PABNA UNIVERSITY OF SCIENCE AND TECHNOLOGY

FACULTY OF ENGINEERING AND TECHNOLOGY

DEPARTMENT OF INFORMATION AND COMMUNICATION ENGINEERING

Assignment
Course Code: STAT-2201
Course Title: Engineering Statistics

Submitted By:

Tahsin Ahmed
Roll: 210601
Session 2020-21
Dept. of Information and Communication Engineering

Submitted To :

Dr. Md. Sarwar Hosain
Associate Professor
Department of Information and Communication Engineering, PUST

Submission date : 23.04.25

Assignment topics: Non-Parametric Test, One Sample and Two Sample Sign Test

Theory:

A **non-parametric test** is a statistical test that **doesn't assume** the data follows any specific distribution (like the normal distribution).

Key Points:

- Used when data is **not normal** or has **outliers**
- Works with **ranks, medians, or signs**
- Good for **small samples** or **ordinal data**

Examples:

- **Sign Test**: Compares medians (1 sample or paired data).
- **Mann-Whitney U Test**: Compares two independent groups.
- **Wilcoxon Signed-Rank Test**: Compares two related groups (like before and after).
- **Kruskal-Wallis Test**: Compares more than two independent groups.
- **Chi-Square Test**: Tests relationships between categorical variables.

One-Sample Sign Test

Purpose:

The **One-Sample Sign Test** is used to test if the **median** of a single sample is equal to a hypothesized value.

Steps:

1. **Subtract** the hypothesized median from each data point.

2. Count how many values are **greater** (positive) and **less** (negative) than the hypothesized median.
3. Ignore values that are **equal** to the hypothesized median.
4. Use a **binomial test** to see if the number of positive and negative signs is significantly different.

Example:

Data: [15, 20, 22, 18, 14, 16]
Hypothesized median: 18

- Subtract 18 from each value:
    - 15 → −3 (negative)
    - 20 → +2 (positive)
    - 22 → +4 (positive)
    - 18 → 0 (ignore)
    - 14 → −4 (negative)
    - 16 → −2 (negative)
- **Positive signs**: 2 (values greater than 18)
- **Negative signs**: 3 (values less than 18)

Now, we use a **binomial test** to check if the number of positive and negative signs is significantly different. If they are not significantly different, the result suggests the median is likely 18.

Two-Sample Sign Test (Paired Data)

Purpose:

The **Two-Sample Sign Test** is used to compare the **medians** of two related samples (for example, measurements before and after an intervention).

Steps:

1. **Calculate the difference** for each pair (After - Before).
2. Count how many differences are **positive** (after > before) and how many are **negative** (after < before).
3. Ignore pairs where the difference is **zero**.
4. Use a **binomial test** to see if the number of positive and negative signs is significantly different.

Example:

Before: [55, 60, 70, 80, 75]
After: [60, 59, 72, 79, 74]

- Calculate the differences:
  - 60 - 55 = +5
  - 59 - 60 = −1
  - 72 - 70 = +2
  - 79 - 80 = −1
  - 74 - 75 = −1
- **Positive signs**: 2
- **Negative signs**: 3

Now, use a **binomial test** to check if the number of positive and negative signs is significantly different. If the negative signs are more frequent, it suggests the "After" group is generally lower than the "Before" group, and we **reject the null hypothesis**.

Comparison: One-Sample vs Two-Sample Sign Test

Feature	One-Sample Sign Test	Two-Sample Sign Test
Purpose	Tests if the **median** of one sample equals a specific value	Tests if there is a **difference** in medians between two related samples
Type of Data	Single sample	Paired/matched samples (e.g., before vs after)
Hypotheses	H_0: Median = given value	H_0: Median difference = 0
Test Based On	Signs (+/−) from comparing values to the hypothesized median	Signs (+/−) of differences between paired values
Distribution Assumption	No assumption (non-parametric)	No assumption (non-parametric)
Example Use	Checking if median salary = $50,000	Comparing blood pressure before and after medication
Ignored Values	Values equal to the median	Differences equal to 0

Advantages of the Sign Test

1. **Simple to Use**
    - Easy to calculate and understand.
2. **No Distribution Assumption**
    - Does not require data to be normally distributed (non-parametric).

3. **Works with Small Samples**
    - Can be used even when the sample size is small.
4. **Robust to Outliers**
    - Only uses signs (+/−), so extreme values don't affect the result.
5. **Handles Ordinal Data**
    - Can be used with ranked or ordinal data, not just numerical.

Disadvantages of the Sign Test

1. **Less Powerful**
    - Not as strong at detecting differences compared to other tests like the Wilcoxon test.
2. **Ignores Magnitude**
    - Only uses the direction (sign) of the difference, not how big the difference is.
3. **Ties Are Ignored**
    - Values equal to the median or with zero difference are ignored, reducing the sample size.
4. **Not Suitable for All Data**
    - Best for simple comparisons; may not be suitable for complex data situation

The one-sample and two-sample sign tests are useful non-parametric methods to compare medians. They are easy to use, work with small samples, and don't require the data to be normally distributed. These tests are great for basic analysis when other methods can't be used.