Chapter 25

Nonparametric Tests

Inference Methods So Far

- Variables have had **Normal distributions**.
- In practice, few variables have true Normal distributions, but our methods have been **robust** (not sensitive to moderate lack of Normality).
- If the data is clearly not Normal, then using the methods of the previous chapters will yield inaccurate results. Other approaches must be investigated.

Options for Non-Normal Data

Suppose plots suggest the data is **non-Normal**.

- If there are **extreme outliers**, investigate their cause before deciding how to proceed; do not remove outliers indiscriminately.
- In some instances, data can be **transformed** to be nearly Normal.
- Perhaps the data follow **another standard distribution**. If so, use methods based on that distribution.
- Finally, there are inference procedures that do not assume any specific form of the distribution of the population. These are called **nonparametric methods**.

Ranking Procedures

- The procedures we will study will replace *t* tests and ANOVA when the Normality condition is not met.
- Analyses will be based on the **ranks** (place in order) of the data and not their actual values.
- Rank tests concern the **center** of the population(s).
  - for Normal distributions, center is described by the mean.
  - for skewed distributions, center is described by the median.
- Ranking the data is essential to finding the median as you need to order the values before you can determine which one is in the center.

Comparison of Testing Procedures

Setting	Normal test	Rank test
One sample	One-sample *t* test Chapter 17	Wilcoxon signed rank test
Matched pairs	Apply one-sample test to differences within pairs	
Two independent samples	Two-sample *t* test Chapter 18	Wilcoxon rank sum test
Several independent samples	One-way ANOVA *F* test Chapter 24	Kruskal-Wallis test

Ranks

- To rank observations, first arrange them **in order** from smallest to largest.
- The **rank** of each observation is its position in this ordered list, starting with rank 1 for the smallest value.
- The **sum of the ranks** for *N* total observations will be:

$$1 + 2 + \cdots + N = \frac{N(N+1)}{2}$$

Case Study

Test Scores

The test scores shown below were recorded by two different professors for two sections of the same course. Test to determine if the distributions are identical or not.

Professor A	74	78	68	72	76	69	71	74
Professor B	75	80	87	81	72	73	80	76

BPS - 5th Ed. Chapter 25 7

Case Study: Test Scores

Ranks

Professor A	Overall Rank	Professor B	Overall Rank
74	7.5	75	9
78	12	80	13.5
68	1	87	16
72	4.5	81	15
76	10.5	72	4.5
69	2	73	6
71	3	80	13.5
74	7.5	76	10.5

The sum of Professor A's ranks is ***W* = 48**.

BPS - 5th Ed. Chapter 25 8

Ties in Ranks

- If two or more values are tied, assign each the **average** of the corresponding ranks. Then continue the ranking with the next value.
  - if the 3rd and 4th values are tied, assign each a rank of 3.5; the 5th value will get a rank of 5. (*Ranks*: 1, 2, 3.5, 3.5, 5).
  - if the 4th, 5th, and 6th values are tied, assign each a rank of 5; the 7th value will get a rank of 7 . (*Ranks*: 1, 2, 3, 5, 5, 5, 7).
- The **sum of the ranks** for *N* total observations remains

$$\frac{N(N+1)}{2}$$

BPS - 5th Ed. Chapter 25 9

Comparing Two Samples

Wilcoxon Rank Sum Test

- The **Wilcoxon Rank Sum** procedure tests the null hypothesis that the two populations have identical distributions. (H_0: two distributions are the same)
- If the two distributions are identical, then samples of the same size should have roughly the same number of small values, the same number of medium values, and the same number of large values. Thus, the sums of ranks for each sample should be roughly the same.
- If instead, the sample sizes differ, then they should have proportionally similar sums of ranks.

BPS - 5th Ed. Chapter 25 10

THE WILCOXON RANK SUM TEST

Draw an SRS of size n_1 from one population and draw an independent SRS of size n_2 from a second population. There are N observations in all, where $N = n_1 + n_2$. Rank all N observations. The sum W of the ranks for the first sample is the **Wilcoxon rank sum statistic.** If the two populations have the same continuous distribution, then W has mean

$$\mu_W = \frac{n_1(N+1)}{2}$$

and standard deviation

$$\sigma_W = \sqrt{\frac{n_1 n_2 (N+1)}{12}}$$

The **Wilcoxon rank sum test** rejects the hypothesis that the two populations have identical distributions when the rank sum W is far from its mean.

