# Chapter 11 – Inference on Two Samples

## OUTLINE

**11.1** Inference about Two Population Proportions: Independent Samples

**11.1A** Using Randomization Tests for Hypothesis Tests about Two Proportions

**11.2** Inference about Two Population Means: Dependent Samples

**11.2A** Using Bootstrapping to Conduct Inference on Two Dependent Means

**11.3** Inference about Two Population Means: Independent Samples

**11.3A** Using Randomization Tests for Hypothesis Tests about Two Means

**11.4** Putting It Together: Which Procedure Do I Use?

## Putting It Together

In Chapters 9 and 10, we discussed inferences regarding a single population parameter. The inferential methods presented in those chapters will be modified slightly in this chapter so that we can compare two population parameters.

Section 11.1 presents inferential methods for comparing two population proportions. That is, inference when the response variable is qualitative with two possible outcomes (success or failure). The first order of business is to decide whether the data are obtained from an independent or dependent sample—simply put, we determine if the observations in one sample are somehow related to the observations in the other. We then discuss methods for comparing two proportions from independent samples. Methods for comparing two proportions from dependent samples are covered in Section 12.3.

Section 11.2 presents inferential methods used to handle dependent samples when the response variable is quantitative. For example, we might want to know whether the reaction time in an individual’s dominant hand is different from the reaction time in the non-dominant hand.

Section 11.3 presents inferential methods used to handle independent samples when there are two levels of treatment and the response variable is quantitative. For example, we might randomly divide 100 volunteers who have a common cold into two groups. The control group would receive a placebo and the experimental group would receive an experimental drug. The response variable might be time until the cold symptoms go away.

We wrap up the chapter with a Putting It Together section. One of the more difficult aspects of inference is determining which inferential method to use. This section helps develop this skill.

## Section 11.1 Inference about Two Population Proportions: Independent Samples

### Objectives

1. Distinguish between Independent and Dependent Sampling
2. Test Hypotheses Regarding Two Population Proportions from Independent Samples
3. Construct and Interpret Confidence Intervals for the Difference between Two Population Proportions
4. Determine the Sample Size Necessary for Estimating the Difference between Two Population Proportions

#### Objective 1: Distinguish between Independent and Dependent Sampling

Objective 1, Page 1

 *Answer the following after watching the video.*

1. Explain why the scenario involving the acne medications is an example of dependent sampling.
2. Explain why the scenario involving fast-food receipts is an example of dependent sampling.
3. What does it mean to say that a sampling method is independent?
4. What does it mean to say that a sampling method is dependent?

Objective 1, Page 3

**Example 1 *Distinguishing between Independent and Dependent Sampling***

Decide whether the sampling method is independent or dependent. Then determine whether the response variable is qualitative or quantitative.

1. Joliet Junior College decided to implement a course redesign of its developmental math program. Students enrolled in either a traditional lecture format course or a lab-based format course in which lectures and homework were done using video and the course management system MyMathLab. There were 1100 students enrolled in the traditional lecture format and 500 enrolled in the lab-based format. Once the course ended, the researchers determined whether the student passed the course. The goal of the study was to determine whether the proportion of students who passed the lab-based format exceeded that of students who passed the lecture format.
2. Do women tend to select a spouse who has an IQ higher than their own? To answer that question, researchers randomly selected 20 women and their husbands. They measured the IQ of each wife–husband team to determine whether there was a significant difference in IQ.

#### Objective 2: Test Hypotheses Regarding Two Population Proportions from Independent Samples

Objective 2, Page 1

**Sampling Distribution of the Difference between Two Proportions (Independent Sample)**

Suppose a simple random sample of size  is taken from a population where  of the individuals have a specified characteristic, and a simple random sample of size  is independently taken from a different population where  of the individuals have a specified characteristic.

The sampling distribution of , where  and , is approximately normal, with mean  and standard deviation , provided that  and  and each sample size is no more than 5% of the population size.

The standardized version of  is then written as

 ,

which has an approximate standard normal distribution.

Objective 2, Page 2

1. What are the three conditions that are necessary to test the difference between two population proportions using independent samples?

Objective 2, Page 2 (continued)

1. State the five steps for testing a hypothesis regarding the difference between two population proportions using independent samples.

