Probabilidad y Estadística

Estudiante: José Luis Calderón Galarza - 2132939
Docente: Luis Ángel Gutiérrez Rodríguez

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.
2. Determine la media, mediana y moda de la variable "Edad".
3. Interprete los resultados obtenidos.

Problem 1: Solución

(a) Clasifique las variables en cualitativas o cuantitativas.

  (i) Cualitativas: Área de trabajo.

  (ii) Cuantitativas: Edad.

(b) Determine la media, mediana y moda de la variable "Edad".

  (i) Media:

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{25 + 30 + 40 + ... + 27}{10} = 35.1$$

  (ii) Mediana:
  Al ser n par:

$$\tilde{X} = \frac{X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}}{2} = \frac{X_{(5)} + X_{(6)}}{2} = \frac{33 + 35}{2} = 34$$

  (iii) Moda:
  No hay valores repetidos, por lo que no hay moda.

(c) Interprete los resultados obtenidos.
La media de edad es de 35.1 años, lo que indica el promedio del grupo. La mediana es 34 años, que representa el punto medio. Como no hay moda, no hay una edad predominante en el conjunto de datos.

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.
2. Interprete la dispersión de los datos.

Problem 2: Solución

(a) Calcule la varianza y la desviación estándar de los datos.

(i) Varianza

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{(70-84.375)^2 + (85-84.375)^2 + ... + (80-84.375)^2}{7}$$
$$= 75.696$$

(ii) Desviacion Estandar

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{(70-84.375)^2 + (85-84.375)^2 + ... + (80-84.375)^2}{7}}$$
$$= 8.7$$

(b) Interprete la dispersión de los datos.
Los datos tienen una dispersión moderada en torno a la media, con una variabilidad de 8.7 puntos.

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?

Problem 3: Solución

(a) ¿Cuál es la probabilidad de que sea programador?

(i) P(P) = 0.6 (probabilidad de que un empleado sea programador).

(ii) P(D) = 0.4 (probabilidad de que un empleado sea diseñador).

(iii) P(IA/P) = 0.7 (probabilidad de que un programador tenga conocimientos de IA).

(iv) P(IA/D) = 0.3 (probabilidad de que un diseñador tenga conocimientos de IA).

$$P(P/IA) = \frac{P(IA/P)P(P)}{P(IA)} = \frac{(0.7)(0.6)}{0.54} = 0.7778$$

$$P(IA) = P(IA/P)P(P) + P(IA/D)P(D) = (0.7)(0.6) + (0.3)(0.4) = 0.54$$

Por lo tanto, la probabilidad de que un empleado con conocimientos de IA sea programador es 0.7778 o 77.78%

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

Problem 4: Solución

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

(a) Calcule la probabilidad de que un lote tenga exactamente 2 defectos.

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

La probabilidad de que un lote tenga exactamente 2 defectos es de 0.224 o 22.4%

(b) Calcule la probabilidad de que un lote tenga al menos 1 defecto.

$$P(X \geq 1) = 1 - P(X = 0) = 1 - 0.0498 = \mathbf{0.9502}$$

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

La probabilidad de que un lote tenga al menos 1 defecto es de 0.9502 o 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.
2. Determine la probabilidad de que X esté entre 40 y 60.
3. Use la función de distribución acumulativa para verificar sus respuestas.

Problem 5: Solución

$$Z = \frac{x-\mu}{\sigma} = \frac{x-50}{10}$$

(a) Determine la probabilidad de que X tome un valor menor que 45.

$$X \sim N(\mu = 50, \sigma = 10)$$

Se estandariza la variable:

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

$$P(Z < -0.5) = P(Z > 0.5) = 0.3085$$

$$P(X < 45) = 0.3085$$

(b) Determine la probabilidad de que X esté entre 40 y 60.
Se estandarizan las variables:

$$Z_1 = \frac{40-50}{10} = -1, Z_2 = \frac{60-50}{10} = 1$$

$$P(40 < X < 60) = P(-1 < Z < 1) = P(Z < 1) - P(Z > 1) = 0.8413 - 0.1587 = 0.6826$$

(c) Use la función de distribución acumulativa para verificar sus respuestas.

La función de distribución acumulativa (FDA) para una variable aleatoria normal se define como:

$$F(x) = P(X \leq x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

donde $\Phi(z)$ representa la función de distribución acumulativa de la normal estándar.

