NONPARAMETRIC STATISTICAL PROCEDURES

OBJECTIVES

In this lecture, you will learn the following items:

- **The difference between Parametric and Non-parametric statistics.**

- **How to rank data.**

- **How to determine counts of observations.**

Introduction

Many of the tests in a traditional, introductory statistics text are based on samples that follow certain assumptions called parameters. Such tests are called ***parametric tests***. Specifically, parametric assumptions include samples that:

- Are randomly drawn from a normally distributed population,
- Consist of independent observations, except for paired values,
- Consist of values on an interval or ratio measurement scale,
- Have respective populations of approximately equal variances,
- Are adequately large, and
- Approximately resemble a normal distribution.

If any of your samples breaks one of these rules, you violate the assumptions of a parametric test. You do have some options. However, you might change the nature of your study so that your data meet the needed parameters.

For instance, if you are using an ordinal or nominal measurement scale, you might redesign your study to use an interval or ratio scale.

If your samples do not resemble a normal distribution, you might have learned a strategy that modifies your data for use with a parametric test.

- First, if you can justify your reasons, you might remove extreme values from your samples called "*outliers*".

For example:

Imagine that you test a group of children and you wish to generalize the findings to typical children in a normal state of mind. After you collect the test results, most children earn scores around 80% with some scoring above and below the average. Suppose, however, that one child scored a 5%. If you find that this child speaks no English because he arrived in your country just yesterday, it would be reasonable to exclude his score from your analysis.

- Second, you might utilize a parametric test by applying a mathematical transformation to the sample values.

For example, you might square every value in a sample.

However, some researchers argue that transformations are a form of data tampering or can distort the results. In addition, transformations do not always work, such as circumstances when data sets have particularly long tails.

- Third, there are more complicated methods for analyzing data that are beyond the scope of most introductory statistics texts. In such a case, you would be referred to a statistician.

Fortunately, there is a family of statistical tests that do not demand all the parameters, or rules, that we listed earlier. They are called nonparametric tests, and this course will focus on several such tests.

State the Null and Research Hypotheses

First, we state the hypotheses for performing the test. The two types of hypotheses are null and alternate.

The null hypothesis (H_0) is a statement that indicates no difference exists between conditions, groups, or variables.

The alternate hypothesis (H_A), also called a research hypothesis, is the statement that predicts a difference or relationship between conditions, groups, or variables.

Set the Level of Risk (or the Level of Significance) Associated with the Null Hypothesis

When we perform a particular statistical test, there is always a chance that our result is due to chance instead of any real difference. For example, we might find that two samples are significantly different.

Imagine, however, that no real difference exists. Our results would have led us to reject the null hypothesis when it was actually true. In this situation, we made a Type I error. Therefore, statistical tests assume some level of risk that we call alpha.

There is also a chance that our statistical results would lead us to not reject the null hypothesis. However, if a real difference actually does exist, then we made a Type II error. We use the Greek letter beta, to represent a Type II error.

After the hypotheses are stated, we choose the level of risk (or the level of significance) associated with the null hypothesis. We use the commonly accepted value of alpha = 0.05. By using this value, there is a 95% chance that our statistical findings are real and not due to chance.

Choose the Appropriate Test Statistic

We choose a particular type of test statistic based on characteristics of the data. For example, the number of samples or groups should be considered. Some tests are appropriate for two samples, while other tests are appropriate for three or more samples.

Measurement scale also plays an important role in choosing an appropriate test statistic. We might select one set of tests for nominal data and a different set for ordinal variables. A common ordinal measure used in social and behavioral science research is the Likert scale.

Compute the Test Statistic

The test statistic, or obtained value, is a computed value based on the particular test you need. Moreover, the method for determining the obtained value is described in each chapter and varies from test to test. For small samples, we use a procedure specific to a particular statistical test. For large samples, we approximate our data to a normal distribution and calculate a z-score for our data.

