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 2013 - 2 **TIPO 1**
DURACIÓN MÁXIMA 2.0 HORAS **31 DE MAYO DE 2013**

NOMBRE____
Apellido paterno **Apellido materno** **Nombre (s)** **Firma**

1. Los datos que se muestran a continuación, representan el costo de energía eléctrica en kW/h durante el mes de abril del 2013, para una muestra aleatoria de 50 departamentos con dos recámaras en el Distrito Federal.

Clase	L_iAp	L_sAp	Fr_i	Fr_s	x_i	Frec.Abs.	Frec.Rel.	Fr.Ac.Abs.	Fr.Ac.Rel.
1	82	102	81.5	102.5	92	4	0.08	4	0.08
2	103	123	102.5	123.5	113	7	0.14	11	0.22
3	124	144	123.5	144.5	134	11	0.22	22	0.44
4	145	165	144.5	165.5	155	11	0.22	33	0.66
5	166	186	165.5	186.5	176	10	0.20	43	0.86
6	187	207	186.5	207.5	197	6	0.12	49	0.98
7	208	228	207.5	228.5	218	1	0.02	50	1
						50			

a) Con los datos agrupados de la muestra, calcular: media, moda, mediana y desviación estándar.
b) Construir el polígono de frecuencias relativas, ubicar en la gráfica las medidas de tendencia central, ¿es posible determinar la simetría de manera empírica?

20 Puntos

Resolución

a) Para las medidas de tendencia central se sabe que,

La media se define por $\overline{X} = \frac{1}{n}\sum_{i=1}^{m} X_i f_i = \sum_{i=1}^{m} X_i f_i^*$

sustituyendo

$$\overline{x} = \frac{1}{50}\left[92(4) + \ldots + 218(1)\right] = \frac{7498}{50} = 149.96$$

La moda es la marca de clase con mayor frecuencia absoluta, esto es debido a que es bimodal

$x_{mo} = 134\ ,\ 155$ o bien $x_{mo} = \frac{134+155}{2} = \frac{289}{2} = 144.5$ o bien

$$x_{mo} = L_{x_{mo\_\inf}} + \left[\frac{a}{a+b}\right] c_{x_{mo}}$$

$$a = f_{x_{mo}} - f_{x_{mo-1}}$$

$b = f_{x_{mo}} - f_{x_{mo+1}}$

$f_{x_{mo}}$: *frecuencia absoluta de la clase que contiene a la moda.*

c_{mo} : *Longitud de la clase que contiene a la moda.*

$L_{Mo}\_\inf$: *Límite inferior de la clase que contiene a la moda.*

sustituyendo

$$x_{mo} = 123.5 + \left[\frac{11-7}{(11-7)+(11-11)}\right](20) = 143.5$$

La mediana es la medida de tendencia central que divide en dos partes iguales a la muestra, entonces

Fr_i - Fr_s	**Fr.Ac.Rel.**
144.5	0.44
$\tilde{x}$	0.5
165.5	0.66

al hacer la interpolación lineal

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{0.66 - 0.44}{165.5 - 144.5} \approx 0.0105$$

sustituyendo en la ecuación de una recta dados dos puntos, se tiene

$y_2 - y_1 = m(x_2 - x_1)$ con $y = 0.5$, $\tilde{x} = mediana$

$0.5 - 0.44 = 0.0105(\tilde{x} - 144.5)$

por lo que $\tilde{x} = 150.227$

La desviación estándar de la muestra, se define como la raíz de la variancia, entonces

$$S_{n-1}^2 = \frac{1}{n-1}\sum_{i=1}^{m}(Xi - \bar{X})^2 f_i$$

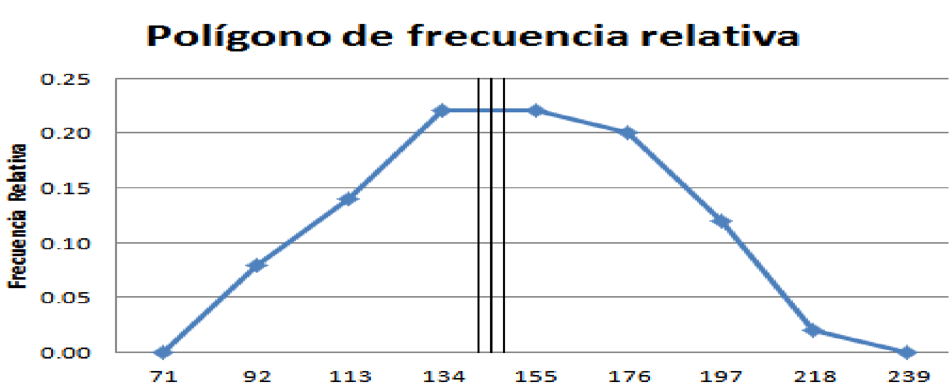
sustituyendo $s_{n-1}^2 = \frac{1}{49}\left[(92 - 149.96)^2(4) + \ldots + (218 - 149.96)^2(1)\right] = \frac{1}{49}(50767.92) \approx 1036.8$

