CHAPTER 8

The Dependent-Sample *t*-Test

Introduction

In the last chapter, we saw that when we have one independent variable with two levels that are independent of one another, we used the independent-sample *t*-test. What happens, however, when the levels of the independent variable represent two different measurements of the same thing? For example, imagine comparing pretest and posttest scores for one group of students. If we did, we would have two measurements of the same group. The independent variable would be "Student Test Scores" and there would be two levels, "Pretest Scores" and "Posttest Scores." In this case, a given student's posttest score would be directly related to his or her pretest score. Because of that, we would say the levels are dependent upon, or influence, one another.

As another example, suppose we are interested in determining if a particular drug has an effect on blood pressure. We would measure a person's blood pressure prior to the study, have the patient take medicine for a period of time, and then measure the blood pressure again. In this case, for a given set of measurements, the "Before Blood Pressure" and the "After Blood Pressure" are the two levels of the independent variable "Blood Pressure." Again, these two levels are related to one another, so we would use the dependent-sample *t*-test (sometimes called the *paired-samples t-test*) to check for significant differences.

That's Great, but How Do We Test Our Hypotheses?

Just like in the prior chapter, here we have an independent variable with two levels and one dependent variable where we have collected quantitative data. Given that, you might be asking, "Why not just use the independent-sample *t*-test?" While that is a logical question, unfortunately, we cannot do that. As we'll soon see, the relationship between the levels of the independent variable creates a problem if we try to do it that way.

Independence versus Dependence

When we talked about the independent-sample *t*-test, the key word was "independent"; scores from one group did not influence scores in the second group. That is not the case when we have two related groups; let's use an example to help us understand what we are getting at.

Using the idea of pretests and posttests we alluded to earlier, let's use the data in Table 8.1 to help us better understand where we are going with this idea.

TABLE 8.1. Student Test Data

Student	Pretest	Posttest
1	52	89
2	48	77
3	51	81
4	45	69
5	50	80
6	60	90
Average	51	81

In this case, the key to understanding what is meant by dependence is based on two things. First, obviously each individual pretest score was made by one given student. Given that, when a student takes the posttest, his score will be "related" to his pretest score. Second, the fact that two sets of scores were collected from one set of students makes them "dependent" on each other; the posttest score should generally be higher than the pretest score for a given student. This might sound somewhat confusing, but we will see more examples that make this concept very clear. The bottom line, for now, is that we have two levels: student pretest scores and student posttest scores.

Computing the t Value for a Dependent-Sample t-Test

Because of the relationship between the two levels of the independent variable, we cannot compute the *t* value using the same formula as the independent-sample *t*-test; if we did, we would dramatically increase the probability of making a Type I error (i.e., rejecting the null hypothesis when we should not). In order to avoid that, we have a different formula to compute the *t* value:

$$t = \frac{\overline{D}}{\sqrt{\left(\dfrac{\sum D^2 - \dfrac{(\sum D)^2}{n}}{n(n-1)}\right)}}$$

Testing a One-Tailed "Greater Than" Hypothesis

To help understand this formula, let's look at my wife's job. She's a media specialist and firmly believes that it doesn't matter what children read (obviously within limits); it is getting them to read anything that makes them better readers.

In order to test her theory, she could start by asking her students to tell her how many hours each week they spend reading outside of school. She could then introduce them to the *Harry Potter* series, books that appeal to children at that age level. After a few weeks she would again ask the children how many hours they read per week. Obviously, she hopes the second number is greater than the first. Her research hypothesis would be:

- *The average number of hours spent reading per week will be significantly greater after allowing students to read books that appeal to them.*

In Table 8.2, we have data for five students with the number of hours weekly they read before the new books were introduced, as well as the number of books they read after the new books were introduced. In order to compute the t value, we need to include two extra columns to create Table 8.3. The first new column, labeled D, shows us the difference between the Before and After scores. The second new column, labeled D^2, shows the squared value of D. At the bottom of each of those columns, you can see we have summed these values.

TABLE 8.2. Number of Books Read

Student	Before	After
1	4	6
2	4	6
3	6	6
4	7	9
5	8	12
Average	5.8	7.8

TABLE 8.3. Computing the Difference Values for Books Read

Student	Before	After	D	D^2
1	4	6	+2	4
2	4	6	+2	4
3	6	6	0	0
4	7	9	+2	4
5	8	12	+4	16
Average	5.8	7.8	$\Sigma D = 10$	$\Sigma D^2 = 28$

Looking back at the equation, you can see the first value we need to compute is D, the average of the difference between scores in each dataset. We can see the sum of the differences (i.e., $\sum D^2$) is 10; to compute the average, we need to divide that by

5, the number of values in our dataset. This leaves us with an average of 2; we can put that into our formula before moving on to the next step.

$$t = \frac{2}{\sqrt{\left(\frac{\sum D^2 - \frac{(\sum D)^2}{n}}{n(n-1)}\right)}}$$

We can now deal with the denominator of the equation. Again, let's go through step by step. First, in the third column of our table we have calculated the difference between both values and summed them (i.e., 10). In the rightmost column, we have taken each of the difference values and squared it. Adding these values gives us the sum of the differences squared (i.e., $\sum D^2$). We can insert these values, along with n (i.e., 5) into the equation. We can finish computing the equation using the following steps.

