

Sebastian Acevedo - 20222020095

Universidad Distrital Francisco José de Caldas

Probabilidad y Estadística

1. Librerías utilizadas

- numpy: Para cálculos matemáticos como media, percentiles, etc.
- pandas: Para crear y manipular la tabla de distribución de frecuencias.
- matplotlib: Para crear la visualización gráfica (histogramas).
- matplotlib.gridspec: Para organizar las subgráficas en una grilla.
- scipy.stats: Para obtener la moda de los datos.

2. Ingreso de datos

Los datos están dados como una cadena de texto. Se procesan con ``split`` y ``map`` para convertirlos en una lista de números flotantes (floats).

3. Creación de intervalos

Se define un valor mínimo inicial (20.12), y con un ancho de intervalo de 2, se crean los límites de clase con ``np.arange``. Estos se usan para agrupar los datos.

4. Cálculo de frecuencias

Con ``np.histogram`` se obtiene la frecuencia absoluta de cada intervalo. Luego se calcula la frecuencia acumulada y las relativas dividiendo por el número total de datos.

5. Tabla de distribución

Se crea un DataFrame de pandas que contiene:

- Intervalos
- Frecuencia absoluta
- Frecuencia acumulada
- Frecuencia relativa
- Frecuencia relativa acumulada

6. Estadísticos descriptivos

Se calculan:

- Media: con ``np.mean``
- Mediana: con ``np.median``
- Moda: con ``stats.mode``
- Q1 y Q3 (cuartiles): con ``np.percentile``

7. Visualización gráfica

Se utiliza matplotlib con gridspec para organizar:

- Una tabla de frecuencias en la parte superior.
- Cuatro histogramas en una cuadrícula 2x2.
 1. Frecuencia absoluta
 2. Frecuencia acumulada
 3. Frecuencia relativa
 4. Frecuencia relativa acumulada

Cada histograma muestra líneas verticales para:

- Media (línea roja)

Sebastian Acevedo - 20222020095

Universidad Distrital Francisco José de Caldas

Probabilidad y Estadística

- Mediana (verde)
- Moda (azul)
- Q1 (naranja)
- Q3 (morado)