Nombre: Bárbara Rossana Pérez Silva
Grupo: 031 **2015576**

1 Tipos de datos y Medidas de tendencia central

En una empresa se han recolectado los siguientes datos de 10 empleados:

Nombre	**Edad (años)**	**Área de trabajo**
Ana	25	Ventas
Luis	30	Administración
Marta	40	Producción
Carlos	35	Ventas
Elena	28	Recursos Humanos
Juan	50	Producción
Sofía	45	Administración
Pedro	38	Ventas
Daniel	33	Producción
Laura	27	Recursos Humanos

1. Clasifique las variables en cualitativas o cuantitativas.
   Cualitativas: Nombres y Áreas de trabajo, al ser esto características no numéricamente posibles.
   Cuantitativas: Edades, ys que son medidas numéricas, estas se pueden calcular con operaciones matemáticas.

2. Determine la media, mediana y moda de la variable "Edad".
   La **media** de un conjunto de datos se calcula con la sumatoria de los elementos a promediar y dividiéndolos entre la cantidad de elementos que se están analizando.

$$\bar{x} = \frac{\sum x_i}{n} = \frac{25+30+40+35+28+50+45+38+33+27}{10} = 35.1$$

   La **mediana** se trata literalmente del elemento que se encuentra a la mitad de los datos que se encuentran ordenados de menor a mayor, o viceversa.

$$25, 27, 28, 30, \mathbf{33, 35,} 38, 40, 45, 50$$

   Como aquí tenemos dos elementos a la mitad de los datos, lo que hacemos es sacar el promedio de estos dos datos:

$$\frac{33+35}{2} = 34$$

La **moda** se refiere a el/los datos que se repiren más veces en el conjunto de datos. En este caso no hay moda unica, por lo que se podría considerar **amodal**.

3. Interprete los resultados obtenidos.
   El promedio de edad de los empleados de la empresa es de 35, esto obtenido con la media. El punto medio de las edades de los 10 empleados es de 34 años, esto significa que la primer mitad de las personas, tendrán una edad menor, o igual, a 34 años, y la segunda mitad será mayor, o igual a esa edad. Por último, analizando las edades podemos notar que, por lo menos en la muestra de 10 empleados, todos tienen diferentes edades.

2 Medidas de dispersión

Dado el siguiente conjunto de datos correspondiente a las calificaciones de 8 estudiantes en un examen:

$$X = 70, 85, 90, 95, 88, 92, 75, 80$$

1. Calcule la varianza y la desviación estándar de los datos.
   La fórmula de la varianza es la siguiente:

$$\sigma^2 = \frac{\sum(X_i - \bar{X})^2}{n}$$

   Para ello necesitamos calcular la media:

$$\bar{X} = \frac{70 + 85 + 90 + 95 + 88 + 92 + 75 + 80}{8} = 84.375$$

$$\sigma^2 = \frac{\sum(X_i - \bar{X})^2}{n}$$

$$\sigma^2 = \frac{(70 - 84.375)^2 + (85 - 84.375)^2 + (90 - 84.375)^2 + (95 - 84.375)^2 + ...}{8}$$

$$\frac{... + (88 - 84.375)^2 + (92 - 84.375)^2 + (75 - 84.375)^2 + (80 - 84.375)^2 +}{8} = 66.23$$

   La desviación estándar se define como la raíz cuadrada de la varianza:

$$\sigma^2 = \sqrt{66.23} = 8.13$$

2. Interprete la dispersión de los datos. Con los datos recabados sobre la información que nos proporcionaron de las calificaciones de estudiantesse puede saber que éstas se encuentran en un rango muy cercano al valor de la media, así como podemos observar al analizar los datos a simple vista.

3 Probabilidades y Teorema de Bayes

Una empresa de tecnología ha identificado que el 60% de sus empleados son programadores, y el 40% son diseñadores. Se sabe que el 70% de los programadores tienen conocimientos de inteligencia artificial (IA), mientras que solo el 30% de los diseñadores tienen estos conocimientos. Si se elige un empleado al azar y se sabe que tiene conocimientos de IA, ¿cuál es la probabilidad de que sea programador?
Aplicando el Teorema de Bayes para la solución de este problema tenemos:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

En donde:
P(A) = probabilidad de que el empleado sea programador = **0.60**
P(B) = probabilidad de que el empleado tenga conocimientos de IA
P(B|A) = Probabilidad de que el empleado sea programador con conocimientos de IA = **0.70**
P(D) = probabilidad de que el empleado sea diseñador = **0.40**
P(B|D) = probabilidad de que el empleado sea diseñador con conocimientos de IA = **0.30**

Calculamos la probabilidad de B:

$$P(B) = P(B|A) \cdot P(A) + P(B|D) \cdot P(D)$$

$$P(B) = (0.70 \cdot 0.60) + (0.30 \cdot 0.40) = 0.54$$

Y ahora, sustituyendo en la fórmula del Teorema de Bayes:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} = \frac{0.70 \cdot 0.60}{0.54} = 0.778$$

Por lo tanto, la probabilidad de que se elija a un empleado al azar y este sea programador con conocimientos de IA es del 77%

4 Distribuciones de probabilidad

Suponga que el número de defectos en un lote de producción sigue una distribución de Poisson con media $\lambda = 3$ defectos por lote.

