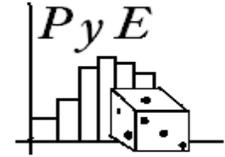




UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE INGENIERÍA
DIVISIÓN DE CIENCIAS BÁSICAS
COORDINACIÓN DE CIENCIAS APLICADAS
DEPARTAMENTO DE PROBABILIDAD Y ESTADÍSTICA
PRIMER EXAMEN FINAL
RESOLUCIÓN



SEMESTRE 2010-1
DURACIÓN MÁXIMA 2.5 HORAS

TIPO 1
DICIEMBRE 1 DE 2009

NOMBRE _____

1. El índice de claridad se determinó en los cielos de Morelos, para cada uno de los 365 días de un año, obteniéndose los siguientes datos.

Límites aparentes	Frecuencia absoluta (f_i)
0.30-0.34	8
0.35-0.39	14
0.40-0.44	28
0.45-0.49	24
0.50-0.54	39
0.55-0.59	51
0.60-0.64	106
0.65-0.69	84
0.70-0.74	11

- a) Determinar las frecuencias relativas y trazar el histograma correspondiente.
b) Los días despejados son aquellos para los que el índice de claridad es por lo menos 0.65. ¿Qué porcentaje de días está despejado?

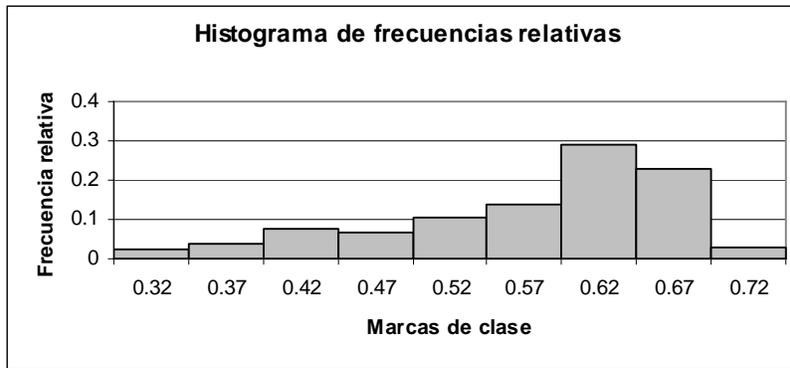
10 Puntos

Resolución

- a) La frecuencia relativa y las marcas de clase, se definen como: $f_i^* = \frac{1}{n} f_i$ y el punto medio de los límites aparentes, respectivamente, sustituyendo se tiene:

Límites aparentes	Frecuencia absoluta (f_i)	Frecuencia relativa (f_i^*)	Marcas de clase (x_i)
0.30-0.34	8	0.0219	0.32
0.35-0.39	14	0.0383	0.37
0.40-0.44	28	0.0767	0.42
0.45-0.49	24	0.0657	0.47
0.50-0.54	39	0.1068	0.52
0.55-0.59	51	0.1397	0.57
0.60-0.64	106	0.2904	0.62
0.65-0.69	84	0.2301	0.67
0.70-0.74	11	0.0301	0.72
	n=365		

El histograma correspondiente es:



- b) Los días despejados son aquellos para los que el índice de claridad es por lo menos 0.65, entonces de la tabla de distribución de frecuencias el porcentaje es: $(0.2301+0.0301)*100=26.02\%$

Límites aparentes	Frecuencia absoluta (f_i)	Frecuencia relativa (f_i^*)	Marcas de clase (x_i)
0.30-0.34	8	0.0219	0.32
0.35-0.39	14	0.0383	0.37
0.40-0.44	28	0.0767	0.42
0.45-0.49	24	0.0657	0.47
0.50-0.54	39	0.1068	0.52
0.55-0.59	51	0.1397	0.57
0.60-0.64	106	0.2904	0.62
0.65-0.69	84	0.2301	0.67
0.70-0.74	11	0.0301	0.72
	n=365		

2. Con base en varios estudios, una compañía ha clasificado de acuerdo con la posibilidad de encontrar petróleo, las formaciones geológicas en tres tipos. La compañía pretende perforar un pozo en un determinado sitio, al que se le asignan las probabilidades de 0.35, 0.4 y 0.25 para los tres tipos de perforación respectivamente. De acuerdo con la experiencia, se sabe que el petróleo se encuentra en un 40% de formaciones de tipo I, en un 20% de formaciones de tipo II y en un 30% del tipo III.
- a) Si la compañía descubre petróleo en ese sitio, determinar la probabilidad de que exista una formación de tipo III.
- b) Determinar la probabilidad de la existencia de una formación del tipo II, si la compañía no encuentra petróleo en ese sitio.

