Chapter 6

Parametric Inferential Statistics

Parametric statistical procedures allow you to draw inferences about populations based on samples of those populations. To make these inferences, you must be able to make certain assumptions about the shape of the distributions of the population samples.

Section 6.1 Review of Basic Hypothesis Testing

The Null Hypothesis

In hypothesis testing, we create two hypotheses that are **mutually exclusive** (i.e., both cannot be true at the same time) and **all inclusive** (i.e., one of them must be true). We refer to those two hypotheses as the **null hypothesis** and the **alternative hypothesis**. The **null hypothesis** generally states that any difference we observe is caused by random error. The **alternative hypothesis** generally states that any difference we observe is caused by a systematic difference between groups.

Type I and Type II Errors

All hypothesis testing attempts to draw conclusions about the real world based on the results of a test (a statistical test in this case). There are four possible combinations of results (see the figure shown to the right).

TEST RESULTS	REAL WORLD: Null Hypothesis True	REAL WORLD: Null Hypothesis False
Reject Null Hypothesis	Type I Error	No Error
Fail to Reject Null Hypothesis	No Error	Type II Error

Two of the possible results are correct test results. The other two results are errors. A **Type I error** occurs when we reject a **null hypothesis** that is, in fact, true, while a **Type II error** occurs when we fail to reject a **null hypothesis** that is, in fact, false.

Significance tests determine the probability of making a **Type I error** (often called *alpha* [α], *Sig.*, or *p*). In other words, after performing a series of calculations, we obtain a probability that the **null hypothesis** is true. If there is a low probability, such as 5 or less in 100 (.05), by convention, we reject the **null hypothesis**. In other words, we typically use the .05 level (or less) as the maximum **Type I error** rate we are willing to accept.

When there is a low probability of a **Type I error**, such as .05, we can state that the **significance** test has led us to "reject the **null hypothesis**." This is synonymous with saying that a difference is "statistically significant." For instance, on a reading test, suppose you found that a random sample of girls from a school district scored higher than a random sample of boys. This result may have been obtained merely because the chance errors associated with random sampling created the observed difference (this is what the **null hypothesis** asserts). If there is a sufficiently low probability that random errors were the cause (as determined by a **significance** test), we can state that the difference between boys and girls is statistically significant.

Significance Levels Versus Critical Values

Most statistics textbooks present hypothesis testing by using the concept of a critical value. With such an approach, we obtain a value for a test statistic and compare it to a critical value we look up in a table. If the obtained value is larger than the critical value, we reject the **null hypothesis** and conclude that we have found a significant difference (or relationship). If the obtained value is less than the critical value, we fail to reject the **null hypothesis** and conclude that there is not a significant difference.

The critical-value approach is well suited to hand calculations. Tables that give critical values for alpha levels of .001, .01, .05, and so on, can be created. It is not practical to create a table for every possible alpha level.

On the other hand, SPSS can determine the exact alpha level associated with any value of a test statistic. Thus, looking up a critical value in a table is not necessary. This, however, does change the basic procedure for determining whether or not to reject the **null hypothesis**.

The section of SPSS output labeled *Sig.* (sometimes *p* or *alpha* [α]) indicates the likelihood of making a **Type I error** if we reject the **null hypothesis**. A value of .05 or less indicates that we should reject the **null hypothesis** (assuming an alpha level of .05). A value greater than .05 indicates that we should fail to reject the **null hypothesis.**

In other words, when using SPSS, we normally reject the **null hypothesis** if the output value under *Sig.* is equal to or smaller than .05, and we fail to reject the **null hypothesis** if the output value is larger than .05.

One-Tailed Versus Two-Tailed Tests

SPSS output generally includes a two-tailed alpha level (normally labeled *Sig.* in the output). A two-tailed hypothesis attempts to determine whether any difference (either positive or negative) exists. Thus, you have an opportunity to make a **Type I error** on either of the two tails of the **normal distribution**.

A one-tailed test examines a difference in a specific direction. Thus, we can make a **Type I error** on only one side (tail) of the distribution. If we have a one-tailed hypothesis, but our SPSS output gives a two-tailed **significance** result, we can take the **significance** level in the output and divide it by two. Thus, if our difference is in the right direction, and if our output indicates a **significance** level of .084 (two-tailed), but we have a one-tailed hypothesis, we can report a **significance** level of .042 (one-tailed).

Phrasing Results

Results of hypothesis testing can be stated in different ways, depending on the conventions specified by your institution. The following examples illustrate some of these differences.

Degrees of Freedom

Sometimes the degrees of freedom are given in parentheses immediately after the symbol representing the test, as in this example:

$t(3) = 7.00, p < .01$

Other times, the degrees of freedom are given within the statement of results, as in this example:

$t = 7.00, df = 3, p < .01$

Significance Level

When you obtain results that are significant, they can be described in different ways. For instance, if you obtained a **significance** level of .006 on a t test, you could describe it in any of the following three ways:

$t(3) = 7.00, p < .05$
$t(3) = 7.00, p < .01$
$t(3) = 7.00, p = .006$

Notice that because the exact probability is .006, both .05 and .01 are also correct.

There are also various ways of describing results that are not significant. For instance, if you obtained a **significance** level of .505, any of the following three statements could be used:

$t(2) = .805, ns$
$t(2) = .805, p > .05$
$t(2) = .805, p = .505$

Statement of Results

Sometimes the results will be stated in terms of the **null hypothesis**, as in the following example:

The null hypothesis was rejected ($t = 7.00, df = 3, p = .006$).

Other times, the results are stated in terms of their level of **significance**, as in the following example:

A statistically significant difference was found: $t(3) = 7.00, p < .01$.

Statistical Symbols

Generally, statistical symbols are presented in *italics*. Prior to the widespread use of computers and desktop publishing, statistical symbols were underlined. Underlining is a signal to a printer that the underlined text should be set in italics. Institutions vary on their requirements for student work, so you are advised to consult your instructor about this.

Section 6.2 Single-Sample *t* Test

Description

The single-sample t test compares the **mean** of a single sample to a known population mean. It is useful for determining if the current set of data has changed from a long-