SOLUCIÓN PRIMER EXAMEN FINAL "PROBABILIDAD Y ESTADÍSTICA"
SEMESTRE 2017-1 **3 de diciembre de 2016**

1. Las lecturas de humedad (en g/m^3) tomadas durante 20 días, se agruparon en la siguiente distribución de frecuencias:

Lecturas de humedad	Frecuencia
10 - 19	3
20 - 29	8
30 - 39	5
40 - 49	3
50 - 59	1

a) Obtener la media, mediana y moda de los datos.
b) Indicar el tipo de sesgo que tienen los datos. **10 puntos**

Solución:

a) Media:

$$\bar{x} = \frac{1}{20}[(14.5)(3) + (24.5)(8) + (34.5)(5) + (44.5)(3) + (54.5)(1)]$$

$$= \frac{600}{20} = 30$$

Por lo tanto, el valor de la media es: 30 g/m^3

Moda:

$x_{mo} = 24.5$ g/m^3 (es el valor de la marca de clase que tiene mayor frecuencia)

Mediana:

$$\tilde{x} = \frac{10-3}{11-3}\,(29.5-19.5) + 19.5 = \frac{7}{8}(10) + 19.5 = 28.25$$

Por tanto, el valor de la mediana es: 28.25 g/m^3

b) Al comparar los valores de tendencia central obtenidos en el inciso anterior, se observa que:

$$24.5 < 28.25 < 30$$

$$x_{mo} < \tilde{x} < \bar{x}$$

Por tanto, los datos tienen sesgo positivo o hacia la derecha.

Si se calcula el coeficiente de sesgo (a_3), se tiene: $a_3 = \frac{m_3}{(s)^3}$,
m_3 es el 3er momento respecto a la media,
s es la desviación estándar

$s = \sqrt{s^2}$; $\quad s^2 = \frac{1}{n-1}\sum_{i=1}^{m}(x_i - \bar{x})^2 f_i$ $\quad$ m= *número de clases*
$x_i = marca\ de\ clase$

$$s^2 = \frac{1}{19}[(14.5-30)^2(3) + (24.5-30)^2(8) + (34.5-30)^2(5) + (44.5-30)^2(3) + (54.5-30)^2(1)] = \frac{1}{19}(2295) \approx 120.789$$

$$\therefore s = \sqrt{120.789} \approx 10.99$$

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

$$= \frac{1}{20}[(14.5-30)^3(3) + (24.5-30)^3(8) + (34.5-30)^3(5) + (44.5-30)^3(3) + (54.5-30)^3(1)] = \frac{11805}{20} = 590.25$$

$$\therefore\ a_3 = \frac{590.25}{(10.99)^3} \approx 0.4446$$

Como $a_3 > 0$, entonces el sesgo de los datos es positivo o hacia la derecha.

2. El 30% de los empleados de una empresa son ingenieros, otro 20% son economistas. El 80% de los ingenieros, ocupan un puesto directivo y el 30% de los economistas también, mientras que del resto de los empleados (otras profesiones), solamente el 10% ocupa un puesto directivo. Si del grupo de directivos, se selecciona al azar a un empleado; ¿cuál es la probabilidad de que sea ingeniero? **10 puntos**

Solución

Sean los eventos:
I: el empleado es ingeniero
E: el empleado es economista
O: el empleado tiene una profesión distinta a ingeniero o economista
D: el empleado ocupa un puesto directivo

Se sabe que:
P(I)= 0.3
P(E)= 0.2
P(O)= 0.5
P(D|I)=0.8
P(D|E)=0.3
P(D|O)= 0.1

La probabilidad de que un empleado que ocupa un puesto directivo sea ingeniero es: P(I |D)

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

P(D)= P(D|I)P(I)+P(D|E)P(E)+ P(D | O)P(O)
$P(D) = (0.8)(0.3) + (0.3)(0.2) + (0.1)(0.5) = 0.35$

Por lo tanto, la probabilidad solicitada es: $P(I|D) = \frac{(0.8)(0.3)}{0.35}$= **0.6857**

3. Sea una variable aleatoria X, que tiene como función de densidad:

$$f(x) = \begin{cases} \dfrac{x+6}{50} & ; \quad -6 \le x \le 4 \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

a) Calcular la función de distribución acumulada de X.
b) Encontrar la probabilidad de que variable aleatoria X se encuentre dentro del intervalo: [-2,0]

15 puntos

Solución:

a) La función de distribución acumulada F(x) se obtiene mediante la relación: F(x)=$\int_{-\infty}^{x} f(x)dx$

En el intervalo: x= (-∞, -6)
F(x)=$\int_{-\infty}^{x} 0\, dx$= 0

En el intervalo: x= [-6,4)
F(x)=$\int_{-\infty}^{-6} 0\, dx + \int_{-6}^{x} \left(\frac{x+6}{50}\right) dx = 0 + \frac{1}{50}\left(\frac{x^2}{2} + 6x\right) |_{-6}^{x}$