Step 1

Step 2

Step 3 (By Hand)

Step 3 (Using Technology)

Step 4

Step 5

Objective 2, Page 3

**Example 2 *Testing a Hypothesis Regarding Two Population Proportions***

In clinical trials of Nasonex®, 3774 adult and adolescent allergy patients (patients 12 years and older) were randomly divided into two groups. The patients in group 1 (experimental group) received of Nasonex, while the patients in group 2 (control group) received a placebo. Of the 2103 patients in the experimental group, 547 reported headaches as a side effect. Of the 1671 patients in the control group, 368 reported headaches as a side effect. It is known that over 10 million Americans who are 12 years and older are allergy sufferers. Is there significant evidence to conclude that the proportion of Nasonex users who experienced headaches as a side effect is greater than the proportion in the control group at the  level of significance?

Objective 2, Page 4

Looking back at the results of Example 2, we notice that the proportion of individuals taking  of Nasonex who experience headaches is statistically significantly greater than the proportion of individuals 12 years and older taking a placebo who experience headaches. While the difference of 4% is statistically significant, it does not have any practical significance.

#### Objective 3: Construct and Interpret Confidence Intervals for the Difference between Two Population Proportions

Objective 3, Page 1

1. State the three conditions that are necessary to construct a  confidence interval for the difference between two population proportions from independent samples.
2. State the formulas for the lower bound and upper bound associated with a  confidence interval for the difference between two population proportions from independent samples.

Objective 3, Page 2

**Example 3 *Constructing a Confidence Interval for the Difference between Two Population Proportions***

The Gallup organization surveyed 1100 adult Americans on May 6–9, 2002, and conducted an independent survey of 1100 adult Americans on May 3–7, 2017. In both surveys Gallup asked the following: “Right now, do you think the state of moral values in the country as a whole is getting better or getting worse?” On May 3–7, 2017, 846 of the 1100 surveyed responded that the state of moral values is getting worse; on May 6–9, 2002, 737 of the 1100 surveyed responded that the state of moral values is getting worse. Construct and interpret a 90% confidence interval for the difference between the two population proportions, .

#### Objective 4: Determine the Sample Size Necessary for Estimating the Difference between Two Population Proportions

Objective 4, Page 1

1. State the formula for the margin of error, *E*, in constructing a  confidence interval for the difference between two population proportions.

Objective 4, Page 2

1. State the formula for the sample size required to obtain a  confidence interval for the difference between two population proportions with a margin of error, *E*, if prior estimates of  and ,  and , are available.
2. State the formula for the sample size required to obtain a  confidence interval for the difference between two population proportions with a margin of error, *E*, if prior estimates of  and  are not available.

Objective 4, Page 3

**Example 4 *Determining Sample Size***

A nutritionist wants to estimate the difference between the proportion of males and females who consume the USDA’s recommended daily intake of calcium.

1. What sample size should be obtained if she wants the estimate to be within 3 percentage points with 95% confidence, assuming that she uses the results of the USDA’s 1994–1996 Diet and Health Knowledge Survey, according to which 51.1% of males and 75.2% of females consume the USDA’s recommended daily intake of calcium.
2. What sample size should be obtained if she wants the estimate to be within 3 percentage points with 95% confidence, assuming that she does not use any prior estimates?

## Section 11.1A Using Randomization Techniques to Compare Two Proportions

### Objectives

1. Use Randomization to Compare Two Population Proportions

Introduction, Page 1

Recall in a completely randomized design, a group of individuals is randomly assigned to two or more treatment groups, the treatment is imposed on the individuals, and a response variable is measured. The methods of this section apply to completely randomized designs where there are two levels of the treatment (or two distinct groups) and the response variable is qualitative with two possible outcomes. Because the response variable is qualitative with two possible outcomes, we analyze the data using proportions.

#### Objective 1: Use Randomization to Compare Two Population Proportions

Objective 1, Page 1

 *Watch the video to gain an understanding of how randomization is used to compare two population proportions*.

1. What is the response variable in the study? Is it qualitative or quantitative?
2. What is the assumption about the group of students in the two courses?

Objective 1, Page 2

1. List the two sample proportions. What is the difference between the two sample proportions?

Objective 1, Page 3

1. What are the two possibilities regarding the difference between the sample proportions?
2. State the two possibilities using the notation of hypothesis tests.

Objective 1, Page 4

1. What is the difference in pass rate after randomly assigning students to a treatment? Does this suggest that students did better in the MRP course or the traditional course?

