Statistics

Statistics is a branch of mathematics that deals with the collection, presentation, analysis, and interpretation of data.

1. Statistical Analysis:

Statistical analysis is the science of collecting, exploring, and presenting large amounts of data to identify patterns and trends. It is also called *quantitative analysis*.

2. Non-Statistical Analysis:

Non-statistical analysis provides generic information that includes text, sound, still images, and moving images. This is also called *qualitative analysis*.

Two Major Categories of Statistics:

1. Descriptive Statistics:

Descriptive statistics helps organize data and focuses on the main characteristics of the data. It provides a summary either numerically or graphically.

Examples: Mean, Median, Mode, Standard Deviation

Types of Descriptive Statistics:

- **Measures of Central Tendency**
- **Measures of Dispersion**
- **Measures of Frequency Distribution**

Measures of Central Tendency:

A central tendency is a single value that represents the center of a data set. The common measures of central tendency include:

- **Mean**: The sum of all values divided by the number of values.
- **Median**: The middle value in an ordered data set.
- **Mode**: The most frequently occurring value in the dataset.

Formula for Mean (Ungrouped Data):

$x^- = \sum x n \bar{x} = \frac{\sum{x}}{n} x^- = n \sum x$

Where:

- $x^- \bar{x} x^-$ = Mean of the dataset
- $\sum x \sum{x} \sum x$ = Sum of all terms

- n = Number of terms

Measures of Dispersion:

Dispersion shows how spread out the data values are. The main measures of dispersion include:

- **Range**: Difference between the highest and lowest values.
  Range=Highest Value−Lowest Value\text{Range} = \text{Highest Value} - \text{Lowest Value}Range=Highest Value−Lowest Value
- **Standard Deviation (SD)**: It shows the average deviation of each data point from the mean.
  Standard Deviation=∑(x−x⁻)2n−1\text{Standard Deviation} = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}Standard Deviation=n−1∑(x−x⁻)2
- **Variance**: The square of the standard deviation.
  Variance=(Standard Deviation)2\text{Variance} = (\text{Standard Deviation})^2Variance=(Standard Deviation)2
- **Mean Deviation**: The average of the absolute deviations from the mean.
  Mean Deviation=∑│X−μ│n\text{Mean Deviation} = \frac{\sum |X - \mu|}{n}Mean Deviation=n∑│X−μ│
- **Quartile Deviation**: Half the difference between the third and first quartile.
  Quartile Deviation=Q3−Q12\text{Quartile Deviation} = \frac{Q3 - Q1}{2}Quartile Deviation=2Q3−Q1

Measures of Frequency Distribution:

Frequency distribution refers to how often each value occurs in a dataset. The main types of frequency distributions include:

- **Ungrouped Frequency Distribution**: Frequency of each value in a dataset.
- **Grouped Frequency Distribution**: Frequency of values grouped into intervals.
- **Relative Frequency Distribution**: Proportion of each value relative to the total frequency.
- **Cumulative Frequency Distribution**: Running total of frequencies less than or equal to each value.

2. Inferential Statistics:

Inferential statistics allows you to generalize from a sample to a population. It applies probability theory to draw conclusions.

Key Techniques in Inferential Statistics:

- **Hypothesis Testing**: A method to test if a hypothesis about a population parameter is supported by sample data.

  - **Null Hypothesis** ($H0H\_0H0$): Assumes no effect or difference.
  - **Alternate Hypothesis** ($H1H\_1H1$): Assumes a significant effect or difference.

- **T-Tests**: Similar to Z-scores, but uses the sample standard deviation. It helps compare means between two groups.
- **ANOVA (Analysis of Variance)**: Used to compare means across multiple groups to see if they differ significantly.

- **Chi-Square Test**: Used for categorical data to assess how likely an observed distribution is due to chance.

Probability:

Probability is a measure of the likelihood of an event occurring, expressed as a number between 0 (impossible) and 1 (certain).

Key Terms in Probability:

- **Experiment**: A process that leads to an uncertain outcome.
- **Outcome**: The result of a single trial of an experiment.
- **Event**: A specific set of outcomes from an experiment.

Types of Probability Distributions:

- **Cumulative Probability Distribution**: Describes the probability that a random variable will take a value less than or equal to a particular value in a continuous range.
- **Discrete Probability Distribution**: Describes probabilities for outcomes that take discrete values (e.g., counts of occurrences).

Here are some reference links for the statistical concepts mentioned in the revised document:

1.

Introduction to Statistics

2.
   1. https://www.scribbr.com/statistics/statistics-introduction/
   2. https://www.britannica.com/science/statistics
   3.

Descriptive Statistics

4.

   1. Measures of Central Tendency: https://www.scribbr.com/statistics/measures-of-central-tendency/
   2. Measures of Dispersion: https://www.scribbr.com/statistics/measures-of-dispersion/
   3. Frequency Distribution: https://www.scribbr.com/statistics/frequency-distributions/

5.

Inferential Statistics

6.

1. Hypothesis Testing: https://www.scribbr.com/statistics/hypothesis-testing/
2. T-Tests: https://www.scribbr.com/statistics/t-test/
3. ANOVA: https://www.scribbr.com/statistics/anova/
4. Chi-Square Test: https://www.scribbr.com/statistics/chi-square-test/

Probability

1. Basic Concepts: https://www.scribbr.com/statistics/probability/
2. Probability Distributions: https://www.scribbr.com/statistics/probability-distributions/