# CHAPTER 13 Two-Sample T Tests

## Chapter Outline

# 13.1 Testing a Hypothesis for Dependent and Independent Samples

## Learning Objectives

- Identify situations that contain dependent or independent samples.
- Calculate the standard deviation for two independent samples.
- Calculate the test statistic to test hypotheses about dependent data pairs.
- Calculate the test statistic to test hypotheses about independent data pairs.

## Introduction

In the previous lessons we learned about hypothesis testing for a single sample mean (comparing a sample mean to a hypothesized population mean). In this chapter we will apply the principals of hypothesis testing to situations involving two samples.

There are many situations in everyday life where we would perform statistical analysis involving two samples. For example, suppose that we wanted to test a hypothesis about the effect of two medications on curing an illness. Or, we may want to test the difference between the means of males and females on the SAT. In both of these cases, we would analyze both samples and the hypothesis would address the difference between two sample means.

In this lesson, we will identify situations with different types of samples, learn to calculate the test statistic, calculate the estimate for population variance for both samples and calculate the test statistic to test hypotheses about the difference of proportions or means between samples.

## Independent Samples

Let's recall what we assumed when we conducted a hypothesis test on a single sample. We knew that we needed to select a random sample from the population, measure that sample statistic and then make an inference about the population based on that sample.

When we work with two independent samples, our null hypothesis assumes that if the samples are selected at random, the samples will vary only by chance and the difference will not be statistically significant. In short, when we have independent samples we assume that the scores of one sample do not affect the other.

## Dependent Samples

Dependent samples are a bit different. Two samples of data are dependent when each score in one sample is paired with a specific score in the other sample. In short, these types of samples are related to each other. Dependent samples can occur in two scenarios. In one, a group may be measured twice such as in a pretest-posttest situation (scores on a test before and after the lesson). The other scenario is one in which an observation in one sample is matched with an observation in the second sample –such as when researchers measure the attitudes or behaviors of twins.

## Testing Hypotheses with Independent Samples

The set up for the independent sample t-test is very similar to the single sample t-test. Now however, instead of one sample mean, we have two. Let's look at the hypotheses for the t-test:

$H_0 : \mu_1 = \mu_2$

$H_A : \mu_1 \neq \mu_2$

Notice our hypothesis statements have two *population* means, denoting the fact that we will be testing whether the means of two separate populations are equal to one another. An equivalent way of writing the hypotheses is as follows:

$H_0 : \mu_1 - \mu_2 = 0$

$H_A : \mu_1 - \mu_2 \neq 0$

Both methods of writing the hypothesis statements are valid.

Let's see how the new hypothesis statements look in our t-statistic formula:

$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{SE_{(\bar{x}_1 - \bar{x}_2)}}$

Where:

$\bar{x}_1 - \bar{x}_2$ is the difference between the sample means

$\mu_1 - \mu_2$ is the difference between the hypothesized population means

$SE_{(\bar{x}_1 - \bar{x}_2)}$ is the standard error of the difference between the sample means

The **standard error of the difference** between the sample means is calculated by:$SE_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

This standard error is called an "unpooled" standard error, because the standard deviation of each sample is considered.

Finally, just like the single sample t-test, we'll need to know the degrees of freedom –so that we can find the correct critical value. Here's the formula for the degrees of freedom when using an unpooled standard error independent samples t-test:

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2}$$

Nice, right? Don't worry, we *will not* be using this formula. Technology automatically calculates this formula for us. When we solve the upooled independent samples t-test by hand, the conservative approach is to use the lowest n of the two groups minus one:

$df = n_{lowest} - 1$

## The Assumptions of the Independent Samples t-test

Just like the single sample t-test, the independent samples t-test has some assumptions that we must consider in order for the test to be valid:

- A random sample of each population is used.
- The random samples are each made up of independent observations
- Each sample is independent of one another
- The population distribution of each population must be nearly normal, or the size of the sample is large.

Notice the "new" assumption –now that we have two means that we are examining, we need to make sure those two means are independent of one another. What dose that "independence" mean? It means that if an observation is assigned to one group, it cannot also be recorded in the other group. Oftentimes, independent groups are things like: males vs. females, Astros vs. Rangers fans, tall vs. short, etc.

### Example: Independent t-test

The head of the English department is interested in the difference in writing scores between freshman English students who are taught by different teachers. The incoming freshmen are randomly assigned to one of two English teachers and are given a standardized writing test after the first semester. We take a sample of eight students from one class and nine from the other. Is there a difference in achievement on the writing test between the two classes?

Here's the data from the two classes:

**TABLE 13.1:**

| | **Class 1** | **Class 2** |
|---|---|---|
| | 35 | 52 |
| | 51 | 87 |
| | 66 | 76 |
| | 42 | 62 |
| | 37 | 81 |
| | 46 | 71 |
| | 60 | 55 |
| | 55 | 67 |
| | 53 | |
| **Mean** | 49.44 | 68.88 |
| **Standard Deviation** | 10.38 | 12.30 |

- **Hypothesis Step 1**: Clearly state the Null and Alternative Hypothesis.

