LESSON: TWO-SAMPLE T-TESTS FOR POPULATION MEANS
This lesson includes an overview of the subject, instructor notes, and example exercises using Minitab.

Two-Sample *t*-Test for Population Means

Lesson Overview

We will construct confidence intervals and conduct hypothesis tests to compare means from two ***independent*** populations. Two sets of sample data are ***independent*** when the observations from one sample have no effect on the other.

Prerequisites

This lesson requires knowledge from the ***Normal Distribution***, ***Sampling Distribution of*** $\bar{X}$, ***Population Mean Inference for Normally Distributed Data and Unknown Population Standard Deviation***, and ***Paired t-Test for Population Means*** lessons. The **paired *t*-test** or ***dependent* sample *t*-test** is used to determine whether the **mean difference** between two sets of observations is some target value. The **two-sample *t*-test** or **two *independent* samples *t*-test** is used to the determine whether the **difference in means** between two sets of independent observations is some target value.

Learning Targets

This lesson teaches students to:

- Recognize independent samples versus paired (dependent) samples.
- Use Minitab to construct a confidence interval and conduct a hypothesis test for the difference in two population means.

Time Required

It will take the instructor at least 60 minutes to do this lesson. One of the most difficult parts of this lesson is carried over from the paired *t*-test, which is determining the difference between independent and paired samples. Once students can recognize two independent groups of data, then they can proceed to construct a confidence interval or perform hypothesis testing in a manner similar to those in the prerequisite inference lessons. Students will have the opportunity to review testing data to determine if it is normally distributed.

Materials Required

- Minitab desktop (20 or higher) or Minitab web app
- Minitab worksheet of sample data, entitled ***TwoSampleTTest_Lesson.mtw***

Assessment

The activity sheet contains exercises for students to assess their understanding of the learning targets for this lesson.

Reference

https://www.ck12.org/statistics/dependent-and-independent-samples/lesson/Testing-a-Hypothesis-for-Dependent-and-Independent-Samples-ADV-PST/

Instructor Notes with Examples

Independent Versus Paired Sample Data

Two sets of sample data are considered to be ***independent*** when the observations in one set have no effect or bearing on the observations from the other set. Unlike paired or dependent observations, entry i in data set 1 is not linked or connected to entry i in data set 2. In fact, entry i may not even exist in both data sets since their sample sizes may be different.

We can sometimes distinguish between independent or paired data according to the way in which the data was collected. ***Independent*** implies that the two samples must be from two completely different populations. Here are some examples of independent samples.

- The math department is interested in the difference in calculus readiness scores for students taught by two different professors. We take a sample of twelve students from one professor's class and nine from the other's class. Each student only studies with members from his own class.
- We want to determine if freshmen and juniors at a certain college differ in their knowledge of current events. We randomly choose n_F = 25 freshmen and n_J = 36 juniors and have them take a current events exam.
- A sociologist is interested in determining if the life expectancy of people in Japan is greater than the life expectancy of people in America. In a sample of 42 Japanese citizens, the mean life expectancy was 84.25 years with a standard deviation of 5.3 years.

In the sample of 53 American citizens, the mean life expectancy was 78.15 years with a standard deviation of 8.1 years.

- Students going to a certain graduate school must take the GRE general exam. Do women have a higher average score than men?
- Each student in your class measures the length of their right ring finger. Is the average length of a male's finger an additional 2 inches longer than a female's finger?

Again, when we have two sets of data, ask students to determine if entry 1 in data set 1 is linked uniquely to entry 1 in data set 2. If it isn't, then we have independent data. Below are some examples for practice identifying independent versus paired data sets.

Example 1. For each of the following scenarios, determine if the two samples are independent or paired.

(a) The weights of marathon runners were taken before and after a run to test if runners lose dangerous levels of fluid.

Solution: Paired. Each runner is weighed before and after a run.

(b) In a small town, a survey is conducted on attitudes towards homeschooling. A random sample of eight parents and a random sample of eight teachers are selected and they answer questions according to an attitude-toward-homeschooling scale.