La raíz cuadrada de la variancia muestral está dada por

$$S_{n-1} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{m}(Xi - \bar{X})^2 f_i} \qquad S_{n-1} = \sqrt{10.36.08} \approx 32.188$$

que es la desviación estándar muestral, hay mucha variabilidad de la muestra.

b)



$$x_{mo} \leq \bar{x} \leq \tilde{x}$$
$$143.50 < 149.96 < 150.23$$

por la posición de las medidas no es posible determinar el sesgo de la muestra de forma empírica, sin embargo, se puede observar de la gráfica que tiene ligero sesgo negativo.

2. Un alumno contesta una pregunta que ofrece cuatro posibles respuestas en un examen de opción múltiple, Supóngase que la probabilidad de que el alumno conozca la respuesta correcta es de 0.8 y la probabilidad de que tenga que contestar al azar es de 0.2. Si el alumno contesta correctamente la pregunta, cuál es la probabilidad de que conteste aleatoriamente.

15 Puntos

Resolución

Sean los eventos:

A que representa el alumno contesta correctamente.

B que representa el alumno contesta de forma aleatoria.

Del enunciado se pide determinar: $P(B|A)$

y se tiene que $P(\overline{B}) = 0.8$ $\quad P(B) = 0.2$ $\quad P(A|\overline{B}) = 1$ $\quad P(A|B) = 0.25$

por el Teorema de Bayes

$$P(B|A) = \frac{P(A \cap B)}{P(A \cap B) + P(A \cap \overline{B})} = \frac{P(B)P(A|B)}{P(B)P(A|B) + P(\overline{B})P(A|\overline{B})}$$

sustituyendo valores

$$P(B|A) = \frac{(0.2)(0.25)}{(0.2)(0.25) + (0.8)(1)} = \frac{0.05}{0.85} \approx 0.059$$

3. En un conmutador telefónico de una empresa comercial, el promedio de llamadas que se reciben en horas hábiles es seis por hora. Determinar la probabilidad de recibir en media hora, entre una y tres llamadas.

15 Puntos

Resolución

Sea X la variable aleatoria que representa el número de llamada que se reciben en una hora en el conmutador.

$$X \sim Poisson\left(\lambda = 6 \ \frac{llamadas}{hora}\right)$$

para media hora del proceso de Poisson

$$X \sim Poisson\left(\lambda = 3\frac{llamadas}{media\ hora}\right)$$

se pide calcular $P(1 \le X \le 3)$ para lo cual el modelo es

$$f_X(x) = \begin{cases} \dfrac{3^x e^{-3}}{x!} & ; \quad x = 0,1,2,3,\ldots \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

sustituyendo

$$P(1 \le X \le 3) = P(X = 1) + P(X = 2) + P(X = 3)$$

$$P(1 \le X \le 3) = \sum_{x=1}^{x=3} \frac{3^x e^{-3}}{x!} = e^{-3} \sum_{x=1}^{x=3} \frac{3^x}{x!} = 12e^{-3} \approx 0.597$$

4. Considere la función de densidad conjunta

$$f_{XY}(x,y)=\begin{cases}\dfrac{16y}{x^3} & ; \quad x>2 \;,\; 0<y<1\\ 0 & ; \quad en\ otro\ caso\end{cases}$$

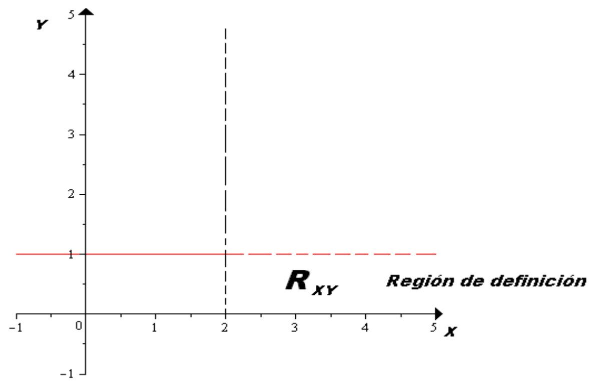
a) Calcular la covarianza. ¿Es determinante el resultado de la covarianza para saber la **independencia** de las variables?

b) Determinar si las variables son estadísticamente independientes.