  We will be testing to see if the mean score of the two classes are equal to one another:

  $H_0 : \mu_1 = \mu_2$

  $H_A : \mu_1 \neq \mu_2$

- **Hypothesis Step 2:** Identify the appropriate significance level and confirm the test assumptions.

  We'll use the standard significance test of 0.05. We were told that students were randomly assigned, and we'll assume that students did not switch classes (for independence), and we'll assume the student score are independent from one another. We'll assume the underlying population of students in each class is nearly normal in the distribution of scores.

- **Hypothesis Step 3**: Analyze the data and generate the test statistic.

  Now we'll use our t-test to get to the analysis.

  First, our standard error of the difference between the sample means:

  $$SE_{(\bar{x}_1-\bar{x}_2)} = \sqrt{\frac{10.38^2}{9} + \frac{12.30^2}{8}}$$

  $$SE_{(\bar{x}_1-\bar{x}_2)} = \sqrt{11.97 + 18.91}$$

  $$SE_{(\bar{x}_1-\bar{x}_2)} = 5.557$$

  Now, we will use that SE in the t-test equation:

  $$t = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{SE_{(\bar{x}_1-\bar{x}_2)}} = \frac{(49.44-68.88)-(0)}{5.557} = -3.498$$

We know that our smallest group has just eight students, so our degrees of freedom is (8-1)=7. The critical value of the t-distribution for 7 degrees of freedom is $\pm$ 2.365.

- **Hypothesis Step 4:** Interpret your results.

Because our calculated t-value is outside the t-critical value (our value falls in the critical region of the t-distribution), we reject our Null hypothesis. We conclude that the populations of students in the two classes significantly differ in their standardized test scores at the end of the semester. As the two classes were randomly assigned, we can plausibly conclude that the difference in the scores was due to the class assignment –which class the students were in (and whatever teaching technique was used).

### Hypotheses with Dependent Samples

As we mentioned earlier, dependent sample are a bit different from the independent samples t-test. Another name for the dependent samples t-test is the paired samples t-test. That name should give you a hint as to how the test is different. In some way, the two variables we will be testing will be paired –or related –to one another. Now, just because we used the word “related” don’t think correlation. This is still a t-test, and we’ll be testing questions about the means of variables.

Let’s see how the dependent samples t-test is different from the indepdent samples. First, the hypothesis statement. In the dependent samples t-test, our hypothesis statement looks like:

$H_0 : \delta = 0$

$H_A : \delta \neq 0$

That symbol is the Greek letter delta –which is the representation of “difference.” So the hypothesis of the dependent samples t-test is that the “difference” between two variables is zero.

Now, that sure sounds a lot like the hypothesis of an independent samples t-test when we test the difference between two population means. The difference is subtle but important. In the independent samples t-test we were testing the difference between two means –we calculated the mean for each group and then compared them. In the dependent samples t-test we are looking at the difference between two variables within a single observation. Let’s put this in context of our previous example of standardized scores.

We were told that the students were tested at the end of the semester –that’s just one time of testing. But what if all the students were tested when they came into the class (sort of a basic knowledge test) and then again at the end of the semester. Assuming the test was the same (or very close) both times the student saw it, then any change in the scores from beginning to end, should represent the knowledge gain over the course of the semester. Because all students have two scores –one at beginning and one at the end of the semester –the dependent samples t-test allows us to take the difference between the two scores for every student. This difference score is only one column of data. What we are really interested in the dependent samples t-test that difference variable

Let’s look at the t-test formula for a better understanding:

$t = \frac{\bar{d} - \delta}{SE_{\bar{d}}}$

Where:

$\bar{d}$ is the average of the difference between paired variables

$SE_{\bar{d}}$ is the standard error of the difference variable

Notice that $\bar{d}$ in the numerator of the formula. That’s the average of new variable of differences. Again, thinking about our example: If there was no knowledge change over the semester, the average of the differences in scores for each student should be zero –that’s why our Null hypothesis statement states that delta (the average difference in scores for the entire population of students) is equal to zero.

The standard error in the formula is the standard error for the variable representing the difference, which is made up of the standard deviation of the differences. This the standard deviation of the differences formula should look a lot like a simple standard deviation:

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

And the formula for the standard error is:

$$SE_{\bar{d}} = \frac{s_d}{\sqrt{n}}$$

Our degrees of freedom are similar to what they were earlier. But this time since the test is only concerned about a new variable of differences for each subject, we can use the (n-1) degrees of freedom formula, where n represents the number of pairs of data. If there were eight students in the class and each was measured twice, we would still have eight difference scores –one for each student.

### The Assumptions of the Dependent Samples t-test

The assumptions of the dependent samples t-test are very much like the single sample t-test assumptions –after all, the test is only concerned with a single variable (the difference).

- The sample of differences (hence the sample of paired observations) is random.
- The paired observations are independent of one another.
- The distribution of population differences must be nearly normal, or the size of the paired observation sample is large.