Solution: Independent. Notice that we have two random samples of 8 participants. Since there is no obvious connection between parents and teachers, these are independent.

(c) A math teacher wants to determine the effectiveness of her "factor by foiling" lesson and gives a pre-exam and a post-exam to 9 students in her class. She is testing to determine if there is mean difference in the scores on the exams.

Solution: Paired. Each student has a pre-exam score and a post-exam score.

(d) The number of years of music education from one random sample of 38 high school seniors from City A and the number of years of music education from a second random sample of 32 high school seniors from City B are recorded. Is there a statistically significant difference between the average years of music education of students in City A and City B?

Solution: Independent. The two random samples of high school seniors have no obvious connections. Plus, the sample sizes being different is an indication that the two samples are not paired.

(e) A college's admissions office tests the difference between the means of males and females on the SAT. They take random samples of a dozen males and a dozen females.

Solution: Independent. The two random samples of students have no obvious connections.

Confidence Intervals

If we have independent samples, we can use the **two-sample *t*-test**, also called the **independent samples *t*-test**, to construct confidence intervals for the true difference in means. The two-sample *t*-test takes sample data from two independent groups to determine the ***difference in means,* $\mu_1 - \mu_2$**. An estimate for the true difference in means is simply $\bar{x}_1 - \bar{x}_2$. The sample variance, $\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}$, is constructed under the assumption that the two groups are independent. It should be noted that there is a different formula for the sample variance if we assume that the two *population* standard deviations are equal. Since we don't even know what the population standard deviations are, the "default" formula for the sample variances is the one provided in this paragraph that assumes that the population variances are unequal.

As long as each population is **normally distributed**, the confidence interval formula for the difference in means is provided in the following text box.

Confidence Interval for Difference in Two Means

A **two-tailed 100(1 – α)% confidence interval for $\mu_1 - \mu_2$** for normally distributed data with σ_1 and σ_2 unknown is:

$$\left(\bar{x}_1 - \bar{x}_2 - t_{\alpha/2,df} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}, \bar{x}_1 - \bar{x}_2 + t_{\alpha/2,df} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right)$$

$\bar{x}_1$ = sample mean of population one, $\bar{x}_2$ = sample mean of population two,
${s_1}^2$ = sample variance of population one, ${s_2}^2$ = sample variance of population two,
n_1 = sample size of population one, n_2 = sample size of population two
$t_{\alpha/2,df}$ = *t*-value corresponding to half of the desired α level with *df* degrees of freedom

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

Example 2 (Clinical Trials). A pilot study was conducted to compare a new post-surgical treatment with the standard of care. Seven subjects were randomly assigned to the new treatment, while seven others were randomly assigned to the standard of care. The recovery times, in days, are given below:

New Treatment	Standard of Care
12	18
13	23
15	24
19	30
20	32
21	35
24	39

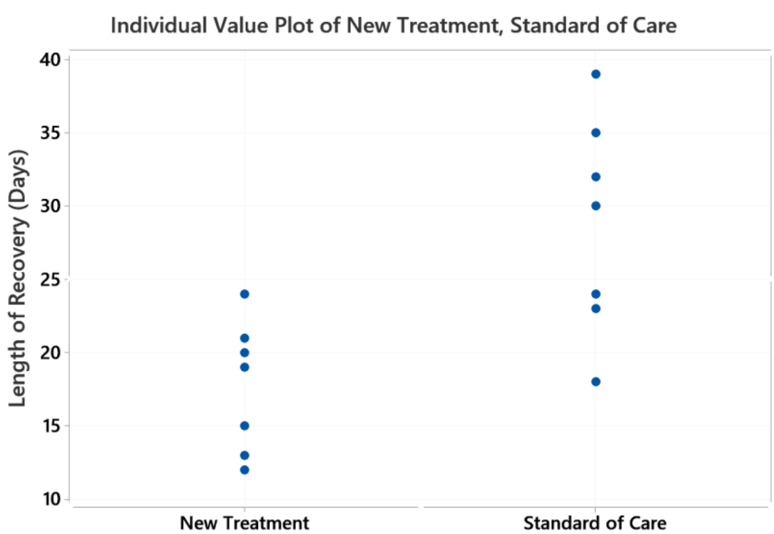
We are interested in the following question: "Is there a difference, on average, between the recovery times (in days) for the two treatments?"