1. $t = \frac{2}{\sqrt{\left(\frac{28 - \frac{100}{5}}{n(n-1)}\right)}}$

2. $t = \frac{2}{\sqrt{\left(\frac{28 - \frac{100}{5}}{n(n-1)}\right)}}$

3. $t = \frac{2}{\sqrt{\left(\frac{28 - 20}{5(5-1)}\right)}}$

4. $t = \frac{2}{\sqrt{\left(\frac{8}{20}\right)}}$

5. $t = \frac{2}{\sqrt{(.4)}}$

6. $t = \frac{2}{.6325}$

7. $t = 3.16$

This leaves us with a computed t value of 3.16. Before we can test the hypothesis, we must determine the critical value of t from the same table we used with the independent-sample *t*-test. This time, however, we will compute our degrees of freedom by subtracting one from the total pairs of data; when we subtract 1 from 5, we are left with 4 degrees of freedom. Using the traditional alpha value of .05, we would refer to our table and find that the critical value of *t* is 2.132.

We can then plot that on our *t* distribution shown in Figure 8.1; remember, we have a one-tailed hypothesis, so the entire *t* value goes on one end. Here our computed value of *t* is greater than our critical value of *t*. Obviously we have rejected the null hypothesis and supported my wife's research hypothesis: children do read more when they are interested in what they are reading.

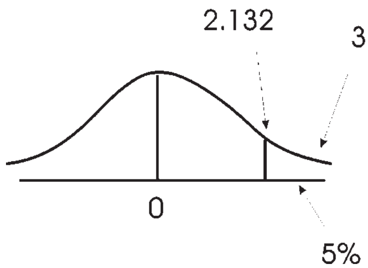

FIGURE 8.1. Comparing the computed and critical values of *t*.

In order to check our work using SPSS, we need to set up our spreadsheet, shown in Figure 8.2, to include two variables, Before and After; we would then include the data for each. Following that, in Figure 8.3, we select Analyze, Compare Means, and Paired Samples T Test. As shown in Figure 8.4, we would then identify the pairs of data we want to compare and click on OK. SPSS would first provide us with Figure 8.5; it verifies what we computed earlier in Table 8.3.

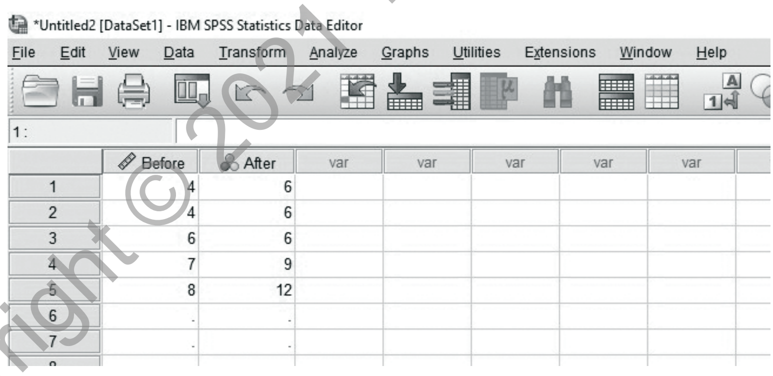

FIGURE 8.2. Before and after data in the Data View spreadsheet.

As we can see in Figure 8.6, the *t* value of 3.16 is exactly what we computed earlier, and our *p* value is less than .05. This means we can support the research hypothesis; kids reading books they enjoy actually did spend significantly more time reading per week than they did when they had no choice in their reading material. What is the bottom line? My wife is happy!

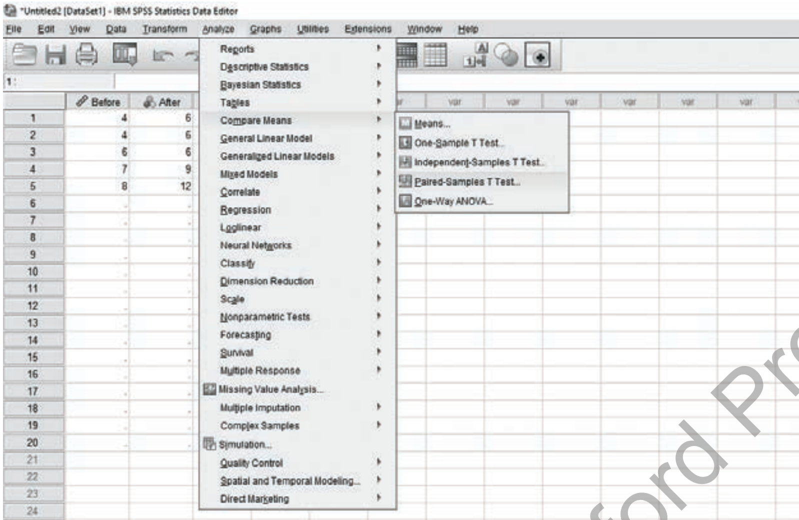

FIGURE 8.3. Using the Compare Means and Paired-Samples T Test command.

