

Probabilidad y Estadística
Primer Examen Final
Resolución
Tipo A

Semestre: 2008-1
Duración máxima: 2.5 h

1. Considerar la siguiente muestra (la resistencia de 50 lotes de algodón, libras necesarias para romper una madeja).

76	100	90	99	97	89	108	94	87	79
101	90	105	83	91	96	81	98	81	98
105	110	91	99	101	94	106	98	93	82
90	86	96	88	97	103	85	106	92	115
97	101	102	96	100	76	96	81	101	93

- a) Completar la tabla de distribución de frecuencias siguiente:

Fronteras de clase		Marca de clase	Frecuencia	Frecuencia acumulada	Frecuencia relativa	Frecuencia relativa acumulada
inf	sup	x_i	f_i	F_i	f_i^*	F_i^*
75.5	81.5					
81.5	87.5					
87.5	93.5					
93.5	99.5					
99.5	105.5					
105.5	111.5					
111.5	117.5					

- b) Dibujar el histograma.

- c) Calcular la media, la moda, la mediana, la desviación estándar y el sesgo.

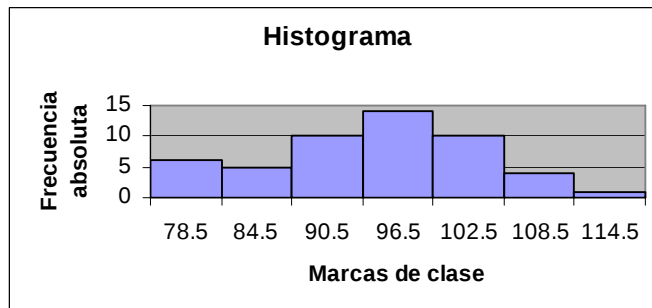
20 puntos

Resolución

- a) La tabla de distribución de frecuencias está dada por:

Fronteras de clase		Marca de clase	Frecuencia	Frecuencia acumulada	Frecuencia relativa	Frecuencia relativa acumulada
inf	sup	x_i	f_i	F_i	f_i^*	F_i^*
75.5	81.5	78.5	6	6	0.12	0.12
81.5	87.5	84.5	5	11	0.10	0.22
87.5	93.5	90.5	10	21	0.20	0.42
93.5	99.5	96.5	14	35	0.28	0.70
99.5	105.5	102.5	10	45	0.20	0.90
105.5	111.5	108.5	4	49	0.08	0.98
111.5	117.5	114.5	1	50	0.02	1
			50			

- b) El histograma es:



c) La media de los datos agrupados está dada por:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^m x_i f_i = \sum_{i=1}^m x_i f_i^* = \left[78.5 \left(\frac{6}{50} \right) + 84.5 \left(\frac{5}{50} \right) + 90.5 \left(\frac{10}{50} \right) + 96.5 \left(\frac{14}{50} \right) + 102.5 \left(\frac{10}{50} \right) + 108.5 \left(\frac{4}{50} \right) + 114.5 \left(\frac{1}{50} \right) \right]$$

$$\bar{x} = 94.46$$

La mediana de los datos agrupados está en:

Fronteras de clase	Frecuencia relativa acumulada
93.5	0.42
$\tilde{x}$	0.50
99.5	0.70

Realizando una interpolación, la ecuación de la recta dados dos puntos está definida por:

$$y - y_0 = \frac{y_1 - y_0}{x_1 - x_0} (x - x_0)$$

la pendiente se obtiene de sustituir:

$$m = \frac{y_1 - y_0}{x_1 - x_0} = \frac{0.70 - 0.42}{99.5 - 93.5} = \frac{0.28}{6} = \frac{7}{150} = 0.047$$

sustituyendo los valores se tiene: $y - 0.7 = 0.047(x - 99.5)$

la recta en forma ordenada al origen es:

