

Nonparametric Tests with Ordinal Data

NOTE TO INSTRUCTOR

In this chapter, students learn how to analyze ordinal data. Students will most likely have difficulty recognizing when they have obtained ordinal data and when they should use the Wilcoxon Signed-Rank, Mann-Whitney U , and Kruskal-Wallis H hypothesis tests. They should have an easier time grasping the use of these tests as they are very similar to the kinds of hypothesis tests that they use with interval data. It would be useful to give examples of different types of data, and discuss which type of test would be appropriate (e.g. parametric test vs. nonparametric test). There are a number of examples in the text that could be used for this purpose. In addition, the classroom exercises have been developed to help students become more comfortable in calculating these new hypothesis tests.

OUTLINE OF RESOURCES

I. Ordinal Data and Correlation

- Discussion Question 18-1 (p. 182)
Discussion Question 18-2 (p. 183)
- Classroom Activity 18-1 Calculating Spearman's Correlation Coefficient (p. 183)

II. Nonparametric Hypothesis Tests

- Discussion Question 18-3 (p. 184)
- Discussion Question 18-4 (p. 184)
Classroom Activity 18-2 Re-evaluation the Candy Test (p. 184)
- Discussion Question 18-5 (p. 185)
- Classroom Activity 18-3 Make It Your Own: First Impressions: The Mann-Whitney U Test (p. 185)
- Discussion Question 18-6 (p. 186)
- Discussion Question 18-7 (p. 187)
- Classroom Activity 18-4: Reinterpreting Product Testing (p. 187)

III. Next Steps: Bootstrapping

IV. Handouts and Transparency Masters

- Handout 18-1: Calculating Spearman's Correlation Coefficient (p. 189)
- Handout 18-2: Re-evaluating the Candy test (p. 190)
- Handout 18-3: Reinterpreting Product Testing (p. 192)

CHAPTER GUIDE

I. Ordinal Data and Correlation

1. The **Spearman rank order correlation coefficient** is a nonparametric statistic that quantifies the association between two ordinal variables.

> Discussion Question 18-1

What is a Spearman correlation coefficient? What would be an everyday example of when you might decide to use it?

Your students' answers should include:

- A Spearman correlation coefficient is a nonparametric statistical test that measures the relationship between two ordinal variables.
 - An everyday example of when you might decide to use it is if you wanted to explore the relationship between your preferred ice cream and the amount of weight you've gained in pounds from consuming it. You would rank-order your preferred flavor of ice cream so that it is an ordinal variable. You would convert the amount of weight you've gained in pounds (interval variable) into an ordinal variable and explore the relationship between these two ranked-ordered data.
2. The Spearman correlation coefficient is also known as Spearman's rho and is symbolized as r_s .
 3. Because both variables need to be ordinal, we may need to convert our original interval data.
 4. To do so, we organize our data from highest to lowest (or lowest to highest, if appropriate) and rank them. If two data points share the same rank, we take the average of the two ranks they would hold if the scores were different.
 5. We can use the Spearman correlation coefficient as an inferential statistic and test the null hypothesis that the correlation is significantly different from 0.
 6. Once we have both sets of ranks, we calculate the difference between each pair of ranks and square the differences. These differences-squared will then be summed.
 7. The formula for Spearman's correlation coefficient is:

$$r_s = 1 - \left[\frac{6(\Sigma D^2)}{N(N^2 - 1)} \right]$$

> Discussion Question 18-2

What steps would you take to calculate Spearman's rho from two interval variables?

Your students' answers should include:

- To calculate Spearman's rho from two interval variables, first, you convert the data from an interval scale of measurement to an ordinal scale of measurement by organizing the data from highest to lowest and ranking them. Next, you calculate the difference between each pair of ranks, square the differences, and then sum them.
8. We interpret Spearman's rho the same as Pearson's correlation coefficient. In other words, the coefficient can range from -1 to 1 , and the sign does not indicate the strength of the relationship.
 9. Also, just like Pearson's correlation coefficient, Spearman's correlation coefficient tells us nothing about causation. The variable only tells us about the magnitude and direction of association between two ordinal variables.

Classroom Activity 18-1**Calculating Spearman's Correlation Coefficient**

When the Pearson correlation was originally discussed in the book, the authors used the correlation between class attendance and grades as an example. In this activity:

- Collect anonymous information about the number of absences the students have had so far in the class and their average test score at this point in the semester. Alternatively, you could ask students to anonymously report the amount of time that they spent studying for the last exam and their score on the exam.
- Present students with the class data, and have them convert this information into ordinal data.
- Finally, have students calculate Spearman's correlation coefficient in small groups. They should use Handout 18-1, found at the end of this chapter, to complete the assignment.