The Effect Size for a Dependent-Sample t-Test

Just as was the case with the independent-sample *t*-test, we can compute an effect size for the dependent-sample *t*-test. Interpreting it is the same as in our earlier example, but the formula is a lot easier to compute; all you do is divide the average mean difference by the standard deviation of the difference.

$$d = \frac{\overline{x}_{\text{difference}}}{s_{\text{difference}}}$$

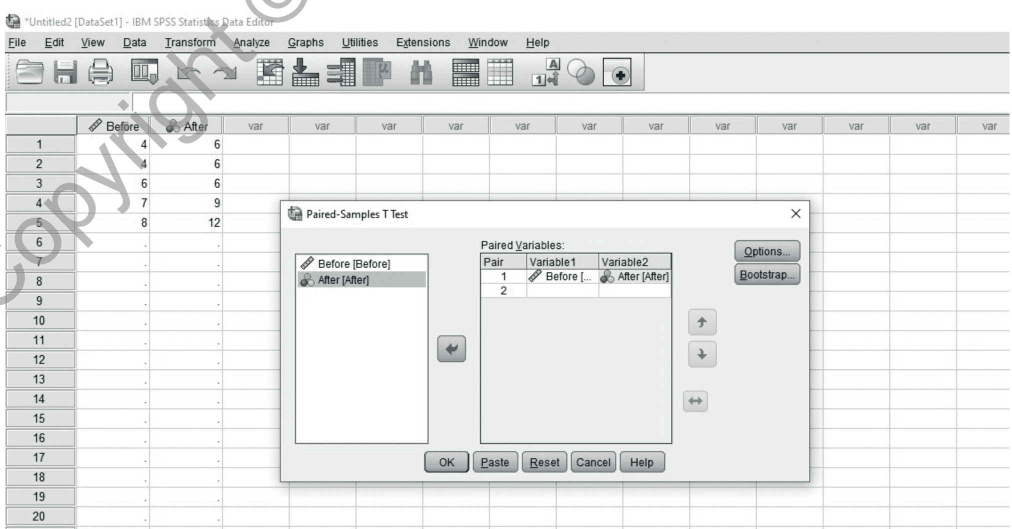

FIGURE 8.4. Creating the pairs of data to be analyzed using the Paired-Samples T Test.

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Before	5.8000	5	1.78885	.80000
	After	7.8000	5	2.68328	1.20000

FIGURE 8.5. Descriptive statistics from the Paired-Samples test.

Paired Samples Test

			Pair 1 After - Before
Paired Differences	Mean		2.00000
	Std. Deviation		1.41421
	Std. Error Mean		.63246
	95% Confidence Interval of the Difference	Lower	.24402
		Upper	3.75598
T			3.162
Df			4
Sig. (2-tailed)			.034

FIGURE 8.6. Inferential statistics from the Paired-Samples test.

Using the values from above and the following two steps, we can compute an effect size of 1.42:

1. $d = \dfrac{2}{1.41}$

2. $d = 1.42$

According to Cohen's standards, this is very large and indicates that the treatment had quite an effect on the dependent variable. Just as we did in the preceding chapter, we will always include the effect size as part of our descriptive statistics.

Testing a One-Tailed "Less Than" Hypothesis

We have seen how well this works with a "greater than" one-tailed hypothesis; now let's look at an example where we are hypothesizing that one average will be significantly less than another. Let's suppose we are working with a track coach at our local high school who is trying to improve his team's times in the 400-meter run. The coach has heard that a diet high in protein leads to more muscle mass and figures this should contribute to his athletes running faster. He decides to test the new diet for 6 weeks and measure the results:

- *A track athlete's time in the 400-meter run will be significantly less after following a high-protein diet for 6 weeks.*

During the diet regimen, the coach collected data shown in Table 8.4; each of the times is measured in seconds.

TABLE 8.4. Race Time Data

Athlete	Before	After
Al	105	98
Brandon	110	105
Charlie	107	100
De'Andre	112	107
Eduardo	101	91
Frank	108	103

Just by looking, it seems that the "after" scores are lower, but let's go ahead and compute our t value. First, since we are looking at a one-tailed "less than" hypothesis, we need to subtract the "Before" from the "After" value. This can be seen in Table 8.5.

TABLE 8.5. Computing the Difference Values for Race Times

Athlete	Before	After	D	D^2
Al	105	98	–7	49
Brandon	110	105	–5	25
Charlie	107	100	–7	49
De'Andre	112	107	–5	25
Eduardo	101	91	–10	100
Frank	108	103	–5	25
Sum			$\Sigma D^2 = 39$	$\Sigma D^2 = 273$

To compute $\overline{x}$, we are again going to divide the sum of the differences (–39) by the number of values in the dataset (6); this gives us an average difference of –6.5. Let's go ahead and enter that into our equation:

$$t = \frac{-6.5}{\sqrt{\left(\frac{\sum D^2 - \frac{(\sum D)^2}{n}}{n(n-1)}\right)}}$$

We continue by inserting the sum of the differences squared (i.e., 273), the sum of the differences (i.e., –39), and n (i.e., 6), into our equation:

$$t = \frac{-6.5}{\sqrt{\left(\frac{273 - \frac{(-39)^2}{6}}{6(6-1)}\right)}}$$

If we simplify that, we come up with the following formulas:

1. $t = \frac{-6.5}{\sqrt{\left(\frac{273 - \frac{1521}{6}}{30}\right)}}$

2. $t = \frac{-6.5}{\sqrt{\left(\frac{19.5}{30}\right)}}$

3. $t = \frac{-6.5}{\sqrt{.065}}$

4. $t = \frac{-6.5}{.806}$

5. $t = -8.06$

We can now help the coach test his hypothesis. First, using our alpha value of .05 and our degrees of freedom of 5, we find that we have a critical *t* value of 2.015. In order to plot this, keep in mind that we have a one-tailed hypothesis, so the entire critical *t* value goes on one end of the distribution as shown in Figure 8.7. In this case, since we have a one-tailed "less than" hypothesis, it needs to go on the left tail of the distribution. Since our computed value of *t* is also negative, it needs to go on the left side of the distribution as well.

