

Matematicas

Capítulo 9

Pagina 423

La Estadística y el análisis de datos

-El análisis de conjuntos de datos(la media, la mediana, la moda)

▶ **MEDIA, MODA Y MEDIANA Super facil | Medidas de tendencia central**

-El promedio ponderado

▶ **¿Cómo se calcula el promedio ponderado?**

-El rango

▶ **COMO CALCULAR EL RANGO DE UN CONJUNTO DE DATOS Super facil - Par...**

-Representación gráfica de los datos(diagramas de barras, gráficas de sectores, diagrama de puntos, diagramas de cajas, histogramas)

▶ **COMO HACER UNA GRAFICA DE BARRAS Super facil**

▶ **Gráfica de Sectores**

▶ **Interpretando datos en gráfica de puntos**

▶ **Diagrama de caja y bigotes. Ejemplo**

▶ **Histograma de Frecuencia Absoluta**

-Relaciones entre conjuntos de datos(gráfica de líneas, diagrama de dispersión)

▶ **EL GRÁFICO DE LÍNEAS**

▶ **Diagrama de Dispersión**



Análisis de conjuntos de datos

El análisis de conjuntos de datos es el proceso de examinar, limpiar y modelar datos con el objetivo de extraer información útil y tomar decisiones fundamentadas. Dentro de este análisis, se utilizan diversas medidas de tendencia central:

1. Media (promedio)

La media aritmética es el valor obtenido al sumar todos los datos de un conjunto y dividir el resultado entre la cantidad total de datos. Se calcula con la fórmula:

$$\text{Media} = \frac{\sum x}{n}$$

donde X representa los valores individuales y n el número total de elementos.

2. Mediana

Es el valor central de un conjunto de datos ordenados de menor a mayor. Si el número de datos es impar, la mediana es el valor del medio. Si es par, se obtiene promediando los dos valores centrales.

3. Moda

La moda es el valor que más se repite en un conjunto de datos. Puede haber:

- **Una única moda (unimodal)**
- **Dos modas (bimodal)**
- **Más de dos modas (multimodal)**
- **Sin moda (cuando todos los valores aparecen con la misma frecuencia)**



Promedio ponderado

El promedio ponderado es una variación de la media aritmética que asigna diferentes pesos a los valores según su importancia o frecuencia. Se calcula con la fórmula:

$$\frac{\sum (X_i \cdot P_i)}{\sum P_i}$$

donde:

- X_i son los valores individuales
- P_i son los pesos asignados a cada valor

Se usa en casos como la calificación de un estudiante, donde algunas actividades tienen más peso que otras.

El rango

El **rango** en estadística es una medida de dispersión que indica la diferencia entre el valor máximo y el valor mínimo de un conjunto de datos. Se calcula con la siguiente fórmula:

♦ Ejemplo:

Si tenemos el conjunto de datos:

3,7,10,15,20

El rango sería:

$$20 - 3 = 17$$

El rango es una medida sencilla que da una idea de la amplitud de los datos, pero no muestra cómo están distribuidos dentro del conjunto. Por eso, a menudo se complementa con otras medidas como la desviación estándar o el rango intercuartílico.



Representación gráfica de los datos

Las representaciones gráficas permiten visualizar patrones y tendencias en los datos. Algunas de las más utilizadas son:

1. Diagramas de barras

Se usan para representar datos categóricos mediante barras de diferentes alturas. Cuanto mayor es la barra, mayor es la frecuencia del dato.

2. Gráficas de sectores (gráficos circulares o de pastel)

Representan proporciones dentro de un conjunto de datos. Cada sector del círculo es proporcional a la frecuencia del dato.

3. Diagrama de puntos

Consiste en marcar un punto por cada valor del conjunto de datos. Se usa en datos discretos y permite identificar repeticiones.

4. Diagramas de cajas y bigotes

Muestran la distribución de un conjunto de datos mediante cinco valores clave:

- Mínimo
- Primer cuartil (Q1)
- Mediana (Q2)
- Tercer cuartil (Q3)
- Máximo

Se usan para identificar la dispersión y la presencia de valores atípicos.

5. Histogramas

Son similares a los diagramas de barras, pero se usan para datos numéricos continuos. Los valores se agrupan en intervalos (clases) y la altura de cada barra representa la frecuencia de los valores en ese rango.



Relaciones entre conjuntos de datos

Para analizar la relación entre dos conjuntos de datos, se usan gráficas específicas:

1. Gráfica de líneas

Conecta puntos mediante líneas para mostrar cómo un conjunto de datos cambia con el tiempo. Se usa para analizar tendencias.

2. Diagrama de dispersión

Representa pares de valores en un plano cartesiano. Es útil para observar relaciones entre dos variables y detectar correlaciones:

- **Positiva** (ambas variables aumentan)
- **Negativa** (una aumenta y la otra disminuye)
- **Nula** (no hay relación evidente)