05. Measure Of Dispersion

Mahir Faysal Tusher

June 2025

Overview

In this discussion, we will cover the **Measure of Dispersion** with a focus on:

- Variance
- Standard Deviation

Variance (Population and Sample)

We first focus on variance, which helps quantify the spread of data around the mean.

Two Sample Distributions

Let us consider two different distributions:

$$\text{Distribution 1: } [2, 2, 4, 4]$$
$$\text{Distribution 2: } [1, 1, 5, 5]$$

Mean Calculation

$$\mu_1 = \frac{2+2+4+4}{4} = \frac{12}{4} = 3$$
$$\mu_2 = \frac{1+1+5+5}{4} = \frac{12}{4} = 3$$

Although both distributions have the same mean, their spread is different.

Variance Formula (Population)

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2 \tag{1}$$

Where:

- x_i = each data point
- μ = population mean
- N = number of data points

Variance Calculation

Distribution 1: $[2, 2, 4, 4]$

$$
\begin{aligned}
(x_i - \mu)^2 &= (2-3)^2 + (2-3)^2 + (4-3)^2 + (4-3)^2 \\
&= (-1)^2 + (-1)^2 + 1^2 + 1^2 = 1 + 1 + 1 + 1 = 4 \\
\sigma_1^2 &= \frac{4}{4} = 1
\end{aligned}
$$

Distribution 2: $[1, 1, 5, 5]$

$$
\begin{aligned}
(x_i - \mu)^2 &= (1-3)^2 + (1-3)^2 + (5-3)^2 + (5-3)^2 \\
&= (-2)^2 + (-2)^2 + 2^2 + 2^2 = 4 + 4 + 4 + 4 = 16 \\
\sigma_2^2 &= \frac{16}{4} = 4
\end{aligned}
$$

Conclusion

- Even though both distributions have the same mean, the spread is different.
- Distribution 1 has a lower variance ($\sigma^2 = 1$), indicating values are closer to the mean.
- Distribution 2 has a higher variance ($\sigma^2 = 4$), indicating a wider spread.
- Variance helps us measure **dispersion** in a dataset.