Chapter 11: Inference on Two Samples

In Chapters 9 and 10, we studied inferential statistics (confidence intervals and hypothesis tests) regarding population parameters of a single population - the average rest heart rate of students in a class, the proportion of ECC who voted, etc.

In Chapter 11, we'll be considering **the relationship between two populations** - means, proportions and standard deviaions. One frequent comparison we want to make between two populations is concerning the proportion of individuals with certain characteristics. For example, suppose we want to determine if college faculty voted at a higher rate than ECC students in the 2008 presidential election. Since we don't have any information from either population, we would need to take samples from each. This isn't an example of a hypothesis test from Section 10.4, about one proportion, it'd be comparing two proportions.

If you're ready to begin, just click on the "start" link below, or one of the section links on the left.

Section 11.1: Inference about Two Proportions

11.1 Inference about Two Proportions
11.2 Inference about Two Means: Dependent Samples
11.3 Inference about Two Means: Independent Samples
11.4 Inference about Two Standard Deviations
11.5 Putting It Together: Which Method Do I Use?

Objectives

By the end of this lesson, you will be able to...

1. test hypotheses regarding two population proportions
2. construct and interpret confidence intervals for the difference between two population proportions
3. determine the sample size necessary for estimating the difference between two population proportions within a specific margin of error

---

In Chapters 9 and 10, we studied inferential statistics (confidence intervals and hypothesis tests) regarding population parameters of a single population - the average rest heart rate of students in a class, the proportion of ECC who voted, etc.

In Chapter 11, we'll be considering **the relationship between two populations** - means, proportions and standard deviaions.

A frequent comparison we want to make between to populations is concerning the proportion of individuals with certain characteristics. For example, suppose we want to determine if college faculty voted at a higher rate than ECC students in the 2008 presidential election. Since we don't have any information from either population, we would need to take samples from each. This isn't an example of a hypothesis test from Section 10.4, about one proportion, it'd be comparing two proportions, so we need some new background.

The information that follows is a bit heavy, but it shows the theoretical background for testing claims and finding confidence intervals for the difference between two population proportions.

The Difference Between Two Population Proportions

In Section 8.2, we discussed the distribution of one sample proportion, $\hat{p}$. What we'll need to do now is develop some similar theory regarding the distribution of the difference in two sample proportions, $\hat{p}_1 - \hat{p}_2$.

The Sampling Distribution of the Difference between Two Proportions

Suppose simple random samples size n_1 and n_2 are taken from two populations. The distribution of $\hat{p}_1 - \hat{p}_2$ where $\hat{p}_1 = \dfrac{x_1}{n_1}$ and $\hat{p}_2 = \dfrac{x_2}{n_2}$, is aproximately normal with mean $\mu_{\hat{p}_1 - \hat{p}_2} = p_1 - p_2$ and standard deviation

$\sigma_{\hat{p}_1 - \hat{p}_2} = \sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}$, provided:

1. $n_1 \hat{p}_1 (1 - \hat{p}_1) \geq 10$
2. $n_2 \hat{p}_2 (1 - \hat{p}_2) \geq 10$
3. both sample sizes are less than 5% of their respective populations.

The standardized version is then

$$z_0 = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}}$$

which has an approximate standard normal distribution.

The thing is, in most of our hypothesis testing, the null hypothesis assumes that the proportions are the same ($p_1 = p_2$), so we can call $p = p_1 = p_2$.

Since $p_1 = p_2$, we can substitute 0 for $p_1 - p_2$, and substitute p for both p_1 and p_2. In that case, we can rewrite the above standardized z the following way:

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}} = \frac{(\hat{p}_1 - \hat{p}_2) - (0)}{\sqrt{\dfrac{p(1-p)}{n_1} + \dfrac{p(1-p)}{n_2}}} = \frac{\hat{p}_1 - \hat{p}_2}{p(1-p)\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$$

Which leads us to our hypothesis test for the difference between two proportions.

Performing a Hypothesis Test Regarding $p_1 - p_2$

Step 1: State the null and alternative hypotheses.

Two-Tailed	Left-Tailed	Right-Tailed
H_0: $p_1 - p_2 = 0$	H_0: $p_1 - p_2 = 0$	H_0: $p_1 - p_2 = 0$
H_1: $p_1 - p_2 \neq 0$	H_1: $p_1 - p_2 < 0$	H_1: $p_1 - p_2 > 0$

Step 2: Decide on a level of significance, α.

