Statistics 2
Chapter 10

Nonparametric Methods Analysis of Ranked Data

Chapter 10

Nonparametric Methods: Analysis of Ranked Data

GOALS

When you have completed this chapter, you will be able to:

ONE

Conduct the sign test for dependent samples using the binomial distribution as the test statistic.

TWO

Conduct the sign test for dependent samples using the normal distribution as the test statistic.

THREE

Conduct a test of hypothesis for the population median.

FOUR

Conduct a test of hypothesis for dependent samples using the Wilcoxon signed-rank test.

Chapter 10

Nonparametric Methods: Analysis of Ranked Data

GOALS

When you have completed this chapter, you will be able to:

FIVE

Conduct the Wilcoxon rank-sum test for independent samples.

SIX

Conduct the Kruskal-Wallis test for several independent samples.

SEVEN

Compute and interpret Spearman’s coefficient of rank correlation.

EIGHT

Conduct a test of hypothesis to determine whether the correlation among the ranks in the population is different from zero.

The Sign Test

- The Sign Test is based on the sign of a difference between two related observations.
- No assumption is necessary regarding the shape of the population.
- The binomial distribution is the test statistic for small samples and the normal approximation to the binomial for large samples.
- The test requires dependent (related) samples.

The Sign Test *continued*

- Procedure:
  - Determine the sign of the difference between related pairs.
  - Determine the number of usable pairs.
  - Compare the number of positive or negative differences to the critical value.
  - If n is the number of usable pairs (without ties), X is the number of pluses or minuses, and the binomial probability $p=.5$, then the formulas for large samples are:

Normal Approximation

- If both $n\pi$ and $n(1-\pi)$ are greater than 5, then the normal approximation can be used with

$$z = \frac{X - n\pi}{\sqrt{n\pi(1-\pi)}}$$

- If the number of pluses or minuses is more than *n*/2, then

$$z = \frac{(X - .5) - .5n}{.5\sqrt{n}}$$

- If the number of pluses or minuses is less than *n*/2, then

$$z = \frac{(X + .5) - .5n}{.5\sqrt{n}}$$

EXAMPLE 1

- The Gagliano Research Institute for Business Studies is comparing the research and development expense (R&D) as a percent of income for a sample of glass manufacturing firms for 1997 and 1998. At the .05 level of significance has the R&D expense declined? Use the sign test.

EXAMPLE 1 *continued*

Company	1997	1998
Savoth Glass	20	16
Ruisi Glass	14	13
Rubin Inc.	23	20
Vaught Bros.	24	17
Lambert Glass	31	22
Pimental Glass	22	20
Olson Glass	14	20
Flynn Glass	18	11

EXAMPLE 1 *continued*

- Step 1: H_0: p=.5 H_1: p<.5
- Step 2: H_0 is rejected if the number of negative signs is 0 or 1.
- Step 3: TS=1 (one negative sign)
- Step 4: H_0 is rejected since there is one negative sign.

Testing a Hypothesis About a Median

- When testing the value of the median, the normal approximation to the binomial distribution is used.
- The z distribution is used as the test statistic.

EXAMPLE 2

- Gordon travel agency claims that their median airfare to all destinations is \$450. This claim is being challenged by a competing agency. A random sample of 300 tickets was selected. Of these, 170 tickets were below \$450. Test the claim at the .05 level of significance.

H_0: median = 450 H_1: median ≠ 450.

- Test statistic: z=2.3671.
- Reject H_0 if 2.3671>1.96.
- Reject H_0 and conclude that the median is not equal to \$450.

Wilcoxon Signed-Rank Test

- If the assumption of normality is violated for the paired-t test, use the Wilcoxon signed-rank test.
- The ordinal scale of measurement is used.
- The observations must be related or dependent.

Wilcoxon Signed-Rank Test

- The steps in conducting the test are:
  - Compute the differences between related observations.
  - Rank the absolute differences from low to high.
  - Return the signs to the ranks and sum positive and negative ranks.
  - Compare the smaller of the two rank sums with the T value.

