CHAPTER

12

Estimation

Tools You Will Need

The following items are considered essential background material for this chapter. If you doubt your knowledge of any of these items, you should review the appropriate chapter or section before proceeding.

- Single-sample *t* statistic (Chapter 9)
- Independent-measures *t* statistic (Chapter 10)
- Related-samples *t* statistic (Chapter 11)

Preview

Suppose you are asked to describe the "typical" college student in the United States. For example, what is the mean age for this population? On the average, how many hours of sleep do they get each night? What is the mean number of times they have pizza per month? What is the mean amount of time spent studying per week? Notice that we are asking questions about values for population parameters (mean age, mean hours of sleep, and so on). Of course, the population of college students in the United States is large—much too large and unmanageable to study in its entirety.

When you attempt to describe the typical college student, you probably will take a look at students you know on your campus. From this group, you can begin to describe what the population of college students might be like. Notice that you are starting with a sample (the students you know) and then you are making some general statements about the population. For example, the mean age is around 21, the mean number of pizzas ordered per month is 10, the mean amount of time spent studying per week is 16 hours, and so on. The process of using sample data to estimate the values for population parameters is called *estimation*. It is used to make inferences about unknown populations and often serves as a follow-up to hypothesis tests.

As the term *estimation* implies, the sample data provide values that are only approximations (estimates) of the population parameters. Many factors can influence these estimates. One obvious factor is sample size. Suppose that you knew only two other students. How might this affect your estimate? What if you knew 100 other students to use in making your estimates? Another factor is the type of estimate used. Instead of estimating the population mean age of students at precisely 21, why not estimate the mean to be somewhere in the interval between 20 and 23 years? Notice that by using an interval you have made your estimate allowing for a *margin of error*. Accordingly, you can be more certain that your estimate contains the population parameter. In this chapter, we address these and many other questions as we closely examine the process of estimation.

12.1 AN OVERVIEW OF ESTIMATION

In Chapter 8, we introduced hypothesis testing as a statistical procedure that allows researchers to use sample data to draw inferences about populations. Hypothesis testing is probably the most frequently used inferential technique, but it is not the only one. In this chapter, we examine the process of estimation, which provides researchers with an additional method for using samples as the basis for drawing general conclusions about populations.

The basic principle underlying all of inferential statistics is that samples are representative of the populations from which they come. The most direct application of this principle is the use of sample values as estimators of the corresponding population values—that is, using statistics to estimate parameters. This process is called *estimation*.

DEFINITION

The inferential process of using sample statistics to estimate population parameters is called **estimation**.

The use of samples to estimate populations is quite common. For example, you often hear news reports such as "Sixty percent of the general public approves of the president's new budget plan." Clearly, the percentage that is reported was obtained from a sample (they don't ask everyone's opinion), and this sample statistic is being used as an estimate of the population parameter.

We already have encountered estimation in earlier sections of this book. For example, the formula for sample variance (Chapter 4) was developed so that the sample value would give an accurate and unbiased estimate of the population variance. Now we examine the process of using sample means as the basis for estimating population means.

PRECISION AND CONFIDENCE IN ESTIMATION

Before we begin the actual process of estimation, a few general points should be kept in mind. First, a sample will not give a perfect picture of the whole population. A sample is expected to be representative of the population, but there always will be some differences between the sample and the entire population. These differences are referred to as *sampling error*. Second, there are two distinct ways of making estimates. Suppose, for example, you are asked to estimate the age of the authors of this book. If you look in the frontmatter of the book, just before the Contents, you will find pictures of Gravetter and Wallnau. We are roughly the same age, so pick either one of us and estimate how old we are. Note that you could make your estimate using a single value (for example, Gravetter appears to be 55 years old) or you could use a range of values (for example, Wallnau seems to be between 50 and 60 years old). The first estimate, using a single number, is called a *point estimate*.

DEFINITION

For a **point estimate**, you use a single number as your estimate of an unknown quantity.

Point estimates have the advantage of being very precise; they specify a particular value. On the other hand, you generally do not have much confidence that a point estimate is correct. For example, most of you would not be willing to bet that Gravetter is exactly 55 years old.

The second type of estimate, using a range of values, is called an *interval estimate*. Interval estimates do not have the precision of point estimates, but they do give you more confidence. For example, it would be reasonably safe for you to bet that Wallnau is between 40 and 60 years old. At the extreme, you would be very confident betting that Wallnau is between 20 and 70 years old. Note that there is a trade-off between precision and confidence. As the interval gets wider and wider, your confidence grows. At the same time, however, the precision of the estimate gets worse. We will be using samples to make both point and interval estimates of a population mean. Because the interval estimates are associated with confidence, they usually are called *confidence intervals*.

DEFINITIONS

For an **interval estimate**, you use a range of values as your estimate of an unknown quantity.

When an interval estimate is accompanied by a specific level of confidence (or probability), it is called a **confidence interval**.

Estimation is used in the same general situations in which we have already used hypothesis testing. In fact, there is an estimation procedure that accompanies each of the hypothesis tests we presented in the preceding chapters. Figure 12.1 shows an example of a research situation in which either hypothesis testing or estimation could be used. The figure shows a population with an unknown mean (the population after treatment). A sample is selected from the unknown population. The goal of estimation is to use the sample data to obtain an estimate of the unknown population mean.

COMPARISON OF HYPOTHESIS TESTS AND ESTIMATION

You should recognize that the situation shown in Figure 12.1 is the same situation in which we have used hypothesis tests in the past. In many ways, hypothesis testing and estimation are similar. They both make use of sample data and either z-scores or t statistics to find out about unknown populations. But these two inferential procedures are

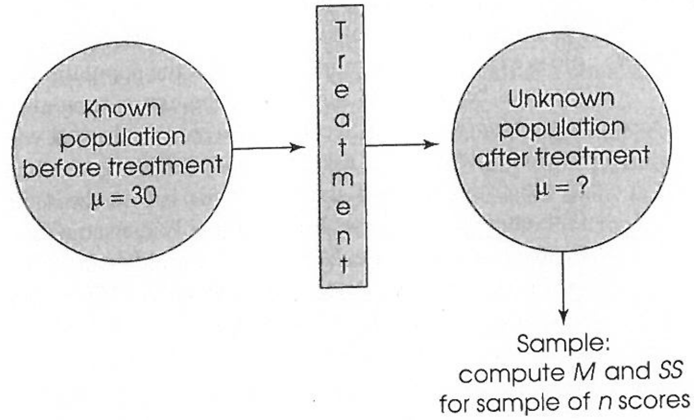


FIGURE 12.1
The basic research situation for either hypothesis testing or estimation. The goal is to use the sample data to answer questions about the unknown population mean after treatment.

designed to answer different questions. Using the situation shown in Figure 12.1 as an example, we could use a hypothesis test to evaluate the effect of the treatment. The test would determine whether the treatment has any effect. Notice that this is a yes-no question. The null hypothesis says, "No, there is no treatment effect." The alternative hypothesis says, "Yes, there is a treatment effect."

