) p_{22} ; 2) $P(X = 2/Y \geq 1)$; 3) $P[Y \leq 2 / 1 \leq X < 3]$.

$$\text{Sol.: 1) } p_{22} = 0,05; \text{ 2) } P(X = 2/Y \geq 1) = 0,3; \text{ 3) } P[Y \leq 2 / 1 \leq X < 3] = 1$$

Y \ X	1	2
1	0,1	0,15
2	0,25	p_{22}
3	0,35	0,1

TEMA 2: CARACTERÍSTICAS DE LAS VARIABLES ALEATORIAS

11º) Si la variable aleatoria U está distribuida uniformemente en $-4 \leq u \leq 4$, determinar $P(|U - 2| < 2)$. [Sol.: $P = 1/2$]

12º) En cierto país la función de densidad de las rentas anuales de las personas que han de pagar impuesto sobre la renta viene dada por $f(r) = A \cdot r^{-4,5}$, $r \geq 900$, siendo cero en el resto

del intervalo. Hallar : **1)** el valor de A; **2)** la renta anual que con probabilidad 0,1 es superada por un contribuyente elegido aleatoriamente.[Sol: **1)** $A = 3,5 \cdot 900^{3,5}$; **2)** $r = 1737,63$]

13º) La distribución de probabilidad de la variable aleatoria discreta X viene dada por $P[X=k]=C$, $k=A+1, A+2, \dots, A+n$, donde A es un entero positivo dado. Hallar la constante C y la media y la varianza de X.[Sol.: $C = \frac{1}{n}$; $E(X) = A + \frac{n+1}{2}$; $\text{var}(X) = \frac{n^2-1}{12}$]

14º) Una variable aleatoria X tiene una función de densidad proporcional a la función $2+bx$, siendo su campo de variación el intervalo (0,1). Calcular el parámetro b si $E[X^2]=17/42$. [Sol.: $b = 3$]

15º) Sea X una variable aleatoria cuya función de densidad es $f(x) = (\alpha-1)x^{-\alpha}$ para $x \geq 1$, siendo cero en el resto del intervalo, en donde α es un parámetro desconocido que se supondrá superior a 3. Hallar: **1)** la expresión de la función de distribución de X; **2)** las expresiones de la media y la varianza de X.[Sol.: **1)** $F(x) = 1 - x^{-\alpha+1}$ para $x \geq 1$ y cero en el resto;

2) $\mu = \frac{\alpha-1}{\alpha-2}$, $\sigma^2 = \frac{\alpha-1}{\alpha-3} - \left(\frac{\alpha-1}{\alpha-2} \right)^2$]

16º) Una variable aleatoria X tiene de media -7 y de varianza 3. Hallar $E[Y]$ siendo $Y = 0,5(1-X+X^2)$. [Sol: $E(Y) = 30$]

17º) Se considera la función f definida por $f(x) = \frac{1}{4} - \lambda x^2$ para $-1 \leq x \leq 1$, siendo cero en el resto del intervalo. **1)** Determinar λ para que sea función de densidad; **2)** Sea X una variable aleatoria de función de densidad f. Calcular $E[X]$ y $\text{Var}[X]$.

18º) Si X es una variable aleatoria de media -1 y desviación típica 2, obtener un límite inferior para $P[|X+1| < 3]$. [Solución: Por la desigualdad de Chebychev: $P[|X+1| < 3] \geq 1 - \frac{4}{9} = \frac{5}{9}$, que es el límite inferior buscado.

19º) Si la variable aleatoria X tiene media 4 y varianza 9, hallar $E[Y]$, siendo $Y = -10 + X^2$. [Sol.: $E(Y) = -10 + E(X^2) = -10 + \text{Var}(X) + (E(X))^2 = -10 + 9 + 16 = 15$]

20º) La distribución de probabilidad de la variable aleatoria discreta X viene dada por $P[X=k] = C$, $k = 1, 2, \dots, 20$, siendo cero en el resto del eje x. Hallar la constante C y la media de la distribución.[Sol.: $C = 1/20$; $\mu = 10,5$]

21º) Si X es una variable aleatoria de media -4 y desviación típica 3, obtener un límite inferior para $P[|X+4| < 6]$.

22º) Se tiene una variable aleatoria X con función de densidad $f(x) = 5e^{-5x}$, $x \geq 0$. Hallar su media y su desviación típica.

23º) Una variable aleatoria X tiene media -1 y varianza 9. Hallar $E[Y]$ siendo $Y = 3 - 2X + 0,5X^2$. [Sol.: $E(Y) = 10$]

24º) El porcentaje de aditivo de un tipo de gasolina es una variable aleatoria X que tiene una función de densidad $f(x) = 20x^3(1-x)$, $0 < x < 1$, y cero en el resto. **1)** Encontrar el porcentaje medio de aditivo. **2)** Si el beneficio viene dado por $B = 18 + 3x$, donde X es el porcentaje de aditivo, hallar $E(B)$. [Sol.: **1)** $E(X) = 2/3$; **2)** $E(B) = 20$]

25º) Sean $X_1, X_2, X_3, \dots, X_n$ variables aleatorias independientes idénticamente distribuidas con función de densidad $f(x) = 1$ para $|x| \leq 1/2$, siendo cero en el resto del intervalo.

Hallar la función generatriz de momentos de $X = (X_1 + X_2 + X_3 + \dots + X_n) \cdot n^{-1/2}$. [Sol.:

$$\psi_{\xi}(t) = \left[\frac{\sqrt{n}}{t} \left(e^{\frac{t}{2\sqrt{n}}} - e^{-\frac{t}{2\sqrt{n}}} \right) \right]^n]$$

26º) Las ventas medias de una cafetería son de 5000 € a la semana con una desviación típica de 100€. Se pide calcular con estos datos:

a) La probabilidad de que las ventas medias semanales sean mayores de 5250 €.

b) Definir el menor intervalo de tal forma que contenga al menos el 95% de las ventas medias semanales. [Sol.: a) de la desigualdad de Chebychev se obtiene que la probabilidad pedida es $\leq 0,16$; b) aplicando la desigualdad de Chebychev se obtiene el intervalo [4552,79; 5447,21]

27º) Una variable aleatoria tiene por función generatriz de momentos $g(t) = \left(\frac{0,4}{1 - 0,6 \cdot e^t} \right)^7$. Hallar su media y su varianza. [Sol.: $\mu = 10,5$; $\sigma^2 = 26,25$]

28º) La variable aleatoria X tiene por función generatriz de momentos $g(t) = \left(\frac{2}{2-t} \right)^4$ para $t < 2$. Hallar la media y la varianza de X. [Sol.: $\mu = 2$, $\sigma^2 = 1$]

29º) Una variable aleatoria X tiene la función de densidad $f(x) =$
 $= \begin{cases} 1+x, & \text{para } -1 < x < 0 \\ 1-x, & \text{para } 0 < x < 1 \\ 0 & \text{en el resto} \end{cases}$ Hallar la función generatriz de momentos. [Sol.: $g(t) =$
 $= \frac{e^t + e^{-t} - 2}{t^2}$]

30º) Una variable aleatoria X tiene por función generatriz de momentos $g(t) =$
 $= \frac{0,2 \cdot e^t}{1 - 0,8 \cdot e^t}$. Deducir su media y su varianza. [Sol.: $\mu = 5$, $\sigma^2 = 20$]