EXAMPLE 3

- Use the Wilcoxon matched-pair signed-rank test to determine if the R&D expenses from EXAMPLE 1 have declined as a percent of income.
  - Step 1: $H_0:$ The R&D expense has stayed the same
  - $H_1:$ The R&D expense has declined
  - Step 2: H_0 is rejected if the smaller of the rank sums is less than or equal to 5.
  - Step 3: TS=5
  - Step 4: H_0 is rejected since the smaller rank sum is 5. The R&D expense has declined

Wilcoxon Rank-Sum Test

- The Wilcoxon Rank-Sum Test is used to determine if two independent samples came from the same or equal populations.
  - No assumption about the shape of the population is required.
  - The data must be capable of being ranked.
  - Each sample must contain at least eight observations.
  - To determine the value of the test statistic W, all data values are ranked from low to high as if they were from a single population.
  - The sum of ranks for each of the two samples is determined.

Wilcoxon Rank-Sum Test

- The smaller of the two sums W is used to compute the test statistic Z from:

$$Z = \frac{W - \dfrac{n_1(n_1 + n_2 + 1)}{2}}{\sqrt{\dfrac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$

EXAMPLE 4

- Hills Community College purchased two cars, a Ford and a Chevy, for the administration's use when traveling. A sample of the repair and maintenance bills for the two cars over the last three years is shown on the next slide. At the .05 level can the university conclude that the sampling populations of cost are the same?

EXAMPLE 4 *continued*

Ford ($)	Rank	Chevy ($)	Rank
25.31	2	14.89	1
33.68	4.5	25.97	3
46.89	6	33.68	4.5
51.83	7	68.98	8
87.65	11	78.23	9
87.9	12	81.75	10
90.89	13	157.9	15
120.67	14		
Total	69.5		50.5

EXAMPLE 4 *continued*

- Step 1: H_0: The populations are the same.
- H_1: The populations are not the same.
- Step 2: H_0 is rejected if z >1.96
- Step 3: TS = z =1.5623
- Step 4: H_0 is not rejected - same distributions

Kruskal-Wallis Test: Analysis of Variance by Ranks

- The Kruskal-Wallis one-way analysis of variance by ranks compares three or more samples to determine if they came from equal populations.
  - The ordinal scale of measurement is required.
  - It is an alternative to the one-way ANOVA.
  - The chi-square distribution is the test statistic.
  - Each sample should have at least five observations.
  - The sample data is ranked from low to

Kruskal-Wallis Test: Analysis of Variance by Ranks

The test statistic is given below:

$$H = \frac{12}{n(n+1)}\left\{\frac{(\Sigma R_1)^2}{n_1} + \frac{(\Sigma R_2)^2}{n_2} + \ldots + \frac{(\Sigma R_k)^2}{n_k}\right\} - 3(n+1)$$

EXAMPLE 5

- Keely Ambrose, director of Human Resources, is studying the percent increase in salary for middle managers at four of its manufacturing plants. A sample of managers is obtained and their percent increase in salary is determined. At the 5% level of significance can Keely conclude that there is a difference in the percent increases.

EXAMPLE 5 *continued*

Milville	Rank	Camden	Rank	Eaton	Rank	White Plains	Rank
2.2	2	1.9	1	3.7	6	5.7	9
3.6	5	2.7	3	4.5	7	6.8	10.5
4.9	8	3.1	4	7.1	13.5	8.9	16
6.8	10.5	6.9	12	9.3	17	11.6	18.5
7.1	13.5	8.3	15	11.6	18.5	13.9	20

Total 39 35 62 74

EXAMPLE 5 *continued*

- Step 1: $H_0:$ The populations are the same.
- $H_1:$ The populations are not the same.
- Step 2: H_0 is rejected if $x^2 > 7.185, df = 3, \alpha = .05$
- Step 3: Test statistic H=5.95
- Step 4: H_0 is not rejected. There is no difference in the populations

Rank-Order Correlation

- Spearman's coefficient of rank correlation is used to explain the degree of the relationship between two sets of data that are at least ordinal level.
- Spearman's coefficient of rank correlation:

$$r_s = 1 - \frac{6 \Sigma d^2}{n(n^2 - 1)}$$

Testing the Significance of r_s

- State the null hypothesis: Rank correlation in population is 0.
- State the alternate hypothesis: Rank correlation in population is not 0.
- The value of the test statistic is computed from the formula:

$$t = r_s \sqrt{\frac{n-2}{1-r_s^2}}$$