Assignment: Descriptive Statistics, Inferential Statistics, Probability, and Frequency Distribution

1. Descriptive Statistics

Descriptive statistics summarize and describe the features of a dataset.

Key Concepts:

1. **Measures of Central Tendency**:
    - **Mean**: The average value.
    - **Median**: The middle value when data is sorted.
    - **Mode**: The most frequent value.
2. **Measures of Dispersion**:
    - **Range**: Difference between the maximum and minimum values.
    - **Variance**: Average of the squared deviations from the mean.
    - **Standard Deviation**: Square root of variance, shows data spread.
3. **Visualization Tools**:
    - **Histogram**: Shows frequency distribution.
    - **Boxplot**: Displays data spread and outliers.
    - **Bar Chart**: Compares categories.

Example:

- Dataset: [5, 7, 8, 8, 10, 15]
    - Mean: 8.83
    - Median: 8
    - Mode: 8
    - Range: 10

---

2. Inferential Statistics

Inferential statistics make predictions or inferences about a population based on a sample.

Key Concepts:

1. **Hypothesis Testing**:
    - **Null Hypothesis (H_0)**: No effect or difference.
    - **Alternative Hypothesis (H_1)**: Indicates an effect or difference.
2. **Confidence Intervals**: A range of values that likely contain the population parameter (e.g., mean).

3. **p-Value**: Probability of observing the data given that the null hypothesis is true.

Example:

- Test whether a coin is fair (50/50 chance of heads/tails).
    - Null Hypothesis: Coin is fair.
    - Alternative Hypothesis: Coin is biased.

---

3. Probability

Probability measures the likelihood of an event occurring.

Key Concepts:

1. **Basic Probability Formula**:
   P(E)=Number of favorable outcomesTotal number of outcomesP(E) = \frac{\text{Number of favorable outcomes}}{\text{Total number of outcomes}}P(E)=Total number of outcomesNumber of favorable outcomes
2. **Types of Events**:
    - **Independent Events**: Occurrence of one doesn't affect the other.
    - **Dependent Events**: One event affects the outcome of another.
    - **Mutually Exclusive Events**: Two events cannot happen simultaneously.
3. **Probability Distributions**:
    - **Normal Distribution**: Bell-shaped curve.
    - **Binomial Distribution**: For discrete outcomes (success/failure).

Example:

- Rolling a die:
    - Probability of rolling a 4: P(4)=16P(4) = \frac{1}{6}P(4)=61

---

4. Frequency Distribution

Frequency distribution shows how often each value occurs in a dataset.

Key Concepts:

1. **Frequency Table**: A table showing frequencies of each value.
2. **Relative Frequency**: Proportion of total frequency.
3. **Cumulative Frequency**: Running total of frequencies.

Visualization Tools:

- **Pie Chart**: Shows proportions.
- **Frequency Polygon**: Line graph connecting frequencies.

Example:

- Dataset: [2, 3, 3, 5, 5, 5, 7]
    - Frequency of 5: 3 times.
    - Relative Frequency: 37≈0.429\frac{3}{7} \approx 0.42973≈0.429

Conclusion

This assignment covers the foundational concepts of statistics and probability. By understanding descriptive and inferential statistics, as well as probability and frequency distribution, we can analyze data effectively and draw meaningful conclusions.

References

1. Descriptive Statistics (Analytics Vidhya)
2. Inferential Statistics (Towards AI)
3. YouTube Video on Statistics
4. Google Drive Resources