1. Calcule la probabilidad de que un lote tenga exactamente 2 defectos. Tenemos la fundion de distribución de Poisson:

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, 2, \ldots$$

Sabemos que la media es $\lambda = 3$ para nuestro evento de k = 2 defectos, tenemos:

$$P(X=2) = \frac{3^2 e^{-3}}{2!} = \frac{9e^{-3}}{2} = 0.224$$

Por lo tanto, la probabilidad de que un lote tenga exactamente 2 defectos es del 22.4%

2. Calcule la probabilidad de que un lote tenga al menos 1 defecto.
   Que tenga al menos un defecto, significa que los defectos pueden ir desde 1 hasta el total del lote, por lo que tenemos:

$$P(X \geq 1) = 1 - P(X=0)$$

Calculamos P(X=0):

$$P(X=0) = \frac{3^0 e^{-3}}{0!} = e^{-3} = 0.0498$$

Por último, sustituimos este valor en la ecuación anterior:

$$P(X \geq 1) = 1 - 0.0498) = 0.9502$$

Por lo tanto, la probabilidad de que un lote tenga al menos 1 defecto es del 95.02

5 Funciones de densidad y distribución acumulativa

Sea X una variable aleatoria con distribución normal de media $\mu = 50$ y desviación estándar $\sigma = 10$.

1. Determine la probabilidad de que X tome un valor menor que 45.
   Para solucionar el problema de que X sea menor que 45 nos apoyamos en la fórmula de disctribución normal estándar N(0,1):

$$Z = \frac{X - \mu}{\sigma} = \frac{45 - 50}{10} = -0.5$$

Teniendo así que:

$$P(X < 45) = P(Z < -0.5) = 0.3085$$

Por lo tanto, la probabilidad de que X tome un valor menor que 45 es del 30.85%

2. Determine la probabilidad de que X esté entre 40 y 60.
   Primero volvemos a calcular el valor de Z con ambas cantidades:

$$Z_1 = \frac{40-50}{10} = -1$$

$$Z_2 = \frac{60-50}{10} = 1$$

Y sabemos que la probabilidad de X se expresa de la siguiente manera:

$$P(40 < X < 60) = P(Z_1 < Z < Z_2) = P(-1 < Z < 1)$$

Calculando con la tabla de distribución normal tenemos:

$$P(Z < 1) = 0.8413 \quad ; \quad P(Z < -1) = 0.1587$$

$$P(Z < 1) - P(Z < -1) = 0.6826$$

Por lo tanto, la probabilidad de que X esté entre 40 y 60 es del 68.26%

3. Use la función de distribución acumulativa para verificar sus respuestas. La función de distribución acumulativa está expresada como:

$$F(x) = P(X \leq x)$$

- Primer problema:

$$F(45) = P(X \leq 45) = P(Z \leq -0.5) = 0.3085$$

- Segundo problema:

$$F(60) - F(40) = P(40 \leq X \leq 60) = 0.6826$$

Y como podemos ver, al comparar ambos resultados con los ya obtenidos, estos son idénticos.

6 Probabilidad condicional

Un dado justo de seis caras se lanza dos veces.

1. ¿Cuál es la probabilidad de obtener un número par en el segundo lanzamiento, dado que en el primero salió un número impar?
   Teniendo en cuenta que contamos con 3 numeros impares 1,3,5 dentro de 6 eventos posibles: $\frac{3}{6} = \frac{1}{2}$, asímismo con los números pares, los cuales son 2,4,6 $\frac{3}{6} = \frac{1}{2}$, una probabilidad de que salga par en la segunda tirada dado que salió impar en la primera es: $P(Par|Impar) = \frac{1}{2}$

2. Interprete los resultados obtenidos.
   El que se aviente un dado y busquemos la probabilidad de obtener un par en el segundo lanzamiento es de $\frac{3}{6} = \frac{1}{2} = 50\%$ dado que estos eventos son independientes entre sí.

7 Distribución binomial

Un examen de opción múltiple tiene 5 preguntas, cada una con 4 posibles respuestas, de las cuales solo una es correcta. Un estudiante responde al azar.