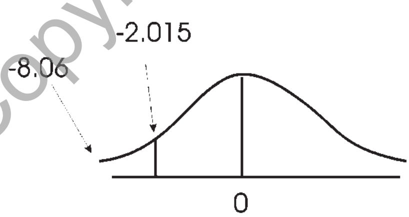

FIGURE 8.7. Using the computed and critical values of *t* to test the hypothesis.

We can clearly see our computed value of *t* of −8.06 is far less than our critical value of *t*; this means we must reject the null hypothesis. It appears that athletes really do run faster if they follow a high-protein diet.

In Figures 8.8 and 8.9, we can confirm what we have done using SPSS. We are further assured of our decision since the *p* value of .000 is less than our alpha value; we can also compute an effect size by dividing the mean difference of our scores by the standard deviation.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Before the Diet	107.1667	6	3.86868	1.57938
	After the Diet	100.6667	6	5.75036	2.34758

FIGURE 8.8. Descriptive statistics from the Paired-Samples test.

Paired Samples Test

			Pair 1 After the Diet - Before the Diet
Paired Differences	Mean		-6.50000
	Std. Deviation		1.97484
	Std. Error Mean		.80623
	95% Confidence Interval of the Difference	Lower	-8.57247
		Upper	-4.42753
T			-8.062
Df			5
Sig. (2-tailed)			.000

FIGURE 8.9. Inferential statistics from the Paired-Samples test.

$$d = \frac{-6.5}{1.97}$$

This yields an effect size of –3.30, but as we did with the independent-sample *t*-test, we must drop the negative sign and use the absolute value. When we do, we see that our effect size means our intervention had a definite effect on our dependent variable.

Testing a Two-Tailed Hypothesis

In order to help us thoroughly understand what we are doing, let's consider a case where we have a two-tailed hypothesis. In this case, we are investigating a drug designed to stop migraine headaches. While we have every indication that it should work, we have discovered that the drug may negatively affect a person's systolic (i.e., the upper number) blood pressure. In some instances, the drug might cause the person's blood pressure to rise; in other cases, it might drop significantly.

Let's create a dataset by taking a patient's blood pressure at the start of our study, administer the migraine drug for 2 weeks and then measure the blood pressure again. After we have completed our study, we wind up with the data in Table 8.6; notice that I've already computed the difference and the squared difference for you.

TABLE 8.6. Computing the Difference Values for Blood Pressure

Patient	Prior	After	D	D^2
1	120	135	15	225
2	117	118	1	1
3	119	131	12	144
4	130	128	–2	4
5	121	121	0	0
6	105	115	10	100
7	128	124	–4	16
8	114	111	–3	9
9	109	117	8	64
10	120	120	0	0
			$\Sigma D = 37$	$\Sigma D^2 = 563$

Just like before, we compute $\overline{D}$ by taking the average of the difference scores (i.e., 37/10 = 3.7) and include it in our formula:

$$t = \frac{3.7}{\sqrt{\left(\dfrac{\sum D^2 - \dfrac{(\sum D)^2}{n}}{n(n-1)}\right)}}$$

We can insert the values for ΣD^2, ΣD, and n into our equation and use the following steps to compute t:

1. $t = \dfrac{3.7}{\sqrt{\left(\dfrac{563 - \dfrac{1369}{10}}{10(10-1)}\right)}}$

2. $t = \dfrac{3.7}{\sqrt{\left(\dfrac{563 - 136.9}{90}\right)}}$

3. $t = \dfrac{3.7}{\sqrt{\left(\dfrac{426.1}{90}\right)}}$

4. $t = \dfrac{3.7}{\sqrt{(4.73)}}$

5. $t = \dfrac{3.7}{2.18}$

Finally, we are left with a t value of 1.70; this matches Figures 8.10 and 8.11 that would be created by SPSS.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Before the Medication	118.3000	10	7.66014	2.42235
	After the Medication	122.0000	10	7.49815	2.37112

FIGURE 8.10. Descriptive statistics from the Paired-Samples test.

Paired Samples Test

			Pair 1 After the Medication - Before the Medication
Paired Differences	Mean		3.70000
	Std. Deviation		6.88073
	Std. Error Mean		2.17588
	95% Confidence Interval of the Difference	Lower	-1.22218
		Upper	8.62218
T			1.700
Df			9
Sig. (2-tailed)			.123

FIGURE 8.11. Inferential statistics from the Paired-Samples test.