$$y = 0.047x - 4.676 + 0.7 = 0.047x - 3.976$$

sustituyendo el punto $P(\tilde{x}, 0.5)$: $0.5 = 0.047\tilde{x} - 3.976$

despejando de la expresión anterior: $\tilde{x} \approx 95.234$

La moda es la marca de clase con mayor frecuencia, entonces: $x_{mo} = 96.5$

Para la variancia de los datos de la muestra, se sabe que:

$$s_{n-1}^2 = \frac{1}{n-1} \sum_{i=1}^m (x_i - \bar{x})^2 f_i$$

sustituyendo:

$$s_{n-1}^2 = \frac{1}{49} \left[(78.5 - 94.46)^2 6 + (84.5 - 94.46)^2 5 + (90.5 - 94.46)^2 10 + (96.5 - 94.46)^2 14 + (102.5 - 94.46)^2 10 + (108.5 - 94.46)^2 4 + (114.5 - 94.46)^2 1 \right]$$

$$s_{n-1}^2 = \frac{1}{49} [4075.92] \approx 83.18$$

La desviación estándar es la raíz cuadrada de la variancia, por lo que:

$$s_{n-1} = \sqrt{s_{n-1}^2} = \sqrt{83.18} \approx 9.12$$

Por la posición de las medidas de tendencia central se puede determinar de manera empírica el sesgo de la muestra, esto es: $\bar{x} < \tilde{x} < x_{mo}$

por lo que: $94.46 < 95.23 < 96.5$

por lo tanto, se puede concluir que la muestra tiene sesgo negativo.

De otra forma el coeficiente de sesgo se puede obtener como el tercer momento con respecto de la media entre la desviación estándar al cubo, esto es:

$$a_3 = \frac{m_3}{(s_{n-1})^3} = \frac{\frac{1}{n} \sum_{i=1}^m (x_i - \bar{x})^3 f_i}{(s_{n-1})^3} = \frac{\sum_{i=1}^m (x_i - \bar{x})^3 f_i^*}{(s_{n-1})^3}$$

sustituyendo valores se tiene:

$$a_3 = \frac{\left[(78.5 - 94.46)^3 \frac{6}{50} + (84.5 - 94.46)^3 \frac{5}{50} + (90.5 - 94.46)^3 \frac{10}{50} + (96.5 - 94.46)^3 \frac{14}{50} + (102.5 - 94.46)^3 \frac{10}{50} + (108.5 - 94.46)^3 \frac{4}{50} + (114.5 - 94.46)^3 \frac{1}{50} \right]}{(\sqrt{83.18})^3}$$

$$a_3 = \frac{[-110.38]}{(\sqrt{83.18})^3} \approx -0.145 < 0$$

por lo tanto se verifica que la muestra tiene sesgo negativo.

2. Un inversionista está pensando en comprar un número muy grande de acciones de una compañía. La cotización de las acciones en la bolsa, durante los seis meses anteriores, es de gran interés para el inversionista. Con base en la información de la bolsa de valores, se observa que la cotización se relaciona con el valor del oro en el mercado internacional. Si el valor del oro aumenta en el mundo, la probabilidad de que aumenten las acciones es de 0.8; si el valor del oro es el mismo, la probabilidad de que las acciones aumenten su valor es de 0.17; si el valor del oro disminuye, la probabilidad es de sólo 0.13. Si para los siguientes meses los analistas asignan las probabilidades de 0.5, 0.3 y 0.2 a los eventos: el oro sube, es el mismo o disminuye en su valor, respectivamente:
- Determinar la probabilidad de que las acciones aumenten su valor en los próximos seis meses.
 - ¿Cuál es la probabilidad de que el valor del oro aumente dentro de seis meses, no obstante que el valor de las acciones aumente?

15 puntos

Resolución

Sean los eventos:

A: El valor del oro aumenta en el mercado internacional.

B: El valor del oro permanece constante en el mercado internacional.

C: El valor del oro disminuye en el mercado internacional.

D: Aumentan las acciones.