II. Nonparametric Hypothesis Tests

1. The **Wilcoxon signed-rank test for matched pairs** is a nonparametric hypothesis test used when there are two groups, a within-groups design, and an ordinal dependent variable. It is symbolized by the capital letter, T .
2. We still use the six steps of hypothesis testing.
3. There are three assumptions for this test. First, the differences between pairs must be able to be ranked. Second, we should use random selection. Lastly, the difference scores should come from a symmetric population distribution.
4. For our hypotheses, we state both the null and research hypotheses in words rather than symbols.

5. We decide on our cutoff level, clarify whether we are using a one- or two-tailed test, and determine the sample size. For our sample size, we look for the number of difference scores that are not 0.
6. We determine our critical value from the table in Appendix B in the text. We reject the null hypothesis if our test statistic is equal to or smaller than the critical value.
7. To calculate the test statistic, we organize our difference scores from highest to lowest in terms of absolute value. We rank the absolute value of the differences. Then we create two columns where one column includes the ranks associated with the positive differences and the other includes the ranks associated with the negative differences.
8. To calculate the test statistic, we sum the ranks for the positive scores and the ranks for the negative scores. The smaller of these is the test statistic, T . In other words, the official formula is: $T = \Sigma R_{\text{smaller}}$.

> Discussion Question 18-3

How does the Wilcoxon signed-rank test for matched pairs differ from a dependent-samples t test?

Your student's answers should include:

- The Wilcoxon signed-rank test is a nonparametric test and a dependent-samples t test is a parametric procedure.
- The Wilcoxon signed-rank test works with ordinal data and the dependent-samples t test works with interval data.

> Discussion Question 18-4

What steps would we take to calculate the Wilcoxon signed-rank test for matched pairs?

Your students' answers should include:

- First, organize your difference scores from highest to lowest.
- Second, rank the absolute value of the differences.
- Third, create two columns, where one column corresponds to the ranks associated with positive differences and the other column corresponds to ranks associated with negative differences.
- Fourth, sum the ranks for the positive scores and sum the ranks for the negative scores.
- Finally, the value of the Wilcoxon signed-rank test is equal to the smaller value of the sum of the ranks.

Classroom Activity 18-2

Re-evaluating the Candy Test

For this activity, students will need the data that they collected in the Classroom Activity in Chapter 10, "The Candy Test." With this data, they will then transform the information to ordinal data and calculate the Wilcoxon signed-rank test for matched pairs using Handout 18-2.

9. The **Mann-Whitney U test** is a nonparametric hypothesis test, symbolized by U and used when there are two groups, a between-groups design and an ordinal dependent variable.
10. With the Mann-Whitney U test, the data must be ordinal, random selection should be used, and ideally, no ranks should be tied.
11. Our null and research hypotheses are also stated in words.
12. We need to use the appropriate table from the appendix to determine critical values. We will want our test statistic to be smaller than the critical value.
13. To calculate the test statistic, we organize our data by raw score from highest to lowest in one column and then by rank in the next column. The rankings also should be separated by group. We then sum the ranks for each group. The official formula for the Mann-Whitney U test is:

$$U_1 = (n_1)(n_2) + \frac{[n_1(n_1 + 1)]}{2 - \Sigma R_1} \quad \text{and} \quad U_2 = (n_1)(n_2) + \frac{[n_2(n_2 + 1)]}{2 - \Sigma R_2}.$$
14. We compare our smaller test statistic to our critical value. If our test statistic is smaller than the critical value, we can reject the null hypothesis.

> Discussion Question 18-5

What steps would you take to calculate the Mann-Whitney U test?

Your students' answers should include:

- The following are the steps that we would take to calculate the Mann-Whitney U test:
 - First, organize your raw scores from highest to lowest in one column and then by rank in another column for the two groups.
 - Second, sum the ranks for the first group and also sum the ranks for the second group.
 - Finally, the value of the Mann-Whitney U test is equal to the smaller value of the sum of the ranks.

Classroom Activity 18-3

Make It Your Own: First Impressions: The Mann-Whitney U Test

You've heard the sayings "put your best foot forward" and "you never get a second chance to make a first impression." Asch (1946) appears to provide support for these assertions. The following exercise, based on Asch's 1946 Experiment VI, presents a series of descriptors for a hypothetical person. Groups A and B are both presented with the exact same descriptors; the difference between groups is the order in which the descriptors are presented. If your results are consistent with Asch (1946), you should find that Group A's impression is significantly more positive (and less negative) than Group B's impression.