20 Puntos

Resolución

La región de definición es



a) La covariancia se define por

$$Cov(X,Y)=E\left([X-\mu_X][Y-\mu_Y]\right)=E(XY)-E(X)E(Y)$$

que es una medida de asociación lineal. Se deben calcular los valores esperados que se requieren, entonces

$$E\left(g(X,Y)\right)=\iint_{R_{XY}} g(x,y)\,f_{XY}(x,y)\,dR$$

sustituyendo en cada caso, se tiene

$$E(XY)=\int_0^1\int_2^\infty xy\left(\frac{16y}{x^3}\right)dxdy=16\int_0^1\int_2^\infty\left(\frac{y^2}{x^2}\right)dxdy=16\int_0^1 y^2\left[-x^{-1}\right]_2^\infty dy=$$

$$E(XY)=-16\int_0^1 y^2\left[0-\frac{1}{2}\right]dy=8\int_0^1 y^2\,dy=\frac{8}{3}\left[y^3\right]_0^1=\frac{8}{3}$$

$$E(X)=\int_0^1\int_2^\infty x\left(\frac{16y}{x^3}\right)dxdy=16\int_0^1\int_2^\infty\left(\frac{y}{x^2}\right)dxdy=16\int_0^1 y\left[-x^{-1}\right]_2^\infty dy=$$

$$E(X)=-16\int_0^1 y\left[0-\frac{1}{2}\right]dy=8\int_0^1 y\,dy=4\left[y^2\right]_0^1=4$$

$$E(Y)=\int_0^1\int_2^\infty y\left(\frac{16y}{x^3}\right)dxdy=16\int_0^1\int_2^\infty\left(\frac{y^2}{x^3}\right)dxdy=16\int_0^1 y^2\left[\frac{x^{-2}}{-2}\right]_2^\infty dy=$$

$$E(Y)=-8\int_0^1 y^2\left[0-\frac{1}{4}\right]dy=2\int_0^1 y^2\,dy=2\left[\frac{y^3}{3}\right]_0^1=\frac{2}{3}$$

sustituyendo en la covariancia se tiene

$$Cov(X,Y)=\frac{8}{3}-(4)\left(\frac{2}{3}\right)=0$$

La covariancia vale cero, pero el resultado No es determinante para concluir si son variables aleatorias conjuntas independientes.

b) Para que las variables aleatorias sean conjuntas independiente debe cumplir la condición

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

de la condición las funciones marginales, se tiene

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\; dy \qquad \text{y} \qquad f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\; dx$$

al sustituir

$$f_X(x) = \int_0^1 \frac{16y}{x^3} dy = \frac{16}{x^3}\left(\frac{y^2}{2}\right)\Bigg|_0^1 = \frac{8}{x^3} \quad ; \quad x > 2$$

por lo que

$$f_X(x) = \begin{cases} \dfrac{8}{x^3} & ; \quad x > 2 \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

$$f_Y(y) = \int_2^{\infty} \frac{16y}{x^3} dx = 16y\left(\frac{x^{-2}}{-2}\right)\Bigg|_2^{\infty} = -8y\left(\frac{1}{x^2}\right)\Bigg|_2^{\infty} = -8y\left(0 - \frac{1}{4}\right) = 2y \quad ; \quad 0 < y < 1$$

$$f_Y(y) = \begin{cases} 2y & ; \quad 0 < y < 1 \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

entonces

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

$$\frac{16y}{x^3} = \left(\frac{8}{x^3}\right)(2y) = \frac{16y}{x^3}$$

El producto de funciones marginales es igual a la función conjunta, por lo tanto se concluye que son variables aleatorias conjuntas independientes, la covariancia es igual a cero, no hay asociación lineal, además, el coeficiente de correlación también es igual a cero.

5. Si la vida media de operación de una pila de linterna es de 24 horas y está distribuida normalmente con una desviación estándar de tres horas. ¿Cuál es la probabilidad de que una muestra aleatoria de 100 pilas, difiera de la verdadera media más de 30 minutos?