(a) Is the data paired or independent?

Solution: Independent. The 14 individuals were randomly assigned to the groups. There are no obvious connections between the subjects in the two groups.

(b) Construct an appropriate graphic to address the question of interest.

Solution: Since we are comparing the means from two groups, an appropriate graphic would be a side-by-side boxplot or individual value plot. Given the small sample sizes, an individual value plot is probably better. Use **Graph > Individual Value Plot > Multiple Y's > Simple** in Minitab desktop or **Graph > Individual Value Plot > Multiple Y Variables > With Categorical Variables** in the Minitab web app**.** It does appear that the mean of the *New Treatment* method is less than the mean of the *Standard Care* treatment.



(c) Is each group normally distributed?

Solution: To save time, we can construct normality plots for the two groups together in one graphic.

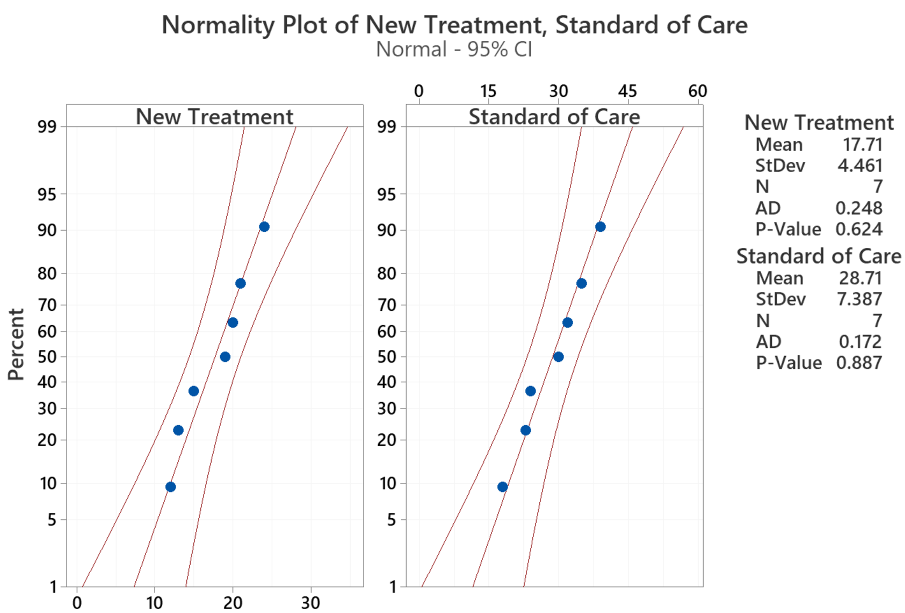
Minitab desktop (20 or higher)

1 Choose **Graph > Probability Plot > Multiple** and click **OK**.
2 In **Graph variables**, enter *'New Treatment' 'Standard of Care'*.
3 Select **Multiple Graphs** and choose **In separate panels of the same graph**.
4 Click **OK** in each dialog box.

Minitab web app

1 Choose **Graph > Probability Plot.**
2 Under **Multiple Y Variables**, choose **Displayed Separately**.
3 In **Y-variables**, enter *'New Treatment' 'Standard of Care'*.
4 Click **OK** in each dialog box.

Since the *p*-values for each normality test is greater than 0.05, the evidence supports both data sets are from normal distributions.



(d) Construct a 95% two-sided confidence interval for the difference in means to answer the original question. Specifically, is there evidence that the average recovery time differs between the two therapies at the $\alpha = 0.05$ significance level?

Solution: We can construct the confidence interval using the formula provided in the text box on page 4. Instead of going through the tedious computations, we'll just go straight to Minitab's two-sample *t*-test.

5 Choose **Stat > Basic Statistics > 2-Sample t**.
6 From the drop-down list, select **Each sample is in its own column**.
7 For **Sample 1**, select the column named *New Treatment*.
8 For **Sample 2**, select the column named *Standard of Care*.
9 Click **OK**.