31º) La variable aleatoria X tiene por función generatriz de momentos $g(t) = \left(\frac{3}{3-t} \right)^3$. Hallar la media y la desviación típica. ¿Qué condición debe cumplir t? [Sol.: $\mu = 1$, $\sigma = \frac{1}{\sqrt{3}}$. Se tiene que cumplir que $t < 3$]

32º) Se tiene una variable aleatoria X con función de densidad $f(x) =$
 $= \begin{cases} 5e^{-5x} & \text{para } x \geq 0 \\ 0 & \text{en el resto} \end{cases}$. Hallar su media, su desviación estándar y su función generatriz de momentos. ¿Qué condición se tiene que cumplir en relación con esta última función? [Sol.: $\mu = 1/5$, $\sigma = 1/5$. Debe suponerse que $t < 5$, obteniéndose entonces que $g(t) = \frac{5}{5-t}$]

33º) La variable aleatoria X tiene por función generatriz de momentos $g(t) = e^{25t + t^2}$. Hallar la media y la varianza de X. [Sol.: $\mu = 25$, $\sigma^2 = 2$]

34º) Sean $X_1, X_2, X_3, \dots, X_{10}$ variables aleatorias independientes idénticamente distribuidas con función de densidad $f(x) = 3e^{-3x}$ para $x \geq 0$, siendo cero en el resto del

intervalo. Hallar la función generatriz de momentos de $X = (X_1 + X_2 + X_3 + \dots + X_n) \cdot 10^{-1/2}$

$$[Sol.: g_X(t) = \left(\frac{3}{3 - \frac{t}{\sqrt{10}}} \right)^{10}, \quad t < 3\sqrt{10}]$$

35º) La variable aleatoria X tiene por función generatriz de momentos $g(t) = e^{20t + 2t^2}$. Hallar la media y la varianza de X . [Sol.: $\mu = 20$, $\sigma^2 = 4$]

36º) Se consideran las variables aleatorias independientes idénticamente distribuidas $X_1, X_2, X_3, \dots, X_{20}$ con función de densidad $f(x) = 4e^{-4x}$ para $x \geq 0$, siendo cero en el resto del intervalo. Hallar la función generatriz de momentos de $X = (X_1 + X_2 + X_3 + \dots + X_n) \cdot 20^{-1/2}$ y el

valor de dicha función para $t = 0$. [Sol.: $g_X(t) = \left(\frac{4}{4 - \frac{t}{\sqrt{20}}} \right)^{20}, \quad t < 4\sqrt{20}; g_X(0) = 1]$

37º) Hallar la función generatriz de momentos de la variable aleatoria X con la siguiente distribución $P(X=1) = 0,3$; $P(X=3) = 0,3$ y $f(x) = 0,04x$, para $4 \leq x \leq 6$. [Sol.: $g(t) = 0,3e^t + 0,3e^{3t} + \frac{0,04(6e^{6t} - 4e^{4t})}{t} + \frac{0,04(e^{6t} - e^{4t})}{t^2}$]

TEMAS 3 Y 4: MODELOS DE PROBABILIDAD

Binomial

38º) Sabemos por los ficheros de demandantes potenciales del producto que fabrica nuestra empresa que un 30% de los mismos no lo compran. Determinar la probabilidad de que al extraer aleatoriamente con reemplazamiento tres fichas de demandantes potenciales, una sea de un no comprador.

Sol.: La variable $X = \text{"nº de clientes no compradores"}$ es $B(3; 0,3) \Rightarrow P[X=1] = 0,441$

39º) Nuestra empresa ofrece dos formas de pago a nuestros clientes a la hora de adquirir nuestro producto: al contado y aplazado. Por la estadística de ventas conocemos que un 25% de las unidades que se venden son abonadas al contado. Calcular la probabilidad de que de las 6 unidades que se han vendido últimamente, tres o más se hayan pagado al contado.

Sol.: La variable $X = \text{"nº de unidades vendidas al contado"}$ es $B(6; 0,25) \Rightarrow P[X \geq 3] = 0,1694$

40º) La probabilidad de que una pieza industrial sea defectuosa es 0,01. Un lote de piezas está compuesto por 5 de ellas. Un lote se rechaza si contiene 1 ó más piezas defectuosas. Calcular la probabilidad de que, de 10 lotes, se rechacen 2 de ellos.

Sol.: La probabilidad de rechazar un lote es $1 - 0,99^5 = 0,049$; la variable $X = \text{"nº de lotes que se rechazan"}$ es $B(10; 0,049) \Rightarrow P[X = 2] = 0,072$.

41º) En cierta ciudad y en invierno llueve un día con probabilidad 0,3. Si son independientes las lluvias en dos días cualesquiera del invierno, se pide: **a)** probabilidad de que en una semana de invierno llueva dos días; **b)** número de días esperado en que llueva en una semana de invierno.

Sol.: **a)** La variable $X = \text{"nº de días que llueve en una semana"}$ es $B(7; 0,3) \leftrightarrow \leftrightarrow P[X = 2] = 0,3177$; **b)** $E(X) = 2,1$.

42º) Justificar razonadamente si al sumar dos variables aleatorias independientes de tipo binomial: $B(5; 0,3)$ y $B(4; 0,3)$ se obtiene otra binomial $B(9; 0,3)$.

Sol.: En efecto, si X_1 y X_2 son las variables, entonces la función generatriz $g_{X_1+X_2}(t) = E(e^{(X_1+X_2)t}) = E(e^{X_1t})E(e^{X_2t}) = (0,7 + 0,3 \cdot e^t)^5 \cdot (0,7 + 0,3 \cdot e^t)^4 = (0,7 + 0,3 \cdot e^t)^9$ que corresponde a una $B(9; 0,3)$.

43º) Dadas dos variables aleatorias X e Y que siguen distribuciones binomiales $B\left(1, \frac{1}{2}\right)$ y $B\left(1, \frac{1}{2}\right)$ respectivamente, ¿es correcto afirmar que $X + Y$ se distribuye según una binomial $B(1, 1)$?

Sol.: No. Se distribuiría $B\left(2, \frac{1}{2}\right)$.

Poisson

44º) La probabilidad de que las unidades que fabricamos de cámaras de vídeo nos las devuelvan por defectuosas sigue la distribución de Poisson de función de cuantía: $\frac{e^{-4} \cdot 4^x}{x!}$.

Determinar la probabilidad de que de las unidades vendidas en el último mes, tres o más sean devueltas por defectuosas.

Sol: $P[X \geq 3] = 1 - P[X \leq 2] = 0,7619$

45º) Una compañía de seguros garantiza pólizas de seguros individuales contra cierto tipo de accidentes. Una encuesta ha permitido establecer que a lo largo de un año una persona tiene una posibilidad entre mil de ser víctima de un accidente cubierto por dichas pólizas y, además, que la compañía podrá vender una media de 4000 pólizas de este tipo al año. Se pide: **a)** probabilidad de que el número de accidentes cubiertos por la póliza no pase de 4 por año; **b)** número de accidentes esperado por año; **c)** probabilidad de que el número de accidentes sea superior a 2 al año.