The goal of estimation, on the other hand, is to determine the value of the population mean after treatment. Essentially, estimation will determine *how much* effect the treatment has (Box 12.1). If, for example, we obtained a point estimate of $\mu = 38$ for the population after treatment, we could conclude that the effect of the treatment is to increase scores by an average of 8 points (from the original mean of $\mu = 30$ to the post-treatment mean of $\mu = 38$).

WHEN TO USE ESTIMATION

There are three situations in which estimation commonly is used:

1. Estimation is used after a hypothesis test when H_0 is rejected. Remember that when H_0 is rejected, the conclusion is that the treatment does have an effect. The next logical question would be, How much effect? This is exactly the question that estimation is designed to answer.
2. Estimation is used when you already know that there is an effect and simply want to find out how much. For example, the city school board probably knows that a special reading program will help students. However, they want to be sure that the effect is big enough to justify the cost. Estimation is used to determine the size of the treatment effect.
3. Estimation is used when you simply want some basic information about an unknown population. Suppose, for example, you want to know about the political attitudes of students at your college. You could use a sample of students as the basis for estimating the population mean.

THE LOGIC OF ESTIMATION

As we have noted, estimation and hypothesis testing are both *inferential* statistical techniques that involve using sample data as the basis for drawing conclusions about an unknown population. More specifically, a researcher begins with a question about an unknown population parameter. To answer the question, a sample is obtained, and a

BOX 12.1 HYPOTHESIS TESTING VERSUS ESTIMATION: STATISTICAL SIGNIFICANCE VERSUS PRACTICAL SIGNIFICANCE

As we already noted, hypothesis tests tend to involve a yes-no decision. Either we decide to reject H_0, or we fail to reject H_0. The language of hypothesis testing reflects this process. The outcome of the hypothesis test is one of two conclusions:

> There is no evidence for a treatment effect (fail to reject H_0)

or

> There is a statistically significant effect (H_0 is rejected)

For example, a researcher studies the effect of a new drug on people with high cholesterol. In hypothesis testing, the question is whether or not the drug has a significant effect on cholesterol levels. Suppose the hypothesis test revealed that the drug did produce a significant decrease in cholesterol. The next question might be, How much of a reduction occurs? This question calls for estimation, in which the size of a treatment effect for the population is estimated.

Estimation can be of great practical importance because the presence of a "statistically significant" effect does not necessarily mean the results are large enough for use in practical applications. Consider the following possibility: Before drug treatment, the sample of patients had a mean cholesterol level of 225. After drug treatment, their cholesterol reading was 210. When analyzed, this 15-point change reached statistical significance (H_0 was rejected). Although the hypothesis test revealed that the drug produced a *statistically significant* change, it may not be *clinically significant*. That is, a cholesterol level of 210 is still quite high. In estimation, we would estimate the population mean cholesterol level for patients who are treated with the drug. This estimated value may reveal that even though the drug does in fact reduce cholesterol levels, it does not produce a large enough change (notice we are looking at a "how much" question) to make it of any practical value. Thus, the hypothesis test might reveal that an effect occurred, but estimation indicates it is small and of little *practical significance* in real-world applications.

---

sample statistic is computed. In general, *statistical inference* involves using sample statistics to help answer questions about population parameters. The general logic underlying the processes of estimation and hypothesis testing is based on the fact that each population parameter has a corresponding sample statistic. In addition, you usually can compute a standard error that measures how much discrepancy is expected, on average, between the statistic and the parameter. For example, a sample mean, M, corresponds to the population mean, μ, with a standard error measured by σ_M or s_M.

In the preceding four chapters we presented four different situations for hypothesis testing: the z-score test, the single-sample t, the independent-measures t, and the repeated-measures t. Although it is possible to do estimation in each of these four situations, we will only consider estimation for the three t statistics. Recall that the general structure of a t statistic is as follows:

$$t = \frac{\begin{matrix}\text{sample mean} \\ \text{(or mean difference)}\end{matrix} - \begin{matrix}\text{population mean} \\ \text{(or mean difference)}\end{matrix}}{\text{estimated standard error}}$$

This general formula is used both for hypothesis testing and for estimation. In each case, the population mean (or mean difference) is unknown. The purpose for a hypothesis test is to evaluate a hypothesis about the unknown population parameter. The purpose for estimation is to determine the value for the unknown population parameter. Because hypothesis testing and estimation have different goals, they will follow different logical paths. These different paths are outlined as follows:

Hypothesis Test	Estimation
Goal: To test a hypothesis about a population parameter—usually the null hypothesis, which states that the treatment has no effect.	*Goal:* To estimate the value of an unknown population parameter—usually the value for an unknown population mean.
A. For a hypothesis test, you begin by hypothesizing a value for the unknown population parameter. This value is specified in the null hypothesis.	A. For estimation, you do not attempt to calculate *t*. Instead, you begin by estimating what the *t* value ought to be. The strategy for making this estimate is to select a "reasonable" value for *t*. (*Note:* You are not just picking a value for *t*; rather, you are estimating where the sample is located in the distribution.)
B. The hypothesized value is substituted into the formula, and *the value* for *t is computed.*	B. As with hypothesis testing, a "reasonable" value for *t* is defined as a high-probability outcome located near the center of the distribution (see Figure 12.2).
C. If the hypothesized value produces a "reasonable" value for *t*, we conclude that the hypothesis was "reasonable," and we fail to reject H_0. If the result is an extreme value for *t*, H_0 is rejected.	C. The "reasonable" value for *t* is substituted into the formula, and *the value for the unknown population parameter is computed.*
D. A "reasonable" value for *t* is defined by its location in a distribution. In general, "reasonable" values are high-probability outcomes in the center of the distribution. Extreme values with low probability are considered "unreasonable" (Figure 12.2).	D. Because you used a "reasonable" value for *t* in the formula, it is assumed that the computation will produce a "reasonable" estimate of the population parameter.