Minitab provides the following information, including the 95% confidence interval:

Two-Sample T-Test and CI: New Treatment, Standard of Care

Method

μ_1: population mean of New Treatment
μ_2: population mean of Standard of Care
Difference: $\mu_1 - \mu_2$
Equal variances are not assumed for this analysis.

Descriptive Statistics

Sample	N	Mean	StDev	SE Mean
New Treatment	7	17.71	4.46	1.7
Standard of Care	7	28.71	7.39	2.8

Estimation for Difference

Difference	95% CI for Difference
-11.00	**(-18.38, -3.62)**

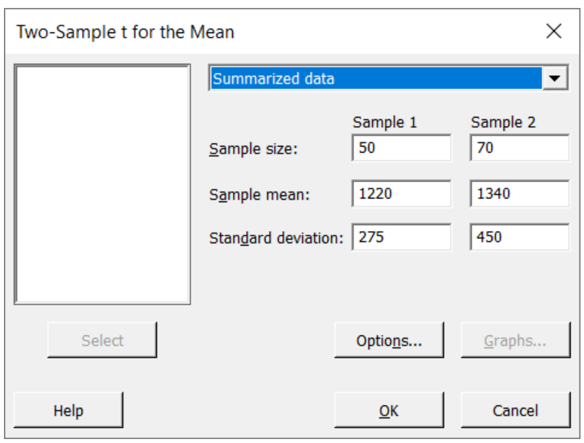
We are 95% certain that the interval (-18.38, -3.62) contains the true difference in the means $\mu_{NT} - \mu_{SC}$, where NT represents *New Treatment* and SC represents *Standard of Care*. Since the interval does not contain 0, this indicates that the difference in means is not 0. We can conclude at $\alpha = 0.05$ that the two therapies are different. Because of the order that we subtracted the means in building the confidence interval, the 95% indicates that the mean of the *New Treatment* is less than the mean of the *Standard of Care*.

Example 3 (Textbook Prices). Our campus bookstore asked a random sample of sophomores and juniors how much they spent on textbooks for the entire academic year. The bookstore believes the two groups of students spend the same amount. While virtually every college textbook today is available digitally, hardback textbooks required for a course can be costly. Fifty sophomores had a mean expenditure of $1220 with a sample standard deviation of $275 and the 70 juniors sampled had a mean expenditure of $1340 with a sample standard deviation of $450. Based on this information is the bookstore's belief accurate?

Solution: Unlike the previous example, we are given summary statistics instead of actual data. The textbook expenditures for the samples of sophomores and juniors are not paired. There is no connection between the two groups; plus, the sample sizes for the two groups are different. Since there are n_S = 50 sophomores and n_J = 70 juniors, the sample sizes are large enough to assume that the distribution of the sample means for both groups are normally distributed.

Since we are not given α, we'll assign it to be 0.05. Since we want to determine if the textbook expenditures are the same, we'll construct a two-sided 95% confidence interval for the difference in the mean expenditures in Minitab.

Since we do not have data in this example, we'll use the 2-Sample *t*'s drop-down box to select **Summarized Data**.



When we click **Options** in the dialogue box, enter *95* as the **Confidence level**. At the top of the dialogue box, Minitab indicates that the *Difference* is computed as (*sample 1 mean*) – (*sample 2 mean*). Do not click the button for **Assume equal variances**. Minitab yields the following information for the confidence interval:

Method

μ_1: population mean of Sample 1
μ_2: population mean of Sample 2
Difference: $\mu_1 - \mu_2$
Equal variances are not assumed for this analysis.

Descriptive Statistics

Sample	N	Mean	StDev	SE Mean

Sample 1	50	1220	275	39
Sample 2	70	1340	450	54

Estimation for Difference

Difference	95% CI for Difference
-120.0	(-251.5, 11.5)

Since the 95% two-sided confidence interval for the difference in the mean textbook expenditures for sophomores and the mean textbook expenditures for juniors contains 0, then our evidence suggests that the expenditures may be the same. We do not have enough evidence to reject that the mean expenditures are the same at α = 0.05. Can we say that the mean amount that juniors pay is $20 more than the mean amount that sophomores pay? No! Our 95% confidence interval's upper bound is $11.50.