Verificamos la primera probabilidad:

$$P(X < 45) = F(45) = \Phi(-0.5) = 0.3085$$

Ahora verificamos la segunda probabilidad:

$$P(40 < X < 60) = F(60) - F(40)$$

$$F(60) = \Phi(1) = 0.8413, \quad F(40) = \Phi(-1) = 0.1587$$

$$P(40 < X < 60) = 0.8413 - 0.1587 = 0.6826$$

Los resultados obtenidos coinciden con los valores calculados previamente, confirmando la validez de las respuestas mediante la función de distribución acumulativa.

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.

Problem 6: Solución

(a) ¿Cuál es la probabilidad de obtener un número par en el segundo lanzamiento, dado que en el primero salió un número impar?

(i) P(A) = 0.5 (probabilidad de que obtener un numero par en un lanzamiento).

(ii) P(B) = 0.5 (probabilidad de que obtener un numero impar en un lanzamiento).

$$P(A/B) = \frac{P(A) \cap P(B)}{P(B)} = \frac{P(A) \cdot P(B)}{P(B)} = P(A) = 0.5$$

Dado que A y B son eventos independientes

$$P(A) \cap P(B) = P(A) \cdot P(B)$$

(b) Interprete los resultados obtenidos.
La probabilidad de que el segundo lanzamiento sea par dado que el primero salio impar es de 0.5 o 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?
2. ¿Cuál es la probabilidad de que acierte al menos una respuesta?

Problem 7: Solución

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

(a) ¿Cuál es la probabilidad de que el estudiante acierte exactamente 3 respuestas?

$$P(x=3) = \binom{5}{3} (0.25)^3 - (1-0.25)^2 = 10(0.015625)(0.5625) = 0.0879$$

La probabilidad de que el estudiante acierte exactamente 3 respuestas es de 0.0879 0 8.79%

(b) ¿Cuál es la probabilidad de que acierte al menos una respuesta?

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

$$P(x = 0) = \binom{5}{0}(0.25)^0 - (1 - 0.25)^5 = (0.75)^5 = 0.2373$$

La probabilidad de que el estudiante acierte al menos una es de 0.7627 o 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.
2. Si se extraen dos bolas sin reemplazo, ¿cuál es la probabilidad de que ambas sean azules?

Problem 8: Solución

(a) Determine la probabilidad de que la bola extraída sea roja.

$$P(Roja) = \frac{Numero\ de\ bolas\ rojas}{Total\ de\ bolas\ en\ la\ urna} = \frac{5}{12} = 0.4167$$

La probabilidad de que la bola extraída sea roja es de 0.4167 o 41.67%

(b) Si se extraen dos bolas sin reemplazo, ¿cuál es la probabilidad de que ambas sean azules?

$$P(AzulyAzul) = \frac{7}{12} \cdot \frac{6}{11} = \frac{7}{22} = 0.3182$$

La probabilidad de que si se extraen dos bolas sin reemplazo ambas sean azules es de 0.3182 o 31.82%

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.

Problem 9: Solución

(a) Calcule la esperanza matemática de la ganancia del jugador.

$$E(x) = \sum_{\forall x} x_i \cdot p_i = \sum_{\forall x} (ganacias) \cdot (probabilidad)$$

Ganacia neta = 1000 - 10 = 990
Probabilidad de ganar = 0.01
Perdida = -10
Probabilidad de perder = 1 - 0.01 = 0.99

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

(b) Interprete el resultado obtenido.
El valor esperado es 0, lo que significa que, en promedio, el jugador no gana ni pierde dinero a largo plazo. Esto sugiere que el juego está diseñado para ser justo, ya que la esperanza matemática es neutral.

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?
2. ¿Cómo se relaciona esto con la Ley de los Grandes Números?

Problem 10: Solución

(a) ¿Cuál es el valor esperado de la frecuencia relativa de obtener cara?

Para una moneda justa, la probabilidad de cara es P(C)=0.5.
En 1000 lanzamientos, la frecuencia relativa esperada es:

$$\frac{numero\ esperado\ de\ caras}{Total\ de\ lanzamientos} = \frac{1000 \cdot 0.5}{1000} = \frac{1}{2} = 0.5$$

(b) ¿Cómo se relaciona esto con la Ley de los Grandes Números?
A medida que el número de lanzamientos aumenta, la frecuencia relativa de obtener cara se acercará a 0.5. En un número pequeño de ensayos, la frecuencia relativa puede variar significativamente, pero con 1000 lanzamientos, se espera que esté muy cerca de 0.5. Esto es precisamente lo que predice la Ley de los Grandes Números.
En conclusión cuantos más lanzamientos se realicen, menor será la desviación de la frecuencia relativa respecto a la probabilidad teórica de 50