15 Puntos

Sea X la variable aleatoria que representa la vida media de operación de una pila de linterna.

$$X_i \sim Normal\left(\mu_{x_i} = 24h\ ,\ \sigma^2_{X_i} = 9h^2\right)$$

por el teorema del límite central, ya que, es una muestra aleatoria grande 100 pilas, cada una idénticamente distribuida, se tiene

$$\bar{X}_i \sim Normal\left(\mu_{\bar{x}} = \mu_{x_i} = 24h\ ,\ \sigma^2_{\bar{X}} = \frac{\sigma^2_{X_i}}{n} = \frac{9h^2}{100}\right)$$

se pide calcular

$$P\left(\left|\bar{X} - \mu_{\bar{X}}\right| > \frac{1}{2}\right) = P\left(\bar{X} - \mu_{\bar{X}} < -\frac{1}{2}\right) + P\left(\bar{X} - \mu_{\bar{X}} > \frac{1}{2}\right) = 1 - P\left(-\frac{1}{2} < \bar{X} - \mu_{\bar{X}} < \frac{1}{2}\right)$$

normalizando y al utilizar tablas de la distribución acumulativa, se tiene

$$P\left(\left|\bar{X} - \mu_{\bar{X}}\right| > \frac{1}{2}\right) \approx 1 - P\left(\frac{-\frac{1}{2}}{\frac{\sigma_X}{\sqrt{n}}} < \frac{\bar{X} - \mu_{\bar{X}}}{\frac{\sigma_X}{\sqrt{n}}} < \frac{\frac{1}{2}}{\frac{\sigma_X}{\sqrt{n}}}\right) = 1 - P\left(\frac{-\frac{1}{2}}{\frac{3}{\sqrt{100}}} < Z < \frac{\frac{1}{2}}{\frac{3}{\sqrt{100}}}\right) = 1 - P\left(-\frac{10}{6} < Z < \frac{10}{6}\right)$$

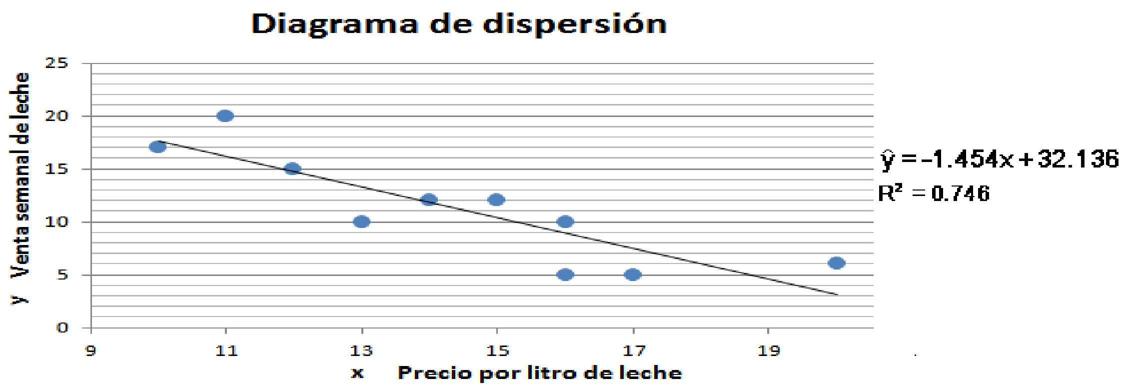
$$P\left(\left|\bar{X} - \mu_{\bar{X}}\right| > \frac{1}{2}\right) = 1 - P(-1.67 < Z < 1.67) = 1 - \left[F_Z(1.67) - F_Z(-1.67)\right] = 1 - (0.9525 - 0.0475) = 0.095$$

6. Los siguientes datos (x, y) representan el precio por litro de leche y la venta semanal de leche en miles de litros, respectivamente, con una muestra de diez observaciones, se obtuvieron los siguientes datos:

(13, 10), (20, 6), (17, 5), (15, 12), (16, 10), (12, 15), (16, 5), (14, 12), (10, 17) y (11, 20)

a) Interpretar el coeficiente de determinación.

b) Estimar la venta semanal en miles de litros de leche para un precio de $18



15 Puntos

Resolución

a) El precio de venta por cada litro de leche, explica muy ligeramente las ventas semanales de leche y el modelo, ya que es el 74.56% que es bajo.

b) El precio es $18 por litro de leche, sí está en el intervalo por lo cual se puede usar la recta para estimar las ventas semanales de dicho producto, por lo que

$\hat{y} = -1.454x + 32.136$

sustituyendo

$\hat{y} = -1.454(18) + 32.136 = 5.964$