Sol.: **a)** La variable $X = \text{"nº de accidentes cubiertos por la póliza"}$ es $B(4000; 0,001) \cong \cong \text{Poisson}(4) \Rightarrow P[X \leq 4] = (\text{de las tablas}) = 0,6288$; **b)** $E(X) = 4$; **c)** $1 - P[X \leq 2] = 0,7619$.

46º) Una compañía de seguros ha comprobado que el 0,005% de la población fallece cada año de un cierto tipo de accidente. Se pide: **a)** ¿cuál es la probabilidad de que la compañía tenga que pagar a más de tres de sus 10.000 asegurados contra tales accidentes en un año determinado?; **b)** ¿cuál es el número de accidentes esperados?

Sol.: La variable $X = \text{"nº de asegurados que fallecen"}$ es $B(10000; 0,00005) \cong \cong \text{Poisson}(0,5) \Rightarrow 1 - P[X \leq 3] = (\text{de las tablas}) = 0,0018$; **b)** $E(X) = 0,5$

47º) En una pequeña ciudad hay dos gasolineras A y B. El número medio de vehículos que llegan a cada una de ellas por hora son 10 y 8 respectivamente, siendo ambos números de llegadas distribuciones de Poisson independientes. Por una causa concreta la gasolinera B cierra, por lo que la gasolinera A debe atender a sus habituales clientes y a los de B. Obtener la distribución del número de llegadas por hora a A el día que B cierra.

Sol.: Si $X_A = \text{"nº de vehículos que llega a A"}$ y $X_B = \text{"nº de vehículos que llega a B"}$, entonces $X = X_A + X_B$ es Poisson(18).

48º) Tenemos dos variables aleatorias independientes X e Y con distribuciones de Poisson de parámetros, respectivamente $\lambda_1 = 2$ y $\lambda_2 = 4$. Indicar razonadamente la función generatriz de momentos de $Z = X + Y$.

Sol: $g_z(t) = e^{6(e^t - 1)}$

Uniforme

49º) Si la variable X está distribuida uniformemente en $-2 \leq x \leq 2$, calcular $P(|X-1| < 0,5)$.

Sol.: $P = 1/4$

50º) Si X es una variable aleatoria distribuida uniformemente en $-2 \leq x \leq 1$, hallar $P(|X-2| \geq 0,5)$.

Sol.: $P = 1$

51º) Suponiendo que la cotización de cierre diaria de las acciones del Banco de Santander Central Hispano tiene una distribución uniforme entre los 20 y 21 euros, se pide:
a) ¿cuál es la probabilidad de que un día la cotización de cierre supere los 20,90 euros?;
b) ¿cuál es el porcentaje de días que presentaron una cotización de cierre entre 20,40 y 20,60 euros?;
c) Si sumáramos dos distribuciones uniformes $U(20; 21)$ y $U(22; 24)$ que fuesen independientes, demuestre si la distribución resultante presenta o no carácter de distribución uniforme;
d) entre los días en los que la cotización de cierre ha sido superior a 20,50 euros, ¿cuál es el porcentaje de los mismos en que la cotización ha oscilado entre 20,80 y 20,90 euros?.

Sol.: **a)** $P = 0,1$; **b)** 20 %; **c)** si X_1 es $U(20, 21)$ y X_2 es $U(22, 24)$, desde luego que $42 \leq X_1 + X_2 \leq 45$. Pero $g_{X_1+X_2}(t) = g_{X_1}(t) \cdot g_{X_2}(t) = \frac{e^{21t} - e^{20t}}{t} \cdot \frac{e^{24t} - e^{22t}}{t}$ que no corresponde a la función generatriz de una $U(42, 45)$; **d)** $P[20,80 \leq X \leq 20,90 / X > 20,50] = \frac{0,1}{0,5} = 0,2 \Rightarrow$ el 20%.

Normal

52º) Las variables aleatorias X_1 y X_2 son independientes y están distribuidas normalmente $N(-10, 3)$ y $N(-14, 4)$ respectivamente. Sea $Z = X_1 - X_2$. Se pide: **a)** ¿qué distribución sigue Z ?; **b)** ¿cuál es su media?; **c)** ¿cuál es su desviación estándar?.

Sol.: Z es $N(4, 5)$

53º) Las variables X_1 , X_2 y X_3 son independientes y respectivamente $N(1, 1)$, $N(0, 1)$ y $N(-1, 2)$. ¿Cómo se distribuye $Z = 2X_1 + X_2 - X_3$?

Sol.: $N(3, 3)$

54º) Sea X una variable aleatoria distribuida normalmente con media -2 y desviación estándar 3. Calcular x_0 tal que $P(X < -x_0) = 0,10$.

Sol.: $x_0 = 5,84$

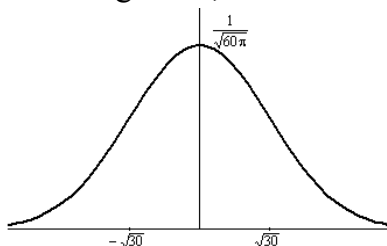
55º) Dos grandes superficies comerciales venden una determinada marca de detergente. Las unidades vendidas semanalmente en el primer establecimiento siguen la ley $N(2.000, 200)$ y en el segundo $N(1.700, 130)$. Calcular la probabilidad de que las ventas de la primera superficie superen en 200 unidades a la segunda, en una determinada semana, bajo el supuesto de independencia.

Sol.: Sean X e Y las unidades vendidas semanalmente en cada establecimiento, respectivamente. Entonces, $X - Y$ es $N(300; 238,54)$. Desde luego $P(X - Y = 200) = 0$ pero si

hacemos una corrección por continuidad (ya que, en realidad, $X - Y$ es una variable discreta), tendremos que $P(X - Y = 200) = P(199,5 < X - Y < 200,5) = 0,0015$

56º) Si tenemos tres distribuciones normales $X_1 : N(5, 2)$; $X_2 : N(-3, 1)$ y $X_3 : N(2, 5)$, ¿cómo se distribuye la variable aleatoria $U = X_1 + X_2 - X_3$? Representar gráficamente la función de densidad de U .

Sol.: $N(0, \sqrt{30})$; la representación gráfica, con diferente escala para cada eje:



57º) Tenemos una variable aleatoria distribuida normalmente, X , de media 16 y varianza 9. Calcular $P(|X - 16| \geq 5)$ y hallar una cota de dicha probabilidad mediante el teorema de Chebyshev.

Sol.: $P = 0,097$; $P \leq 9/25$

Teorema Central del Límite

58º) Una empresa aseguradora de automóviles estima que la probabilidad de que un automóvil asegurado tenga al menos un accidente al año es 0,01. Si su cartera de clientes consta de 8.000 vehículos, se pide: **a)** distribución del número de automóviles accidentados de sus clientes al año; **b)** número esperado de automóviles accidentados en el año.

Sol.: **a)** La variable $X =$ "nº de automóviles accidentados" es $B(8000; 0,01) \leftrightarrow \leftrightarrow P[X = x] = \binom{8000}{x} \cdot 0,01^x \cdot 0,99^{8000-x} \leftrightarrow F(x) = \sum \binom{8000}{i} \cdot 0,01^i \cdot 0,99^{8000-i}$. Pero (teorema de

Moivre) X es aproximadamente $N(80; 8,9)$ y en este caso $P[X = x] = P[x - \frac{1}{2} < X < x + \frac{1}{2}] =$

$$= F\left(\frac{x + \frac{1}{2} - 80}{8,9}\right) - F\left(\frac{x - \frac{1}{2} - 80}{8,9}\right), \text{ donde } F \text{ es la función de distribución de la } N(0, 1);$$

b) $E(X) = 80$.

