Measures

R4 Cheng

December 24, 2024

For Continuous Data

1. Central Tendency
2. Variability or Dispersion
3. Skewness
4. Kurtosis

Central Tendency

Common central tendency measures: mean, median, mode (the most frequent value)

Mean

Sample Mean:

$$\bar{x} = \frac{\sum x_i}{n}$$

Population Mean:

$$\mu = \frac{\sum X_i}{N}$$

Median

Sample Median: $\tilde{x}$

Population Median: η (eta)

Dispersion or Variability

4 common measures of dispersion:

1. Range: $R = \max - \min$
2. Variance: population $\sigma^2 = \frac{\sum (X_i - \mu)^2}{N}$, sample $s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n-1}$
3. Standard Deviation: population $\sigma = \sqrt{\sigma^2}$, sample $s = \sqrt{s^2}$

4. Coefficient of Variation (CV): $CV = \frac{s}{\bar{x}} \times 100\%$

Remark. *Why use $n-1$? becuase it is proved to be more accurate.*

Disadvantages of Range: sensitive to outliers

Variance means the distance from the mean