$= \frac{1}{50}\left[\frac{x^2}{2} + 6x - \left(\frac{(-6)^2}{2} + 6(-6)\right)\right]$

$= \frac{1}{50}\left[\frac{x^2}{2} + 6x - (18 - 36)\right] = \frac{1}{50}\left(\frac{x^2}{2} + 6x + 18\right) = \frac{x^2}{100} + \frac{6x}{50} + \frac{18}{50}$

En el intervalo [4,∞)
F(x)= $\int_{-\infty}^{-6} 0\, dx + \int_{-6}^{4} \left(\frac{x+6}{50}\right) dx + \int_{4}^{x} 0 dx = 1$

$\therefore$ *la función de distribución acumulada de X es*:

$$\boldsymbol{F(x)} = \begin{cases} \mathbf{0} \; ; \; \boldsymbol{x} < \mathbf{-6} \\ \dfrac{\boldsymbol{x^2}}{\mathbf{100}} + \dfrac{\mathbf{6}\boldsymbol{x}}{\mathbf{50}} + \dfrac{\mathbf{18}}{\mathbf{50}} \; ; \; \mathbf{-6} \le \boldsymbol{x} \le \mathbf{4} \\ \mathbf{1} \; ; \; \boldsymbol{x} > \mathbf{4} \end{cases}$$

b) Solicitan obtener : $P\ (-2\ \le X\ \le 0)$

P(-2 $\le X \le$ 0)=$\int_{-2}^{0} \left(\frac{x+6}{50}\right) dx = \frac{1}{50}\left(\frac{x^2}{2} + 6x\right)\Big|_{-2}^{0} = \mathbf{0.2}$

Por lo tanto, la probabilidad de que variable aleatoria X se encuentre dentro del intervalo [- 2, 0], es: **0.2**

4. Una caja contiene: 3 fusibles de la marca A, 2 de la marca B y 3 fusibles de la marca C. Se eligen al azar dos fusibles para un automóvil.
Si X es el número de fusibles marca A y; Y el número de fusibles marca B que se seleccionan, encontrar:
a) La función de distribución de probabilidad conjunta de X y Y.
b) Obtener la probabilidad de que en la selección realizada haya un fusible marca B y un fusible marca C.

20 puntos

Solución
a) Se tienen 3 fusibles marca A, 2 fusibles marca B y 3 fusibles marca C.
En total 8 fusibles.
Se eligen al azar 2 de los 8 fusibles

Los valores que puede tomar X son: 0,1,2
Los valores que puede tomar Y son: 0,1,2

Por tanto la función de probabilidad conjunta de X,Y, tendrá la siguiente forma:

$f_{XY}(x,y)$		x		
		0	1	2
y	0			
	1			
	2			

$$P(X=0,Y=0)=\frac{\binom{3}{0}\binom{2}{0}\binom{3}{2}}{\binom{8}{2}}=\frac{3}{28}$$

$$P(X=0,Y=1)=\frac{\binom{3}{0}\binom{2}{1}\binom{3}{1}}{\binom{8}{2}}=\frac{6}{28}$$

$$P(X=0,Y=2)=\frac{\binom{3}{0}\binom{2}{2}\binom{3}{0}}{\binom{8}{2}}=\frac{1}{28}$$

$$P(X=1,Y=0)=\frac{\binom{3}{1}\binom{2}{0}\binom{3}{1}}{\binom{8}{2}}=\frac{9}{28}$$

$$P(X=1,Y=1)=\frac{\binom{3}{1}\binom{2}{1}\binom{3}{0}}{\binom{8}{2}}=\frac{6}{28}$$

P(X=1, Y=2) = 0

$$P(X=2,Y=0)=\frac{\binom{3}{2}\binom{2}{0}\binom{3}{0}}{\binom{8}{2}}=\frac{3}{28}$$

P(X=2,Y=1)=0
P(X=2,Y=2)=0

∴ La función de probabilidad conjunta de X,Y es:

$f_{XY}(x,y)$		x		
		0	1	2
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$
	1	$\frac{6}{28}$	$\frac{6}{28}$	0
	2	$\frac{1}{28}$	0	0

b) La probabilidad de que en la selección realizada haya un fusible marca B y un fusible marca C es: P(X=0, Y=1) = $f_{XY}(0,1) = \frac{6}{28}$

5. La revista "Ingeniería y Tecnología", comienza una campaña telefónica de difusión con el propósito de aumentar el número de suscriptores. Con base en la experiencia, se sabe que una de cada 20 personas que recibe la llamada se suscribe a la revista. Si en un determinado día 25 personas reciben la llamada telefónica de la revista:

a) ¿Cuál es la probabilidad de que por lo menos dos de ellas se suscriban a la publicación?
b) ¿Cuál es la probabilidad de que la segunda persona que se suscriba sea la doceava que recibe la llamada? **20 puntos**

Solución :

a) Sea X el número de personas que se suscriben a la publicación al recibir la llamada.
p = 0.05, es la probabilidad de que la persona que recibe la llamada se suscriba a la publicación.
Se realizan 25 llamadas.
Por lo tanto: X~binomial (n=25, p=0.05)

