INGENIERÍA EN INFORMÁTICA DE GESTIÓN
ESTADÍSTICA
junio, 98

TEORÍA (1.5 puntos)

Cuestión 1. Sean A y B dos sucesos tales que $P(A) = 1/3$, $P(B) = 1/5$ y $P(A|B) + P(B|A) = 2/3$. Calcular $P(A \cup B)$.

Cuestión 2. Sean X_1, X_2, X_3, X_4, X_5 variables aleatorias independientes normales con media 10 y varianza 81. Sean U y V normales estandard independientes e independientes de las Xi, $i = 1, 2, 3, 4, 5$. Determinar la distribución, la media y la desviación típica de la variables $S = X_1 + X_2 + X_3 + X_4 + X_5$ y $T = S + U + V$.

Cuestión 3. Deseamos estimar la media poblacional de una cierta variable. Para ello tomamos dos muestras con el mismo tamaño muestral pero diferentes desviaciones típicas. La muestra 1 tiene una desviación típica de 2.3 y la segunda de 4.7. ¿Con qué muestra conseguiremos un menor error en la estimación por intervalo de confianza de la media poblacional?. Justifícalo.

Ejercicio 1 (1pto):

Dos empresas fabrican microprocesadores cada una de un tipo (Tipo I y II) pero son tan parecidos que sólo se pueden distinguir con la ayuda de un microscopio. La forma habitual de intentar distinguirlos es observando la presencia o ausencia de un anillo magnético. El 90% de los microprocesadores del tipo I y el 20% de los del tipo II tienen el anillo. Se sabe que en la provincia en la que se trabaja el 70% son del tipo I. Se pide:

a) Supongamos que nos encontramos con un microprocesador con anillo y decidimos que es del tipo I. ¿Con que probabilidad estaremos en lo cierto?

b) Si todos los microprocesador con anillo son clasificados como del tipo I y los que no lo tienen como del tipo II, ¿qué proporción de microprocesadores estará correctamente clasificado?

Ejercicio 2 (2 puntos):

En un quiosco de periódicos se supone que el número de ventas diarias se distribuye normalmente con media 30 y varianza 2. Determinar:

a) Probabilidad de que en un dia se vendan entre 13 y 31 periódicos.

b) Determinar el máximo número de periódicos que se venden en el 90% de las ocasiones.

c) Supongamos que en una ciudad hay 10 quioscos del mismo tipo y con las mismas características. Además se consideran independientes. Determinar la probabilidad de que más de dos quioscos vendan entre 13 y 31 periódicos.

Ejercicio 3 (2.25 puntos):

Se realizaron investigaciones con el fin de estudiar la relación entre la elevación de la temperatura de las celdas solares en ºC por encima de la temperatura ambiente (y) y la cantidad de aislamiento en megawatts por centímetro cuadrado (x)

x	9	25	20	12	15	22	14	16	24	25	15	12	10	10
y	25	70	50	30	45	60	28	50	68	68	20	21	20	20

a) Calcula la tabla de frecuencias conjuntas. ¿Qué porcentaje de valores tienen entre 16 y 20 grados y un aislamiento de más de 50 megawatts por centímetro? Calcular el intervalo donde se encuentra el 50% de las temperaturas centrales observadas.

b) Para un aislamiento de 62 ¿qué temperatura cabe esperar? Discute la fiabilidad de tu respuesta.

c) Construye un intervalo de confianza para la diferencia de medias de la temperatura y el aislamiento a un nivel de confianza del 95% , suponiendo poblaciones normales con varianzas iguales. ¿Podemos suponer con esta confianza que la media de y supera a la de x en 10?

Ejercicio 4 (2.25 puntos):

La variable aleatoria X, tiene como función de densidad:

$f(x) = x$ si $0 < x < 1$; $f(x) = a - x$ si $1 < x < 2$ y 0 en el resto.

a) Determinar el valor de a para que sea función de densidad. Probabilidad de que la variable aleatoria tome un valor entre 0.5 y 1.5. $E(X)$.

b) Sea la variable $Y = \theta X$, donde θ es un parámetro desconocido que se desea estimar en base a una muestra. Probar si el estimador $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ es un estimador insesgado de θ.

ESTADÍSTICA EN INGENIERÍA TÉCNICA EN INFORMÁTICA DE GESTIÓN
CONVOCATORIA SEPTIEMBRE 1999

1. (2 ptos) En una empresa disponen de ordenadores PC normales y portátiles. Los precios de 10 ordenadores PC normales seleccionados al azar, en miles de pesetas, son:

$132, 148, 102, 184, 141, 156, 118, 131, 166, 172$

Se pide:

a) Calcular los correspondientes intervalos de confianza para la media y desviación típica con una confianza del 95%, suponiendo normalidad en los precios anteriores.

b) Se sabe que la media total de los precios de todos los ordenadores es de 158000 ptas y que en la empresa disponen de 140 ordenadores normales y 10 portátiles. ¿Cuál es la media del precio de los portátiles?