Objective 1, Page 5

1. What is the difference in pass rate for each of the two random assignments using StatCrunch’s Urn applet? Did either result in a difference of pass rates as extreme, or more extreme, than the observed result of 0.209?

Objective 1, Page 6

1. After 2000 more random assignments, what proportion of the times did we observe with 19 or more students passing the MRP course? Based on these results, what is the appropriate decision about the null hypothesis?

Objective 1, Page 9

Rather than using the Urn applet to do the random assignment, we can use the Randomization Test for Two Proportions applet in StatCrunch.

Objective 1, Page 10

1. Examine the graph of the outcomes in the random assignments of students. Where is the graph centered? Why should this not be surprising?
2. What is the shape of the distribution of the difference in pass rates?

Objective 1, Page 11

1. Suppose we wanted to determine if there was a difference in the pass rates of the MRP course versus the Traditional course. What would the null and alternative hypotheses be?

Objective 1, Page 12

12) State the five steps for testing hypotheses regarding two proportions using random assignment.

Step 1

Step 2

Step 3

Step 4

Step 5

Objective 1, Page 13

**Example 2 *Testing a Hypothesis Regarding Two Population Proportions***

Zoloft and Trintellix are drugs meant to alleviate symptoms associated with major depressive disorder. In clinical trials of 3066 patients taking Zoloft, 790 reported nausea as a side effect. In clinical trials of 1013 patients taking Trintellix, 216 reported nausea as a side effect. Is there significant evidence to conclude a difference in the proportion of patients who report nausea as a side effect for these two drugs? Source: [www.zoloft.com](http://www.zoloft.com/) and us.[trintellix.com](https://www.trintellixhcp.com/)

## Section 11.2 Inference about Two Population Means: Dependent Samples

### Objectives

1. Test Hypotheses for a Population Mean from Matched-Pairs Data
2. Construct and Interpret Confidence Intervals about a Population Mean Difference of Matched-Pairs Data

#### Objective 1: Test Hypotheses for a Population Mean from Matched-Pairs Data

Objective 1, Page 1

Inference on matched-pairs data is similar to inference regarding a single population mean.

Recall that if the population from which the sample was drawn is normally distributed with no outliers or the sample size is large  we said that



follows Student’s *t*-distribution with  degrees of freedom.

Objective 1, Page 2

 *Watch the video that explains the procedure for hypothesis tests regarding the difference between two dependent means.*

1. What are the four conditions that must be satisfied before testing a hypothesis regarding the difference between two population means using dependent samples?

Objective 1, Page 2 (continued)

1. State the five steps for testing a hypothesis regarding the difference between two population means using dependent samples.

Step 1

Step 2

Step 3 (By Hand)

Step 3 (Using Technology)

Step 4

Step 5

Objective 1, Page 3

1. What tool is used to determine if the differenced data come from a population that is normally distributed?
2. What tool is used to determine if the differenced data contain outliers?

Objective 1, Page 4

**Example 1 *Testing a Hypothesis for Matched-Pairs Data***

Professor Andy Neill measured the time (in seconds) required to catch a falling meter stick for 12 randomly selected students' dominant hand and non-dominant hand. Professor Neill wants to know if the reaction time in an individual's dominant hand is less than the reaction time in his or her non-dominant hand. A coin flip is used to determine whether reaction time is measured using the dominant or non-dominant hand first. Conduct the test at the  level of significance. The data obtained are presented in Table 1. (Data from Professor Andy Neill, Joliet Junior College)

**Table 1**

| Student | Dominant Hand, | Non-dominant Hand, |
| --- | --- | --- |
| 1 | 0.177 | 0.179 |
| 2 | 0.210 | 0.202 |
| 3 | 0.186 | 0.208 |
| 4 | 0.189 | 0.184 |
| 5 | 0.198 | 0.215 |
| 6 | 0.194 | 0.193 |
| 7 | 0.160 | 0.194 |
| 8 | 0.163 | 0.160 |
| 9 | 0.166 | 0.209 |
| 10 | 0.152 | 0.164 |
| 11 | 0.190 | 0.210 |
| 12 | 0.172 | 0.197 |

#### Objective 2: Construct and Interpret Confidence Intervals about a Population Mean Difference of Matched-Pairs Data

Objective 2, Page 1

1. State the formulas for the lower bound and upper bound associated with a  confidence interval for the population mean difference .