15 Puntos

Resolución

Sean:

I el evento que representa una formación geológica del tipo I.

II el evento que representa una formación geológica del tipo II.

III el evento que representa una formación geológica del tipo III.

A el evento que representa encontrar petróleo.

- a) La probabilidad de encontrar petróleo es la probabilidad total, entonces:

$$P(A) = P(A \cap I) + P(A \cap II) + P(A \cap III)$$

$$P(A) = P(I)P(A|I) + P(II)P(A|II) + P(III)P(A|III)$$

$$P(A) = (0.35)(0.4) + (0.4)(0.2) + (0.25)(0.3) = 0.295$$

Para determinar la probabilidad de que haya una formación geológica de tipo III, dado que la compañía encuentra petróleo en ese sitio, se utiliza el Teorema de Bayes, entonces:

$$P(III|A) = \frac{P(A \cap III)}{P(A)}$$

esto es:

$$P(III|A) = \frac{P(III)P(A|III)}{P(I)P(A|I) + P(II)P(A|II) + P(III)P(A|III)}$$

sustituyendo:

$$P(III|A) = \frac{(0.25)(0.3)}{(0.35)(0.4) + (0.4)(0.2) + (0.25)(0.3)} = 0.2542$$

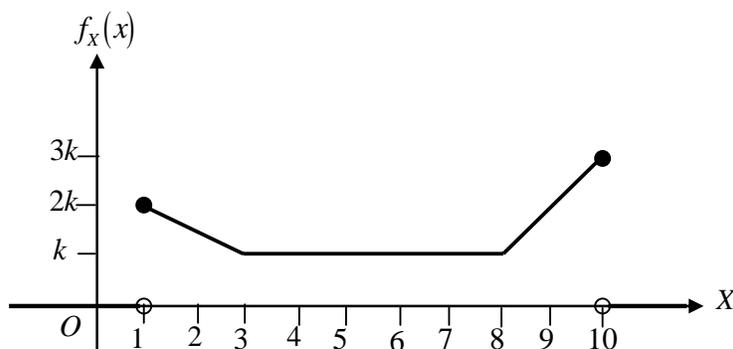
- b) Se calcula la probabilidad de la existencia de una formación del tipo II, si se sabe que la compañía no encuentra petróleo en ese sitio, del Teorema de Bayes, esto es:

$$P(II|\bar{A}) = \frac{P(\bar{A} \cap II)}{P(\bar{A})} = \frac{P(\bar{A} \cap II)}{1 - P(A)}$$

utilizando la regla de la multiplicación:

$$P(II|\bar{A}) = \frac{P(II)P(\bar{A}|II)}{1 - 0.295} = \frac{(0.4)(0.8)}{0.705} = 0.4539$$

3. Dada la función de densidad de la gráfica siguiente:



Determinar: el valor de k , la forma analítica de $f_x(x)$, la función de distribución que muestra el comportamiento acumulado $F_x(x)$ y la mediana \tilde{x} .

15 Puntos

Resolución

El valor de k se obtiene por la propiedad:

$$1 = \int_{-\infty}^{\infty} f_x(x) dx$$

el valor de k en este caso, se puede obtener de forma geométrica, entonces:

$$1 = bh + \frac{bh}{2} + \frac{bh}{2} = 9k + \frac{2k}{2} + \frac{2(2k)}{2} = 12k$$

por lo tanto:

$$k = \frac{1}{12} \approx 0.0833$$

Para determinar la forma analítica de la función de densidad, se utiliza la forma de la recta dados dos puntos, la función constante y otra vez, la forma de la recta dados dos puntos, entonces:

$$f_X(x) = \begin{cases} -\frac{1}{24}x + \frac{5}{24} & ; \quad 1 \leq x \leq 3 \\ \frac{1}{12} & ; \quad 3 \leq x \leq 8 \\ \frac{1}{12}x - \frac{7}{12} & ; \quad 8 \leq x \leq 10 \\ 0 & ; \quad \text{en otro caso} \end{cases}$$