2. (2 ptos) Una función continua viene definida por

$$f(x) = \begin{cases} -\frac{49}{16}x + 2 & 0 < x < b \\ \frac{1}{4} & b \leq x < 2 \end{cases}$$

Se pide:

a) Calcular el valor de b.

b) Calcular la función de distribución y utilizándola determinar $P(2.3 < x < 2.8)$ y $P(x > 1.2)$.

c) Calcular los estadísticos mediana y varianza.

d) Calcular los siguientes percentiles: 21, 89.

e) Calcular las siguientes probabilidades $P(-1 < x < 1/2)$ y $P(x = 1/3)$.

3. (2 ptos) Se plantean las siguientes cuestiones:

Cuestión 1. Sea X una variable aleatoria de forma que $E(X) = 1/\beta$ y $Var(X) = 1/\beta^2$. Se desea estimar el parámetro $\theta = 1/\beta^2$ y para ello tomamos una muestra de tamaño n, $X_1, \ldots, X_n$. Como estimador elegimos $\hat{\theta} = \sum_{i=1}^{n} \frac{X_i^2}{2n}$. Comprobar si es insesgado.

Cuestión 2. Sea una población normal con media μ y varianza $\sigma^2 = 4$. Extraida una muestra de tamaño 16, ¿cuál es la probabilidad de que $\overline{x}$ sea mayor que $\mu + 2$?.

4. (2 ptos) En un grupo de 30 empresas se estudiaron las variables $X =$"número de horas trabajadas", $Y =$"salario mensual (miles de pesetas)" que presentaron la siguiente distribución:

x / y	80-110	110-140	140-170	170-200
31-33	4	5	2	0
33-35	1	2	4	3
35-37	0	1	3	5

Se pide:

a) Basándote en estos datos, ¿podemos afirmar que ambas variables están relacionadas? Justifica numéricamente la respuesta.

b) ¿Cuál es la probabilidad de que un empleado que trabaja más de 33 horas semanales cobre 170000 ptas o más?

c) Por los resultados de una encuesta sabemos que un trabajador se considera “satisfecho” si trabaja menos de 33 horas semanales o cobra más de 110000 ptas mensuales. ¿Cuál es la probabilidad de que un empleado de este colectivo se sienta “satisfecho”?

INGENIERIA TECNICA EN INFORMATICA DE GESTION (ITIG)
ESTADÍSTICA (F04). 22 de Junio de 2000

1. (2.5 puntos) (a) Dadas dos muestras de una misma población con n_1 y n_2 datos y sus correspondientes varianzas muestrales comprobar si el siguiente estimador de la varianza poblacional común a ambas muestras es insesgado:

$$SC = \frac{n_1 s_{n_1}^2 + n_2 s_{n_2}^2}{n_1 + n_2 - 2}$$

(b) Determinar el estimador máximo versímil para una muestra aleatoria simple de tamaño n procedente de una distribución de la forma:

$$f(x;\theta) = (\theta+1)x^{\theta} \qquad 0 < x < 1.$$

(c) Supongamos la recta de regresión $y = a + bx$. Determinar la forma de la recta de regresión cuando trabajamos con las variables tipificadas z_y, z_x, donde $z_x = \frac{x-\overline{x}}{s_x}$, $z_y = \frac{y-\overline{y}}{s_y}$.

2. (2 puntos) Con el objeto de establecer un plan de producción, una empresa ha estimado que la demanda aleatoria de sus potenciales clientes se comportará semanalmente con arreglo a la ley de probabilidad siguiente:

$$f(x) = \begin{cases} k(4x - 2x^2) & 0 < x < 2 \\ 0 & \text{en otro caso} \end{cases}$$

donde x viene expresada en millones de unidades. Se pide:

(a) Determinar el valor de la constante desconocida

(b) ¿Qué cantidad deberá tener dispuesta a la venta, al comienzo de cada semana, para poder satisfacer la demanda en dicho periodo, con una probabilidad de 0.5?

(c) Consideramos ahora la demanda en el periodo de una mes, es decir, sea $Y = 4 * X$. Determinar el valor esperado de la demanda durante un mes concreto.

3. (2 puntos) Una compañía dedicada al transporte público explota tres líneas periféricas de una gran ciudad de forma que el 60% de los autobuses cubren el servicio de la primera línea, el 30% cubren el de la segunda y el resto el de la tercera. Se sabe que la probabilidad de que se averíe diariamente un autobús es 2% en la primera línea, 4% en la segunda y un 1% en la tercera. Determinar:

(a) Probabilidad de que un autobús sufra una avería un determinado día

(b) Sabiendo que un autobús ha sufrido una avería, probabilidad de que preste servicio en la primera línea.