You can see from the descriptive statistics that the average blood pressure before the medication is 118.3, with a slightly higher blood pressure of 122 after the medication has been taken. If we compute our effect size of .538 based on these values, we can see that the intervention had a moderate influence on our dependent variable.

We can now plot our computed *t* value and the appropriate critical value. Remember, since we are dealing with a two-tailed hypothesis, it is necessary to divide alpha by 2 and find the critical value for alpha = .025. Using the table, along with 9 degrees of freedom, our critical value of *t* is 2.262. As shown in Figure 8.12, we would then mark that off on both ends of our normal curve.

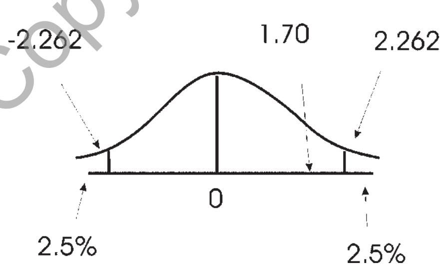

FIGURE 8.12. Using the computed and critical values of *t* to test the hypothesis.

Since our computed value is within the range of the positive and negative critical values, we do not reject the null hypothesis; even though the average "After" blood pressure rose slightly, it wasn't a significant difference. This is verified by the two-tailed *p* value of .123. In other words, after all of this, we have shown that the drug manufacturers have nothing to worry about. Apparently, their new migraine medicine doesn't significantly affect the average systolic blood pressure.

Let's Move Forward and Use Our Six-Step Model

Since we are now proficient in doing the calculations by hand, let's use our six-step approach to work with a middle school math teacher. In this case, the teacher has noticed that many of her students are very apathetic toward math. This, she knows, can contribute to low levels of achievement. The teacher, realizing that something must be done, begins to investigate ways to get her students more interested in their studies.

During her investigation, the teacher finds a new software package that seems to be exactly what she needs. The particular package starts by asking students about their personal interests and activities, and then it creates word problems based on that information. By tailoring the math lesson to each individual student, the teacher hopes to foster more interest in her subject. She hopes, of course, this will lead to less apathy and higher achievement.

This really excites the teacher and a plan is immediately put into effect. The teacher plans to measure student interest in math and then use the software for 10 weeks. At the end of the 10 weeks, she plans on using the same instrument to see if there has been any change in their feelings toward math.

STEP 1

Identify the Problem

There is definitely a straightforward, important problem in the teacher's room, and she has the expertise, resources, and time to investigate it. It would be easy for her to collect numeric data to determine levels of apathy toward math, and there appear to be no problems with the ethicality of her investigation. Here is her problem statement:

- *Math teachers have found many students are apathetic toward their subject matter; this, they feel, may lead to lower achievement. This study will investigate using a software package that tailors the text of word problems to each student's given interests and activities. This, they believe, will lead to lower apathy and higher achievement.*

STEP 2

State a Hypothesis

From the text of the scenario it is easy to develop the hypothesis the teacher will be investigating:

- *Levels of student apathy will be significantly lower after using the new software system for 10 weeks than they were prior to using the software.*

As we can see, we have a directional hypothesis because we have stated that students will have lower levels of apathy after the intervention than they did before it. Again, we stated it in this manner because it reflects the situation that the teacher wants to investigate. The corresponding null hypothesis would read:

- *There will be no significant difference in levels of student apathy before using the software and levels of student apathy after using the software.*

STEP 3

Identify the Independent Variable

We have two groups we can easily identify—students before using the software and the same students after using the software. Given that, we can call the independent variable "Student" and use the two groups as our levels. Since we are measuring the same group at two different points in time (i.e., their pretest scores and their posttest scores), we will use a dependent-sample *t*-test.

STEP 4

Identify and Describe the Dependent Variable

Our dependent variable is student apathy, and we can see the scores below in Table 8.7. The first column, "Student ID Number," is the unique number identified for each student. The second column shows the student's apathy score at the start of the 10 weeks, and the third column shows the same student's apathy score at the end of the 10 weeks.

TABLE 8.7. Student Apathy Scores

Student ID number	Starting apathy score	Ending apathy score
1	62	50
2	67	55
3	67	55
4	60	68
5	69	60
6	66	60
7	70	65
8	72	65
9	68	70

SPSS would produce Figure 8.13, our descriptive statistics. In this case, we used nine pairs of data representing "Starting Apathy" and "Ending Apathy" scores. In the Mean column, we see two values. The Starting Apathy scores show that students had an average score of 66.78 on the questionnaire administered at the start of the 10 weeks. The Ending Apathy scores, collected at the end of the 10 weeks, show an average of 60.89. For each of the mean values, we also see the standard deviation and the standard error of the mean.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Ending Apathy	60.8889	9	6.67915	2.22638
	Starting Apathy	66.7778	9	3.76755	1.25585

FIGURE 8.13. Descriptive statistics from the Paired-Samples test.

STEP 5

Choose the Right Statistical Test

In this case, we have one independent variable with two levels that are related; we say the levels are related because the quantitative data collected for one level is related to the quantitative data in the other level. It is clear that we need to use a dependent-sample *t*-test.

STEP 6

Use Data Analysis Software to Test the Hypothesis

Our software would compute the results shown in Figure 8.14.