Determine the Value Needed for Rejection of the Null Hypothesis Using the Appropriate Table of Critical Values for the Particular Statistic

For small samples, we reference a table of critical values located in Appendix B. Each table provides a critical value to which we compare a computed test statistic. Finding a critical value using a table may require you to use such data characteristics as the degrees of freedom, number of samples, and/or number of groups. In addition, you may need the desired level of risk, or alpha. For large samples, we determine a critical region based on the level of risk (or the level of significance) associated with the null hypothesis. We will determine if the computed z-score falls within a critical region of the distribution.

Compare the Obtained Value with the Critical Value

Comparing the obtained value with the critical value allows us to identify a difference or relationship based on a particular level of risk. Once this is accomplished, we can state whether we must reject or must not reject the null hypothesis. While this type of phrasing may seem unusual, the standard practice in research is to state results in terms of the null hypothesis.

Some of the critical value tables are limited to particular sample or group size(s). When a sample size exceeds a table's range of value(s), we approximate our data to a normal distribution. In such cases, we use Standard Normal Table to establish a critical region of z-scores. Then, we calculate a z-score for our data and compare it with a critical region of z-scores. For example, if we use a two-tailed test with = 0.05, we do not reject the null hypothesis if the z-score is between -1.96 and $+1.96$. In other words, we do not reject if the null hypothesis if $-1.96 < z < 1.96$.

Interpret the Results

We can now give meaning to the numbers and values from our analysis based on our context. If sample differences were observed, we can comment on the strength of those differences. We can compare the observed results with the expected results. We might examine a relationship between two variables for its relative strength or search a series of events for patterns.

Reporting the Results

Communicating results in a meaningful and comprehensible manner makes our research useful to others. There is a fair amount of agreement in the research literature for reporting statistical results from parametric tests. Unfortunately, there is less agreement for nonparametric tests. We have attempted to use the more common reporting techniques found in the research literature.

RANKING DATA

Many of the nonparametric procedures involve ranking data values. Ranking values is really quite simple. Suppose that you are a math teacher and wanted to find out if students score higher after eating a healthy breakfast. You give a test and compare the scores of four students who ate a healthy breakfast with four students who did not. Table 1 shows the results.

Table 1

Students who ate breakfast	Students who skipped breakfast
87	93
96	83
92	79
84	73

To rank all of the values from Table 1 together, place them all in order in a new table from smallest to largest (see Table 2). The first value receives a rank of 1, the second value receives a rank of 2, and so on.

Table 2

Value	Rank
73	1
79	2
83	3
84	4
87	5
92	6
93	7
96	8

Notice that the values for the students who ate breakfast are in bold type. On the surface, it would appear that they scored higher. However, if you are seeking statistical significance, you need some type of procedure.

RANKING DATA WITH TIED VALUES

The aforementioned ranking method should seem straightforward. In many cases, however, two or more of the data values may be repeated. We call repeated values ties, or tied values. Say, for instance, that you repeat the preceding ranking with a different group of students. This time, you collected new values shown in Table 3.

Table 3

Students who ate breakfast	Students who skipped breakfast
90	75
85	80
95	55
70	**90**

Rank the values as in the previous example. Notice that the value of 90 is repeated. This means that the value of 90 is a tie. If these two student scores were different, they would be ranked 6 and 7. In the case of a tie, give all of the tied values the average of their rank values. In this example, the average of 6 and 7 is 6.5 (Table 4).

Table 4

Value	Rank ignoring tied values	Rank accounting for tied values
55	1	1
70	2	2
75	3	3
80	4	4
85	5	5
90	**6**	**6.5**
90	**7**	**6.5**
95	8	8

Most nonparametric statistical tests require a different formula when a sample of data contains ties. It is important to note that the formulas for ties are more algebraically complex. What is more, formulas for ties typically produce a test statistic that is only slightly different from the test statistic formulas for data without ties. It is probably for this reason that most statistics texts omit the formulas for tied values.

As you will see, however, we include the formulas for ties along with examples where applicable.

