UNIVERSIDAD AUTONOMA DE NUEVO LEON

FACULTAD DE CIENCIAS FISICO MATEMATICAS

INTELIGENCIA ARTIFICIAL

ACT 7: LABORATORIO DE REPASO DE PROBABILIDAD Y ESTADÍSTICA

Elaborado por:
Angel Francisco Hernández Gámez
2011704 Grupo: 31

Tipos de datos y medidas de tendencia central

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

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

1. CLASIFIQUE LAS VARIABLES EN CUALITATIVAS O CUANTITATIVAS.

Nombre: Variable cualitativa (nominal).

Edad: Variable cuantitativa (de razón, ya que tiene sentido el cero y se pueden realizar operaciones aritméticas).

Área de trabajo: Variable cualitativa (nominal).

2. DETERMINE LA MEDIA, MEDIANA Y MODA DE LA VARIABLE "EDAD".

A) CALCULO DE LA MEDIA:

$$\text{Media} = \frac{25+30+40+35+28+50+45+38+33+27}{10} = 35.1 \text{ años}$$

B) CALCULO DE LA MEDIANA:

Ordena las edades: 25, 27, 28, 30, 33, 35, 38, 40, 45, 50

Al haber 10 valores, la mediana es el promedio del 5º y 6º valor:

$$\frac{33+35}{2} = 34 \text{ años}$$

C) MODA:

Como cada edad solo aparece una sola vez, no existe un valor para la moda.

3. INTERPRETE LOS RESULTADOS OBTENIDOS:

La media (35.1 años) representa el promedio de edades, mientras que la mediana (34 años) indica que la mitad de los empleados tiene menos de 34 años y la otra mitad más. La ausencia de moda revela que no hay una edad que se repita con mayor frecuencia.

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.

Primero calculemos la media:

$$\text{Media} = \frac{70 + 85 + 90 + 95 + 88 + 92 + 75 + 80}{8} = \frac{675}{8} = 84.375$$

Ahora para la varianza se utiliza la fórmula:

$$\text{Varianza} = \frac{\sum_{i=1}^{n} \left(x_i - \text{media}\right)^2}{n}$$

Realizamos los cálculos:

$$\begin{aligned}
&(70 - 84.375)^2 \approx 206.64,\\
&(85 - 84.375)^2 \approx 0.39,\\
&(90 - 84.375)^2 \approx 31.64,\\
&(95 - 84.375)^2 \approx 112.89,\\
&(88 - 84.375)^2 \approx 13.14,\\
&(92 - 84.375)^2 \approx 58.14,\\
&(75 - 84.375)^2 \approx 87.89,\\
&(80 - 84.375)^2 \approx 19.14,
\end{aligned}$$

$$\sum_{i=1}^{8} (x_i - \text{media})^2 \approx 529.875,$$

$$\text{Varianza} \approx \frac{529.875}{8} \approx 66.23.$$

$$\text{Desviación estándar} = \sqrt{66.23} \approx 8.14$$

2. INTERPRETE LA DISPERSIÓN DE LOS DATOS.

La dispersión nos indica que, en promedio, los datos varían alrededor de la media en unos 8.14 puntos, lo que sugiere una variabilidad moderada en las calificaciones.

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 en 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?

Se tiene que:

- P(Programador) = 0.6
- P(Diseñador) = 0.4
- P(IA|Programador) = 0.70
- P(IA|Diseñador) = 0.30

Entonces:

$$
\begin{aligned}
P(\text{Programador} \mid \text{IA}) &= \frac{P(\text{IA} \mid \text{Programador}) \cdot P(\text{Programador})}{P(\text{IA} \mid \text{Programador}) \cdot P(\text{Programador}) + P(\text{IA} \mid \text{Diseñador}) \cdot P(\text{Diseñador})} \\
&= \frac{0.70 \times 0.6}{0.70 \times 0.6 + 0.30 \times 0.4} \\
&= \frac{0.42}{0.42 + 0.12} \\
&= \frac{0.42}{0.54} \\
&\approx 0.7778 \quad (\text{77.78\%})
\end{aligned}
$$

Por lo tanto, hay un 77.78% de probabilidad de que si se selecciona un empleado al azar sea programador y tenga conocimientos de IA.

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.

Utiliza la fórmula de Poisson:

$$P(X = 2) = \frac{e^{-3} \cdot 3^2}{2!} = \frac{e^{-3} \cdot 9}{2} \approx \frac{0.04979 \times 9}{2} \approx 0.2240 \quad \text{(22.40\%)}$$

Por lo que hay un 22.40% de probabilidad de que hayan exactamente 2 defectos en un lote.