Because the goal of the estimation process is to compute a value for the unknown population mean or mean difference, it usually is easier to regroup the terms in the *t* formula so that the population value is isolated on one side of the equation. In algebraic

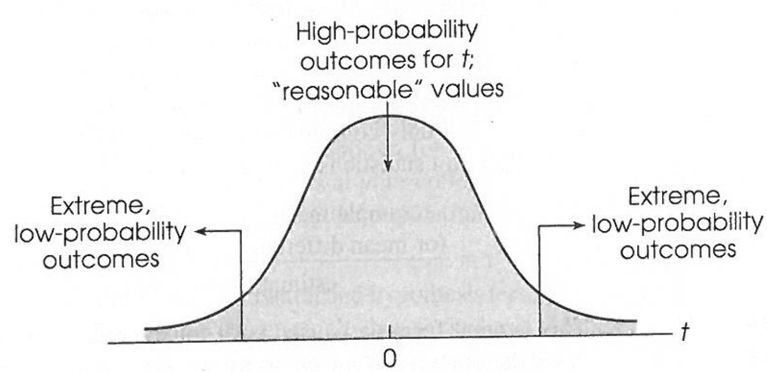


FIGURE 12.2

For estimation or hypothesis testing, the distribution of *t* statistics is divided into two sections: the middle of the distribution, consisting of high-probability outcomes that are considered "reasonable," and the extreme tails of the distribution, consisting of low-probability, "unreasonable" outcomes.

terms, we are solving the equation for the unknown population parameter. The result takes the following form:

$$\begin{matrix}\text{population mean} \\ \text{(or mean difference)}\end{matrix} = \begin{matrix}\text{sample mean} \\ \text{(or mean difference)}\end{matrix} \pm t(\text{estimated standard error}) \quad (12.1)$$

This is the general equation that we will use for estimation. Consider the following two points about Equation 12.1:

1. On the right-hand side of the equation, the values for the sample mean and the estimated standard error can be computed directly from the sample data. Thus, only the value of t is unknown. If we can determine this missing value, then we can use the equation to calculate the unknown population mean.
2. Although the specific value for the t statistic cannot be determined, we do know what the entire distribution of t statistics looks like. We can use the distribution to *estimate* what the t statistic ought to be.
   a. For a point estimate, the best bet is to use $t = 0$, the exact center of the distribution. There is no reason to suspect that the sample data are biased (either above average or below average), so $t = 0$ is a sensible value. Also, $t = 0$ is the most likely value, with probabilities decreasing steadily as you move away from zero toward the tails of the distribution.
   b. For an interval estimate, we will use a range of t values around zero. For example, to be 90% confident that our estimation is correct, we will simply use the range of t values that forms the middle 90% of the distribution. Note that we are estimating that the sample data correspond to a t statistic somewhere in the middle 90% of the t distribution.

Once we have estimated a value for t, then we have all the numbers on the right-hand side of the equation and we can calculate a value for the unknown population mean. Because one of the numbers on the right-hand side is an estimated value, the population mean that we calculate is also an estimated value.

LEARNING CHECK

1. Estimation is used to determine whether or not a treatment effect exists. (True or false?)
2. Estimation is primarily used to estimate the value of sample statistics. (True or false?)
3. In general, as confidence increases, the precision of an interval estimate decreases. (True or false?)
4. Explain why it would *not* be sensible to use estimation after a hypothesis test in which the decision was "fail to reject the null hypothesis."

ANSWERS

1. False. Estimation is used to determine *how much* effect the treatment has.
2. False. Estimation procedures are used to estimate the value for unknown population parameters.
3. True. There is a trade-off between precision and confidence.
4. If the decision is fail to reject H_0, then you have concluded that there is not a significant treatment effect. In this case, it would not make sense to estimate how much effect there is.

12.2 ESTIMATION WITH THE *t* STATISTIC

In the preceding three chapters, we introduced three different versions of the t statistic: the single-sample t in Chapter 9, the independent-measures t in Chapter 10, and the repeated-measures t in Chapter 11. Although the three t statistics were introduced in the context of hypothesis testing, they all can be adapted for use in estimation. As we saw in the previous section, the general form of the t equation for estimation is as follows:

$$\begin{matrix}\text{population mean} \\ \text{(or mean difference)}\end{matrix} = \begin{matrix}\text{sample mean} \\ \text{(or mean difference)}\end{matrix} \pm t(\text{estimated standard error})$$

With the single-sample t, we will estimate an unknown population mean, μ, using a sample mean, M. The estimation formula for the single-sample t is

$$\mu = M \pm ts_M \tag{12.2}$$

With the independent-measures t, we will estimate the size of the difference between two population means, $\mu_1 - \mu_2$, using the difference between two sample means, $M_1 - M_2$. The estimation formula for the independent-measures t is

$$\mu_1 - \mu_2 = M_1 - M_2 \pm ts_{(M_1 - M_2)} \tag{12.3}$$

Finally, the repeated-measures t statistic will be used to estimate the mean difference for the general population, μ_D, using the mean difference for a sample, M_D. The estimation formula for the repeated-measures t is

$$\mu_D = M_D \pm ts_{M_D} \tag{12.4}$$

To use the t statistic formulas for estimation, we must determine all of the values on the right-hand side of the equation (including an estimated value for t) and then use these numbers to compute an estimated value for the population mean or mean difference. Specifically, you first compute the sample mean (or mean difference) and the estimated standard error from the sample data. Next, you estimate a value, or a range of values, for t. More precisely, you are estimating where the sample data are located in the t distribution. These values complete the right-hand side of the equation and allow you to compute an estimated value for the mean (or the mean difference). The following examples demonstrate the estimation procedure with each of the three t statistics.

ESTIMATION OF μ FOR SINGLE-SAMPLE STUDIES

In Chapter 9, we introduced single-sample studies and hypothesis testing with the t statistic. Now we will use a single-sample study to estimate the value for μ, using point and interval estimates.