Objective 2, Page 2

**Example 2 *Constructing a Confidence Interval for Matched-Pairs Data***

Using the data from Table 1, construct a 95% confidence interval estimate of the mean difference, .

**Table 1**

| Student | Dominant Hand, | Non-dominant Hand, |
| --- | --- | --- |
| 1 | 0.177 | 0.179 |
| 2 | 0.210 | 0.202 |
| 3 | 0.186 | 0.208 |
| 4 | 0.189 | 0.184 |
| 5 | 0.198 | 0.215 |
| 6 | 0.194 | 0.193 |
| 7 | 0.160 | 0.194 |
| 8 | 0.163 | 0.160 |
| 9 | 0.166 | 0.209 |
| 10 | 0.152 | 0.164 |
| 11 | 0.190 | 0.210 |
| 12 | 0.172 | 0.197 |

Data from Professor Andy Neill, Joliet Junior College

## Section 11.2A Using Bootstrapping to Conduct Inference on Two Dependent Means

### Objectives

1. Test Hypotheses about Two Dependent Means Using the Bootstrap Method

Introduction, Page 1

1. How are the pairs selected in a matched-pair experimental design?
2. In the matched-pairs design, how many levels of treatment are there?

#### Objective 1: Test Hypotheses about Two Dependent Means Using the Bootstrap Method

Objective 1, Page 1

Earlier, we presented the Bootstrap method for testing hypotheses about a single mean. We can use this method to test hypotheses on matched-pairs data. To use the Bootstrap, compute the difference in each matched pair. As with Bootstrapping on a single mean, the differenced data must first be adjusted to build the null model so that the data come from a population whose mean difference is that stated in the null hypothesis. Then find bootstrap resamples based on the differenced data.

1. List the three ways the null and alternative hypotheses can be structured when analyzing matcher pairs data.

Objective 1, Page 2

**Example 1 *Using Bootstrapping to Test Hypotheses about Matched-Pairs Data***

Professor Andy Neill measured the time (in seconds) required to catch a falling meter stick for 12 randomly selected students' dominant hand and non-dominant hand. Professor Neill wants to know if the reaction time in an individual's dominant hand is less than the reaction time in his or her non-dominant hand. A coin flip is used to determine whether reaction time is measured using the dominant or non-dominant hand first. The data obtained are presented in Table 1.

**Table 1**

| Student | Dominant Hand, | Non-dominant Hand, |
| --- | --- | --- |
| 1 | 0.177 | 0.179 |
| 2 | 0.210 | 0.202 |
| 3 | 0.186 | 0.208 |
| 4 | 0.189 | 0.184 |
| 5 | 0.198 | 0.215 |
| 6 | 0.194 | 0.193 |
| 7 | 0.160 | 0.194 |
| 8 | 0.163 | 0.160 |
| 9 | 0.166 | 0.209 |
| 10 | 0.152 | 0.164 |
| 11 | 0.190 | 0.210 |
| 12 | 0.172 | 0.197 |

Data from Professor Andy Neill, Joliet Junior College

Objective 1, Page 3

1. List the six steps for using the bootstrap to test hypotheses about matched-pairs data.

Step 1

Step 2

Step 3

Step 4

Step 5

Step 6

## Section 11.3 Inference about Two Population Means: Independent Samples

### Objectives

1. Test Hypotheses Regarding Two Population Means from Independent Samples
2. Construct and Interpret Confidence Intervals about the Difference of Two Independent Means

#### Objective 1: Test Hypotheses Regarding Two Population Means from Independent Samples

Objective 1, Page 1

1. For the study about the new experimental drug, what is the response variable? Is it qualitative or quantitative?
2. If we let  represent the mean time until cold symptoms go away for the individuals taking the drug and  represent the mean time until cold symptoms go away, what are the null and alternative hypotheses?

Objective 1, Page 2

**Sampling Distribution of the Difference of Two Means: Independent Samples with Population Standard Deviations Unknown (Welch's *t*)**

Suppose that a simple random sample of size  is taken from a population with unknown mean  and unknown standard deviation . In addition, a simple random sample of size  is taken from a second population with unknown mean  and unknown standard deviation . If the two populations are normally distributed or the sample sizes are sufficiently large , then



approximately follows Student's *t*-distribution with the smaller of  or  degrees of freedom, where  is the sample mean and  is the sample standard deviation from population 1, and  is the sample mean and  is the sample standard deviation from population 2.

