Probabilidad y estadística

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.
   El nombre y área de trabajo son variables cualitativas y la edad es una variable cuantitativa.

2. Determine la media, mediana y moda de la variable "Edad".

$$\bar{x} = \frac{25 + 30 + 40 + 35 + 28 + 50 + 45 + 38 + 33 + 27}{10} = \frac{351}{10} = 35.1$$

$$25,27,28,30,33,35,38,40,45,50$$

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

$$\breve{x} = amodal$$

3. Interprete los resultados obtenidos

$$\therefore \textit{El promedio de edad de los } 10 \textit{ empleados de la empresa es de } 35 \textit{ años,}$$
$$\textit{la mediana es de } 34 \textit{ años y no se repite ninguna edad entre los empleados}$$

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.

$$\bar{x} = 85$$

$$s^2 = \frac{\sum(x-\bar{x})^2}{n-1}$$

$$= \frac{(70-84.4)^2}{7} + \frac{(85-84.4)^2}{7} + \frac{(90-84.4)^2}{7} + \frac{(95-84.4)^2}{7} + \frac{(88-84.4)^2}{7} + \frac{(92-84.4)^2}{7} + \frac{(75-84.4)^2}{7} + \frac{(80-84.4)^2}{7} = 75.69$$

$$s = \sqrt{75.69} = 8.7$$

2. Interprete la dispersión de los datos.

$\therefore$ *Una desviación estándar de 8.7 significa que, en promedio, las calificaciones de los estudiantes están aproximadamente 8.7 puntos por encima o por debajo de la media. La varianza de 75.69 nos dice que hay una dispersión moderada*

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?

$$A \rightarrow Que\ sea\ programador$$
$$B \rightarrow Que\ tenga\ conocimientos\ de\ IA$$

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

$$P(A) = 0.6$$

$$P(B|A) = 0.7$$

$$P(B) = P(B|A)P(A) + P(B|A')P(A')$$

$$P(A') = 0.4\ y\ P(B|A') = .3$$

$$P(B) = (0.7)(0.6) + (0.3)(0.4) = 0.54$$

$$P(A|B) = \frac{(0.7)(0.6)}{0.54} \approx 0.7778$$

$$\therefore \textit{La probabilidad de que un empleado sea programador dado que tiene conociemientos de IA es de } 77.78\%$$

4. **Distribuciones de probabilidad**

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

   1. Calcule la probabilidad de que un lote tenga exacta mente 2 defectos.

$$x \rightarrow \textit{Número de artículos defectuosos}$$
$$x \sim \textit{Poisson con } \lambda = 3$$

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{3^x e^{-3}}{x!}$$

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

$$\therefore \textit{La probabilidad de que un lote tenga exactamente 2 defectuosos es de } 22.4\%$$

   2. Calcule la probabilidad de que un lote tenga al menos 1 defecto.

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

$$P(x \geq 1) = 1 - \frac{3^0 e^{-3}}{0!} = 1 - 0.0497 = 0.9502$$

$$\therefore \textit{La probabilidad de que un lote tenga al menos un defectuoso es de } 95.02\%$$

5. **Funciones de densidad y distribución acumulativa**

   Sea X una variable aleatoria con distribución normal de media μ = 50 y desviación estándar σ = 10.

   1. Determine la probabilidad de que X tome un valor menor que 45.

$$x \sim N(50,10)$$
$$P(x < 45) = Z = \frac{X - \mu}{\sigma} = \frac{45 - 50}{10} = -0.5$$

$$P(Z < -0.5) \approx 0.3085$$

$$\therefore \textit{La probablidad de que x tome un valor menos a 45 es de } 30.85\%$$