Method

1. Divide the class into two groups, A and B. Let them know that this exercise is about how we form impressions about other people. Let the

students know that, while one group is observing the stimuli, the other group must keep their eyes closed and their heads down.

2. Group A: Have group B cover their eyes and put their heads down on their desks. Cover Transparency Master 18-1A.
3. Reveal one word at a time, starting with the first word. Show each word for approximately two seconds before revealing the next word in the list.
4. Repeat step 2 with group B viewing Transparency Master 18-1B and group A covering their eyes.
5. Hand out Transparency Master 18-1C to collect your data. The dependent measures here are the number of positive descriptors selected.
6. If your students are like Asch's participants, then the findings will be that your students are influenced by the order of adjectives in which someone is described. Show your class the two lists. Comment that the words are the same for both groups; the manipulation is the order of presentation.
7. Transparency Master 18-2 contains the six steps of hypothesis testing. (All transparencies can be found at the end of this chapter. You may display the transparencies on an overhead projector by photocopying them onto acetate, or you may use PowerPoint by scanning the transparency masters into your computer.)

Source

Asch, S. E. (1946). Forming impressions of personality. *The Journal of Abnormal and Social Psychology*, 41(3), 258-290.

15. The **Kruskal-Wallis H** test is a nonparametric hypothesis test, symbolized by H , used when there are more than two groups, a between-groups design, and an ordinal dependent variable.
16. There are two assumptions for this test. First, the data must be ordinal. Next, we should use random selection.
17. As with our previous nonparametric hypothesis tests, we state the null and research hypotheses in words rather than symbols.
18. We can use the chi-square table to determine critical values because the distributions are similar.
19. To calculate the test statistic, we organize our data by raw score from highest to lowest in one single column and then by rank in the second column. We then separate the ranks by group. We then take the mean of the ranks in each group. We also need to calculate the grand mean. The formula for the test statistic is: $H = [(12/N)(N + 1)][\sum n(M - GM)^2]$.
20. After calculating our test statistic, we compare it to the critical value. If the test statistic is larger than the critical value, we can reject the null hypothesis.

> Discussion Question 18-6

What steps would we take to calculate the Kruskal-Wallis H Test?

Your students' answers should include:

- First, organize your raw scores from highest to lowest in one single column and then by rank in another column. Second, separate the ranks by group.
- Finally, take the mean of the ranks for each group and also calculate a grand mean.

> Discussion Question 18-7

What would be an example of when we would use a Wilcoxon signed-rank test for matched pairs? A Mann-Whitney U test? A Kruskal-Wallis H Test?

Your students' answers should include:

- We can use a Wilcoxon signed-rank test for matched pairs if we wanted to compare data on accomplishment-related national pride for two different time periods (for year 1995–1996 vs. year 2003–2004).
- We can use a Mann-Whitney U test if we wanted to compare data on accomplishment-related national pride between communist and noncommunist countries.
- We can use a Kruskal-Wallis H test if we wanted to compare data on accomplishment-related national pride for three countries from Asia (Japan, South Korea, and Taiwan), three from Europe (Finland, Portugal, and France), and three from South America (Venezuela, Chile, and Uruguay).

Classroom Activity 18-4**Reinterpreting Product Testing**

For this activity, students will need the data that they collected from the Classroom Activity in Chapter 12, "Snack Food/Drink Product Testers." Alternatively, you could use the Classroom Activity, "Yoga and Stress" for this same activity if it is preferred. In small groups, students will use the data collected from the activity to calculate a Kruskal-Wallis H test using Handout 18-3.

III. Next Steps: Bootstrapping

1. **Bootstrapping** is a statistical process by which the original sample data are used to represent the entire population, and we repeatedly take samples from the original sample data to form a confidence interval.
2. We only bootstrap when circumstances have constrained our choices and the information is potentially very important.

Additional Reading

Agresti, A. (2002). *Categorical Data Analysis*. New York: Wiley.

Bradley, J. V. (1968). *Distribution Free Statistical Tests*. Englewood Cliffs, NJ: Prentice-Hall.

Delucchi, K. L. (1983). The use and misuse of chi-square: Lewis and Burke revisited. *Psychological Bulletin*, 81, 95-106.

Lewis, D., and Burke, C. J. (1949). The use and misuse of chi-square. *Psychological Bulletin*, 46, 433-489.