(c) Supongamos que la empresa cuenta con 100 autobuses con las mismas características que los anteriores. Probabilidad de que al menos 4 autobuses sufran una avería un cierto día.

4. (2 puntos) La DGT registró durante 5 minutos las velocidades, en km/hora, a la que circulaban los vehículos por una determinada autopista (A) y por una carretera nacional (B), obteniendo los siguientes registros:

A: $98, 115, 95, 137, 122, 141, 101, 97, 98, 110, 133, 127, 140, 135, 122, 95$

B: $78, 105, 75, 107, 102, 71, 101, 67, 88, 100, 93, 82, 100, 105, 72, 85$

Se pide:

(a) Determinar si existe relación lineal entre A y B y en su caso calcular la recta de regresión de A frente a B. Calcular un estadístico de bondad de ajuste.

(b) Existen diferencias significativas entre A y B, con una significación del 5%? Podemos asegurar que la media de B es 80, con un nivel de confianza del 95%?

(c) Con los datos anteriores de A, calcular $P(\overline{x} - \mu < 2)$.

INGENIERIA TECNICA EN INFORMATICA DE GESTION (ITIG)
ESTADÍSTICA (F04). 16 de Septiembre de 2000

1. (2.5 puntos) (a) Dada una colección de variables aleatorias $\{X_i\}_{i=1}^{n}$ con $Cov(X_i, X_j) = -0.3$, ¿podríamos aplicar el Teorema Central de Límite para obtener la $P(\sum X_i < K)$, siendo K una constante?

(b) Supongamos una muestra de tamaño 100 con las siguientes características: Mo (Moda)=5, Q_{50}(mediana)=5.2, a_2(momento ordinario de orden 2)=26 y a_1(momento ordinario de orden 1)=5.1. Aplicar el estadístico necesario para determinar si esta muestra es o no simétrica.

(c) Dada una muestra de una distribución normal con varianza $\sigma^2 = 4$ y media poblacional desconocida, determinar el tamaño muestral necesario para que un intervalo sobre la media desconocida al 95% tenga una amplitud de 2, teniendo en cuenta que la amplitud de un intervalo es la diferencia entre los extremos superior e inferior.

(d) Sea $X \to Bi(n = 3, p = 0.5)$. Determinar la expresión de la Función Generatriz de Momentos dada por $\varphi_X(t) = E(e^{tX})$. Calcular $E(X)$ a través de la función anterior.

2. (2.5 puntos) Una tienda de ordenadores decide premiar a sus clientes con un ordenador PC bajo las siguientes condiciones. Se ponen en una urna 8 papeletas correspondientes a 3 PC, 3 WS (WorkStations) y 2 Mc (Macs). Cada cliente realizará sin devolución 4 extracciones. Se definen las siguientes variables: $X =$"número de PCs extraídos", $Y =$"número de Mcs extraídos". La tienda concede el premio si en un intento se sacan 2 PCs y 1 Mc. Se pide:

(a) Distribución de probabilidad individual de X e Y.

(b) Distribución de probabilidad conjunta.

(c) Determinar las siguientes probabilidades: $P(\text{"premio"})$, $P(X = 1|Y = 1)$.

3. (2 puntos) El beneficio aleatorio que una empresa, dedicada a la prestación de un servicio público, puede obtener a lo largo de 1 año sigue la distribución de probabilidad (con x en millones de ptas)

$$F(x) = \begin{cases} \frac{1}{2}\exp\left(-\left(\frac{x}{50}\right)^2\right) & x \leq 0 \\ 1 - \frac{1}{2}\exp\left(-\left(\frac{x}{50}\right)^2\right) & x > 0 \end{cases}$$

Se pide:

(a) Probabilidad de que el beneficio sea mayor que 50 millones. Probabilidad de que el beneficio sea menor que -50 millones (el menos se entiende por pérdida). Probabilidad de que el beneficio se encuentre entre 100 y 200 millones.

(b) Determinar la función de densidad de la variable beneficio.

(c) Supongamos que en un cierto mes, el beneficio (en millones) ha sido de $Y = (1/6)X + 8$. Determinar $E(Y)$, $Var(Y)$.

4. (1.5 puntos) Para comparar la velocidad del servicio de 2 servidores A y B, hemos anotado el tiempo de espera (en segundos) para conectar con ciertas direcciones en cada uno de ellos. Los resultados han sido:

A: $5, 10, 20, 4, 3, 30, 12, 9, 5, 24, 9, 7$

B: $12, 7, 10, 8, 7, 16, 7, 20, 10, 11, 17, 15$

Se pide:

(a) Justifica si existe relación lineal entre ambas variables.

(b) Tomando como varianzas poblacionales las correspondientes muestrales, averigua si existen diferencias significativas entre el tiempo medio de espera de ambos servidores con una significación del 5%.