Del enunciado se tienen los siguientes datos:

$$P(A) = 0.5 \qquad P(B) = 0.3 \qquad P(C) = 0.2$$

$$P(D|A) = 0.8 \qquad P(D|B) = 0.17 \qquad P(D|C) = 0.13$$

- a) Se pide calcular la probabilidad total para:

$$P(D) = P(A \cap D) + P(B \cap D) + P(C \cap D)$$

que equivale a:

$$P(D) = P(A)P(D|A) + P(B)P(D|B) + P(C)P(D|C)$$

sustituyendo los valores se tiene:

$$P(D) = (0.5)(0.8) + (0.3)(0.17) + (0.2)(0.13)$$

$$P(D) = 0.477$$

- b) Para el cálculo de la probabilidad se utiliza el Teorema de Bayes:

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

sustituyendo los valores se tiene:

$$P(A|D) = \frac{(0.5)(0.8)}{0.477} \approx 0.8386$$

3. El PH con que se mide la acidez del agua, es importante en los estudios de lluvia ácida. Para determinado lago en Veracruz, se llevan a cabo mediciones testigo de acidez para que se pueda notar cualquier cambio originado por la lluvia ácida. El PH de las muestras del agua es una variable aleatoria X , cuya función de densidad de probabilidad es:

$$f_x(x) = \begin{cases} \frac{3}{8}(7-x)^2 & ; \quad 5 < x < 7 \\ 0 & ; \quad \text{en otro caso} \end{cases}$$

- a) Calcular la función de distribución acumulativa $F_x(x)$.

- b) Calcular la probabilidad de que el PH sea menor que seis en una muestra de agua de este lago.

15 puntos

Resolución

Sea X la variable aleatoria que representa el PH en la muestra de agua de dicho lago.

- a) La función de distribución que muestra el comportamiento acumulado está definida por:

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

sustituyendo para la función de densidad:

$$F_X(x) = \int_{-\infty}^x f_X(t) dt = \int_5^x \frac{3}{8}(7-t)^2 dt$$

por cambio de variable:

$$F_X(x) = \frac{3}{8} \int_5^x (7-t)^2 dt$$

$$u = 7 - t$$

$$du = -dt$$

$$dt = -du$$

$$F_X(x) = -\frac{3}{8} \int_5^x u^2 du = -\frac{3}{8} \left(\frac{1}{3} u^3 \right) \Big|_5^x = -\frac{1}{8} (u^3) \Big|_5^x$$

$$F_X(x) = -\frac{1}{8} (7-t)^3 \Big|_5^x = -\frac{1}{8} [(7-x)^3 - (7-5)^3]$$

$$F_X(x) = -\frac{1}{8} [(7-x)^3 - 8] = 1 - \frac{1}{8} (7-x)^3, \quad 5 < x < 7$$

Por lo tanto la función de distribución acumulada es:

$$F_X(x) = \begin{cases} 0 & ; \quad x \leq 5 \\ 1 - \frac{1}{8} (7-x)^3 & ; \quad 5 < x < 7 \\ 1 & ; \quad x \geq 7 \end{cases}$$

- b) Se pide calcular $P(X \leq 6)$, se sabe que:

$$P(X \leq 6) = F_X(X = 6) = 1 - \frac{1}{8} (7-6)^3 = 1 - \frac{1}{8} = \frac{7}{8}$$

4. Se ha observado que el tiempo necesario para que una persona se traslade de una ciudad a otra, tiene una distribución normal con media de 15 horas y una variancia de dos horas. Calcular la probabilidad de que el recorrido tarde:
- Cuando mucho 16 horas.
 - Menos de 13.5 horas o más de 17 horas.