Siegel S. (1956). *Nonparametric Statistics*. New York: McGraw-Hill.

Von Eye, A., Bogat, G. A., and Von Weber, S. F. (2005). Goodness of fit for categorical variables. In Everitt, B. S., and Howell, D. C., *Encyclopedia of Statistics for the Behavioral Sciences*. Chichester, ENGL: Wiley.

Zumbo, B. D., and Zimmerman, D. W. (2000). Scales of measurement and the relation between parametric and nonparametric statistical tests. In Thompson, B. (Ed.), *Advances in Social Science Methodology*, 6. Greenwich, CT: JAI Press.

Online Resources

The following Web site has a nice explanation of the mathematics underlying chi square, including applets to perform the calculations:

<http://math.hws.edu/javamath/ryan/ChiSquare.html>

This Web site is the home of the Journal of Nonparametric Statistics:

<http://www.tandf.co.uk/journals/titles/10485252.html>

The following resource provides the basic content for the *Journal of Nonparametric Statistics*:

<http://www.ingentaconnect.com/content/tandf/gnst>

This is a great tutorial for a number of different statistical techniques, including the Mann-Whitney U test:

<http://www.texasoft.com/tutindex.html>

HANDOUT 18.2: RE-EVALUATING THE CANDY TEST

Directions: With the data that was collected in the Classroom Activity, “The Candy Test” in Chapter 10, answer each of the questions below.

1. Transform the data collected in the activity to ordinal data in the space below.

2. Calculate the Wilcoxon signed-rank test for matched pairs.

HANDOUT 18.2 (continued)

3. Interpret the findings from the nonparametric test. Is the test significant? How do you know? What does this tell us about our classroom activity?

4. Do your findings differ from those in Chapter 10? Why or why not?

HANDOUT 18.3: REINTERPRETING PRODUCT TESTING

Directions: With the data that was collected in the Classroom Activity from Chapter 12, answer each of the questions below.

1. In the space below, transform the interval data to ordinal data.

2. Calculate the Kruskal-Wallis H test using this ordinal data.

HANDOUT 18.3 (continued)

3. Interpret the findings from the nonparametric test. Is the test significant? How do you know? What does this tell us about the classroom activity?

4. Do your findings differ from those in Chapter 10? Why or why not?

TRANSPARENCY MASTER 18-1A
(for use with Classroom Activity 18-3)

Group A

INTELLIGENT

GENEROUS

CRITICAL

IMPETUOUS

OBSTINATE

RESENTFUL

JEALOUS

CRUEL

TRANSPARENCY MASTER 18-1B
(for use with Classroom Activity 18-3)

Group B

CRUEL

JEALOUS

RESENTFUL

OBSTINATE

IMPETUOUS

CRITICAL

GENEROUS

INTELLIGENT

TRANSPARENCY MASTER 18-1C
(for use with Classroom Activity 18-3)

Response Sheet for Make It Your Own: First Impressions

Directions: Answer the following question by circling your response:
Based on the descriptors just viewed, I believe that a person with this kind of personality is probably . . .

- | | |
|-----------------|--------------|
| 1. UNSELFISH | SELFISH |
| 2. UNHAPPY | HAPPY |
| 3. IRRITABLE | GOOD NATURED |
| 4. SOCIABLE | UNSOCIABLE |
| 5. POPULAR | UNPOPULAR |
| 6. UNRELIABLE | RELIABLE |
| 7. GOOD LOOKING | UNATTRACTIVE |
| 8. DISHONEST | HONEST |
| 9. WISE | CUNNING |

Group: _____

Positive descriptors: _____

Negative descriptors: _____

TRANSPARENCY MASTER 18-2
 (for use with Classroom Activity 18-3)

The Six Steps of Hypothesis Testing for Mann-Whitney U Test

Step 1. Determine the assumptions.

1. The data must be ordinal.
2. Participants were randomly selected.
3. Ideally, there are no ranks tied.

Step 2. State the null and research hypotheses.

H_0 There will be no impression difference based on presentation order.

H_1 First impressions make a difference.

Step 3. Determine the characteristics of the comparison distributions.

The sample sizes for each group.

Step 4. Determine the critical values, or cutoffs.

$$p < 0.05$$

Step 5. Calculate the test statistics.

$$U_1 = (n_1)(n_2) + \frac{n_1(n_1 + 1)}{2} - \Sigma R_1$$

$$U_2 = (n_1)(n_2) + \frac{n_2(n_2 + 1)}{2} - \Sigma R_2$$

Step 6. Make a decision.

