07. Standard Deviation

Mahir Faysal Tusher

June 2025

Population Statistics

We denote:

- N: Population size
- μ: Population mean
- σ^2: Population variance
- σ: Population standard deviation

Population Mean

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$$

Population Variance

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

Population Standard Deviation

$$\sigma = \sqrt{\sigma^2}$$

Interpretation:

- Variance measures the *spread* or *dispersion* of the data.
- Standard deviation tells us how far a typical data point is from the mean.

For example, if:

- Mean = 3
- Standard deviation = 1

- Then:
  - A data point of 4 is 1 standard deviation to the right of the mean.
  - A data point of 2 is 1 standard deviation to the left.
  - A point like 4.5 would be 1.5 standard deviations to the right.

Sample Statistics

We denote:

- n: Sample size
- $\bar{x}$: Sample mean
- s^2: Sample variance
- s: Sample standard deviation

Sample Mean

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Sample Variance

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

Note: We use $n-1$ in the denominator because of **Bessel's correction**, which corrects bias in the estimation of population variance from a sample.

Sample Standard Deviation

$$s = \sqrt{s^2}$$

Summary of Concepts

- **Mean** (both population and sample): Central tendency of data.
- **Variance**: Measures average squared deviation from the mean.
- **Standard Deviation**: Square root of variance; indicates how much data typically deviates from the mean.
- **Bessel's Correction**: Adjusts sample variance by dividing by $n-1$ instead of n.

Important Terminologies

- Mean, Median, Mode (central tendency)
- Variance, Standard Deviation (dispersion)
- Population vs. Sample (scope)

Understanding these formulas and their implications is crucial for performing correct statistical analysis.