EXAMPLE 12.1

For several years researchers have noticed that there appears to be a regular, year-by-year increase in the average IQ for the general population. This phenomenon is called the Flynn effect after the researcher who first reported it (Flynn, 1984, 1999), and it means that psychologists must continuously update IQ tests to keep the population mean at $\mu = 100$. To evaluate the size of the effect, a researcher obtained a 10-year-old IQ test that was standardized to produce a mean IQ of $\mu = 100$ for the population 10 years ago. The test was then given to a sample of $n = 64$ of today's 20-year-old adults. The average score for the sample was $M = 107$ with a standard deviation of $s = 12$. The researcher would like to use the data to estimate how much IQ scores have changed during the past 10 years. Specifically, the researcher would like to make a point estimate and an 80% confidence interval estimate of the population

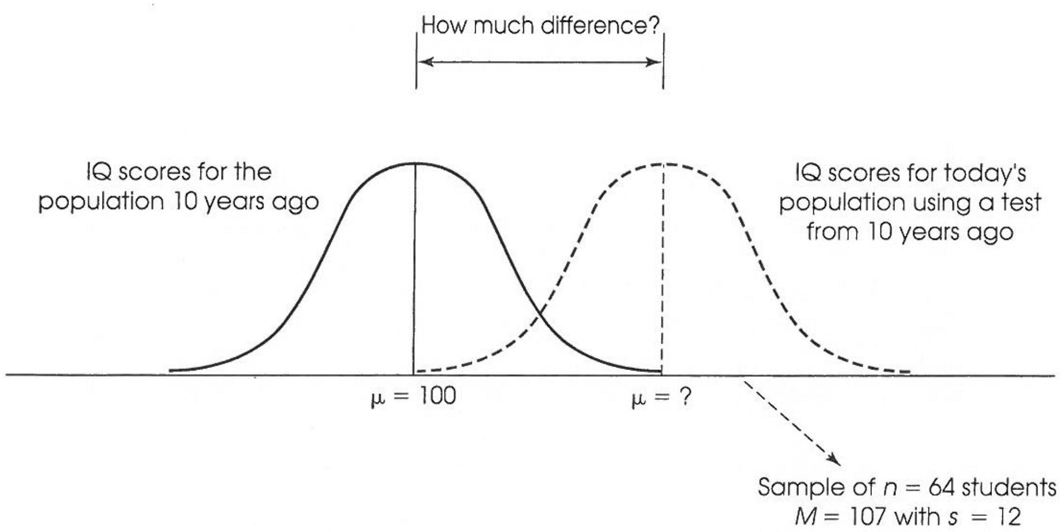


FIGURE 12.3

The structure of the research study described in Example 12.1. The goal is to use the sample to estimate the population mean IQ for students taking a 10-year-old test. We can then estimate how much IQ scores have changed during the past 10 years.

mean for people taking a 10-year-old IQ test. The structure for this research study is shown in Figure 12.3.

In this example, we are using a single sample to estimate the mean for a single population. In this case, the estimation formula for the single-sample *t* is

$$\mu = M \pm ts_M$$

To use the equation, we must first compute the estimated standard error and then determine the estimated value(s) to be used for the *t* statistic.

Compute the estimated standard error, s_M To compute the estimated standard error, it is first necessary to calculate the sample variance. For this example, we are given a sample standard deviation of $s = 12$, so the sample variance is $s^2 = (12)^2 = 144$. The estimated standard error is

$$s_M = \sqrt{\frac{s^2}{n}} = \sqrt{\frac{144}{64}} = \frac{12}{8} = 1.50$$

The point estimate As noted earlier, a point estimate involves selecting a single value for *t*. Because the *t* distribution is always symmetrically distributed with a mean of zero, we will always use $t = 0$ as the best choice for a point estimate. Using the sample data and the estimate of $t = 0$, we obtain

$$\begin{aligned}\mu &= M \pm ts_M \\ &= 107 \pm 0(1.50) \\ &= 107\end{aligned}$$

This is our point estimate of the population mean. Note that we simply have used the sample mean, M, to estimate the population mean, μ. The sample is the only information that we have about the population, and it provides an unbiased estimate of the population mean (Chapter 7, page 202). That is, on average, the sample mean provides an accurate representation of the population mean. Based on this point estimate, our conclusion is that today's population would have a mean IQ of $\mu = 107$ on an IQ test from 10 years ago. Thus, we are estimating that there has been a 7-point increase in IQ scores (from $\mu = 100$ to $\mu = 107$) during the past decade.

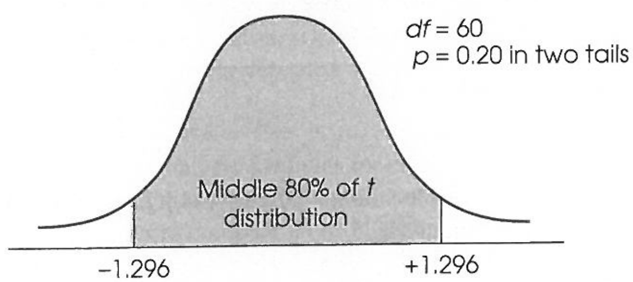
To have 80% in the middle there must be 20% (or .20) in the tails. To find the t values, look under two tails, .20 in the t table.

The interval estimate For an interval estimate, select a range of t values that is determined by the level of confidence. In this example, we want 80% confidence in our estimate of μ. Therefore, we will estimate that the t statistic is located somewhere in the middle 80% of the t distribution. With $df = n - 1 = 63$, the middle 80% of the distribution is bounded by t values of $+1.296$ and -1.296 (using $df = 60$ from the table). These values are shown in Figure 12.4. Using the sample data and this estimated range of t values, we obtain

$$\mu = M \pm t(s_M) = 107 \pm 1.296(1.50) = 107 \pm 1.944$$

FIGURE 12.4

The 80% confidence interval with $df = 60$ is constructed using t values of $t = -1.296$ and $t = +1.296$. The t values are obtained from the table using 20% (0.20) as the proportion remaining in the two tails of the distribution.

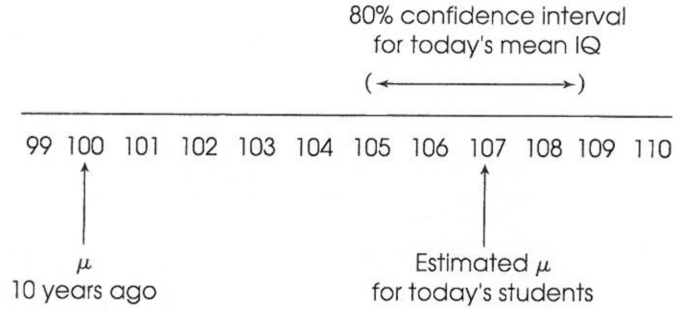


At one end of the interval we obtain 108.944 (107 + 1.944), and at the other end of the interval we obtain 105.056 (107 − 1.944). Our conclusion is that today's population would have a mean IQ between 105.056 and 108.944 if they used an IQ test from 10 years ago. In other words, we are concluding that the mean IQ has increased over the past 10 years, and we are estimating with 80% confidence that the size of the increase is between 5 and 9 points. The confidence comes from the fact that the calculation was based on only one assumption. Specifically, we assumed that the t statistic was located between $+1.296$ and -1.296, and we are 80% confident that this assumption is correct because 80% of all the possible t values are located in this interval. Finally, note that the confidence interval is constructed around the sample mean. As a result, the sample mean, $M = 107$, is located exactly in the center of the interval.

