**Measures of Dispersion: Variance and Standard Deviation**

**Variance and Standard Deviation** are two important measures of dispersion that describe how spread out the values in a dataset are.

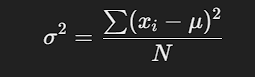
## **1. Variance (σ² or s²)**

### **Definition:**

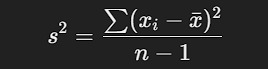
Variance measures the average squared deviation of each data point from the mean. A **higher variance** means greater spread in the data.

### **Formula:**

For a **population (σ²):**



For a **sample (s²):**



Where:

* xi= Individual data points
* μ = Population mean
* xˉ = Sample mean
* N = Population size
* n = Sample size

### **Example Calculation:**

Dataset: **{4, 8, 6, 5, 10}**

1. Mean xˉ=(4+8+6+5+10)/5=6.6
2. Compute squared deviations:
   * (4 - 6.6)² = 6.76
   * (8 - 6.6)² = 1.96
   * (6 - 6.6)² = 0.36
   * (5 - 6.6)² = 2.56
   * (10 - 6.6)² = 11.56
3. Sum = **23.2**
4. Sample Variance (n=5): s2=23.25−1=5.8
5. Population Variance (N=5): σ2=23.25=4.64

## **2. Standard Deviation (σ or s)**

### **Definition:**

Standard deviation is the **square root of variance** and provides a measure of spread in the same units as the data.

### **Formula:**



### **Example Calculation (Using Previous Variance):**

* Sample Standard Deviation:



* Population Standard Deviation:



## **Comparison: Variance vs. Standard Deviation**

## **Python Code to Calculate Variance & Standard Deviation**

import numpy as np

data = [4, 8, 6, 5, 10]

# Sample Variance & Standard Deviation

sample\_variance = np.var(data, ddof=1) # ddof=1 for sample

sample\_std\_dev = np.std(data, ddof=1)

# Population Variance & Standard Deviation

population\_variance = np.var(data) # Default is population

population\_std\_dev = np.std(data)

print(f"Sample Variance: {sample\_variance}")

print(f"Sample Standard Deviation: {sample\_std\_dev}")

print(f"Population Variance: {population\_variance}")

print(f"Population Standard Deviation: {population\_std\_dev}")

## **Conclusion**

* **Variance measures the spread by averaging squared deviations.**
* **Standard deviation is its square root, making interpretation easier.**
* **Standard deviation is widely used in statistics, finance, and data science.**