Let's look at an example of the dependent samples t-test in action to put it all together:

A math teacher wants to determine the effectiveness of her statistics lesson. She gives a simple skills test to nine students before the start of class (a pre-test) and the same skills test to the same students at the end of class (a post-test).

Here's the data (with some calculations):

**TABLE 13.2:**

| **Student** | **Pre-test** | **Post-test** | **Difference** |
|---|---|---|---|
| 1 | 78 | 80 | 2 |
| 2 | 67 | 69 | 2 |
| 3 | 56 | 70 | 14 |
| 4 | 78 | 79 | 1 |
| 5 | 96 | 96 | 0 |
| 6 | 82 | 84 | 2 |
| 7 | 84 | 88 | 4 |
| 8 | 90 | 92 | 2 |
| 9 | 87 | 92 | 5 |
| | | **Mean** | 3.56 |
| | | **s** | 4.19 |

- **Hypothesis Step 1**: Clearly state the Null and Alternative Hypothesis.

  The statistics instructor is interested in the improvement over the semester, for each of her students. Since there are two measures for each student, and those measures are paired, we'll need to use the dependent samples t-test. We'll assume that the normal state of affaires would be that there's no change due to chance (so our delta is equal to zero). The Null and Alternative would be:

$H_0 : \delta = 0$

$H_A : \delta \neq 0$

- **Hypothesis Step 2**: Identify the appropriate significance level and confirm the test assumptions.

  We'll assume the standard significance level of 0.05. We'll assume that the students in her class are random, that they the students are independent of one another, and that the distribution of all possible difference scores in the population are nearly normally distributed.

- **Hypothesis Step 3**: Analyze the data.

  We have the mean and standard deviation for the data –not for each time variable (pre and post), but for the difference between the two variables for each student (the column on the right). This column of differences is what we'll use in the test. First, let's calculate the Standard Error:

  $SE_{\bar{d}} = \frac{s_d}{\sqrt{n}}$

  $SE_{\bar{d}} = \frac{4.19}{\sqrt{9}} = 1.40$

  Now, we'll use that in our dependent samples t-test:

  $t = \frac{\bar{d}-\delta}{SE_{\bar{d}}} t = \frac{3.56-0}{1.40} = 2.54$

  With this test, we have nine students, so we have (9-1) = 8 degrees of freedom. The t-critical value for 8 degrees of freedom is $\pm$ 2.306.

- **Hypothesis Step 4:** Interpret your results.

  As our calculated t-statistic is greater than our t-critical value (our value lies in the critical region), we reject our Null hypothesis and conclude that there was in fact a change in student performance from the Pre-test to the Post-test.

### Lesson Summary

In addition to testing single samples associated with a mean, we can also perform hypothesis tests with two samples. We can test two independent samples (which are samples that do not affect one another) or dependent samples which assume that the samples are related to each other.

When testing a hypothesis about two independent samples, we follow a similar process as when testing one random sample. However, when computing the test statistic, we need to calculate the estimated standard error of the difference between sample means.

We carry out the test on the means of two independent samples in a similar way as the testing of one random sample. However, we use the following formula to calculate the test statistic:

$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{SE.(\bar{x}_1 - \bar{x}_2)}$ with the standard error defined above.

We can also test the likelihood that two dependent samples are related. To calculate the test statistic for two dependent samples, we use the formula:

$$t = \frac{\bar{d} - \delta}{SE_{\bar{d}}}$$

## Review Questions

1. In hypothesis testing, we have scenarios that have both dependent and independent samples. Give an example of an experiment with (1) dependent samples and (2) independent samples.
2. True or False: When we test the difference between the means of males and females on the SAT, we are using independent samples.
3. A study is conducted on the effectiveness of a drug on the hyperactivity of laboratory rats. Two random samples of rats are used for the study and one group is given Drug A and the other group is given Drug B and the number of times that they push a lever is recorded. The following results for this test were calculated:

### TABLE 13.3:

| | **Drug A** | **Drug B** |
|---|---|---|
| $X$ | 75.6 | 72.8 |
| $n$ | 18 | 24 |
| $s^2$ | 12.25 | 10.24 |
| $s$ | 3.5 | 3.2 |

3. a. Does this scenario involve dependent or independent samples? Explain.
   b. What would the hypotheses be for this scenario?
   c. Calculate the estimated standard error for this scenario.
   d. What is the test statistic and at an alpha level of .05 what conclusions would you make about the null hypothesis?
4. A survey is conducted on attitudes towards drinking. A random sample of eight married couples is selected, and the husbands and wives respond to an attitude-toward-drinking scale. The scores are as follows:

### TABLE 13.4:

| **Husbands** | **Wives** |
|---|---|
| 16 | 15 |
| 20 | 18 |
| 10 | 13 |
| 15 | 10 |
| 8 | 12 |
| 19 | 16 |
| 14 | 11 |
| 15 | 12 |

1. (a) What would be the hypotheses for this scenario?
   (b) Calculate the estimated standard deviation for this scenario.
   (c) Compute the standard error of the difference for these samples.
   (d) What is the test statistic and at an alpha level of .05 what conclusions would you make about the null hypothesis?