When the statistical tests in this book are explained using the computer program SPSS (Statistical Package for Social Scientists), there is no mention of any special treatment for ties. That is because SPSS automatically detects the presence of ties in any data sets and applies the appropriate procedure for calculating the test statistic.

COUNTS OF OBSERVATIONS

Some nonparametric tests require counts (or frequencies) of observations. Determining the count is fairly straightforward and simply involves counting the total number of times a particular observations is made. For example, suppose you ask several children to pick their favorite ice cream flavor given three choices: vanilla, chocolate, and strawberry. Their preferences are shown in Table 5.

To find the counts for each ice cream flavor, list the choices and tally the total number of children who picked each flavor. In other words, count the number of children who picked chocolate. Then, repeat for the other choices, vanilla and strawberry. Table 6 reveals the counts from Table 5.

Table 5

Participant	Flavor
1	Chocolate
2	Chocolate
3	Vanilla
4	Vanilla
5	Strawberry
6	Chocolate
7	Chocolate
8	Vanilla

Table 6

Flavor	Count
Chocolate	4
Vanilla	3
Strawberry	1

SUMMARY

In this lecture, we described differences between parametric and nonparametric tests. We also addressed assumptions by which non-parametric tests would be favorable over parametric tests. Then, we presented an overview of the nonparametric procedures included in this book. We also described the step-by-step approach we use to explain each test. Finally, we included explanations and examples of ranking and counting data, which are two tools for managing data when performing particular nonparametric tests.

The lectures that follow will present step-by-step directions for performing these statistical procedures both by manual, computational methods and by computer analysis using SPSS. In the next lecture, we address procedures for comparing data samples with a normal distribution.

PRACTICE QUESTIONS

1. Male high school students completed the 1-mile run at the end of their 9th grade and the beginning of their 10th grade. The following values represent the differences between the recorded times. Notice that only one student's time improved (−2 : 08). Rank the values in Table 7 beginning with the student's time difference that displayed improvement.

Table 7

Participant	Value	Rank
1	0:36	
2	0:28	
3	1:41	
4	0:37	
5	1:01	
6	2:30	
7	0:44	
8	0:47	
9	0:13	
10	0:24	
11	0:51	
12	0:09	
13	−2:08	
14	0:12	
15	0:56	

2. The values in Table 1.10 represent weekly quiz scores on math. Rank the quiz scores.

Table 8

Participant	Score	Rank
1	100	
2	60	
3	70	
4	90	
5	80	
6	100	
7	80	
8	20	
9	100	
10	50	

3. Using the data from the previous example, what are the counts (or frequencies) of passing scores and failing scores if a 70 is a passing score?

SOLUTIONS TO PRACTICE QUESTIONS

1. The value ranks are listed in Table 9 Notice that there are no ties.

Table 9

Participant	Value	Rank
1	0:36	7
2	0:28	6
3	1:41	14
4	0:37	8
5	1:01	13
6	2:30	15
7	0:44	9
8	0:47	10
9	0:13	4
10	0:24	5
11	0:51	11
12	0:09	2
13	−2:08	1
14	0:12	3
15	0:56	12

2. The value ranks are listed in Table 10 Notice the tied values. The value of 80 occurred twice and required averaging the rank values of 5 and 6.

Table 10

Participant	Score	Rank
1	**100**	**9**
2	60	3
3	70	4
4	90	7
5	**80**	**5.5**
6	**100**	**9**
7	**80**	**5.5**
8	20	1
9	**100**	**9**
10	50	2

3. Table 11 shows the passing scores and failing scores using 70 as a passing score. The counts (or frequencies) of passing scores is $n_{\text{passing}} = 7$. The counts of failing scores is $n_{\text{failing}} = 3$.

Table 11

Participant	Score	Pass/Fail
1	100	Pass
2	60	Fail
3	70	Pass
4	90	Pass
5	80	Pass
6	100	Pass
7	80	Pass
8	20	Fail
9	100	Pass
10	50	Fail