Objective 1, Page 3

1. What are the four conditions that must be satisfied before testing a hypothesis regarding the difference between two population means using independent samples?
2. State the five steps for testing a hypothesis regarding the difference between two population means using independent samples.

Step 1

Step 2

Step 3 (By Hand)

Step 3 (Using Technology)

Step 4

Step 5

Objective 1, Page 4

1. What tool is used to determine if the sample is drawn from a population that is normally distributed?
2. What tool is used to determine if the sample contains outliers?

Objective 1, Page 5

**Example 1 *Testing a Hypothesis Regarding Two Independent Means***

In the Spacelab Life Sciences 2 payload, 14 male rats were sent to space. Upon their return, the red blood cell mass (in milliliters) of the rats was determined. A control group of 14 male rats was held under the same conditions (except for space flight) as the space rats, and their red blood cell mass was also determined when the space rats returned. The project, led by Dr. Paul X. Callahan, resulted in the data listed in Table 2. Does the evidence suggest that the flight animals have a different red blood cell mass from that of the control animals at the  level of significance?

**Table 2**

| Flight | Control |
| --- | --- |
| 8.59 | 8.65 |
| 8.64 | 6.99 |
| 7.43 | 8.40 |
| 7.21 | 9.66 |
| 6.87 | 7.62 |
| 7.89 | 7.44 |
| 9.79 | 8.55 |
| 6.85 | 8.70 |
| 7.00 | 7.33 |
| 8.80 | 8.58 |
| 9.30 | 9.88 |
| 8.03 | 9.94 |
| 6.39 | 7.14 |
| 7.54 | 9.14 |

Data from NASA Life Sciences Data Archive

Objective 1, Page 6

1. Using the smaller of  or  for the degrees of freedom is conservative. State the formula for degrees of freedom that is used by computer software for increased precision.

#### Objective 2: Construct and Interpret Confidence Intervals about the Difference of Two Independent Means

Objective 2, Page 1

1. State the formulas for the lower bound and upper bound associated with a  confidence interval for the difference of two means.

Objective 2, Page 2

**Example 2 *Constructing a Confidence Interval for the Difference of Two Independent Means***

Recently, a device called low-level laser therapy has evolved as a potential solution to hair loss. A randomized, placebo-controlled experiment was conducted in which 65 females were randomly assigned to one of two treatment groups. The 43 subjects in Group 1 were exposed to a 9-beam lasercomb treatment, while the 22 subjects in Group 2 received a placebo treatment for a total of 16 weeks. The response variable in the study was hair density (measured in hair count per square centimeter). For the subjects in Group 1, the mean change in hair density was 20.2 with a standard deviation of 11.2. For the subjects in Group 2, the mean change in hair density was 2.8 with a standard deviation of 16.5. Estimate the mean difference in hair density between Group 1 and Group 2 by constructing a 95% confidence interval. Note: Analysis of the sample data suggest the sample data come from populations that are normally distributed. Source: Jimenez, JJ, et. al. "Efficacy and safety of low-level laser device in the treatment of male and female pattern hair loss: a multicenter, randomized, sham device-controlled, double-blind study." American Journal of Clinical Dermatology Volume 15, Issue 2(2014).

Objective 2, Page 4

Statistical software provides an option for two types of two-sample *t*-tests: one that assumes equal population variances (pooling) and another that does not assume equal population variances. Welch's *t*-statistic does not assume that the population variances are equal.

Because testing the equality of variances is so volatile, we are content to use Welch's *t*. Welch's *t*-test is more conservative than the pooled *t*. The price that must be paid for the conservative approach is that the probability of a Type II error is higher with Welch's *t* than with the pooled *t* when the population variances are equal. However, the two tests typically provide the same conclusion.

## Section 11.3A Using Randomization Techniques to Compare Two Independent Means

### Objectives

1. Use Randomization to Compare Two Population Means

#### Objective 1: Use Randomization to Compare Two Population Means

Objective 1, Page 1

1. For the study, what is the response variable? Is it qualitative or quantitative?

Objective 1, Page 3

1. View the dot plot by gender. Does it suggest that females are spending more time on homework, on average?

Objective 1, Page 3

1. What are the two possible explanations for the sample mean difference of 19.2 minutes?

Objective 1, Page 4

1. If  is the mean time spent on homework by female students and  is the mean amount of time spent on homework by male students, what are the null and alternative hypotheses?