We can see a mean difference of –5.889, which would allow us to compute an effect size of .856. While these numbers look promising, do not let them take your attention away from testing our hypothesis using the *p* value. In this case, we have to divide the (Sig. 2-tailed) value of .033 by 2; this gives us a one-tailed *p* value of .0165 so we will reject the null hypothesis; the ending apathy score is significantly lower than apathy at the outset of the study; hopefully this will lead to higher achievement in math!

Paired Samples Test

			Pair 1 Ending Apathy - Starting Apathy
Paired Differences	Mean		-5.88889
	Std. Deviation		6.88194
	Std. Error Mean		2.29398
	95% Confidence Interval of the Difference	Lower	-11.17882
		Upper	-.59896
T			-2.567
Df			8
Sig. (2-tailed)			.033

FIGURE 8.14. Inferential statistics from the Paired-Samples test.

The Case of the Unexcused Students

In this case, let's try to help one of our friends: the principal of a local high school. It seems, with every year that passes, that students miss more and more days due to unexcused absences. The principal, being a believer in the tenets of behavioral psychology, decides to try motivating the students with a reward. The students are told that, if they can significantly decrease the number of times they are absent during the term, they will be rewarded with a party. Let's use our six-step approach to see if we can help him investigate whether or not his plan will work.

STEP 1

Identify the Problem

The principal is faced with somewhat of a two-pronged problem. Students are missing more and more days of school; since absences are directly related to lower achievement, it would be in the best interest of the students to be at school as often as possible. At the same time, many states base school funding partially on student attendance rates; it would be in the principal's best interests (and career aspirations) to get students to be in school as often as possible. This is clearly a problem that can be investigated using inferential statistics.

- *The school in question is experiencing a problem with low student attendance. Since both achievement and school funding are related to student attendance, it is imperative that action be taken to address the problem. The principal will investigate whether an extrinsic reward will help increase attendance.*

STEP 2

State a Hypothesis

Here is the research hypothesis that corresponds to the principal's plan:

- *Students will have significantly fewer absences after the party program than they did before the party program.*

We can see this is a directional hypothesis in that we are suggesting the total number of absences will be less than they were before starting the study. We can state the null hypothesis in the following manner:

- *There will be no significant difference in the number of absences prior to the party program and after the party program is announced.*

STEP 3

Identify the Independent Variable

We could easily identify the independent variable and its levels by noting our interest in one set of students at two distinct points in time: before the promise of the party and after the party plan was announced.

STEP 4

Identify and Describe the Dependent Variable

We know we must collect the number of absences for each student, both in the semester before the plan and at the end of the semester in which the plan was implemented. In Table 8.8, we can see we have a row that contains the number of absences for students A through H before the start of the study and another row that shows the number of absences for the same students at the end of the term. SPSS would calculate the descriptive statistics shown in Figure 8.15.

TABLE 8.8. Absence Data

	A	B	C	D	E	F	G	H
Before the plan	8	7	6	5	6	8	4	3
After the plan	7	6	7	4	6	7	4	3

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Ending Absences	5.50	8	1.604	.567
	Starting Absences	5.88	8	1.808	.639

FIGURE 8.15. Descriptive statistics from the Paired-Samples test.

Things are looking good; the average number of absences after the implementation of the party program is less than the average number before starting the program. We have to be careful, though; the values are very close. We should run the appropriate statistical test and see what the output tells us.

STEP 5

Choose the Right Statistical Test

Here we have the perfect scenario for a dependent-sample *t*-test because there is one independent variable with two related levels and one dependent variable representing quantitative data.

STEP 6

Use Data Analysis Software to Test the Hypothesis

We can use SPSS to easily compute the inferential statistics we would need to test our hypothesis; these statistics are shown in Figure 8.16.

We have a very small standard deviation of the difference

Paired Samples Test

			Pair 1 Ending Absences - Starting Absences
Paired Differences	Mean		-.375
	Std. Deviation		.744
	Std. Error Mean		.263
	95% Confidence Interval of the Difference	Lower	-.997
		Upper	.247
T			-1.426
Df			7
Sig. (2-tailed)			.197

FIGURE 8.16. Inferential statistics from the Paired-Samples test.

(i.e., .744) and an even smaller mean difference score (i.e., –.375); this gives us an effect size of .504; things are not looking good for the principal but let's move forward.

Knowing that we have a one-tailed hypothesis, we would divide our Sig. (2-tailed) value and arrive at a one-tailed p value of .0985; this means we cannot reject our null hypothesis. We could verify this by comparing our computed value of t, –1.416, to our critical t value, 2.365. Although the principal tried, the offer of a party at the end of the term just was not enough of a stimulus to get students to come to school.

The Case of Never Saying Never

Never one to give up, the principal of the high school decides to try again. Thinking back over the original plan, the principal decides the offer of a party just was not enough of a reward to get students to change their behavior. The principal brainstorms with some teachers, and they develop what they feel is a better plan. The literature suggests that the traditional school day, because it starts so early in the morning, is not conducive to getting teenagers to come to school. Knowing that, the principal suggests starting school at noon and running it until 6:00 P.M. This, the principal feels, will allow the students to sleep in and still have time in the evening for all of their activities. Upon announcing the plan to the public, the principal is inundated with criticism as many feel forcing children to stay later in the day will cause even more absenteeism. The principal, vowing at least to test the plan, concedes they may be right but continues to plan for setting up the new schedule.