10 puntos

Resolución

Sea T la variable aleatoria que representa el tiempo de traslado de una ciudad a otra.

$$T \sim \text{Normal}(\mu = 15, \sigma^2 = 2)$$

- a) Se pide la probabilidad de que cuando mucho el traslado tarde 16 horas, entonces:

$$P(T \leq 16) = P(T < 16) \approx P\left(Z \leq \frac{T - \mu}{\sigma}\right) = P\left(Z \leq \frac{16 - 15}{\sqrt{2}}\right)$$

de tablas de la función de distribución acumulativa normal estándar:

$$P(T \leq 16) = P\left(Z \leq \frac{1}{\sqrt{2}}\right) = P(Z \leq 0.71) = P(Z < 0.71) = F_Z(0.71) = 0.7611$$

- b) Se quiere calcular la probabilidad de que un viaje tarde cuando mucho 13.5 horas o al menos 17 horas, entonces:

$$P(T \leq 13.5 \cup T \geq 17) = P(T < 13.5 \cup T > 17) = P(T \leq 13.5) + P(T \geq 17) \approx$$

$$\approx P\left(Z \leq \frac{13.5-15}{\sqrt{2}}\right) + P\left(Z \geq \frac{17-15}{\sqrt{2}}\right) = P\left(Z \leq \frac{-1.5}{\sqrt{2}}\right) + P\left(Z \geq \frac{2}{\sqrt{2}}\right) = P(Z \leq -1.06) + P(Z \geq 1.41) =$$

$$= F_Z(-1.06) + (1 - F_Z(1.41))$$

usando la tabla de la función de distribución acumulativa normal estándar se tiene:

$$P(T \leq 13.5) + P(T \geq 17) = F_Z(-1.06) + (1 - F_Z(1.41)) = 0.1446 + 0.0793 = 0.2239$$

5. Un ingeniero para su empresa de fabricación de computadoras compra, a un proveedor, grandes cantidades de un cierto componente electrónico y ha adoptado un plan para aceptar cada uno de los envíos de componentes, el cual consiste en inspeccionar una muestra aleatoria de 10 componentes. Si el comprador encuentra a lo más dos componentes defectuosos en la muestra, acepta el lote enviado por el proveedor. Se sabe por registros de la empresa que los envíos de este proveedor traen el 16% de componentes defectuosos.

- ¿Cuál es la probabilidad de que el lote sea aceptado?
- ¿Cuál es el promedio de los componentes defectuosos que deberá esperar el ingeniero siempre que revise una muestra de 10 componentes?

10 puntos

Resolución

Sea X la variable aleatoria que representa el número de componentes defectuosos en la muestra.

$$X \sim \text{Binomial}(n=10, p=0.16)$$

- La probabilidad de que el lote sea aceptado, es la probabilidad de que X sea menor o igual a dos.

$$P(X \leq 2) = P(X=0) + P(X=1) + P(X=2)$$

$$P(X \leq 2) = \binom{10}{0}(0.16)^0(0.84)^{10} + \binom{10}{1}(0.16)^1(0.84)^9 + \binom{10}{2}(0.16)^2(0.84)^8$$

$$P(X \leq 2) \approx 0.7936$$

- El promedio de componentes defectuosos es la media, esto es:

$$\mu = np = 10(0.16) = 1.6$$

El ingeniero espera obtener dos componentes defectuosos al revisar una muestra de diez.

6. Sean X el precio de un producto en dólares y Y las ventas totales, la función de densidad de probabilidad está dada por:

$$f_{XY}(x,y) = \begin{cases} 5xe^{-xy} & ; 0.20 < x < 0.40, \quad y > 0 \\ 0 & ; \text{en otro caso} \end{cases}$$

- Determinar la probabilidad de que el precio sea menor de 30 centavos y las ventas sobrepasen las 20000 unidades.

- Dadas las funciones marginales

$$f_X(x) = \begin{cases} 5 & ; 0.20 < x < 0.40 \\ 0 & ; \text{en otro caso} \end{cases}$$

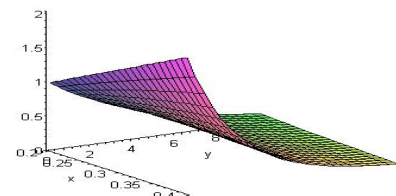
$$f_Y(y) = \begin{cases} \frac{5}{y^2}e^{-0.2y}(1+0.2y) - \frac{5}{y^2}e^{-0.4y}(1+0.4y) & ; y > 0 \\ 0 & ; \text{en otro caso} \end{cases}$$

¿son variables aleatorias conjuntas independientes?