La función de distribución que muestra el comportamiento acumulado, está definida por:

$$F_X(x) = P(X \leq x) = P(X < x) = \int_{-\infty}^x f_X(t) dt$$

sustituyendo por intervalos:

en $1 \leq x \leq 3$

$$F_X(x) = \int_1^x \left(-\frac{1}{24}t + \frac{5}{24} \right) dt = -\frac{1}{48}x^2 + \frac{5}{24}x - \frac{3}{16}$$

en $3 \leq x \leq 8$

$$F_X(x) = \frac{1}{4} + \int_3^x \frac{1}{12} dt = \frac{1}{12}x$$

en $8 \leq x \leq 10$

$$F_X(x) = \frac{2}{3} + \int_8^x \left(\frac{1}{12}t - \frac{7}{12} \right) dt = \frac{1}{24}x^2 - \frac{7}{12}x + \frac{8}{3}$$

por lo tanto, la función acumulativa está dada por:

$$F_X(x) = \begin{cases} 0 & ; \quad x \leq 1 \\ -\frac{1}{48}x^2 + \frac{5}{24}x - \frac{3}{16} & ; \quad 1 \leq x \leq 3 \\ \frac{1}{12}x & ; \quad 3 \leq x \leq 8 \\ \frac{1}{24}x^2 - \frac{7}{12}x + \frac{8}{3} & ; \quad 8 \leq x \leq 10 \\ 1 & ; \quad x \geq 10 \end{cases}$$

Para determinar la mediana, se sabe que $P(X \leq \tilde{x}) = P(X < \tilde{x}) = \frac{1}{2}$, sustituyendo e igualando en la función de distribución, en el tercer intervalo:

$$\frac{1}{2} = \frac{1}{12}\tilde{x}$$

despejando:

$$\tilde{x} = \frac{12}{2} = 6$$

4. A un Centro de llamadas telefónicas de una empresa eléctrica, llegan en promedio tres llamadas por minuto. Supóngase que dichas llamadas siguen un proceso de Poisson.
- Calcular la probabilidad de recibir menos de tres llamadas en dos minutos.
 - Obtener la probabilidad de que el tiempo entre dos llamadas consecutivas sea mayor a dos minutos.
 - Determinar el tiempo promedio entre llamadas y su desviación estándar.
 - Calcular la probabilidad de recibir al menos dos llamadas en un minuto.

15 Puntos

Resolución

Sea X la variable aleatoria que representa el número de llamadas por minuto que llegan al Centro de llamadas.

$$X \sim \text{Poisson} \left(\lambda = 3 \left[\frac{\text{llamadas}}{\text{minuto}} \right] \right)$$

- a) En dos minutos se tiene:

$$X \sim \text{Poisson} (\lambda = 6)$$

Se pide calcular la probabilidad de recibir menos de tres llamadas en dos minutos, $P(X < 3)$, esto es:

$$P(X < 3) = P(X = 0) + P(X = 1) + P(X = 2)$$

sustituyendo:

$$P(X < 3) = \frac{(6)^0 e^{-6}}{0!} + \frac{(6)^1 e^{-6}}{1!} + \frac{(6)^2 e^{-6}}{2!}$$

$$P(X < 3) = e^{-6} + 6e^{-6} + 18e^{-6} = e^{-6} (1 + 6 + 18) = 25e^{-6} \approx 0.0619$$

- b) Sea Y la variable aleatoria que representa el tiempo entre dos llamadas consecutivas que llegan un Centro de llamadas.

$$Y \sim \text{Exponencial} \left(\lambda = 3 \left[\frac{\text{llamadas}}{\text{minuto}} \right] \right)$$