59º) Los gastos de transporte que realiza una oficina oscilan uniformemente entre 100.000 y 140.000 pts. al mes. Se pregunta: **a)** ¿cuál es la probabilidad de que en un mes determinado el gasto en transporte sea exactamente 120.000 pts.?; **b)** calcule la desviación típica del gasto mensual; **c)** estimándose que el gasto en transporte es excesivo, se pretende llevar a cabo un control para comprobar la necesidad de dicho gasto. Para ello se observa aleatoriamente el gasto mensual durante tres años. ¿Cuál es la probabilidad de que el gasto mensual medio, durante esos tres años, sea superior a 130.000 pts.?

Sol.: a) $P = 0$; b) $\sigma = \frac{20000}{\sqrt{3}}$; c) Si $X_i = \text{"gasto de transporte en el mes } i"$ \Rightarrow

$X = \frac{1}{36} \sum_{i=1}^{36} X_i$ se distribuye aproximadamente $N\left(120000, \frac{10000}{3\sqrt{3}}\right)$, luego $P[X > 130000] = P[Z > 3\sqrt{3}] \cong 0$

60º) Las variables aleatorias X_1, X_2, \dots, X_n son independientes y tienen una distribución uniforme definida por la función de densidad $f(x) = \begin{cases} 1/2, & 0 \leq x \leq 2 \\ 0, & \text{en el resto} \end{cases}$.

Consideremos la variable aleatoria $Y_n = \frac{1}{n} \sum_{i=1}^n X_i$. Hallar la probabilidad de que Y_n sea mayor que 0,9 para $n = 36$.

Sol.: Y_n es aproximadamente $N\left[1, \frac{1}{6\sqrt{3}}\right] \Rightarrow P[Y_n > 0,9] = 0,8485$

61º) Un concesionario de automóviles vende vehículos de la misma marca. Sabiendo que la probabilidad de que este tipo de vehículos esté funcionando 4 años después es de 0,6, determinar la probabilidad de que, de 5.000 automóviles vendidos, más de 3.000 estén en servicio dentro de 4 años.

Sol.: $X = \text{"nº de vehículos que están funcionando después de cuatro años"}$ se distribuye $B(5000; 0,6) \cong N(3000, \sqrt{1200})$; haciendo la corrección por continuidad: $P[X > 3000,5] = 0,4942$

62º) La demanda aleatoria del producto que fabrica nuestra empresa oscila entre 10 y 20 unidades por día. Determinar la probabilidad de que en un periodo de 200 días, el número de unidades demandadas sea mayor de 3.000, bajo el supuesto de que las demandas en los distintos días son independientes entre sí.

Sol.: Consideraremos que la variable $X_i = \text{"nº de unidades demandadas el día } i"$ se distribuye de manera uniforme entre 10 y 20, por lo que $\mu = 15$ y $\sigma^2 = \frac{100}{12}$ luego la variable X

$= \sum_{i=1}^{200} X_i$ es aproximadamente $N\left(3000, \frac{100\sqrt{2}}{2\sqrt{3}}\right)$; efectuando la corrección por continuidad,

$P[X > 3000] = P[X \geq 3000,5] = P[Z \geq 0,0122] = 0,4951$

63º) En una empresa de construcción el número mensual medio de días de baja de los trabajadores debidas a accidentes laborales por obra, se distribuye según una Poisson con varianza 15,6. Calcular la probabilidad de que, en cuatro obras independientes, se acumulen más de 68 días, en total, de baja en el periodo de un mes.

Sol.: Si X_i es Poisson (15,6) entonces $X = \sum_{i=1}^4 X_i$ es Poisson(62,4) que es aproximadamente $N(62,4; \sqrt{62,4})$, luego, efectuandola correspondiente corrección por continuidad, $P[X > 68,5] = P[Z > 0,77] = 0,22$.

64º) Una cadena comercial que tiene 81 establecimientos ha comprobado que el volumen de ventas diario de cada uno, en millones de pesetas se ajusta a la función de densidad $f(x) = \frac{x}{50}$ para $0 \leq x \leq 10$. Suponiendo que las ventas de los distintos establecimientos son independientes, se desea conocer. **1º)** La probabilidad de que la venta

total de la cadena en un día determinado sea mayor de 650 millones de pesetas. 2º) El volumen de venta total que es superado con una probabilidad del 30%.

Sol.: 1º) Si X_i = “millones de pesetas de venta diaria del establecimiento i” \Rightarrow

$$\mu = \int_0^{10} \frac{x^2}{50} dx = \frac{20}{3} \quad \text{y} \quad \sigma = \frac{5\sqrt{2}}{3} \quad \text{luego} \quad X = \sum_{i=1}^{81} X_i \text{ es } N(540, 15\sqrt{2}) \Rightarrow P[X > 650] =$$

$$P\left[Z > \frac{110}{15\sqrt{2}}\right] \cong 0; \quad 2^\circ) P[X > x] = 0,3 \Rightarrow x = 551 \text{ millones de pesetas.}$$

Chi-cuadrado

65º) Una variable aleatoria χ^2 tiene 10 grados de libertad. Hallar la media, la varianza y la probabilidad de que dicha variable aleatoria sea mayor que 9,342.

Sol.: $\mu=10$, $\sigma^2 = 20$, $P(\chi^2 \geq 9,342) = 0,5$

t-Student

66º) Se consideran dos variables aleatorias independientes X e Y. La variable X tiene una distribución normal $N(0,1)$. La variable Y tiene una distribución χ^2 con 4 grados de libertad. Hallar en $P\left[\frac{2X}{\sqrt{Y}} \leq m\right] = 0,05$ el valor de m.

Sol.: $m = -2,132$

67º) Si U y V son dos variables aleatorias independientes tales que la variable U tiene una distribución normal $N(0,1)$ y V tiene una distribución χ^2 con 25 grados de libertad, hallar $P\left[\frac{5U}{\sqrt{V}} \leq -1,058\right]$.

Sol.: $p = 0,15$

F-Snedecor

68º) Las variables aleatorias X e Y son independientes y tienen distribuciones χ^2 con 30 y 10 grados de libertad respectivamente. Hallar en $P\left[\frac{X}{3Y} > m\right] = 0,05$ el valor de m.

Sol.: $m = 2,70$

69º) Se consideran dos variables aleatorias χ^2 , U y V independientes, con 5 y 30 grados de libertad, respectivamente. Hallar $P\left[6\frac{U}{V} < 3,70\right]$.

Sol.: $p = 0,99$

70º) Hallar: 1) t_1 en $P(t \leq t_1) = 0,90$, siendo $n = 9$ y 2) f_1 en $P(F > f_1) = 0,01$, siendo n_1 y n_2 iguales a 12.

Sol.: 1) $t_1 = 1,383$; 2) $f_1 = 4,16$

71º) Se consideran dos variables aleatorias χ^2 , U y V, independientes, con 5 y 30 grados de libertad, respectivamente. Hallar $P\left[6\frac{U}{V} \leq 2,53\right]$.