Step 3: Compute the test statistic, $z = \dfrac{\hat{p}_1 - \hat{p}_2}{\hat{p}(1-\hat{p})\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$.

Step 4: Determine the *P*-value.

Step 5: Reject the null hypothesis if the *P*-value is less than the level of significance, α.

Step 6: State the conclusion.

A note about the difference between two proportions: As with the previous two sections, the order in which the proportions are placed is not important. The important thing is to note clearly in your work what the order is, and then to construct your alternative hypothesis accordingly.

Hypothesis Testing Regarding $p_1 - p_2$ Using StatCrunch

1. Go to **Stat** > **Proportions** > **Two Sample** > **with summary**
2. Enter the number of successes and observations for each sample and press **Next**.
3. Set the null proportion difference and the alternative hypothesis.
4. Click on **Calculate**.

The results should appear.

Example 1

Problem: Suppose a researcher believes that college faculty vote at a higher rate than college students. She collects data from 200 college faculty and 200 college students using simple random sampling. If 167 of the faculty and 138 of the students voted in the 2008 Presidential election, is there enough evidence at the 5% level of significance to support the researcher's claim?

Solution:

First, we need to check the conditions. Both sample sizes are clearly less than 5% of their respective populations. In addition,

$$n_1 \hat{p}_1 (1 - \hat{p}_1) = 200 \left(\frac{167}{200} \right) \left(1 - \frac{167}{200} \right) \approx 27.6 \geq 10$$

$$n_2 \hat{p}_2 (1 - \hat{p}_2) = 200 \left(\frac{138}{200} \right) \left(1 - \frac{138}{200} \right) \approx 42.8 \geq 10$$

So our conditions are satisfied.

Step 1:

Let's take the two portions in the order we receive them, so
$p_1 = p_f$ (faculty) and $p_1 = p_s$ (students)

Our hypotheses are then:
H_0: $p_f - p_s = 0$
H_1: $p_f - p_s > 0$ (since the researcher claims that faculty vote at a *higher* rate)

Step 2: $\alpha = 0.05$ (given)

Step 3: (we'll use StatCrunch)

Step 4: Using StatCrunch:

Hypothesis test results:
p_1 : proportion of successes for population 1
p_2 : proportion of successes for population 2
$p_1 - p_2$: difference in proportions
H_0 : $p_1 - p_2 = 0$
H_A : $p_1 - p_2 > 0$

Difference	Count1	Total1	Count2	Total2	Samp	Z-Stat	P-value
$p_1 - p_2$	167	200	138	200		3.4073462	0.0003

(Trimmed to fit on this page.)

Step 5: Since the *P*-value $< \alpha$, we reject the null hypothesis.

Step 6: Based on these results, there is very strong evidence (certainly enough at the 5% level of significance) to support the researcher's claim.

Confidence Intervals about the Difference Between Two Proportions

We can also find a confidence interval for the difference in two population proportions.

In general, a **(1-α)100% confidence interval for $p_1 - p_2$** is

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Note: The following conditions must be true:

1. $n_1 \hat{p}_1(1-\hat{p}_1) \geq 10$
2. $n_2 \hat{p}_2(1-\hat{p}_2) \geq 10$
3. both sample sizes are less than 5% of their respective populations.

Confidence Intervals About p_1-p_2 Using StatCrunch

1. Go to **Stat** > **Proportions** > **Two Sample** > **with summary**
2. Enter the number of successes and observations for each sample and press **Next**.
3. Click the Confidence Interval button, and set the confidence level.
4. Click on **Calculate**.

The results should appear.

Example 2

Problem: Considering the data from Example 1, find a 99% confidence interval for the difference between the proportion of faculty and the proportion of students who voted in the 2008 Presidential election.

Solution: From Example 1, we know that the conditions for performing inference are met, so we'll use StatCrunch to find the confidence interval.

99% confidence interval results:

p_1 : proportion of successes for population 1
p_2 : proportion of successes for population 2
p_1 - p_2 : difference in proportions

Difference	Count1	Total1	Count2	Total2	.	L. Limit	U. Limit
p_1 - p_2	168	200	137	20 /346		0.047218394	0.26278162

(Timmed to fit on this page.)

So we can say that we're 99% confident that the difference between the proportion of faculty who vote and the proportion of students who vote is between 3.7% and 25.3%.