15 puntos

Resolución

La gráfica de la función de densidad de probabilidad es:



- a) La probabilidad pedida es que el precio sea menor de 30 centavos y las ventas sobrepasen las 20000 unidades, entonces:

$$\begin{aligned}
 P(X < 0.3 \cap Y > 2) &= \int_{0.2}^{0.3} \int_2^{\infty} 5xe^{-xy} dy dx = \int_{0.2}^{0.3} \lim_{R \rightarrow \infty} \int_2^R 5xe^{-xy} dy dx = \\
 &= -5 \int_{0.2}^{0.3} \lim_{R \rightarrow \infty} e^{-xy} \Big|_2^R dx = -5 \int_{0.2}^{0.3} \lim_{R \rightarrow \infty} [e^{-xR} - e^{-2x}] dx = -5 \int_{0.2}^{0.3} -e^{-2x} dx = 5 \int_{0.2}^{0.3} e^{-2x} dx = \\
 &= -\frac{5}{2} e^{-2x} \Big|_{0.2}^{0.3} = -\frac{5}{2} [e^{-2(0.3)} - e^{-2(0.2)}] = \frac{5}{2} [e^{-2(0.2)} - e^{-2(0.3)}] \approx 0.3038
 \end{aligned}$$

- b) Para que las variables aleatorias conjuntas sean independientes, es suficiente que se cumpla:

$$f_{XY}(x, y) = f_X(x) f_Y(y)$$

sustituyendo las funciones marginales, se tiene:

$$5xe^{-xy} \neq 5 \left(\frac{5}{y^2} e^{-0.2y} (1 + 0.2y) - \frac{5}{y^2} e^{-0.4y} (1 + 0.4y) \right)$$

por lo tanto no son independientes; es decir, son variables aleatorias conjuntas dependientes.

7. Una compañía fabrica focos que tienen un periodo de vida distribuido aproximadamente normal, con media igual a 800 horas y desviación estándar de 40 horas. Calcular la probabilidad de que una muestra aleatoria de 16 focos tenga una vida promedio de menos de 775 horas.

15 puntos

Resolución

Sea X la variable aleatoria que representa el periodo de vida, en horas, de un foco.

$$X \sim Normal(\mu = 800, \sigma^2 = (40)^2)$$

$$\text{con } n = 16, \quad X_i \sim Normal(\mu = 800, \sigma^2 = (40)^2), \quad i = 1, 2, 3, \dots, 16$$

se quiere calcular la probabilidad de que la vida promedio de los 16 focos, sea cuando mucho de 775 horas, entonces:

$$P(\bar{X} \leq 775) = P(\bar{X} < 775)$$

como se conoce la variancia, se sabe que:

$$\bar{X} \sim Normal\left(\mu_{\bar{X}} = \mu_X, \sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{n}}\right)$$

sustituyendo:

$$\bar{X} \sim Normal\left(\mu_{\bar{X}} = 800, \sigma_{\bar{X}} = \frac{40}{\sqrt{16}}\right)$$

o bien:

$$\bar{X} \sim Normal(\mu_{\bar{X}} = 800, \sigma_{\bar{X}} = 10)$$

la probabilidad pedida es:

$$P(\bar{X} \leq 775) = P(\bar{X} < 775) \approx P\left(\frac{\bar{X} - \mu_{\bar{X}}}{\frac{\sigma_{\bar{X}}}{\sqrt{n}}} < \frac{775 - 800}{\frac{40}{\sqrt{16}}}\right) = P\left(Z < \frac{-25}{10}\right) =$$

De tablas de la función de distribución acumulativa normal estándar:

$$= P(Z < -2.50) = F_Z(-2.50) = 0.0062$$