La probabilidad de que por lo menos dos de las 25 personas que reciben la llamada se suscriban a la publicación es:

$$P(X\geq 2)= 1- P(X<2)= 1- P(X=0)+P(X=1)$$
$$= 1- \binom{25}{0}(0.05)^0(0.95)^{25} - \binom{25}{1}(0.05)^1(0.95)^{24}$$
$$=1- 0.2774-0.3650= \mathbf{0.3576}$$

b) Sea X el número de llamadas que deben realizarse para tener 2 que se suscriban a la revista.
X ~ pascal (r=2, p= 0.05), su función de probabilidad es:

$$f(x) = \binom{x-1}{r-1} p^r q^{x-r}$$

r = número de éxitos deseados.

Por lo tanto, la probabilidad de que la segunda persona que se suscriba a la revista sea la doceava que recibe la llamada es:

$$P(X=12) = \binom{11}{1}(0.05)^2(0.95)^{10} \approx \mathbf{0.01646}$$

6. Las edades de los habitantes de la Ciudad de México tienen una distribución normal con varianza $\sigma^2= 6$. Se toma una muestra aleatoria con 25 observaciones, encuentre la probabilidad de que la muestra tenga una varianza s^2 mayor que 9.1.

15 puntos

Solución:

Sea X la variable aleatoria que representa las edades de los habitantes de la Cd. de México. $X \sim N(\sigma^2= 6)$

Se toma una muestra aleatoria de tamaño n=25

Se solicita calcular para la muestra: $P(s^2>9.1)$

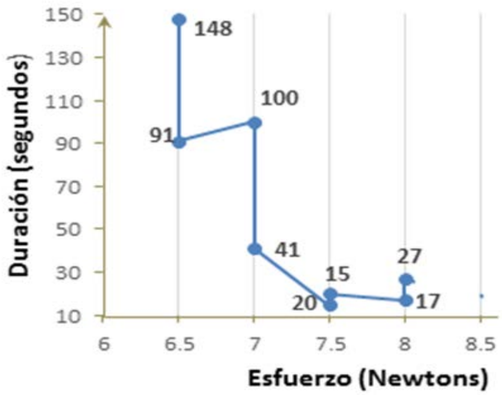
$$P(s^2>9.1)= P\left(\chi^2> \frac{(9.1)(24)}{6}\right) = P(\chi^2>36.4)$$

tomando en cuenta que $\chi 2$ *tiene 24 grados de libertad*

Se obtiene que $P(x^2>36.4) = \mathbf{0.05}$, dicho valor es la probabilidad de que la varianza de la muestra sea mayor que 9.1

7. Se estudia empíricamente la relación que existe entre el nivel de esfuerzo aplicado a una probeta de un material plástico y el tiempo que transcurre antes de su fractura, se hizo una prueba a 9 probetas y se recabaron los siguientes datos (gráfica), indique el valor del coeficiente de correlación de los datos recabados e interprete el valor obtenido.

10 puntos



Solución:

	valores				
	x	y	x^2	y^2	xy
	6.5	148	42.25	21904	962
	6.5	91	42.25	8281	591.5
	7	100	49	10000	700
	7	41	49	1681	287
	7.5	20	56.25	400	150
	7.5	15	56.25	225	112.5
	8	27	64	729	216
	8	17	64	289	136
Sumas:	**58**	**459**	**423**	**43509**	**3155**

El coeficiente de correlación r se obtiene mediante la expresión:

$$r=\frac{SSxy}{\sqrt{SSxxSSyy}}$$

$$SSxy= \sum_{i=1}^{n}(x_i)(y_i) - \frac{\sum_{i=1}^{n}(xi)\sum_{i=1}^{n}(yi)}{n} = (3155) - \frac{(58)(459)}{8} = -172.75$$

$$SSxx= \sum_{i=1}^{n}(x_i)^2 - \frac{\left(\sum_{i=1}^{n}x_i\right)^2}{n} = (423) - \frac{58^2}{8} = 2.5$$

$$SSyy= \sum_{i=1}^{n}(y_i)^2 - \frac{\left(\sum_{i=1}^{n}y_i\right)^2}{n} = (43509) - \frac{459^2}{8} = 17173.875$$

$$\therefore r=\frac{-172.75}{\sqrt{(2.5)(17173.875)}} = \mathbf{-0.8337}$$

El coeficiente de correlación proporciona el grado de asociación lineal que existe entre las variables. En este caso al ser negativo significa que las variables se relacionan de forma inversa (a medida que el valor de X aumenta el valor de Y disminuye), además como el valor de r es cercano a -1 , significa que existe una importante relación lineal entre las variables estudiadas. Es decir, el nivel de esfuerzo aplicado a la probeta y el tiempo que transcurre antes de su fractura se encuentran inversamente relacionados, a mayor esfuerzo aplicado, menor tiempo de duración de la probeta.