Determining the Sample Size Needed

In Section 9.3, we learned how to find the necessary sample size if a specific margin of error is desired. We can do a similar analysis for the difference in two proportions. From the confidence interval formula, we know that the margin of error is:

$$E = z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

If we assume that $n_1 = n_2 = n$, we can solve for n and get the following result:

The **sample size required** to obtain a $(1-\alpha)100\%$ confidence interval for p_1-p_2 with a margin of error E is:

$$n = n_1 = n_2 = [\hat{p}_1(1-\hat{p}_1) + \hat{p}_2(1-\hat{p}_2)]\left(\frac{z_{\alpha/2}}{E}\right)^2$$

rounded up to the next integer, if $\hat{p}_1$ and $\hat{p}_2$ are esimates for p_1 and p_2, respectively.

If no prior estimate is available, use $\hat{p}_1 = \hat{p}_2 = 0.5$, which yields the following formula:

$$n = n_1 = n_2 = 0.5\left(\frac{z_{\alpha/2}}{E}\right)^2$$

again rounded up to the next integer.

Note: As in Section 9.3, the desired margin of error should be expressed as a decimal.

Let's try one.

Example 3

Suppose we want to study the success rates for students in Mth098 Intermediate Algebra at ECC. We want to compare the success rates of students who place directly into Mth098 with those who first took Mth096 Beginning Algebra. From past experience, we know that a typical success rate for students in this class is about 65%. How large of a sample size is necessary to create a 95% confidence interval for the difference of the two passing rates with a maximum error of 2%?

[reveal answer]

Section 11.2: Inference about Two Means: Dependent Samples

Objectives

By the end of this lesson, you will be able to...

1. distinguish between independent and dependent samples
2. test hypotheses regarding matched-pairs data
3. construct and interpret confidence intervals about the population mean difference of matched-pairs data

---

In the next two sections, we'll focus on the relationship between two means. Before we begin, we need to discuss the two possible situations.

Independent vs. Dependent Samples

In general, there are two possible situations regarding two population means. Consider the following examples:

Example 1

Problem: Suppose we measure the thickness of plaque (mm) in the carotid artery of 10 randomly selected patients with mild atherosclerotic disease. Two measurements are taken, thickness before treatment with Vitamin E (baseline) and after two years of taking Vitamin E daily. (Source: UCLA Dept. of Statistics)

Discussion: In this example, we would be comparing the mean plaque thickness before Vitamin E with the mean thickness after - so the same 10 patients would be in the sample "before" and the sample "after". When the individuals slected are paired in this manner, we call the samples **dependent**.

Example 2

Problem: Nine observations of surface soil pH were made at two different locations. Does the data suggest that the true mean soil pH values differ for the two locations? (Source: UCLA Dept. of Statistics)

Discussion: Unlike in Example 1, the samples in this example are completely unrelated (*two different locations*). In examples like this, where the indivuaduals selected have no relation to each other, we call the samples **independent**.

In general, two samples are **dependent** if the individuals in one sample determine the individuals in the other sample. (i.e. matched-pair design) Two samples are **independent** when the individuals in one sample *do not* determine the individuals in the other sample. (i.e. completely randomized)

This will be very important as we progress, because we will need to distinguish between whether the samples are dependent or independent, because the statistical methods will be very different.

The Difference, d

When considering dependent samples, we analyze the difference, d, in each matched pair. For example, suppose we consider the thickness of plaque (mm) in the carotid artery, referenced in Example 1. If an individual had a maximal thickness of 0.92mm before the Vitamin E treatment, and 0.95mm after the treatment, the difference, d, for that individual would be

$$d = 0.95 - 0.92 = 0.03mm$$

So in general, for this experiment, we would define the difference, d, to be:

$$d = \textit{thickness after Vit. E treatment - thickness before treatment}$$

Before we can perform any inferential statistics, we need to know the distribution of d.

The Distribution of the Difference, d

Suppose the following are true concerning a sample:

1. the sample is obtained using simple random sampling, and
2. the sample data are matched pairs, and
3. the sample has no outliers and the population from which the sample is drawn is normally distributed, or the sample size is large (n≥30).

Then the test statistic

$$t_0 = \frac{\bar{d} - \mu_{\bar{d}}}{s_d / \sqrt{n}}$$

follows the *t*-distribution with n-1 degrees of freedom.

With that in mind, we can now perform statistical inference like hypothesis tests and confidence intervals. We'll start with hypothesis tests, and follow the same steps we did when we were analyzing the population mean, when σ was unknown.