Objective 1, Page 5

1. What is the sample mean difference (females minus males) for the set of randomly assigned data?

Objective 1, Page 6

1. What is the sample mean difference (females minus males) for the set of randomly assigned data from StatCrunch?

Objective 1, Page 7

1. Only 31 out of 5000 random assignments of gender to study time produced a sample mean difference of 19.2 minutes or higher. What is the *P*-value for this hypothesis test?
2. Based on the observed results, what is our decision about the statement in the null hypothesis? What is the conclusion for this test?

Objective 1, Page 8

1. What is the shape of the distribution of randomized differences?
2. Where is the distribution of randomized differences centered? Why should this not be surprising?

Objective 1, Page 9

1. State the five steps for testing hypotheses regarding two independent means using random assignment.

Step 1

Step 2

Step 3

Step 4

Step 5

Objective 1, Page 10

**Example 1 *Testing Hypotheses Regarding Two Independent Means***

In the Spacelab Life Sciences 2 payload, 14 male rats were sent to space. Upon their return, the red blood cell mass (in milliliters) of the rats was determined. A control group of 14 male rats was held under the same conditions (except for space flight) as the space rats, and their red blood cell mass was also determined when the space rats returned. The project, led by Dr. Paul X. Callahan, resulted in the data listed in Table 2. Does the evidence suggest that the flight animals have a different red blood cell mass from that of the control animals at the  level of significance?

**Table 2**

| Flight | Control |
| --- | --- |
| 8.59 | 8.65 |
| 8.64 | 6.99 |
| 7.43 | 8.40 |
| 7.21 | 9.66 |
| 6.87 | 7.62 |
| 7.89 | 7.44 |
| 9.79 | 8.55 |
| 6.85 | 8.70 |
| 7.00 | 7.33 |
| 8.80 | 8.58 |
| 9.30 | 9.88 |
| 8.03 | 9.94 |
| 6.39 | 7.14 |
| 7.54 | 9.14 |

Data from NASA Life Sciences Data Archive

## Section 11.4 Putting It Together: Which Procedure Do I Use?

### Objective

1. Determine the Appropriate Hypothesis Test to Perform

#### Objective 1: Determine the Appropriate Hypothesis Test to Perform

Objective 1, Page 1

 *Answer the following after watching the video.*

1. How can you determine whether a scenario calls for inference using proportions or inference using means?
2. Explain how to determine whether two samples are dependent or independent?

Objective 1, Page 2

**Flowchart for Determining Which Type of Test to Perform**

Flowchart describing the determining which type of test to perform. What parameter is addressed in the hypothesis? If it is proportion, p, then consider if it is dependent or independent sampling? If it is dependent sampling, then provided the samples are obtained randomly and the total number of observations where the outcomes differ is at least 10, use the normal distribution with z subscript 0 equals fraction numerator absolute value of f subscript 12 minus f subscript 21 end absolute value minus 1 divided by denominator square root of f subscript 12 plus f subscript 21 end root end fraction. If it is independent sampling, then provided np hat open parenthesis 1 minus p hat close parenthesis greater than or equals to 10 for each sample and the sample size is no more than 5 percent of the population size, use the normal distribution with z subscript 0 equals p hat subscript 1 minus p hat subscript 2 divided by denominator square root of p hat open parenthesis 1 minus p hat end root square root of 1 divided by n subscript 1 plus 1 divided by n subscript 2 end root where p hat equals fraction numerator x subscript 1 plus x subscript 2 divided by denominator n subscript 1 plus n subscript 2 end fraction. If the parameter addressed is mean, mu, then consider if it is dependent or independent sampling? If it is independent sampling, then provided each sample size is greater than 30 or each population is normally distributed, use Student's t-distribution: t subscript 0 equals fraction numerator open parenthesis x bar subscript 1 minus x bar subscript 2 close parenthesis minus open parenthesis mu subscript 1 minus mu subscript 2 close parenthesis divided by denominator square root of s subscript 1 squared divided by n subscript 1 plus s subscript 2 squared divided by n subscript 2 end root end fraction. If it is dependent sampling, then provided each sample size is greater than 30 or the differences come from a population that is normally distributed, use Student's t-distribution with n minus 1 degrees of freedom with t subscript 0 equals fraction numerator d bar minus mu subscript d divided by s subscript d divided by square root of n end fraction.