Sol.: 0,95

TEMAS 5, 6 Y 7: MUESTREO Y ESTIMACIÓN PUNTUAL.

72º) Consideremos una variable aleatoria X representativa de una población, cuya función de densidad es $f(x) = \begin{cases} 3k, & 4 \leq x \leq 6 \\ 0 & \text{en el resto} \end{cases}$, donde k es una constante a determinar.

Supongamos extraída una muestra aleatoria simple de tamaño 50. Se pide: **a)** calcular la media y la varianza de la “**media muestral**”; **b)** calcular la media de la “**varianza muestral**”.

Sol.: $k = \frac{1}{6}$, $\mu = E(X) = 5$ y $\sigma^2 = \text{Var}(X) = \frac{76}{3} - 25 = \frac{1}{3}$ luego: **a)** $E(\bar{X}) = 5$;
 $\text{Var}(\bar{X}) = \frac{1}{150}$; **b)** $E(S^2) = \sigma^2 = \frac{1}{3}$

73º) De una población normal $N(\mu, 1)$ se obtienen muestras de tamaño 2; como estimadores de μ se consideran los siguientes:

$$\hat{\mu}_1 = \frac{2}{3}X_1 + X_2; \quad \hat{\mu}_2 = \frac{2}{5}X_1 + \frac{4}{5}X_2; \quad \hat{\mu}_3 = \frac{X_1 + X_2}{2}$$

Se pide: **a)** determinar si son o no estimadores insesgados; **b)** hallar su varianza; **c)** estudiar su eficiencia; **d)** estudiar su distribución en el muestreo.

Sol.: **a)** $E(\hat{\mu}_1) = \frac{5}{3}\mu \Rightarrow$ no es insesgado; $E(\hat{\mu}_2) = \frac{5}{3}\mu \Rightarrow$ no es insesgado; $E(\hat{\mu}_3) = \mu \Rightarrow$ sí es insesgado; **b)** $\text{Var}(\hat{\mu}_1) = \frac{13}{9}\sigma^2 = \frac{13}{9}$; $\text{Var}(\hat{\mu}_2) = \frac{4}{5}\sigma^2 = \frac{4}{5}$; $\text{Var}(\hat{\mu}_3) = \frac{1}{2}\sigma^2 = \frac{1}{2}$; **c)** $\hat{\mu}_3$ es el más eficiente por ser insesgado y poseer la menor varianza; **d)** $\hat{\mu}_1$ es $N\left(\frac{5}{3}\mu, \frac{\sqrt{13}}{3}\right)$; $\hat{\mu}_2$ es $N\left(\frac{5}{3}\mu, \frac{2}{\sqrt{5}}\right)$; $\hat{\mu}_3$ es $N\left(\mu, \frac{1}{\sqrt{2}}\right)$

74º) Supongamos que el Banco de España decide efectuar una investigación sobre los rendimientos obtenidos por la banca española con un determinado producto financiero. Para ello selecciona una muestra aleatoria simple de 9 bancos, y además dispone de la información de que los rendimientos de producto en cuestión, en todo el conjunto bancario, se distribuye según una distribución normal de media 6% y de desviación típica del 3%. Sobre la base de ello se pide: **a)** ¿Cuál es la probabilidad de que el rendimiento medio muestral se mantenga entre el 5% y el 7%?. **b)** ¿Cuál es la probabilidad de que la varianza muestral sea superior a 9?. **c)** El valor de K tal que $P[S^2 > K] = 0,98$. **d)** Suponiendo, ahora, que la desviación típica para todo el conjunto bancario fuera desconocida, y conociésemos que la desviación típica de la muestra de 9 bancos es del 2%, se pide obtener la probabilidad de que la media muestral sea superior al 8%.

Sol.: **a)** 0,6827; **b)** 0,4335; **c)** $K = 2,2865$; **d)** 0,0085

75º) Como estimador del parámetro a de la función de densidad $f(x; a) = a \cdot e^{-ax}$, para $x \geq 0$, en muestras aleatorias simples de tamaño n , se considera el estadístico: $R = \frac{n}{\sum_{i=1}^n X_i}$.

Demuéstrese que es un estimador suficiente.

Sol.: $f(x_1, x_2, \dots, x_n, a) = a^n \cdot e^{-a \sum x_i} = a^n \cdot e^{-\frac{a}{R} n} = \left(a \cdot e^{-\frac{a}{R}} \right)^n \cdot 1$. Llamando $g(R, a) = \left(a \cdot e^{-\frac{a}{R}} \right)^n$ y $h(x_1, x_2, \dots, x_n) = 1$, del teorema de factorización de Fisher-Neyman se deduce que R es suficiente para estimar a .

76º) La estimación de un parámetro a partir de una muestra se puede comparar al tiro al blanco con fusil. En este paralelismo:

- El centro de la diana representa el verdadero valor del parámetro.
- Cada disparo representa una estimación (muestra) concreta.
- El fusil es el estimador (es decir, la fórmula de estimación).

En el planteamiento de este símil, ¿cuándo diremos que el fusil será eficiente?

Sol.: Debe ser en primer lugar insesgado, es decir, el valor esperado debe ser el centro de la diana (es decir, cuando se apunte al centro se espera que dé en el centro) y además la varianza del estimador debe ser la mínima, es decir, el fusil no tiene que tener ninguna desviación.

77º) ¿Guarda alguna relación el concepto de estimador eficiente y la cota de Cramer-Rao.

Sol.: La cota de Cramer-Rao proporciona un límite inferior para la varianza del estimador. Si un estimador es eficiente, su varianza es la mínima y coincide con la cota de Cramer-Rao

78º) De una población binomial se extrae una muestra $X_i = \{1, 1, 0, 1, 0, 0, 1, 0, 1, 0\}$. Estímese el parámetro “p” por el método de máxima verosimilitud.

$$\text{Sol.: } p = \frac{1}{2}$$

79º) De una población, determinada por una variable aleatoria con función de densidad $f(x) = \theta e^{-\theta x}$, $0 < x < \infty$, se extrae la muestra $X_i = \{1; 1,5; 2; 2,3; 3; 3,1; 3,7; 3,9; 4\}$. Estímese el parámetro desconocido por el método de los momentos.

$$\text{Sol.: estimador } \hat{\theta} = \frac{1}{\bar{X}}; \text{ estimación } \hat{\theta} \cong 0,3673$$

80º) La distribución de probabilidad de una variable aleatoria es $P[X = x] = \frac{\theta^x}{x!} e^{-\theta}$, $x = 0, 1, 2, 3, \dots$. Determinar el estimador de máxima verosimilitud del parámetro θ , para muestras de tamaño n .

$$\text{Sol.: } \hat{\theta} = \bar{X}$$

81º) Estimar por el método de la máxima verosimilitud el parámetro “a” de una distribución de ingresos económicos cuya función de densidad es $f(x) = a \cdot e^{-ax}$, donde $x > 0$ y $a > 0$.

$$\text{Sol.: } \hat{a} = \frac{n}{\sum X_i}$$

82º) El número de piezas defectuosas fabricadas al día por parte de una empresa sigue el modelo de Poisson $P[X = x] = \frac{e^{-\lambda} \lambda^x}{x!}$, $x = 0, 1, 2, \dots$. Estimar λ por el método de los momentos y de la máxima verosimilitud.