Hypothesis Testing

When two populations are independent, a two-sample *t*-test is used to perform hypothesis testing on the difference in population means. For this hypothesis test, we typically assume that there is no difference between the means of the two independent samples. That is, our null hypothesis is:

$$H_0: \mu_1 = \mu_2 \text{ or } H_0: \mu_1 - \mu_2 = 0$$

We can also test if the difference between the means is some constant *c*. In introductory statistics classes, we are most often concerned with equality of the two means. In this lesson, I'll use examples testing whether there is or is not a difference in the means. I'll include an example of the difference being a constant in the assessment materials.

The estimate for the difference in the population means $\mu_1 - \mu_2$ is the difference in the sample means $\bar{x}_1 - \bar{x}_2$. Assuming unknown and unequal population variations, the standard deviation for the difference in means is $\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$. The simplicity in the standard deviation formula is due to the independence of the two samples. As long as each population is normally distributed, we can follow the same basic standardized test statistic formula that we used in the 1-sample *t*-test lesson to determine t_0:

$$t_0 = \frac{\text{(sample mean)} - \text{(hypothesized mean)}}{\text{(standard error of mean)}} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

If our null hypothesis is that the population means are equal, then this formula reduces down to:

$$t_0 = \frac{(\bar{x}_1 - \bar{x}_2) - (0)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

In order to determine the *p*-value for this standardized test statistic, we need to know the degrees of freedom (*df*) associated with t_0 for the difference in two independent population means. The degrees of freedom formula used by Minitab is

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

Hypothesis test for the difference in two independent population means: $\mu_1 - \mu_2$

The null and alternative hypotheses for testing whether the difference in population means is 0, or if the population means are equal, are:

Null hypothesis:	$H_0: \mu_1 - \mu_2 = 0$	$H_0: \mu_1 = \mu_2$
Alternative hypothesis:	$H_a: \mu_1 - \mu_2 \neq 0$	$H_a: \mu_1 \neq \mu_2$

Note: The alternative hypothesis can also be > or <. If we want to determine if the difference in population means is some constant *c*, H_0 becomes $\mu_1 - \mu_2 = c$.

Assuming that both populations are normally distributed and that the null hypothesis $H_0: \mu_1 - \mu_2 = 0$ is true, the standardized test statistic is

$$t_0 = \frac{(\bar{x}_1 - \bar{x}_2) - (0)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t_{df}, \text{where}$$

$\bar{x}_1$ = sample mean of population one, $\bar{x}_2$ = sample mean of population two,
${s_1}^2$ = sample variance of population one, ${s_2}^2$ = sample variance of population two,
n_1 = sample size of population one, n_2 = sample size of population two

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

Example 4 (Years of School Music Education). The number of years of music education from one random sample of 38 high school seniors from City A and the number of years of music education from a second random sample of 32 high school seniors from City B are recorded. The average years of music education for the sample from City A is 5.2 years with a standard deviation of 1.7 years. The average years of music education for the sample from City B is 3.5 years with a standard deviation of 2.1 years. Is there a statistically significant difference between the average years of music education in City A and City B?