STEP 1

Identify the Problem

Obviously, the problem is exactly the same as what the principal addressed before. His plan worked to a minor degree, but as we have talked about, that difference may just be caused by error due to the exact students he worked with, the type of year the study was conducted, and so on. Given that, it's time to move forward and try something else. Remember, his students' future, not to mention his job, may hang in the balance! While the problem is the same, the problem statement will change slightly:

- *The school in question is experiencing a problem with low student attendance. Since both achievement and school funding are related to student attendance, it is imperative that action be taken to address the problem. The principal will investigate whether a change in the school's schedule will help increase attendance.*

STEP 2

State a Hypothesis

Based on this story, we can clearly see a hypothesis has formed. Since the principal is not sure if attendance will go up or down, the hypothesis must be stated as two-tailed (Step 1):

- *Students attending high school under the new schedule will have a significantly different number of absences than they did when they were under the old schedule.*

The null hypothesis would read:

- *There will be no significant difference in absences between the new schedule and the old schedule.*

STEP 3

Identify the Independent Variable

Just as in the prior example, the independent variable is the student body. Remember, however, we are measuring them at two different points in time (i.e., before the new schedule and after the new schedule).

STEP 4

Identify and Describe the Dependent Variable

The dependent variable is the number of times the students were absent; let's use the data in Table 8.9 to compute the descriptive statistics shown in Figure 8.17.

TABLE 8.9. Absence Data

	A	B	C	D	E	F	G	H
Old schedule	4	5	6	5	6	7	6	7
New schedule	5	6	7	7	6	8	6	8

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Old Schedule	5.75	8	1.035	.366
	New Schedule	6.63	8	1.061	.375

FIGURE 8.17. Descriptive statistics from the Paired-Samples test.

STEP 5

Choose the Right Statistical Test

This scenario calls for a dependent sample *t*-test since the two levels of the independent variable are related, and we use the dependent variable to collect parametric data.

STEP 6

Use Data Analysis Software to Test the Hypothesis

Using the results shown in Figure 8.18, we could compute a fairly large effect size (i.e., 1.37) but that might be the only good news. First, we can clearly see the mean number of absences went up during the trial period. This bad news is compounded when we look at the two-tailed *p* value, .006, and see that it is clearly less

Paired Samples Test

			Pair 1 Old Schedule - New Schedule
Paired Differences	Mean		-.875
	Std. Deviation		.641
	Std. Error Mean		.227
	95% Confidence Interval of the Difference	Lower	-1.411
		Upper	-.339
T			-3.862
Df			7
Sig. (2-tailed)			.006

FIGURE 8.18. Inferential statistics from the Paired-Samples test.

than our alpha value of .05. We have to reject our null hypothesis; there was a significant difference in the number of absences. Unfortunately for the principal, this means the new schedule caused a significant increase in the number of absences!

Just in Case—A Nonparametric Alternative

Just as was the case with the independent-sample *t*-test, there are instances where either the data distribution is not conducive to using parametric statistics or we have collected data that are ordinal (ranked) in nature. In cases where this happens and the levels of the independent variable are dependent on one another, we have to use the nonparametric *Wilcoxon t-test*. Setting up, running, and interpreting the output of the Wilcoxon test is very similar to that of the dependent-sample *t*-test. Again, this is not something that happens often, and, like the Mann–Whitney U test, this test is something you need to keep in the back of your mind for those rare instances.

Summary

The dependent-sample *t*-test, much like its independent counterpart, is easy to understand, both conceptually and from an applied perspective. The key to using both of these inferential techniques is to keep in mind that they can be used only when you have one independent variable with two levels and when the dependent variable measures quantitative data that is fairly normally distributed. Again, the labels “independent” and “dependent” describe the relationship between the two levels of the independent variable that are being compared.

As I said earlier, these two inferential tests are used widely in educational research. What happens, though, when you have more than one independent variable, more than two levels of an independent variable, or even multiple dependent variables? These questions, and more, will be answered in the following chapters.

Do You Understand These Key Words and Phrases?

dependent-sample *t*-test

Wilcoxon *t*-test

Do You Understand These Formulas?

Effect size for a dependent-sample *t*-test:

$$d = \frac{\overline{x}_{\text{difference}}}{s_{\text{difference}}}$$

t score for the dependent-sample t-test:

$$t = \frac{\bar{D}}{\sqrt{\left(\frac{\sum D^2 - \frac{(\sum D)^2}{n}}{n(n-1)}\right)}}$$

Quiz Time!

As usual, before we wind up this chapter, let's take a look at a couple of case studies. Read through these and answer the questions that follow. If you need to check your work, the answers are at the end of the chapter.

The Case of Technology and Achievement

Proponents of technology in the classroom have suggested that supplying children with laptop computers will significantly increase their grades.

1. What is the hypothesis being tested?
2. What is the independent variable and its levels?
3. What is the dependent variable?
4. Based on Figures 8.19 and 8.20, what decision should they make?

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Laptops Used	85.2667	15	3.80726	.98303
	No Laptops	86.5333	15	4.79385	1.23777

FIGURE 8.19. Descriptive statistics from the Paired-Samples test.