2. Determine la probabilidad de que X este entre 40 y 60.

$$Z_{40} = \frac{X-\mu}{\sigma} = \frac{40-50}{10} = -1$$

$$Z_{60} = \frac{X-\mu}{\sigma} = \frac{60-50}{10} = 1$$

$$P(40 \le x \le 60) = P(Z < 1) - P(Z < -1) = (1 - 0.1587) - 0.1587 = 0.8413 - 0.1587 = 0.6826$$

$$\therefore \textit{La probablidad de que x este entre } 40 \textit{ y } 60 \textit{ es de } 68.26\%$$

3. Use la función de distribución acumulativa para verificar sus respuestas.

$$\int_{-\infty}^{45} \frac{1}{10\sqrt{2\pi}} e^{-\frac{(x-50)^2}{2(100)}} = 0.3085$$

$$\int_{40}^{60} \frac{1}{10\sqrt{2\pi}} e^{-\frac{(x-50)^2}{2(100)}} = 0.6826$$

$$\therefore \textit{Por tanto las respuestas son correctas}$$

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?
2. Interprete los resultados obtenidos

$$A \rightarrow \textit{número par}$$

$$B \rightarrow \textit{número impar}$$

$$P(A) = \frac{1}{2} \ \ y\ P(B) = \frac{1}{2}$$

$$P(A \cap B) = \frac{1}{2} * \frac{1}{2} = \frac{1}{4}$$

$$P(A|B) = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{2}{4} = 0.5$$

$$\therefore \textit{La probabilidad de obtener un número par en el segundo lanzamiento dado que en el primero salió un número impar es de } 50\%$$

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?

$$x \to Número\ de\ respuestas\ correctas$$

$$x \sim B\ \ n = 5\ \ y\ \theta = 0.25$$

$$f(x) = \binom{n}{x} \theta^x (1-\theta)^{n-x}$$

$$P(x=3) = \binom{5}{3} (0.25)^3 (1-0.25)^2 = 0.0878$$

$$\therefore La\ probabilidad\ de\ que\ un\ estudainte\ acierte\ exactamente\ 3\ respuestas\ es\ de\ 8.78\%$$

2. ¿Cuál es la probabilidad de que acierte al menos una respuesta?

$$P(x \geq 1) = 1 - P(x=0) = 1 - 0.2373 = 0.7627$$

$$\therefore La\ probabilidad\ de\ que\ un\ estudainte\ acierte\ al\ menos\ una\ respuesta\ es\ de\ 76.27\%$$

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.

$$P\ (que\ la\ bola\ extraída\ sea\ roja) = \frac{5}{12} \approx 0.4167$$

$$\therefore La\ probabilidad\ de\ que\ la\ bola\ extraída\ sea\ roja\ es\ de\ 41.67\%$$

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

$$P\ (de\ que\ ambas\ sean\ azules\ ) = \frac{7}{12} * \frac{6}{11} = \frac{7}{22} \approx 0.3181$$

$$\therefore La\ probabilidad\ de\ que\ ambas\ sean\ azules\ es\ de\ 31.81\%$$

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.
2. Interprete el resultado obtenido

$$x \rightarrow que\ la\ persona\ gané$$

$$x = 0.01\ y\ x' = 0.99$$

$$E(x) = (990 * 0.01) + (-10 * 0.99) = 9.9 - 9.9 = 0$$

$$\therefore En\ promedio, el\ jugador\ ni\ gana\ ni\ pierde\ dinero\ a\ largo\ plazo.$$

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?

$$P(cara) = 0.5$$
$$E(x = cara) = 0.5$$
$$\therefore El\ valor\ esperado\ de\ la\ frecuencia\ relativa\ de\ obtener\ cara\ es\ de\ 50\%$$

2. ¿Cómo se relaciona esto con la Ley de los Grandes Números
   La frecuencia relativa se define como el número de caras obtenidas dividido entre el número total de lanzamientos. Dado que la moneda es justa y las probabilidades de obtener cara o cruz son iguales, el valor esperado de la frecuencia relativa de cara es 0.5, lo que significa que, en promedio, se espera que la mitad de los lanzamientos den como resultado cara.