Estadística para Ciencia de los Datos
Tarea 1
Juan Pablo Fernandez Delgado

1) Una agencia de bienes raíces tiene una base de datos de propiedades vendidas en la cuál rastrearon la cantidad de días que duró en concretarse la venta, además de su precio de venta (ver tabla). Basado en los datos responda los siguientes apartados. Responda todos los puntos utilizando ecuaciones únicamente.

	Menos de 30 días	De 31 a 90	Más de 90	Total
Menos de $50,000	50	40	10	100
50,000 a 99,999	20	150	80	250
100,000 a 149,999	20	280	100	400
Más de 150,000	10	10	30	50
Total	100	480	220	800

Si A es el evento de vender una casa en más de 90 días, estime la probabilidad de A

a) $P(A: \text{más de } 90) \ \frac{\sum_{i=1}^{m} n_{ij}}{N}$ donde j: 3 (columna de mas 90)
m: 4 (cantidad filas)
N: 800 (total casas)

$P(A: \text{más de } 90) \ \frac{(10+80+100+30)}{800}$

$\frac{220}{800} = \frac{11}{40} = 0{,}275$

R/ La probabilidad de vender una casa en más de 90 días es de 0.275 ó 27%.

Si B es el evento de vender una casa en menos de 50,000, estime la probablidad de B (5 puntos)

P(B: menos de$50.000): $\dfrac{\sum_{j=1}^{3}}{N}$ donde i = 1 (filas menos $50.000
1:3 (cant columnas)
N: 800 (total casas)

P(B: menos de $50000):

$$\frac{(50+40+10)}{800}$$

$$\frac{100}{800} = \frac{1}{8} : 0.125$$

R/ La probabilidad de vender una casa de menos de $50000 es de 0,125 o 12,5 %

¿Cuál es la probabilidad de que A y B ocurran juntos?

P(A: más de 90, B: menos de 50.000) : $\dfrac{n_{ij}}{N}$

$\dfrac{10}{800} : \dfrac{1}{8} : 0.0125$ donde i: 1 (filas menos de 50.000]
j: 3 (columna más 90)
N: 800 (total casas)

R/ La probabilidad de que ocurran juntas es de 0.0125 ó 1.25%

Si una casa se define que tiene un precio de menos de $50,000, ¿cuál es la probabilidad que tarde 90 o menos días en venderse?

$$P(A:\text{ mas de } 90, B:\text{ menos de } 50.000) = \frac{n_{ij}}{\sum_{j=1} n_{ij}}$$

$$\frac{10}{(50+40+10)} = \frac{10}{100} = \frac{1}{10} = 0,1$$

La probabilidad es de 0.1 ó 10%.

¿Se puede considerar que los eventos A y B son independientes? (Demuestrelo matemáticamente)

Para considerar que son independientes se tiene las siguientes condiciones

1. $P(A) = P(A|B)$
2. $P(A,B) = P(A)\,P(B)$

$$P(A) = P(A|B) \qquad\qquad P(A,B) = P(A)P(B)$$

$$\frac{11}{40} \neq \frac{1}{10} \qquad\qquad \frac{1}{80} = \frac{11}{40} \times \frac{1}{8}$$

$$\frac{1}{80} \neq \frac{11}{320}$$

Se puede decir que las variables son dependientes

2. Hay tres aulas: en la N° 1 hay 8 mujeres y 4 hombres, en la N°2 hay 10 mujeres y 20 hombres, y en la N°3 hay 6 mujeres y 10 hombres. Se escoge un aula al azar y luego de ella se extrae al azar una persona. Si se sabe que la persona extraída es mujer, pero no se sabe de cuál aula proviene, ¿Cuál es la probabilidad de que ese persona provenga del aula N°3?

	H	M	Total
A_1	4	8	12
A_2	20	10	30
A_3	10	6	16
	34	24	58

Probabilida que sea mujer

$$P(A: mujer) \quad \frac{\sum_{i=1}^{n} n_{ij}}{N} = \frac{8+10+6}{58} = \frac{12}{29} = 0,40$$

La probabilidad que sea mujer es de 0,40 ó 40%.

Probabilidad Aula 3

$$P(A: Aula\ 3) \quad \frac{\sum_{i=1}^{n} n_{ij}}{N} = \frac{10+6+16}{58} =$$

$$\frac{32}{58} = \frac{16}{29} = 0,55$$

R/ La probabilidad de que sea Aula 3 es de 0,55 ó 55%.

3. Se lanzan dos dados. Sean: A= *suma de puntos impar*, B= *1 en el primer dado*, C= *suma de punto es igual a 7*. ¿Son A y B independientes?, ¿Son A y C independientes? , ¿Son B y C independientes?

4. La distribución de probabilidad de una variable aleatoria Y está dada por la siguiente tabla. Suponga que se realizó un experimento donde se obtuvieron 4 millones de valores observados de Y. Con base en estos valores y las distribución de probabilidad de Y encuentre su valor esperado.

El conjunto de y toma valores $\{0, 1, 2, 3\}$

Calculamos la esperanza

$$E(x) = \sum x e^{x} P(x)$$

$$= 0 \cdot \frac{1}{8} + 1 \cdot \frac{1}{4} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{4}$$

$$\frac{6}{8} + \frac{4}{4} =$$

$$\frac{6}{8} = 0,75$$

R/ El valor de la esperanza es de 0,75 ó 75%.