Sol.: por ambos métodos se obtiene que $\hat{\lambda} = \bar{X}$

83º) Una variable aleatoria tiene por función de densidad $f(x, \theta) = \frac{1}{6 \cdot \theta^4} e^{-\frac{x}{\theta}} \cdot x^3$, para $x \geq 0$. Se extrae una muestra aleatoria simple de tamaño 2. Se pide: **a)** hallar el estimador de θ por el método de la máxima verosimilitud y por el método de los momentos; **b)** demostrar si es insesgado y eficiente el estimador obtenido por cada uno de los dos métodos anteriores.

Sol.: **a)** Por ambos métodos se obtiene que $\hat{\theta} = \frac{X_1 + X_2}{8}$; **b)** $E(\hat{\theta}) = \frac{E(X_1) + E(X_2)}{8} = \frac{4\theta + 4\theta}{8} = \theta \Rightarrow$ es insesgado; calculamos $E(X^2) = 20\theta^2 \Rightarrow \text{Var}(X) = 20\theta^2 - 16\theta^2 = 4\theta^2$, luego $\text{Var}(\theta) = \frac{8\theta^2}{64} = \frac{\theta^2}{8}$. Por otra parte, cota de Cramer-Rao = $\frac{1}{2E\left[\left(\frac{\partial \ln\left(\frac{1}{6\theta^4} e^{-\frac{x}{\theta}} \cdot x^3\right)}{\partial \theta}\right)^2\right]} = \frac{\theta^2}{8}$, luego θ es eficiente.

84º) Consideremos una variable aleatoria X de una determinada población y una muestra aleatoria simple de tamaño 16. Se pide: **a)** Si X es $N(\mu, 2)$, determinar $P[S^2 \leq 2,14]$, siendo S^2 la varianza muestral; **b)** si X siguiera una distribución de Poisson de parámetro λ , obtener el estimador máximo verosímil de λ .

Sol.: **a)** $P = 0,08$; **b)** $\lambda = \bar{X}$

85º) Dada la función de densidad $f(x; a) = \frac{x}{a^2} e^{-\frac{x}{a}}$, con $x \geq 0$; $a \geq 0$, se pide: **a)** calcular el estimador de máxima verosimilitud de a en muestras de tamaño n ; **b)** ¿es insesgado?

Sol.: **a)** $L(X_1, X_2, \dots, X_n, a) = \frac{\prod_{i=1}^n X_i}{a^{2n}} e^{-\frac{\sum_{i=1}^n X_i}{a}} \rightarrow \ln L = \sum_{i=1}^n \ln X_i - 2n \ln a - \frac{\sum_{i=1}^n X_i}{a} \rightarrow$

$$\frac{\partial \ln L}{\partial a} = -\frac{2n}{a} + \frac{\sum_{i=1}^n X_i}{a^2} = 0 \rightarrow \hat{a} = \frac{1}{2n} \sum_{i=1}^n X_i \quad (\text{Se trata de un máximo pues})$$

$$\frac{\partial^2 \ln L}{\partial a^2} = \frac{2}{a^3} \left(an - \sum_{i=1}^n X_i \right) < 0 \quad \text{para } a = \frac{1}{2n} \sum_{i=1}^n X_i; \quad \textbf{b)} \text{ Calculemos en primer lugar } E(X):$$

$$E(X) = \frac{1}{a^2} \int_0^\infty x^2 e^{-\frac{x}{a}} dx = 2a \quad (\text{integrando dos veces por partes}). \text{ Luego } E(\hat{a}) =$$

$$= \frac{1}{2n} \sum_{i=1}^n E(X_i) = \frac{1}{2n} \sum_{i=1}^n 2a = \frac{2na}{2n} = a, \text{ por tanto sí es insesgado.}$$

86º) Una empresa ha decidido lanzar al mercado una determinada marca de coche. Al objeto de planificar su producción, supone que el coche que va a ofrecer puede ser adquirido

por el 10% o por el 20% de los habitantes de una gran ciudad. A tal efecto, consultados 10 habitantes, sólo 3 de ellos se muestran dispuestos a la adquisición del coche. Si suponemos que la muestra de las 10 consultas obtenidas es aleatoria simple, se pide conocer qué porcentaje de los dos contemplados será tomado en consideración por la empresa si la elección entre ambos se efectúa con base en el criterio de máxima verosimilitud.

Sol.: Si $p = 0,1 \Rightarrow P[X = 3] = \binom{10}{3} \cdot 0,1^3 \cdot 0,9^7 = 0,05$; si $p = 0,2 \Rightarrow P[X = 3] = \binom{10}{3} \cdot 0,2^3 \cdot 0,8^7 = 0,20$ luego es más verosímil suponer que $p = 0,2$.

TEMA 8: ESTIMACIÓN POR INTERVALOS DE CONFIANZA

87º) Obtener un intervalo de confianza del 99% para la media de una población normal, siendo la media muestral 10, la desviación estándar poblacional 4 y el tamaño de la muestra 49 (Sol.: [8,5351 , 11,4649])

88º) En una población normal tal que, para una muestra de tamaño 10, la varianza muestral es 4 y la media muestral es 11, encontrar un intervalo de confianza del 90% para la media de la población. (Sol.: [9,778 , 12,22])

89º) De una población normal se extrae una m.a.s. de tamaño 25, calculándose la desviación estándar muestral que es 4. Hallar un intervalo de confianza del 90% para la desviación estándar de dicha población (tómense áreas iguales en los extremos de la distribución correspondiente). (Sol.: [3,31 , 5,37])

90º) Obtener un intervalo de confianza del 95% para la media de una población normal, siendo la media muestral 12, la desviación estándar poblacional 3 y el tamaño de la muestra 36. (Sol.: [11,02 , 12,98])

91º) En una población normal se extrae una muestra aleatoria simple de tamaño 12, calculándose la desviación estándar muestral, que es 3. Hallar un intervalo de confianza del 90% para la varianza de dicha población (tómense áreas iguales en los extremos de la distribución correspondiente). (Sol.: [5,49 , 23,6])

92º) Obtener un intervalo de confianza del 95% para la media de una población normal, siendo la media muestral 6, la desviación estándar poblacional 4 y el tamaño de la muestra 25 (Sol.: [4,432 , 7,568])

93º) De una población normal se extrae una m.a.s. de tamaño 16, calculándose la desviación estándar muestral que es 4. Hallar un intervalo de confianza del 95% para la varianza de dicha población (tómense áreas iguales en los extremos de la distribución correspondiente). (Sol.: [9,31 , 40,88])

94º) En una población normal tal que, para una muestra de tamaño 10, la varianza muestral es 4 y la media muestral es 5, encontrar un intervalo de confianza del 90% para la media de la población. (Sol.: [3,778 , 6,222])

95º) Obtener un intervalo de confianza del 95% para la media de una población normal, siendo la media muestral 10, la desviación estándar poblacional 4 y el tamaño de la muestra 25 (Sol.: [8,432 , 11,568])

96º) Dadas la media muestral, cuyo valor es 4 y la desviación estándar muestral que es igual a 3, estando la variable aleatoria distribuida normalmente, determinar los límites de confianza del 95% para μ con una muestra de tamaño 10. (Sol.: [1,738 , 6,262]) [jun-94-1ª]