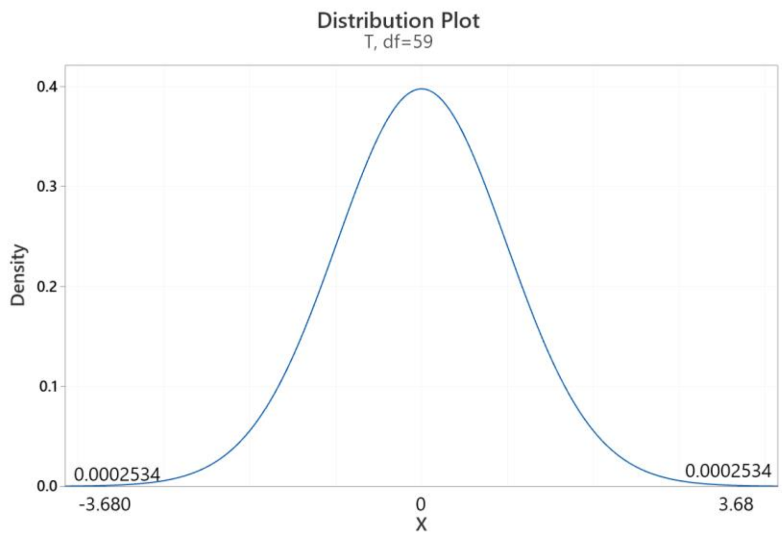
Solution: The two groups of data are independent. There is no obvious connection between students at schools from City A and City B. Since the sample sizes of both populations are larger than 30, we can assume the distributions of the sample means for each city are normal. We can use the standardized test statistic formula for the difference in two population means to obtain:

$$t_0 = \frac{(\bar{x}_A - \bar{x}_B) - (0)}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}} = \frac{(5.2 - 3.5) - (0)}{\sqrt{\frac{1.7^2}{38} + \frac{2.1^2}{32}}} = 3.68 \sim t_{59}$$

This one time, we'll go ahead and compute the degrees of freedom for this example "by-hand."

$$df = \text{ floor}\left(\frac{\left(\frac{1.7^2}{38} + \frac{2.1^2}{32}\right)^2}{\frac{\left(\frac{1.7^2}{38}\right)^2}{37} + \frac{\left(\frac{2.1^2}{32}\right)^2}{31}} \right) \approx \text{floor}(59.48) = 59.$$

For a *t* distribution with *df* = 59, the *p*-value can be determined using Minitab's **Graph > Probability Distribution Plot**. It is approximately 0.0005068.



Minitab provides the same results with its **2-Sample t** procedure:

1. Choose **Stat > Basic Statistics > 2-Sample t**.
2. Select **Summarized Data**. Enter values for the *Sample size, Sample mean, and Sample standard deviation* for both samples.
3. Click **OK**.

Two-Sample T-Test and CI

Method

μ_1: population mean of Sample 1
μ_2: population mean of Sample 2
Difference: $\mu_1 - \mu_2$
Equal variances are not assumed for this analysis.

Descriptive Statistics

Sample	N	Mean	StDev	SE Mean
Sample 1	38	5.20	1.70	0.28
Sample 2	32	3.50	2.10	0.37

Estimation for Difference

Difference	95% CI for Difference
1.700	(0.775, 2.625)

Test

Null hypothesis	H_0: $\mu_1 - \mu_2 = 0$
Alternative hypothesis	H_1: $\mu_1 - \mu_2 \neq 0$

T-Value	DF	P-Value
3.68	59	0.001

Since the test statistic's *p*-value is less than $\alpha = 0.05$, then we have enough evidence to reject H_0, indicating that there is a difference in the number of years of music educations between City A and City B. Notice that our "by-hand" results match up exactly with Minitab's **2-Sample t** procedure.

Example 5. The ***Survey of Study Habits and Attitudes (SSHA)*** is a psychological test designed to measure the motivation, study habits, and attitudes toward learning of college students. These factors, along with ability, are important in explaining success in school. Scores on the SSHA range from 0 to 200. A selective private college gives the SSHA to a random sample of both male and female first-year students. The scores for the sample of $\boldsymbol{n_W = 18}$ **women** are:

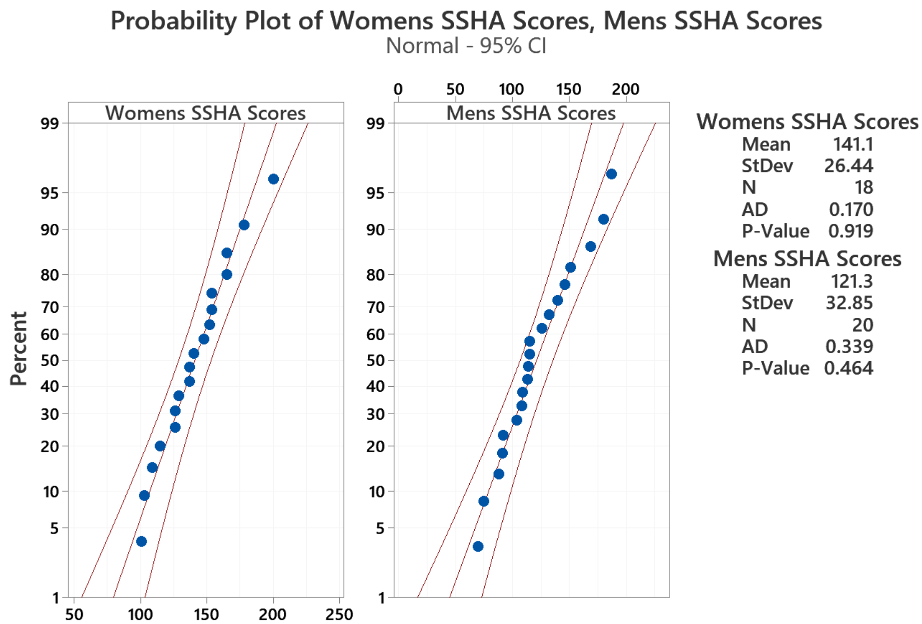
154	109	137	115	152	140	154	178	101
103	126	126	137	165	165	129	200	148

The scores for the sample of $\boldsymbol{n_M = 20}$ **men** are:

108	140	114	91	180	115	126	92	169	146
109	132	75	88	113	151	70	115	187	104

Most studies claim that the true mean SSHA score for men is lower than the true mean SSHA score for a comparable population of women. Test this claim using the above data. Assume that the unknown population variances for the men and women's scores are NOT equal.

Solution: The two groups of data are independent. There is no obvious connection between the women and men at this school. Since the sample sizes of both populations are less than 30, we cannot simply assume the distributions of the sample means are normally distributed. To assess normality of each sample, we need a probability plot for each individually. (See page 6, Example 2 for Minitab instructions.) As neither plot shows an obvious trend away from the theoretical normality plot line, the data is consistent with being drawn from normal populations. Also, the *p*-values for the normality tests are greater than $\alpha = 0.05$.



Let μ_M represent the true mean SSHA score for first-year college men at this given school and μ_W represent the true mean SSHA score for first-year college women at this given school. We are testing:

H_0: $\mu_M = \mu_W$, which can also be written as $\mu_M - \mu_W = 0$
H_a: $\mu_M < \mu_W$, which can also be written as $\mu_M - \mu_W < 0$

Since the independent data sets appear to come from normal populations and their population standard deviations are unknown, the appropriate hypothesis test to perform is a two-sample *t*-test. In this example, we'll go directly to Minitab to obtain the standardized test statistic and its *p*-value.

1. Choose **Stat > Basic Statistics > 2-Sample t**.
2. Select **Each sample is in its own column**. Select *Mens SSHA Scores* as Sample 1 and select *Womens SSHA Scores* as Sample 2.
3. Click **OK**.

Minitab returns the following output:

Two-Sample T-Test and CI: Mens SSHA Scores, Womens SSHA Scores

Method

μ_1: population mean of Mens SSHA Scores
μ_2: population mean of Womens SSHA Scores
Difference: $\mu_1 - \mu_2$
Equal variances are not assumed for this analysis.

Descriptive Statistics

Sample	N	Mean	StDev	SE Mean
Mens SSHA Scores	20	121.3	32.9	7.3
Womens SSHA Scores	18	141.1	26.4	6.2

Estimation for Difference

Difference	95% Upper Bound for Difference
-19.81	-3.53

Test

Null hypothesis H_0: $\mu_1 - \mu_2 = 0$

Alternative hypothesis H_1: $\mu_1 - \mu_2 < 0$

T-Value	DF	P-Value
-2.06	35	0.024

At significance level $\alpha = 0.05$, our data supports the claim that the true mean SSHA score for men is lower than the true mean SSHA score for a comparable population of women.