Figure 12.5 provides a visual presentation of the results from Example 12.1. The original population mean from 10 years ago is shown along with the two estimates (point and interval) of today's mean. The estimates clearly indicate that there has been an increase in IQ scores over the past 10 years, and they provide a clear indication of how large the increase is.

FIGURE 12.5

A representation of the estimates made in Example 12.1. Ten years ago the mean IQ was $\mu = 100$. Based on a sample of today's students, the mean has increased to $\mu = 107$ (point estimate) or somewhere between 105.056 and 108.944 (interval estimate).



Note that the population mean, μ, is a constant value. Because the mean does not change, it is incorrect to think that sometimes μ is in a specific interval and sometimes it is not. Instead, the intervals change from one sample to another, so that μ is in some of the intervals and is not in others. The probability that any individual confidence interval actually contains the mean is determined by the level of confidence (the percentage) used to construct the interval.

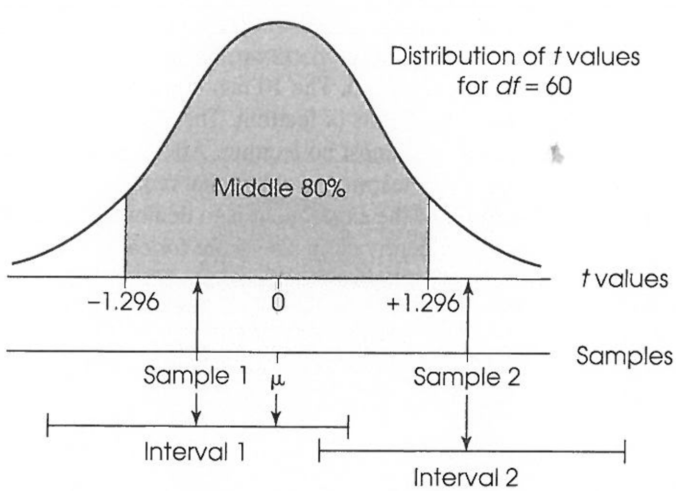
Interpretation of the confidence interval In the preceding example, we computed an 80% confidence interval to estimate an unknown population mean. We obtained an interval ranging from 105.056 to 108.944, and we are 80% confident that the unknown population mean is located within this interval. You should note, however, that the 80% confidence applies to the *process* of computing the interval rather than the specific end points for the interval. For example, if we repeated the process over and over, we could eventually obtain hundreds of different samples (each with $n = 64$ scores) and we could calculate hundreds of different confidence intervals. However, each interval is computed using the same procedure. Specifically, each interval is centered around its own sample mean, and each interval extends from a value corresponding to $t = -1.296$ at one end to $t = +1.296$ at the other end.

Figure 12.6 shows the distribution of t values with the middle 80% highlighted. To emphasize the fact that you can calculate a t value for each sample, we have added a line representing all the different samples. For example, a sample with a mean equal to μ would produce a t value of $t = 0$. Two other samples are shown in the figure.

1. Sample 1 is an example of a sample with a t value located inside the boundaries of $t = \pm 1.296$. Note that 80% of all the possible samples will be located between

FIGURE 12.6

Interpretation of the 80% confidence for an 80% confidence interval. Of all the possible samples, 80% will have t scores located in the middle 80% of the distribution and will produce confidence intervals that overlap and contain the population mean. Thus, 80% of all the possible confidence intervals will contain the true value for μ.



these boundaries because 80% of all the possible t statistics are between ± 1.296. Also note that the confidence interval for sample 1 overlaps the population mean in the center of the distribution. Thus, the confidence interval for sample 1 includes the true population mean.

2. Sample 2, on the other hand, corresponds to a t value located outside the ± 1.296 boundaries. Note that only 20% of all the possible samples will be outside the boundaries. Also note that the confidence interval for sample 2 does not contain the population mean.

Out of all the possible samples of $n = 64$ people, 80% will be similar to sample 1. That is, they will correspond to t values between ± 1.296 and will produce confidence intervals that contain the population mean. The other 20% of the possible samples will be similar to sample 2. They will correspond to t values outside the ± 1.296 boundaries, and they will produce confidence intervals that do not contain μ. Thus, out of all the different confidence intervals that we could calculate, 80% will actually contain the population mean and 20% will not.

Note that the population mean, μ, is a constant value. Because the mean does not change, it is incorrect to think that sometimes μ is in a specific interval and sometimes it is not. Instead, the intervals change from one sample to another, so that μ is in some of the intervals and is not in others. The probability that any individual confidence interval actually contains the mean is determined by the level of confidence (the percentage) used to construct the interval.

ESTIMATION OF $\mu_1 - \mu_2$ FOR INDEPENDENT-MEASURES STUDIES

The independent-measures t statistic uses the data from two separate samples to evaluate the mean difference between two populations. In Chapter 10, we used this statistic to answer a yes-no question: Is there any difference between the two population means? With estimation, we ask, *How much* difference? In this case, the independent-measures t statistic is used to estimate the value of $\mu_1 - \mu_2$. The following example demonstrates the process of estimation with the independent-measures t statistic.

EXAMPLE 12.2

Recent studies have allowed psychologists to establish definite links between specific foods and specific brain functions. For example, lecithin (found in soybeans, eggs, and liver) has been shown to increase the concentration of certain brain chemicals that help regulate memory and motor coordination. This experiment is designed to demonstrate the importance of this particular food substance.

The experiment involves two separate samples of newborn rats (an independent-measures experiment). The 10 rats in the first sample are given a normal diet containing standard amounts of lecithin. The 5 rats in the other sample are fed a special diet, which contains almost no lecithin. After 6 months, each of the rats is tested on a specially designed learning problem that requires both memory and motor coordination. The purpose of the experiment is to demonstrate the deficit in performance that results from lecithin deprivation. The score for each animal is the number of errors it makes before it solves the learning problem. The data from this experiment are as follows:

Regular Diet	No-Lecithin Diet
$n = 10$	$n = 5$
$M = 25$	$M = 33$
$SS = 250$	$SS = 140$

Because we fully expect that there will be a significant difference between these two treatments, we will not do the hypothesis test (although you should be able to do it). We want to use these data to obtain an estimate of the size of the difference between the two population means; that is, how much does lecithin affect learning performance? We will use a point estimate and a 95% confidence interval.