Se pide calcular la probabilidad de que el tiempo entre dos llamada consecutivas sea mayor de dos minutos, $P(Y > 2)$, esto es:

$$P(Y > 2) = 1 - F_Y(Y = 2) = e^{-3(2)} = e^{-6} \approx 0.0025$$

- c) El tiempo promedio entre dos llamadas es la media (valor esperado) de la variable aleatoria con distribución exponencial, esto es:

$$E(Y) = \frac{1}{\lambda} = \frac{1}{3}$$

La variancia de la variable aleatoria con distribución exponencial, es:

$$\text{Var}(Y) = \frac{1}{\lambda^2} = \frac{1}{3^2} = \frac{1}{9}$$

Entonces la desviación estándar, es:

$$\sigma_x = \sqrt{\text{Var}(Y)} = \sqrt{\frac{1}{\lambda^2}} = \sqrt{\frac{1}{3^2}} = \frac{1}{3}$$

- d) En un minuto se tiene:

$$X \sim \text{Poisson} (\lambda = 3)$$

Se pide calcular la probabilidad de recibir al menos dos llamadas en un minutos, $P(X \geq 2)$, esto es:

$$P(X \geq 2) = P(X = 2) + P(X = 3) + P(X = 4) + \dots +$$

sustituyendo:

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - \left[\frac{(3)^0 e^{-3}}{0!} + \frac{(3)^1 e^{-3}}{1!} \right]$$

$$P(X \geq 2) = 1 - [e^{-3} + 3e^{-3}] = 1 - e^{-3}[1 + 3] = 1 - 4e^{-3} \approx 0.8009$$

5. Sean las variables aleatorias conjuntas X y Y con función de probabilidad:

$f_{XY}(x, y)$		y			
		-3	-1	1	3
x	1	0	$\frac{1}{4}$	$\frac{1}{4}$	0
	9	$\frac{1}{4}$	0	0	$\frac{1}{4}$

- Obtener $E(XY)$, $E(X)$ y $E(Y)$
- Calcular $Cov(X, Y)$
- ¿Son X y Y variables aleatorias independientes?

15 Puntos

Resolución

$$a) E(XY) = \sum_{\forall y} \sum_{\forall x} xy f_{XY}(x, y)$$

sustituyendo:

$$E(XY) = (1)(-1)\left(\frac{1}{4}\right) + (1)(1)\left(\frac{1}{4}\right) + (9)(-3)\left(\frac{1}{4}\right) + (9)(3)\left(\frac{1}{4}\right) = 0$$

El valor esperado de X , se define por:

$$E(X) = \sum_{\forall x} x f_X(x)$$

sustituyendo:

$$E(X) = (1)(0.5) + (9)(0.5) = 5$$

El valor esperado de Y , se define de forma análoga:

$$E(Y) = \sum_{\forall y} y f_Y(y)$$

sustituyendo:

$$E(Y) = (-3)(0.25) + (-1)(0.25) + (1)(0.25) + (3)(0.25) = 0$$

- La covarianza se define como:
 $Cov(X, Y) = E(XY) - E(X)E(Y)$
 Sustituyendo los resultados del inciso a):
 $Cov(X, Y) = 0 - (5)(0) = 0$
- Las variables aleatorias conjuntas son independientes si:
 $f_{XY}(x, y) = f_X(x) f_Y(y)$

Para ello, se calculan las funciones marginales, entonces:

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

x	1	9
$f_X(x)$	0.5	0.5

y	-3	-1	1	3
$f_Y(y)$	0.25	0.25	0.25	0.25

sustituyendo:

$$f_{XY}(X=1, Y=-3) = 0 \neq \left(f_X(X=1) = \frac{1}{2} \right) \left(f_Y(Y=-3) = \frac{1}{4} \right) = \frac{1}{8}$$

y observa que $f_{XY}(x, y) \neq f_X(x) f_Y(y)$

Por lo que se concluye que las variables aleatorias son dependientes.