1. ¿Cuál es la probabilidad de que el estudiante acierte exactamente 3 respuestas?
   Siguiendo la distribución binomial tenemos:

   $$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \ldots\ n$$

   En donde: n = número de preguntas = 5
   k = número de respuestas correctas = 3
   p = probabilidad de acertar = $\frac{1}{4}$
   Sustituyendo en la fórmula:

   $$P(X=3) = \binom{5}{3} (\frac{1}{4})^3 (\frac{3}{4})^2 = 0.0879$$

   Por lo tanto, la probabilidad de que el estudiante acierte exactamente 3 respuestas es del 8.79%

2. ¿Cuál es la probabilidad de que acierte al menos una respuesta?
   Que acierte al menos una significa que podría llegar a acertar desde 1 hasta 5 preguntas:

   $$P(X \geq 1) = 1 - P(X=0)$$

   Así que tenemos que calcular primero cuando no acierte ninguna pregunta, X = 0:

   $$P(X=0) = \binom{5}{0} (\frac{1}{4})^0 (\frac{3}{4})^5 = 0.2373$$

   Y sustituyendo en la ecuación anterior tenemos:

   $$P(X \geq 1) = 1 - 0.2373 = 0.7627$$

   Por lo tanto, la probabilidad de que acierte al menos una respuesta es del 76.27%

8 Regla de Laplace

Una urna contiene 5 bolas rojas y 7 bolas azules. Se extrae una bola al azar.

1. Determine la probabilidad de que la bola extraída sea roja.
   La regla de Laplace nos dice que consiste en el cociente entre los resultados probables y los resultados posibles de un experimento con una variable aleatoria.
   Sabiendo que el total de bolas es de 12 y las soluciones factibles son 5, entonces:

$$P(bola/roja) = \frac{5}{12} = 0.417$$

Por lo tanto, la probabilidad de que la bola extraída sea roja es del 41.7%.

2. Si se extraen dos bolas sin reemplazo, ¿cuál es la probabilidad de que ambas sean azules?
   Tenemos que:

$$P(primera/bola/azul) = \frac{7}{12}$$

$$P(segunda/bola/azul) = \frac{6}{11}$$

   Y para obtener la probabilidad final de que ambas bolas sean azules:

$$P(azules) = \frac{7}{12} \cdot \frac{6}{11} = \frac{42}{132} = 0.318$$

   Por lo tanto, la probabilidad de que ambas sean azules al momento de sacar 2 bolas es del 31.8%

9 Esperanza matemática

Suponga que una persona juega una lotería donde el premio es de 1000 dólares con una probabilidad de 0.01, y el costo del boleto es de 10 dólares.

1. Calcule la esperanza matemática de la ganancia del jugador.
   Primero tenemos que exponer los posibles resultados:

   (a) La persona gana los 1000 dólares (ganancia) con una probabilidad del 0.01 (probabilidad).

   (b) La persona pierde 10 dólares (ganancia) con una probabilidad del 0.99 (probabilidad).

   La esperanza matemática se calcula como:

$$E[X] = (1000 - 10)(0.01) + (-10)(0.99)$$

$$E[X] = (990)(0.01) + (-10)(0.99)$$

$$E[X] = 9.9 - 9.9 = 0$$

   Por lo tanto, la esperanza matemática de la ganancia del jugador es 0.

2. Interprete el resultado obtenido.
   E(X) = 0 nos dice que la persona no gana ni pierde dinero a largo plazo, entonces si la persona vuelve a jugar es posible que en un algún punto gane y sus ganancias y perdidas se equilibren.

10 Ley de los grandes números

Un experimento consiste en lanzar una moneda justa 1000 veces y calcular la frecuencia relativa de obtener cara.

1. ¿Cuál es el valor esperado de la frecuencia relativa de obtener cara?
   Al estar hablando de una moneda justa esto quiere decir que la probabilidad de obtener cara es de $P(cara) = \frac{1}{2}$. La frecuencia relativa de obtener cara se define como:

   $$f = \frac{X}{n}$$

   Siendo:
   X = número de caras obtenidas
   n = número de lanzamientos

   $$E(X) = n \cdot P(cara) = 1000 \cdot .5 = 500$$

   Aplicando los datos que nos dieron:

   $$f = \frac{E[X]}{n} = \frac{500}{1000} = 0.5$$

   Por lo tanto, el valor esperado de la frecuencia relativa de obtener cara es del 50

2. ¿Cómo se relaciona esto con la Ley de los Grandes Números?
   La Ley de los Grandes Números establece que, a medida que el número de ensayos en un experimento aumenta, la frecuencia relativa de un evento se aproxima a su probabilidad teórica, en nuestro caso, la probabilidad teórica se refiere al evento de obtener cara en el lanzamiento de una moneda, siendo este de probabilidad del $\frac{1}{2}$, confirmando el hecho de que la frecuencia de arrojar la moneda 1000 veces está relacionada a la Ley de los Grandes Números.