The basic equation for estimation with an independent-measures experiment is

$$\mu_1 - \mu_2 = (M_1 - M_2) \pm ts_{(M_1-M_2)}$$

The first step is to obtain the known values from the sample data. The sample mean difference is easy; one group averaged $M = 25$, and the other averaged $M = 33$, so there is an 8-point difference. Note that it is not important whether we call this a +8 or a −8 difference. In either case, the size of the difference is 8 points, and the regular diet group scored lower. Because it is easier to do arithmetic with positive numbers, we will use

$$M_1 - M_2 = 8$$

Compute the standard error To find the standard error, we first must pool the two variances:

$$s_p^2 = \frac{SS_1 + SS_2}{df_1 + df_2} = \frac{250 + 140}{9 + 4} = \frac{390}{13} = 30$$

Next, the pooled variance is used to compute the standard error:

$$s_{(M_1-M_2)} = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}} = \sqrt{\frac{30}{10} + \frac{30}{5}} = \sqrt{3 + 6} = \sqrt{9} = 3$$

Recall that this standard error combines the error from the first sample and the error from the second sample. Because the first sample is much larger, $n = 10$, it should have less error. This difference shows up in the formula. The larger sample contributes an error of 3 points, and the smaller sample contributes 6 points, which combine for a total error of 9 points under the square root.

Sample 1 has $df = 9$, and sample 2 has $df = 4$. The *t* statistic has $df = 9 + 4 = 13$.

Estimate the value(s) for *t* The final value needed on the right-hand side of the equation is *t*. The data from this experiment would produce a *t* statistic with $df = 13$. With 13 degrees of freedom, we can sketch the distribution of all the possible *t* values. This distribution is shown in Figure 12.7. The *t* statistic for our data is somewhere in this distribution. The problem is to estimate where. For a point estimate, the best bet

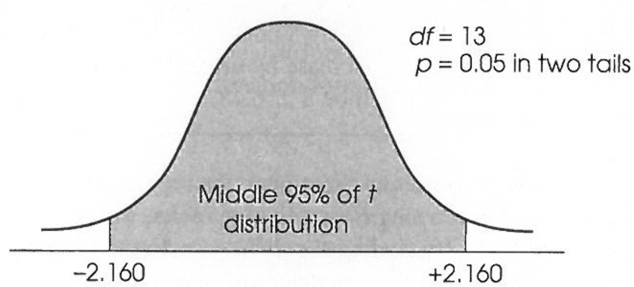


FIGURE 12.7
The distribution of *t* values with $df = 13$. Note that *t* values pile up around zero and that 95% of the values are located between −2.160 and +2.160.

is to use $t = 0$. This is the most likely value, located exactly in the middle of the distribution. To gain more confidence in the estimate, you can select a range of t values. For 95% confidence, for example, you would estimate that the t statistic is somewhere in the middle 95% of the distribution. Checking the table, you find that the middle 95% is bounded by values of $t = +2.160$ and $t = -2.160$.

Using these t values and the sample values computed earlier, we now can estimate the magnitude of the performance deficit caused by lecithin deprivation.

Compute the point estimate For a point estimate, use the single-value (point) estimate of $t = 0$:

$$\begin{aligned}\mu_1 - \mu_2 &= (M_1 - M_2) \pm ts_{(M_1-M_2)}\\ &= 8 \pm 0(3)\\ &= 8\end{aligned}$$

Note that the result simply uses the sample mean difference to estimate the population mean difference. The conclusion is that lecithin deprivation produces an average of 8 more errors on the learning task. (Based on the fact that the non-deprived animals averaged around 25 errors, an 8-point increase would mean a performance deficit of approximately 30%.)

Construct the interval estimate For an interval estimate, or confidence interval, use the range of t values. With 95% confidence, at one extreme,

$$\begin{aligned}\mu_1 - \mu_2 &= (M_1 - M_2) + ts_{(M_1-M_2)}\\ &= 8 + 2.160(3)\\ &= 8 + 6.48\\ &= 14.48\end{aligned}$$

and at the other extreme,

$$\begin{aligned}\mu_1 - \mu_2 &= (M_1 - M_2) - ts_{(M_1-M_2)}\\ &= 8 - 2.160(3)\\ &= 8 - 6.48\\ &= 1.52\end{aligned}$$

This time we conclude that the effect of lecithin deprivation is to increase errors, with an average increase somewhere between 1.52 and 14.48 errors. We are 95% confident of this estimate because our only estimation was the location of the t statistic, and we used the middle 95% of all the possible t values.

Note that the result of the point estimate is to say that lecithin deprivation will increase errors by *exactly* 8 points. To gain confidence, you must lose precision and say that errors will increase by *around* 8 points (for 95% confidence, we say that the average increase will be 8 ± 6.48).

ESTIMATION OF μ_D FOR REPEATED-MEASURES STUDIES

Finally, we turn our attention to the repeated-measures study. Remember that this type of study has a single sample of subjects, which is measured in two different treatment conditions. By finding the difference between the score for treatment 1 and the score for treatment 2, we can determine a difference score for each subject.

$$D = X_2 - X_1$$

The mean for the sample of D scores, M_D, is used to estimate the population mean μ_D, which is the mean for the entire population of difference scores.

EXAMPLE 12.3

A research study has demonstrated that self-hypnosis can be an effective treatment for allergies (Langewitz, Izakovic, & Wyler, 2005). The researchers recruited a sample of patients with moderate to severe allergic reactions. The patients were then trained to focus their minds on a specific place, such as a ski slope in the middle of winter, where allergies did not bother them. The participants who practiced this self-hypnosis therapy for the full 2 years of the study were then tested for allergic reactions to pollen under two different conditions: once without self-hypnosis and once after using the self-hypnosis therapy. Hypothetical data, similar to the actual research results, show that allergic reactions averaged $M = 71$ without self-hypnosis and $M = 50$ after using the self-hypnosis therapy. For this sample of $n = 16$ patients, the difference scores averaged $M_D = 21$ points lower when the patients were using self-hypnosis, with $SS = 1215$. We will use these results to estimate how much effect self-hypnosis therapy would have on allergy symptoms for the general population. Specifically, we will make a point estimate and a 90% confidence interval estimate for the population mean difference, μ_D.

You should recognize that this study requires a repeated-measures t statistic. For estimation, the repeated-measures t equation is as follows:

$$\mu_D = M_D \pm ts_{M_D}$$

The sample mean is $M_D = 21$, so all that remains is to compute the estimated standard error and estimate the appropriate value(s) for t.