A note about the difference, d: Many students find it confusing how to determine which value should go first when setting up the difference. In reality, it doesn't matter! The important thing is to note clearly in your work how *d* is set up, and then to construct your alternative hypothesis accordingly.

Performing a Hypothesis Test Regarding *d*

Step 1: State the null and alternative hypotheses.

Two-Tailed	Left-Tailed	Right-Tailed
$H_0: \mu_d = 0$	$H_0: \mu_d = 0$	$H_0: \mu_d = 0$
$H_1: \mu_d \neq 0$	$H_1: \mu_d < 0$	$H_1: \mu_d > 0$

Step 2: Decide on a level of significance, α.

Step 3: Compute the test statistic, $t_0 = \frac{\bar{d} - \mu_{\bar{d}}}{s_d / \sqrt{n}}$.

Step 4: Determine the *P*-value.

Step 5: Reject the null hypothesis if the *P*-value is less than the level of significance, α.

Step 6: State the conclusion.

Hypothesis Testing Regarding d Using StatCrunch

1. Enter the data.
2. Go to **Stat** > **t-Statistics** > **paired**.
3. Select the columns the 1st and 2nd samples are in and click **Next**.
4. Set the null mean difference and the alternative hypothesis.
5. Click on **Calculate**.

The results should appear.

Note: StatCrunch does not display the variable correctly. It will display $\mu_1 - \mu_2$ rather than μ_d.

Example 3

Problem: Suppose we want to determine if a diet drug is effective. To determine it's effectiveness, we randomly select 10 volunteers, and measure their weight before the diet drug treatment, and again one month later. The results are shown below.

	A	B	C	D	E	F	G	H	I	J
Before	190	211	198	203	262	224	251	238	219	255
After	184	204	197	208	246	221	256	225	211	243

Is there evidence to support the company's claim that the diet drug does cause weight loss at the 5% level of significance?

Solution:

First, we need to calculate the differences. It's important to always write down which direction you want to define the difference, d. In this case, we'll use:

$$d = \text{Before} - \text{After}$$

	A	B	C	D	E	F	G	H	I	J
Before	190	211	198	203	262	224	251	238	219	255
After	184	204	197	208	246	221	256	225	211	243
Difference	6	7	1	-5	16	3	-5	13	8	12

We need to then determine if the differences are normally distributed, since our sample size is less than 30.
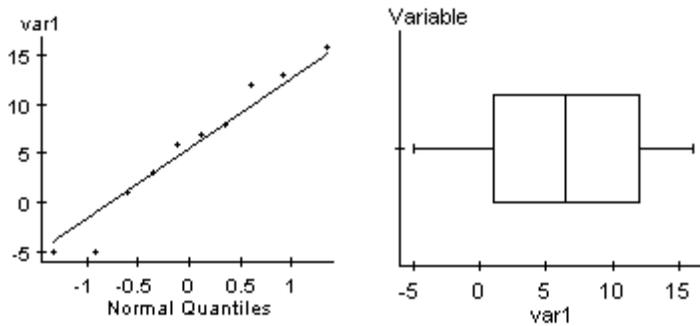
We can see that the Q-Q plot is fairly linear and the boxplot shows no outliers, so it's reasonable to say that the differences are normally distributed.

Step 1:
H_0: μ_d = 0
H_1: μ_d > 0
(Since the company wants to show that the average weight loss is positive.)

Step 2: α = 0.05 (given)

Step 3: (we'll use StatCrunch)

Step 4: Using StatCrunch:

Hypothesis test results:
$\mu_1 - \mu_2$: mean of the paired difference between Before and After
$H_0 : \mu_1 - \mu_2 = 0$
$H_A : \mu_1 - \mu_2 > 0$

Difference	Sample Diff.	Std. Err.	DF	T-Stat	P-value
Before - After	5.6	2.2715633	9	2.4652627	0.0179

Differences stored in column, Differences.

Step 5: Since the *P*-value < α, we reject the null hypothesis.

Step 6: Based on these results, it would appear that there *is* enough evidence to support the claim that the drug causes weight loss.

> **Note:** If we had first set up the difference as d = After - Before, the alternative hypothesis would then be H_1: μ_d < 0 (since the company claims the "after" is less than the "before").

Confidence Intervals about the Mean Difference

Since the distribution of $\bar{d}$ follows the *t*-distribution, we can also create a confidence interval for the population mean difference.