6. Dados los siguientes números aleatorios con distribución uniforme entre 0 y 1, generar un número aleatorio con distribución normal con parámetros $\mu_X = 2$ y $\sigma_X = 4$

0.89	0.01	0.95	0.45	0.59	0.87
0.19	0.27	0.64	0.75	0.25	0.63

15 Puntos

Resolución

Utilizando el TLC se tiene que:

$$X = \mu_X + \sigma_X \left(\sum_{i=1}^{12} R_i - 6 \right)$$

sustituyendo:

$$\sum_{i=1}^{12} r_i = 6.49$$

$$X = 2 + 4(6.49 - 6) = 3.96$$

7. La siguiente tabla muestra datos sobre el desgaste de acero dulce y y la viscosidad del aceite x .

y [10^{-4} mm^3]	240	181	193	155	172	110	113	75	94
x	1.6	9.4	15.5	20	22	35.5	43	40.5	33

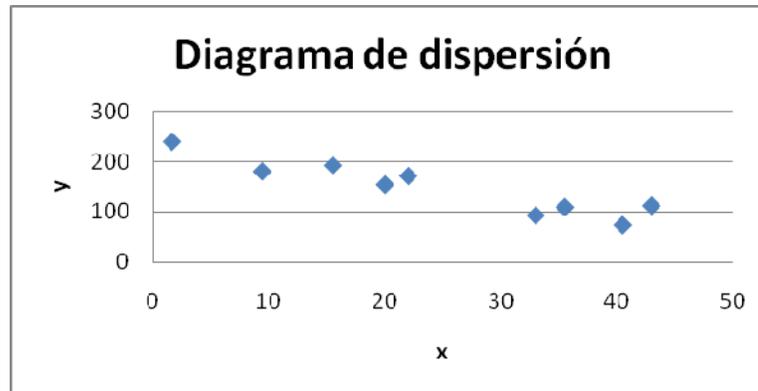
Las sumas relevantes, se muestran a continuación:

Sumas	
y	1333
x	220.5
xy	26864.4
x^2	7053.67
y^2	220549

- a) Trazar la gráfica de la dispersión de los datos. ¿Parece conveniente el uso de un modelo de regresión lineal?
- b) Ajustar un modelo de regresión lineal simple.
- c) Determinar el valor que se espera del desgaste cuando la viscosidad es 30
- d) ¿Puede considerarse válido el modelo? Justificar su respuesta.

15 Puntos
Resolución

a)



b) El modelo es:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^9 x_i y_i - \frac{\sum_{i=1}^9 x_i \sum_{i=1}^9 y_i}{9}}{\sum_{i=1}^9 x_i^2 - \frac{(\sum_{i=1}^9 x_i)^2}{9}} = \frac{26864.4 - \frac{(220.5)(1333)}{9}}{7053.67 - \frac{(220.5)^2}{9}} \approx -3.5086$$

y

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\hat{\beta}_0 = \frac{1333}{9} - (-3.5086) \left(\frac{220.5}{9} \right) \approx 234.0718$$

El modelo queda:

$$\hat{y} = 234.0718 - 3.5086x$$

- c) Para el valor que se espera del desgaste cuando la viscosidad es 30, sustituyendo en el modelo:
 $\hat{y}(30) = 234.0718 - 3.5086(30) \approx 128.8138$
- d) Para determinar si el modelo es válido debe obtenerse el coeficiente de determinación.
 El coeficiente de correlación, está definido por:

$$r = \frac{SS_{xy}}{\sqrt{SS_{xx} SS_{yy}}}$$

$$SS_{xx} = \sum_{i=1}^9 x_i^2 - \frac{\left(\sum_{i=1}^9 x_i\right)^2}{9} = 7053.67 - \frac{(220.5)^2}{9} = 1651.42$$

$$SS_{yy} = \sum_{i=1}^9 y_i^2 - \frac{\left(\sum_{i=1}^9 y_i\right)^2}{9} = 220549 - \frac{(1333)^2}{9} \approx 23116.8889$$

$$SS_{xy} = \sum_{i=1}^9 x_i y_i - \frac{\sum_{i=1}^9 x_i \sum_{i=1}^9 y_i}{9} = 26864.4 - \frac{(220.5)(1333)}{9} = -5794.1$$

sustituyendo:

$$r = \frac{SS_{xy}}{\sqrt{SS_{xx} SS_{yy}}} = \frac{-5794.1}{\sqrt{(1651.42)(23116.8889)}} \approx -0.9378$$

Entonces el coeficiente de determinación será:

$$r^2 = R^2 = \frac{SS_{xy}^2}{SS_{xx} SS_{yy}} = \frac{(-5794.1)^2}{(1651.42)(23116.8889)} \approx 0.8794$$

El ajuste es regular y puede considerarse válido el modelo dependiendo del error que se esté dispuesto a cometer.