Compute the standard error To find the standard error, we first must compute the sample variance:

$$s^2 = \frac{SS}{n-1} = \frac{1215}{15} = 81$$

Now the estimated standard error is

$$s_{M_D} = \sqrt{\frac{s^2}{n}} = \sqrt{\frac{81}{16}} = \frac{9}{4} = 2.25$$

To complete the estimate of μ_D, we must identify the value of t. We will consider the point estimate and the interval estimate separately.

Compute the point estimate To obtain a point estimate, a single value of t is selected to approximate the location of M_D. Remember that the t distribution is symmetrical and bell-shaped with a mean of zero (see Figure 12.7). Because $t = 0$ is the most frequently occurring value in the distribution, this is the t value used for the point estimate. Using this value in the estimation formula gives

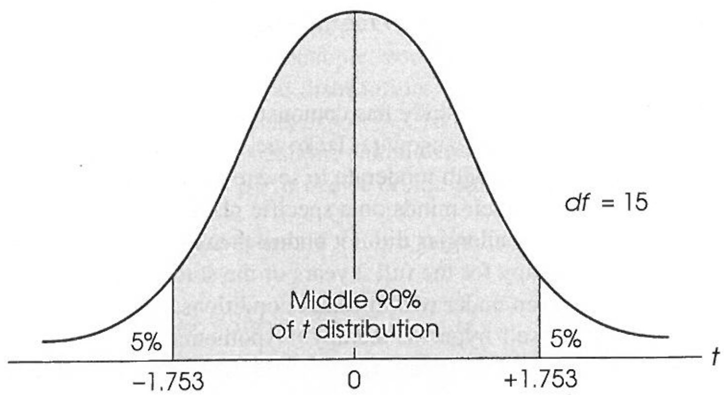
$$\mu_D = M_D \pm ts_{M_D} = 21 \pm 0(2.25) = 21$$

For this example, our best estimate is that the self-hypnosis will lower allergy symptoms in the general population by an average of $\mu_D = 21$ points. As noted several times before, the sample mean, $M_D = 21$, provides the best point estimate of μ_D.

Construct the interval estimate We also want to make an interval estimate in order to be 90% confident that the interval contains the value of μ_D. To get the

FIGURE 12.8

The t values for the 90% confidence interval are obtained by consulting the t tables for $df = 15$, $p = 0.10$ for two tails.



interval, it is necessary to determine what t values form the boundaries of the middle 90% of the t distribution. To use the t distribution table, we first must determine the proportion associated with the tails of this distribution. With 90% in the middle, the remaining area in both tails must be 10%, or $p = .10$. Also note that our sample has $n = 16$ scores, so the t statistic will have $df = n - 1 = 15$. Using $df = 15$ and $p = 0.10$ for two tails, you should find the values $+1.753$ and -1.753 in the t table. These values form the boundaries for the middle 90% of the t distribution (Figure 12.8). We are confident that the t value for our sample is in this range because 90% of all the possible t values are there. Using these values in the estimation formula, we obtain the following: On one end of the interval,

$$\begin{aligned}\mu_D &= M_D - ts_{M_D} \\ &= 21 - 1.753(2.25) \\ &= 21 - 3.94 \\ &= 17.06\end{aligned}$$

and on the other end of the interval,

$$\begin{aligned}\mu_D &= 21 + 1.753(2.25) \\ &= 21 + 3.94 \\ &= 24.94\end{aligned}$$

Therefore, the researchers can conclude that self-hypnosis therapy would reduce allergy symptoms in the general population by an average of around 21 points. They can be 90% confident that the average reduction is between 17.06 and 24.94 points. Given that the allergy symptoms for untreated patients averaged $M = 71$, this is a reduction of around 30%.

LEARNING CHECK

1. A researcher would like to determine the average reading ability for third-grade students in the local school district. A sample of $n = 25$ students is selected and each student takes a standardized reading achievement test. The average score

for the sample is $M = 72$ with $SS = 2400$. Use the sample results to construct a 99% confidence interval for the population mean.

2. A psychologist studies the change in mood from the follicular phase (prior to ovulation) to the luteal phase (after ovulation) during the menstrual cycle. In a repeated-measures study, a sample of $n = 9$ women take a mood questionnaire during each phase. On average, the participants show an increase in dysphoria (negative moods) of $M_D = 18$ points with $SS = 152$. Determine the 95% confidence interval for population mean change in mood.

3. In families with several children, the first-born tend to be more reserved and serious, whereas the last-born tend to be more outgoing and happy-go-lucky. A psychologist is using a standardized personality inventory to measure the magnitude of this difference. Two samples are used: 8 first-born children and 8 last-born children. Each child is given the personality test. The results are as follows:

First-born	Last-born
$M = 11.4$ $SS = 26$	$M = 13.9$ $SS = 30$

   a. Use these sample statistics to make a point estimate of the population mean difference in personality for first-born versus last-born children.
   b. Make an interval estimate of the population mean difference so that you are 80% confident that the true mean difference is in your interval.

ANSWERS

1. The sample variance is $s^2 = 100$ and the estimated standard error is $s_M = 2$ points. With $df = 24$ and 99% confidence, the t statistic should be between $+2.797$ and -2.797. With 99% confidence, we estimate that the population mean is between 66.406 and 77.594.
2. $s^2 = 19$, $s_{M_D} = 1.45$, $df = 8$, $t = \pm 2.306$; estimate that μ_D is between 14.66 and 21.34.
3. a. For a point estimate, use the sample mean difference: $M_1 - M_2 = 2.5$ points.
   b. Pooled variance = 4, estimated standard error = 1, $df = 14$, $t = \pm 1.345$. The 80% confidence interval is 1.16 to 3.85.

12.3 A FINAL LOOK AT ESTIMATION

FACTORS AFFECTING THE WIDTH OF A CONFIDENCE INTERVAL

Two characteristics of the confidence interval should be noted. First, notice what happens to the width of the interval when you change the level of confidence (the percent confidence). To gain more confidence in your estimate, you must increase the width of the interval. Conversely, to have a smaller interval, you must give up confidence. This is the basic trade-off between precision and confidence that was discussed earlier. In the estimation formula, the percentage of confidence influences the width of the interval by way of the t value. The larger the level of confidence (the percentage), the larger the t value and the larger the interval. This relationship can be seen in Figure 12.8. In the figure, we identified the middle 90% of the t distribution in order to find a 90%

confidence interval. It should be obvious that if we were to increase the confidence level to 95%, it would be necessary to increase the range of t values and thereby increase the width of the interval.