2. CALCULE LA PROBABILIDAD DE QUE UN LOTE TENGA AL MENOS 1 DEFECTO.

$$\begin{aligned} & P(X \geq 1) = 1 - P(X = 0), \\ & P(X = 0) = \frac{e^{-3} \cdot 3^0}{0!} = e^{-3} \approx 0.04979, \\ \therefore \quad & P(X \geq 1) \approx 1 - 0.04979 \approx 0.9502 \quad \text{(95.02\%)}. \end{aligned}$$

Entonces, la probabilidad de que un lote tenga al menos 1 defecto es del 95.02%

Funciones de densidad y distribución acumulativa

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

1. DETERMINE LA PROBABILIDAD DE QUE *X* TOME UN VALOR MENOR DE 45.

Estandarizamos: $Z = \dfrac{45 - 50}{10} = -0.5$

Consultando la tabla Z: $P(Z < -0.5) \approx 0.3085 \quad$ (30.85%)

2. DETERMINE LA PROBABILIDAD DE QUE *X* ESTÉ ENTRE 40 Y 60.

$$\begin{aligned} \text{Para } X = 40: \quad Z = \frac{40 - 50}{10} = -1, \\ \text{Para } X = 60: \quad Z = \frac{60 - 50}{10} = 1, \end{aligned}$$

Por lo tanto, $P(40 < X < 60) = P(-1 < Z < 1) \approx 0.6827 \quad (68.27\%)$.

3. USE LA FUNCIÓN DE DISTRIBUCIÓN ACUMULATIVA PARA VERIFICAR SUS RESPUESTAS.

Para verificar los resultados usando la función de distribución acumulativa (CDF) de la normal, seguimos estos pasos:

1. Definimos la función de distribución acumulativa para una variable $X \sim N(\mu, \sigma)$:

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

donde $\Phi(z)$ es la CDF de la distribución normal estándar.

2. Para $X < 45$:
   a. Calculamos el valor z:

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

   b. Usamos una tabla de la normal o calculadora para obtener:

$$\Phi(-0.5) \approx 0.3085$$

Esto significa que $P(X < 45) \approx 30.85\,\%$, que coincide con el cálculo anterior.

3. Para $40 < X < 60$:
   a. Para $X = 60$, calculamos z:

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

   Obtenemos: $\Phi(1) \approx 0.8413$

   b. Para X = 40, calculamos z:

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

   Obtenemos: $\Phi(-1) \approx 0.1587$

   c. La probabilidad de que X esté entre 40 y 60 es:

$$P(40 < X < 60) = \Phi(1) - \Phi(-1)$$
$$= 0.8413 - 0.1587 \approx 0.6826 \quad \text{(o 68.26\%)}$$

Este es un valor consistente con el resultado obtenido previamente.

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?**

Debido a la independencia de los lanzamientos

$$P(\text{par en 2do} \mid \text{impar en 1er}) = P(\text{par en 2do}) = \frac{1}{2} = 0.5$$

2. **INTERPRETE LOS RESULTADOS OBTENIDOS.**

El resultado confirma que el hecho de obtener un número impar en el primer lanzamiento no afecta la probabilidad de obtener un número par en el segundo.

Distribución binomial

Un examen de opción multiple 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?**

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

$$\binom{5}{3} = 10$$

$$P(X=3) = 10 \times (0.015625) \times (0.5625) \approx 0.0879 \quad \text{(8.79\%)}$$

2. **¿CUÁL ES LA PROBABILIDAD DE QUE ACIERTE AL MENOS UNA RESPUESTA?**

$$\begin{aligned} & P(\text{al menos 1}) = 1 - P(0 \text{ aciertos}), \\ & P(0 \text{ aciertos}) = \left(\frac{3}{4}\right)^5 \approx 0.2373, \\ \therefore \quad & P(\text{al menos 1}) \approx 1 - 0.2373 \approx 0.7627 \quad \text{(76.27\%)} \end{aligned}$$

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.**

$$P(\text{roja}) = \frac{5}{12} \approx 0.4167 \quad \text{(41.67\%)}$$

2. **SI SE EXTRAEN DOS BOLAS SIN REEMPLAZO, ¿CUÁL ES LA PROBABILIDAD DE QUE AMBAS SEAN AZULES?**

$$
\begin{aligned}
P(\text{primera azul}) &= \frac{7}{12}, \\
P(\text{segunda azul} \mid \text{primera azul}) &= \frac{6}{11}, \\
P(\text{ambas azules}) &= \frac{7}{12} \times \frac{6}{11} = \frac{42}{132} = \frac{7}{22} \approx 0.3182 \quad \text{(31.82\%)}
\end{aligned}
$$

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.

Datos:

- Premio = $1000 (probabilidad 0.01)
- Costo del boleto = $10

Se define la ganancia neta:

- Si se gana: ganancia = 1000 – 10 = $990
- Si se pierde: ganancia = –10

La esperanza matemática es:

$$
\begin{aligned}
E(X) &= 0.01 \times 990 + 0.99 \times (-10) \\
&= 9.9 - 9.9 \\
&= 0
\end{aligned}
$$

2. INTERPRETE EL RESULTADO OBTENIDO.

El juego es justo en términos de esperanza, ya que en promedio, el jugador no gana ni pierde dinero, aunque el riesgo (varianza) puede ser considerable.

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?

En una moneda justa, la probabilidad de cara es 0.5, por lo que el valor esperado de la frecuencia relativa es 0.5.

2. ¿CÓMO SE RELACIONA ESTO CON LA LEY DE LOS GRANDES NÚMEROS?

La ley establece que a medida que aumenta el número de lanzamientos, la frecuencia relativa se aproxima cada vez más a la probabilidad teórica (en este caso, 0.5).