97º) Elegida una muestra de 100 pilas se observa una duración media de 158 horas con una desviación típica muestral de 30 horas. Hallar un intervalo de confianza de nivel 0,99 para la duración media μ . (Sol.: [150 , 166])

98º) Conocidos $\bar{X} = 3$; $S = 2$ y $n = 12$, estando X distribuida normalmente, encontrar los límites de confianza del 90% para μ . (Sol.: [1,92 , 4,08])

99º) Conocidos $\bar{X} = 10$; $S = 3$ y $n = 10$, estando X distribuida normalmente, encontrar los límites de confianza del 95% para μ . (Sol.: [7,738 , 12,262])

100º) Obtener un intervalo de confianza del 95% para la media de una población normal, siendo la media muestral 10, la desviación estándar poblacional 4 y el tamaño de la muestra 16 (Sol.: [8,04 , 11,96])

101º) Las ventas anuales de cierta empresa se distribuyen Normal con media μ (desconocida) y desviación típica 2 unidades monetarias (u.m.). En los últimos años se han observado las siguientes ventas en u.m.: 12, 13, 10 y 13. Construir un intervalo de confianza para el parámetro μ al nivel de confianza del 90 %.

Sol.: [10,35 ; 13,65]

102º) La distribución de la "calidad de cierto producto" es una población $N(\alpha, 1)$. Estimar el intervalo de confianza, del 95 %, para el parámetro α si la muestra aleatoria simple de calidades observadas es: 1, 1, 2 y 1.

Sol.: [0,27 ; 2,23]

103º) El peso en toneladas del rendimiento agrario de ciertas tierras cultivadas, sigue una distribución $N(\alpha, 3)$. Los últimos rendimientos independientes observados han sido 48, 53 y 49. Obtener un intervalo de confianza del 95%, de nivel de confianza, para la media α .

Sol.: [46,60 ; 53,39]

104º) Debido a las tendencias en materia sanitaria, una determinada fábrica de cigarrillos ha establecido un menor contenido de alquitrán en sus productos. Con el fin de comprobar que se han cumplido los objetivos se procede a realizar un experimento con 20 cigarrillos elegidos al azar de lotes diferentes. Se tienen los siguientes datos muestrales, procedentes de una distribución normal, para el contenido del alquitrán: $\bar{x} = 22$ mg; $s = 4$ mg. Se pide un intervalo de confianza del 90% para el contenido medio de alquitrán de la citada marca. [Sol.: el intervalo [20,45 , 23,55]].

105º) Una determinada población sigue una distribución normal $N(\mu, \sigma)$ y por procedimiento de muestreo aleatorio simple se extraen muestras de tamaño 20.

Para una muestra concreta de tamaño 20, se ha obtenido una Media Muestral de valor 12 y una Varianza Muestral de valor 0,81. Se pide: **A)** Estimar un intervalo de confianza del 95 % para la "Media Poblacional" μ ; **B)** Estimar un intervalo de confianza del 98 % para la "Varianza Poblacional" σ^2 . [Sol.: $11,58 < \mu < 12,42$ y $0,43 < \sigma^2 < 2,01$]

106º) El precio del kilo de merluza en las pescaderías de Madrid sigue una distribución normal; se toma una muestra aleatoria de 10 pescaderías y se observan los siguientes precios del artículo:

1270 1230 1350 1240 1300
1400 1250 1260 1200 1000

Obtener, para un nivel de confianza del 95 %: **a)** Un intervalo de confianza para la media poblacional; **b)** Un intervalo de confianza para la varianza poblacional. [Sol.: **a)** $1178,11 < \mu < 1321,89$; **b)** $4778,42 < \sigma^2 < 33666,67$]

TEMAS 9 Y 10: CONTRASTE DE HIPÓTESIS

107º) Se quiere contrastar la hipótesis nula $H_0 (\theta = 1)$, frente a la alternativa $H_1 (\theta = 3)$, en una población que se distribuye según una normal $N(\theta, 1)$, utilizando muestras de tamaño dos cuyos resultados han sido $(x_1 = 2, x_2 = 3)$. Determinar la mejor región crítica de tamaño $\alpha = 0,05$ para efectuar el contraste y llevarlo a cabo.

$$\text{Sol.: } \frac{L_0}{L_1} = \frac{e^{-\frac{1}{2} \sum_{i=1}^2 (x_i - 1)^2}}{e^{-\frac{1}{2} \sum_{i=1}^2 (x_i - 3)^2}} = e^{\frac{1}{2} \sum_{i=1}^2 [(x_i - 1)^2 - (x_i - 3)^2]} = e^{\frac{1}{2} \sum_{i=1}^2 (-4x_i + 8)} = e^{2 \sum_{i=1}^2 (-x_i + 2)} \leq k, \text{ dentro de}$$

C (región crítica), es decir: $2 \sum_{i=1}^2 (-X_i + 2) \leq \ln k \Leftrightarrow \bar{X} \geq \frac{1}{4} \ln k - 2 = \bar{x}_c$. Bajo la hipótesis H_0 ,

\bar{X} es $N\left(1, \frac{1}{\sqrt{2}}\right)$ luego $0,05 = P[\bar{X} > \bar{x}_c] = P\left[Z > \frac{\bar{x}_c - 1}{1/\sqrt{2}}\right] \rightarrow \frac{\bar{x}_c - 1}{1/\sqrt{2}} = 1,65 \Rightarrow \bar{x}_c = 2,17$. En la muestra efectuada $\bar{x} = 2,5$, luego debe rechazarse H_0 .

108º) De una población normal se obtiene una muestra de cinco individuos cuyos ingresos anuales, en millones, son $X_i = \{3; 2,5; 4; 2; 4,5\}$. Contrastar la hipótesis de que la media de la población sea $\mu = 3$, con un nivel de significación del 5 %.

Sol.: Sea $H_0: \mu = 3$ y $H_1: \mu \neq 3$. Bajo la hipótesis nula, la variable $\frac{\bar{X} - 3}{S} \sqrt{5}$ es t_4 y de las tablas se obtiene la región crítica de tamaño 0,05: $|t_4| \geq 2,276$. Para la muestra efectuada, $\frac{\bar{x} - 3}{s} \sqrt{5} = 0,4313$ luego se acepta H_0 .

109º) Un comerciante vende naranjas cuyo peso individual es una variable aleatoria normal cuya media es 180 gr según nos asegura. Un cliente pesa las naranjas que le ha comprado y sus pesos individuales han resultado ser (en gramos): 170, 150, 190 y 160. ¿Aceptará el cliente la hipótesis del comerciante a partir de la información proporcionada por tal muestra, con un nivel de significación del 5%?

Sol.: Sea $H_0: \mu = 180$ y $H_1: \mu \neq 180$. Bajo la hipótesis nula, la variable $\frac{\bar{X} - 180}{S} \sqrt{3}$ es t_3 y de las tablas se obtiene la región crítica de tamaño 0,05: $|t_3| \geq 3,182$. Para la muestra efectuada, $\frac{\bar{x} - 180}{s} \sqrt{3} = -1,268$ luego se acepta H_0 .