Second, note what would happen to the interval width if you had a different sample size. This time the basic rule is as follows: The bigger the sample (n), the smaller the interval. This relationship is straightforward if you consider the sample size as a measure of the amount of information. A bigger sample gives you more information about the population and allows you to make a more precise estimate (a narrower interval). The sample size controls the magnitude of the standard error in the estimation formula. As the sample size increases, the standard error decreases, and the interval gets smaller.

With t statistics, the sample size has an additional effect on the width of a confidence interval. Remember that the exact shape of the t distribution depends on degrees of freedom. As the sample size gets larger, df also gets larger, and the t values associated with any specific percentage of confidence get smaller. This fact simply enhances the general relationship that the larger the sample, the smaller the confidence interval.

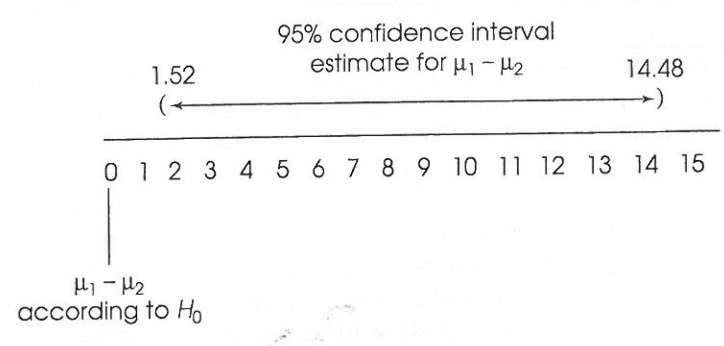
ESTIMATION, EFFECT SIZE, AND HYPOTHESIS TESTS

The process of estimation, especially the estimation of mean differences, provides a relatively simple and direct method for evaluating effect size. For example, the outcome of the study in Example 12.3 indicates that self-hypnosis can reduce allergy symptoms by an estimated 21 points, or about 30%. In this case, the estimation process produces a very clear and understandable indication of how large the treatment effect actually is.

In addition to describing the size of a treatment effect, estimation can be used to get an indication of the "significance" of the effect. Example 12.2 presented an independent-measures research study examining the effect of lecithin on problem-solving performance for rats. Based on the results of this study, it was estimated that the mean difference in performance produced by lecithin was $\mu_1 - \mu_2 = 8$ points. The 95% confidence interval estimated the mean difference to be between 1.52 points and 14.48 points. The confidence interval estimate is shown in Figure 12.9. In addition to the confidence interval for $\mu_1 - \mu_2$, we have marked the spot where the mean difference is equal to zero. You should recognize that a mean difference of zero is exactly what would be predicted by the null hypothesis if we were doing a hypothesis test. You also should realize that a zero difference ($\mu_1 - \mu_2 = 0$) is *outside* the 95% confidence interval. In other words, $\mu_1 - \mu_2 = 0$ is not an acceptable value if we want 95% confidence in our estimate. To conclude that a value of zero is *not acceptable* with 95% confidence is equivalent to concluding that a value of zero is *rejected* with 95% confidence. This conclusion is equivalent to rejecting H_0 with $\alpha = .05$. On the other hand, if a mean difference of zero was included within the 95% confidence interval, then we would have to conclude that $\mu_1 - \mu_2 = 0$ is an acceptable value, which is the same as failing to reject H_0.

FIGURE 12.9

The 95% confidence interval for the population mean difference ($\mu_1 - \mu_2$) from Example 12.2. Note that $\mu_1 - \mu_2 = 0$ is excluded from the confidence interval, indicating that a zero difference is not an acceptable value (H_0 would be rejected in a hypothesis test).



LEARNING CHECK

1. If all other factors are held constant, an 80% confidence interval will be wider than a 90% confidence interval. (True or false?)
2. If all other factors are held constant, a confidence interval computed from a sample of $n = 25$ will be wider than a confidence interval from a sample of $n = 100$. (True or false?)
3. A 99% confidence interval for a population mean difference (μ_D) extends from -1.50 to $+3.50$. If a repeated-measures hypothesis test with two tails and $\alpha = .01$ were conducted using the same data, the decision would be to fail to reject the null hypothesis. (True or false?)

ANSWERS

1. False. Greater confidence requires a wider interval.
2. True. The smaller sample will produce a wider interval.
3. True. The value $\mu_D = 0$ is included within the 99% confidence interval, which means that it is an acceptable value with $\alpha = .01$ and would not be rejected.

SUMMARY

1. Estimation is a procedure that uses sample data to obtain an estimate of a population mean or mean difference. The estimate can be either a point estimate (single value) or an interval estimate (range of values). Point estimates have the advantage of precision, but they do not give much confidence. Interval estimates provide confidence, but you lose precision as the interval grows wider.
2. Estimation and hypothesis testing are similar processes: Both use sample data to answer questions about populations. However, these two procedures are designed to answer different questions. Hypothesis testing will tell you whether or not a treatment effect exists (yes or no). Estimation will tell you how much treatment effect there is.
3. The estimation process begins by solving the t-statistic equation for the unknown population mean (or mean difference).

$$\begin{array}{c}\text{population mean}\\ \text{(or mean difference)}\end{array} = \begin{array}{c}\text{sample mean}\\ \text{(or mean difference)}\end{array} \pm t(\text{estimated standard error})$$

   Except for the value of t, the numbers on the right-hand side of the equation are all obtained from the sample data. By using an estimated value for t, you can then compute an estimated value for the population mean (or mean difference). For a point estimate, use $t = 0$. For an interval estimate, first select a level of confidence and then look up the corresponding range of t values from the t-distribution table. For example, for 90% confidence, use the range of t values that determine the middle 90% of the distribution.
4. For a single-sample study, the mean from one sample is used to estimate the mean for the corresponding population.

$$\mu = M \pm ts_M$$

   For an independent-measures study, the means from two separate samples are used to estimate the mean difference between two populations.

$$(\mu_1 - \mu_2) = (M_1 - M_2) \pm ts_{(M_1 - M_2)}$$

   For a repeated-measures study, the mean from a sample of difference scores (D values) is used to estimate the mean difference for the general population.

$$\mu_D = M_D \pm ts_{M_D}$$

5. The width of a confidence interval is an indication of its precision: A narrow interval is more precise than a wide interval. The interval width is influenced by the sample size and the level of confidence.
   a. As sample size (n) gets larger, the interval width gets smaller (greater precision).
   b. As the percentage confidence increases, the interval width gets larger (less precision).