BPS - 5th Ed. Chapter 25 11

Case Study: Test Scores

Computations

$$\mu_W = \frac{n_1(N+1)}{2} = \frac{8(16+1)}{2} = 68$$

$$\sigma_W = \sqrt{\frac{n_1 n_2 (N+1)}{12}} = \sqrt{\frac{(8)(8)(16+1)}{12}} = 9.52$$

Standardizing gives the **Wilcoxon test statistic** as:

$$z = \frac{W - \mu_W}{\sigma_W} = \frac{48-68}{9.52} = -2.10$$

BPS - 5th Ed. Chapter 25 12

Case Study: Test Scores

1. ***Hypotheses:***
   H_0: Professors A & B have same grade distributions
   H_a: Professors A & B have different grade distributions
2. ***Test Statistic:***
   $$z = \frac{W - \mu_W}{\sigma_W} = \frac{48 - 68}{9.52} = -2.10$$
3. ***P-value:***
   P-value = 2 *P*(*Z* < -2.10) = 0.0358 *(using a computer)*
   P-value = 2 *P*(*Z* < -2.10) = 0.0358 *(Table A)*
4. ***Conclusion:***
   Since the *P*-value is smaller than $\alpha = 0.05$, there is strong evidence that the two professors have different grade distributions.

BPS - 5th Ed. Chapter 25 13

Matched Pairs

Wilcoxon Signed Rank Test

- The **Wilcoxon Signed Rank** procedure tests the null hypothesis that the two populations have identical distributions (H_0: two distributions are the same) against the alternate hypothesis that one is systematically higher (H_a: scores are systematically higher in distribution 2)
- If the two distributions are identical, then differences should be centered around 0. Thus, the sums of ranks of the positive differences should be close to that of the negative scores.

BPS - 5th Ed. Chapter 25 14

Case Study

Tell me a Story

A study of early childhood education asked kindergarten students to tell fairy tales that had been read to them earlier in the week. Each child told two stories. The first had been read to them and the second had been read but also illustrated with pictures. An expert listened to a recording of the children and assigned a score for certain uses of language.

BPS - 5th Ed. Chapter 25 15

Case Study

Tell me a Story

Story 2	0.77	0.49	0.66	0.28	0.38
Story 1	0.40	0.72	0.00	0.36	0.55
Difference	0.37	-0.23	0.66	-0.08	-0.17

H_0: scores have the same distributions
H_a: scores are systematically higher in Story 2

BPS - 5th Ed. Chapter 25 16

Case Study: Tell me a Story

Ranks

Absolute difference	0.08	0.17	0.23	0.37	0.66
Rank	1	2	3	4	5

The sum of ranks of positive differences is ***W*+ = 9**.

$$\mu_{W^+} = \frac{(5)(6)}{4} = 7.5 \qquad \sigma_{W^+} = \sqrt{\frac{(5)(6)(11)}{24}} = 3.708$$

$$z = \frac{W^+ - \mu_{W^+}}{\sigma_{W^+}} = \frac{9 - 7.5}{3.708} = 0.27$$

This gives a one-sided p-value of 0.394. Not enough significant evidence. Maybe sample size is too small.

BPS - 5th Ed. Chapter 25 17

THE WILCOXON SIGNED RANK TEST FOR MATCHED PAIRS

Draw an SRS of size n from a population for a matched pairs study and take the differences in responses within pairs. Rank the absolute values of these differences. The sum W^+ of the ranks for the positive differences is the **Wilcoxon signed rank statistic.** If the distribution of the responses is not affected by the different treatments within pairs, then W^+ has mean

$$\mu_{W^+} = \frac{n(n+1)}{4}$$

and standard deviation

$$\sigma_{W^+} = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

The **Wilcoxon signed rank test** rejects the hypothesis that there are no systematic differences within pairs when the rank sum W^+ is far from its mean.

BPS - 5th Ed. Chapter 25 18