

**Centro de Enseñanza Técnica Industrial**

**Desarrollo de Software**

**Actividad 2 - Clase 7**

**Jesús Alberto Aréchiga Carrillo**

**22310439 6N**

**Profesor**

**Clara Margarita Fernández Riveron**

**marzo de 2025**

**Guadalajara, Jalisco**

## Introducción

La varianza es una medida estadística fundamental que describe la dispersión de un conjunto de datos en torno a su media. A diferencia de la desviación media, la varianza da un peso mayor a las desviaciones más grandes al elevarlas al cuadrado, lo que permite obtener una perspectiva más robusta sobre la variabilidad de los datos. Su importancia radica en que facilita la comprensión de cuán “extendidos” o concentrados se encuentran los valores de una muestra o población, siendo además una base para otros conceptos estadísticos clave como la desviación estándar, la covarianza y el análisis de regresión.

## Ejercicio:

Programar en Python una calculadora de las medidas de dispersión, para grupos de datos de 100 valores, además que grafique los datos y valores obtenidos.

import numpy as np

import matplotlib.pyplot as plt

from statistics import mean, median, mode, pvariance, pstdev

# 100 datos enteros aleatorios entre 1 y 100 como ejemplo

data = np.random.randint(1, 101, size=100).tolist()

# Medidas de tendencia central

media = mean(data)

mediana = median(data)

# En caso de que existan valores con frecuencia repetida

try:

    moda = mode(data)

except:

    moda = None

# Medidas de dispersión

rango\_datos = max(data) - min(data)

varianza = pvariance(data)  # varianza poblacional

desviacion\_estandar = pstdev(data)  # desviación estándar poblacional

# Rango intercuartílico (IQR)

q1 = np.percentile(data, 25)

q3 = np.percentile(data, 75)

iqr = q3 - q1

# -----------------------------

# 3. Impresión de resultados

# -----------------------------

print("RESULTADOS ESTADÍSTICOS:")

print(f" - Media:                 {media:.2f}")

print(f" - Mediana:               {mediana:.2f}")

print(f" - Moda:                  {moda if moda is not None else 'No definida'}")

print(f" - Rango:                 {rango\_datos}")

print(f" - Varianza:              {varianza:.2f}")

print(f" - Desviación Estándar:   {desviacion\_estandar:.2f}")

print(f" - Cuartil 1 (Q1):        {q1:.2f}")

print(f" - Cuartil 3 (Q3):        {q3:.2f}")

print(f" - Rango Intercuartílico: {iqr:.2f}")

# Gráfica de datos

plt.figure(figsize=(8, 5))

plt.hist(data, bins=10, edgecolor='black', alpha=0.7)

# Trazamos líneas verticales para la media, mediana y moda

plt.axvline(media, color='red', linestyle='dashed', linewidth=2, label=f'Media = {media:.2f}')

plt.axvline(mediana, color='green', linestyle='dashed', linewidth=2, label=f'Mediana = {mediana:.2f}')

if moda is not None:

    plt.axvline(moda, color='blue', linestyle='dashed', linewidth=2, label=f'Moda = {moda}')

plt.title('Distribución de Datos (100 valores)')

plt.xlabel('Valor')

plt.ylabel('Frecuencia')

plt.legend()

plt.tight\_layout()

plt.show()

A screen shot of a computer

AI-generated content may be incorrect.

A graph with blue and red lines

AI-generated content may be incorrect.

## Conclusiones:

La varianza constituye una herramienta esencial en el análisis de datos, puesto que ayuda a cuantificar de forma precisa el grado de variabilidad dentro de un conjunto. Al resumir la información en un solo valor, permite comparar distintos grupos de datos y sienta las bases para técnicas de inferencia y modelado estadístico más avanzadas. Comprender y calcular la varianza resulta indispensable para todo aquel que busque extraer conclusiones rigurosas y fiables a partir de información cuantitativa.