110º) En una población $N(\mu, 1)$, se pretende contrastar la hipótesis $H_0: \mu = 2$, frente a la hipótesis alternativa $H_1: \mu = 6$; para efectuar el contraste nos hemos decidido por la siguiente prueba:

- Aceptar $H_0: \mu = 2$, cuando la media muestral sea menor que 4.
- Aceptar $H_1: \mu = 6$, cuando la media de la muestra sea mayor que 4.

Suponiendo que dicha contrastación se va a realizar a través de la inferencia implícita en muestras aleatorias simples de tamaño 10, se pide: **1)** ¿cuál es el valor del nivel de significación α ?; **2)** ¿cuál es la probabilidad de cometer error de tipo II?; **3)** ¿cuál es la potencia del contraste?; **4)** efectuar el contraste suponiendo que la muestra concreta obtenida ha

ido
(1, 3, 5, 4, 4, 3, 4, 5, 2, 3).

Sol.: **1)** Bajo la hipótesis nula, \bar{X} es $N\left(2, \frac{1}{\sqrt{10}}\right) \Rightarrow \alpha = P(\bar{X} > 4) \cong 0$; **2)** Bajo la hipótesis alternativa, \bar{X} es $N\left(6, \frac{1}{\sqrt{10}}\right) \Rightarrow \beta = P(\bar{X} < 4) \cong 0$; **3)** potencia del contraste = $1 - \beta \cong 1$; **4)** puesto que $\bar{x} = 3,4$, se acepta H_0 .

111º) Una variable aleatoria X puede tener una de las dos distribuciones de probabilidad siguientes:

$$\left. \begin{array}{l} f(x) = 2 - 2x \\ h(x) = -6x^2 + 6x \end{array} \right\} \text{ siendo } 0 \leq x \leq 1$$

Mediante el lema de Neyman-Pearson, y con una muestra aleatoria de tamaño uno, determine la mejor región crítica de tamaño 0,1, si se considera como hipótesis nula la función $f(x)$ y como hipótesis alternativa la función $h(x)$. Calcule la potencia del contraste.

$$[\text{Sol.: } \frac{2 - 2x}{-6x^2 + 6x} \leq K \text{ dentro de } R \Leftrightarrow \frac{2}{6x} \leq K \Leftrightarrow x \geq K']$$

$$0,1 = \int_x^1 (2 - 2t) dt = 1 - 2x + x^2 = (1 - x)^2 \Leftrightarrow x = 1 - \sqrt{0,1} \cong 0,6838, \text{ luego la mejor región}$$

crítica es el intervalo $[0,6838, 1]$.

$$\beta = P(0 < x < 0,6838 \mid X \text{ se distribuye con función de densidad } h(x)) = \int_0^{0,6838} (-6x^2 + 6x) dx = 0,7632. \text{ Luego la potencia de contraste } 1 - \beta = 0,2367.$$

112º) En la función de densidad $f(x; a) = ae^{-ax}$ para $x > 0$, se ha contrastado la hipótesis nula $H_0(a = a_0)$ frente a la hipótesis alternativa $H_1(a = a_1)$, con un nivel de significación del 15%, mediante una muestra aleatoria de tamaño 1 y una potencia de 0,92, resultando como mejor región crítica $x < 0,054$ (siendo x el valor muestral). Calcúlese los valores del parámetro “a” bajo las dos hipótesis (a_0, a_1).

$$\text{Sol.: } 0,15 = \int_0^{0,054} a_0 e^{-a_0 x} dx = 1 - e^{-0,054 a_0} \Rightarrow a_0 \cong 3$$

$$0,92 = 1 - e^{-0,054 a_1} \Rightarrow a_1 \cong 46,77$$

113º) Una población normal tiene varianza $\sigma^2 = 9$. Encontrar la probabilidad β para la hipótesis $H_0: \mu = 1$, frente a $H_1: \mu = 2$ a un nivel de significación del 10% para una muestra de tamaño 25 (β , probabilidad de cometer error de tipo II). [Sol.: $P(Z < -0,387) = 0,3483$]

114º) En una distribución normal de media cero se efectúan dos hipótesis sobre la varianza. Una hipótesis nula, que su valor es igual a 16, y una hipótesis alternativa, que es igual a 4. Obténgase la mejor región crítica para $n = 10$ y un nivel de significación $\alpha = 0,10$.

¿Rechazaría la hipótesis de que $\sigma = 4$ si $\sum_{i=1}^{10} x_i^2 = 85,9$? [Sol.: la mejor región crítica es

$$\sum_{i=1}^{10} x_i^2 \leq 77,84; \text{ no se rechazaría la hipótesis}]$$

115º) En una población normal de desviación estándar 4, hallar el tamaño de la muestra aleatoria simple para contrastar la hipótesis nula de que la media es cero frente a la alternativa de que la media es -3 , siendo la probabilidad del error de tipo I igual a 2,5% y la probabilidad del error de tipo II igual al 10%. [Sol.: $n = 19$]

116º) Supongamos que X_1, X_2, \dots, X_n es una muestra aleatoria de una población con distribución normal con media $\mu = 10$. Encontrar la mejor región crítica para contrastar $H_0 (\sigma^2 = 16)$ respecto a $H_1 (\sigma^2 = 36)$. Aplicarla al caso $n = 12$ y $\alpha = 0,05$. [Sol.: *la mejor región*

crítica es $\sum_{i=1}^{12} (x_i - 10)^2 \geq 336,48$]

117º) En una población normal de media cero y varianza σ^2 se quiere contrastar la hipótesis $H_0 (\sigma^2 = 4)$ frente a la alternativa $H_1 (\sigma^2 = 6,4)$. Si se toma una muestra aleatoria de tamaño 10 y $\sum_{i=1}^{10} x_i^2 = 45,2$, ¿debería aceptarse la hipótesis H_0 a un nivel de significación del

5%? Justifique la respuesta. [Sol.: *se acepta H_0 pues $R: \sum_{i=1}^{10} x_i^2 \geq 73,228$]*

118º) En una población normal $(\mu, 1)$ se desea contrastar la hipótesis $H_0 (\mu = 1)$ frente a la alternativa $H_1 (\mu = 4)$ con un nivel de significación del 5% y una potencia de contraste del 95%. Determinar el tamaño de la muestra. [Sol.: $n = 2$]

119º) Si en una población normal de media cero se quiere contrastar la hipótesis nula de que la desviación estándar es 10 frente a la alternativa de que es 5 (supuestos una muestra y un nivel de significación), ¿cuál sería la expresión de la mejor región crítica para efectuar dicho contraste?. [Sol.: $\frac{1}{100} \sum_{i=1}^n x_i^2 \leq k$, obteniéndose k de la distribución χ^2 con n grados de libertad]

120º) En una población normal de media 0 y desviación estándar σ se quiere contrastar la hipótesis $\sigma = 4$ frente a la alternativa $\sigma = 2$. Se toma una muestra de tamaño 10, siendo

$\sum_{i=1}^{10} x_i^2 = 125,82$. Determinar la mejor región crítica de tamaño 0,05 para efectuar dicho contraste e indicar si se rechaza la hipótesis $\sigma = 4$. [Sol.: *se acepta pues se rechazaría si*

$\sum_{i=1}^{10} x_i^2 \leq 63,04$]