Paired Samples Test

			Pair 1 Laptops Used - No Laptops
Paired Differences	Mean		-1.26667
	Std. Deviation		6.08824
	Std. Error Mean		1.57198
	95% Confidence Interval of the Difference	Lower	-4.63822
		Upper	2.10489
T			-.806
Df			14
Sig. (2-tailed)			.434

FIGURE 8.20. Inferential statistics from the Paired-Samples test.

The Case of Worrying about Our Neighbors

Citizens are concerned that the annexation of property adjoining their town will decrease the value of their homes. The city government insists there will be no change.

1. What is the hypothesis being investigated?
2. What is the independent variable and its levels?
3. What is the dependent variable?
4. Based on Figures 8.21 and 8.22, are the citizens' concerns warranted?

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Before Annex	$116,027.73	15	$10,229.535	$2,641.255
	After Annex	$106,555.87	15	$8,521.641	$2,200.278

FIGURE 8.21. Descriptive statistics from the Paired-Samples test.

Paired Samples Test

			Pair 1 After Annex - Before Annex
Paired Differences	Mean		$-9,471.867
	Std. Deviation		$14,517.440
	Std. Error Mean		$3,748.387
	95% Confidence Interval of the Difference	Lower	$-17,511.357
		Upper	$-1,432.377
t			-2.527
df			14
Sig. (2-tailed)			.024

FIGURE 8.22. Inferential statistics from the Paired-Samples test.

The Case of SPAM

In order to attract new customers, a local entrepreneur is advertising an Internet service that guarantees customers will receive fewer unsolicited emails per month than they would using other services. After enrolling for the service, several customers complained that, contrary to the advertisements, they were actually getting more SPAM!

1. What is the hypothesis the entrepreneur is stating?
2. What is the independent variable and its levels?
3. What is the dependent variable?
4. Based on Figures 8.23 and 8.24, should the customers complain?

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	SPAM Before	46.00	16	5.465	1.366
	SPAM After	65.88	16	5.976	1.494

FIGURE 8.23. Descriptive statistics from the Paired-Samples test.

Paired Samples Test

			Pair 1 SPAM After - SPAM Before
Paired Differences	Mean		19.875
	Std. Deviation		9.444
	Std. Error Mean		2.361
	95% Confidence Interval of the Difference	Lower	14.843
		Upper	24.907
T			8.418
Df			15
Sig. (2-tailed)			.000

FIGURE 8.24. Inferential statistics from the Paired-Samples test.

The Case of "We Can't Get No Satisfaction"

The board of directors at a large corporation recently hired a new President, and, after a few weeks, the corporation's administration felt they should survey the management staff to see if they were satisfied with their choice. Based on the results, managers seemed happy, leaving the administration feeling they had found exactly the right person for the job. After three months, however, the complaints starting piling in—something was going seriously wrong, management seemed very upset! Based on that, the members of the management team were asked to complete the same satisfaction survey again. To the board of director's dismay, the management's satisfaction was lower; the question is, however, was it significantly lower?

1. What is the hypothesis the board of directors is stating?
2. What is the independent variable and its levels?
3. What is the dependent variable?
4. Based on Figures 8.25 and 8.26, what should the board of directors do?

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Satisfaction Before	60.1333	15	6.63181	1.71233
	Satisfaction After	42.2667	15	5.57375	1.43914

FIGURE 8.25. Descriptive statistics from the Paired-Samples test.

Paired Samples Test

			Pair 1 Satisfaction Before - Satisfaction After
Paired Differences	Mean		17.86667
	Std. Deviation		7.70776
	Std. Error Mean		1.99013
	95% Confidence Interval of the Difference	Lower	13.59825
		Upper	22.13508
T			8.978
df			14
Sig. (2-tailed)			.000

FIGURE 8.26. Inferential statistics from the Paired-Samples test.

The Case of "Winning at the Lottery"

I've often told friends that playing the lottery is a pass-time for people who don't understand probability. To investigate this idea, let's imagine that I can identify a group of 20 people who will tell me how much they spend on lottery tickets each week. I would then ask them to watch a brief video on probability as it relates to gambling. My thought is that this would lead to them to spending less money the following week.

1. What is the hypothesis I am testing?
2. What is the independent variable and its levels?
3. What is the dependent variable?
4. Based on Figures 8.27 and 8.28, what will I learn?

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Lottery Money Before	$15.80	20	$11.67	$2.61
	Lotter Money After	$12.00	20	$7.67	$1.72

FIGURE 8.27. Descriptive Statistics from the Paired-Samples Test.

Paired Samples Test

			Pair 1 Lottery Money Before – Lottery Money After
Paired Differences	Mean		$3.80
	Std. Deviation		$5.97
	Std. Error Mean		$1.33
	95% Confidence Interval of the Difference	Lower	$1.59825
		Upper	$6.59409
T			2.847
Df			19
Sig. (2-tailed)			.010

FIGURE 8.28. Inferential statistics from the Paired-Samples Test.

IBM, the IBM logo, ibm.com, and SPSS are trademarks or registered trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "IBM Copyright and trademark information" at www.ibm.com/legal/copytrade.shtml.