> In general, a **(1-α)100% confidence interval for μ_d** is
>
> $$\bar{d} \pm t_{\alpha/2} \cdot \frac{s_d}{\sqrt{n}}$$

where $t_{\alpha/2}$ is computed with n-1 degrees of freedom.

Note: The sample size must be large (n≥30) with no outliers or the population must be normally distributed.

Confidence Intervals About μ_d Using StatCrunch

1. Enter the data.
2. Go to **Stat** > **t-Statistics** > **paired**.
3. Select the columns the 1st and 2nd samples are in and click **Next**.
4. Select "Confidence Interval" and select the confidence level.
5. Click on **Calculate**.

The results should appear.

Note: StatCrunch does not display the variable correctly. It will display μ_1 - μ_2 rather than μ_d.

Example 4

Problem: Consider the weight loss data from Example 3.

	A	B	C	D	E	F	G	H	I	J
Before	190	211	198	203	262	224	251	238	219	255
After	184	204	197	208	246	221	256	225	211	243

Find a 90% confidence interval for the population mean difference.

Solution:

From Example 3, we know that the differences are normally distributed with no outliers, so we can find the confidence interval.

$$d = Before - After$$

	A	B	C	D	E	F	G	H	I	J
Before	190	211	198	203	262	224	251	238	219	255
After	184	204	197	208	246	221	256	225	211	243
Difference	6	7	1	-5	16	3	-5	13	8	12

Using StatCrunch:

90% confidence interval results:

μ_1 - μ_2 : mean of the paired difference between Before and After

Difference	Sample Diff.	Std. Err.	DF	L. Lim	U. Lim
Before - After		5.6	2.2715633	9	1.4359679	9.764032

So we can say that we're 90% confident that the mean weight loss (difference) is between 1.4 and 9.8.

Section 11.3: Inference about Two Means: Independent Samples

Objectives

By the end of this lesson, you will be able to...

1. test hypotheses regarding the difference of two independent means
2. construct and interpret confidence intervals regarding the difference of two independent means

---

In 2005, Larry Summers, then President of Harvard, gave a speech at the NBER Conference on Diversifying the Science and Engineering Workforce. In that speech, he made some very controversial remarks regarding differences in the genders. In particular,

It does appear that on many, many different human attributes-height, weight, propensity for criminality, overall IQ, mathematical ability, scientific ability-there is relatively clear evidence that whatever the difference in means-which can be debated-there is a difference in the standard deviation, and variability of a male and a female population.

Suppose we wanted to do a comparison between the genders. In Section 11.2, we looked at comparing two means with matched-pairs data - dependent samples. What if there isn't a relationship between the two samples? We certainly can't pair them up then, and find the *mean difference*. What we need to anlyze instead is the *difference of two means*. First, we need to know something about it's distribution.

The Difference of Two Independent Means

The Distribution of the Difference of Two Means, d

Suppose a simple random sample of size n_1 is taken from a population with unknown mean μ_1 and unknown standard deviation σ_1. In addition, a simple random sample of size n_2 is taken from a second population with unknown mean μ_2 and unknown standard deviation σ_2. If the two populations are normally distributed or the sample sizes are sufficiently large ($n_1, n_2 \geq 30$), then

$$t_0 = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

*approximately** follows the *t*-distribution with the smaller of n_1-1 or n_2-1 degrees of freedom.

Note: There is no exact method for comparing two means with unequal populations, but this statistic is a close approximation. It is known as **Welch's approximate t**, in honor of English statistician Bernard Lewis Welch (1911-1989).

Now that we have the distribution of the difference between two means, we can perform statistic inference (hypothesis testing and confidence intervals).

Performing a Hypothesis Test Regarding the Difference Between Two Independent Means

Step 1: State the null and alternative hypotheses.

	Two-Tailed	Left-Tailed	Right-Tailed
H_0:	$\mu_1 - \mu_2 = 0$	$\mu_1 - \mu_2 = 0$	$\mu_1 - \mu_2 = 0$
H_1:	$\mu_1 - \mu_2 \neq 0$	$\mu_1 - \mu_2 < 0$	$\mu_1 - \mu_2 > 0$

Step 2: Decide on a level of significance, α.

Step 3: Compute the test statistic, $t_0 = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$.

Step 4: Determine the *P*-value.

Step 5: Reject the null hypothesis if the *P*-value is less than the level of significance, α.

Step 6: State the conclusion.

Hypothesis Testing Regarding $\mu_1 - \mu_2$ Using StatCrunch

1. Enter the data. (Note: If you're copying from another file, be careful - put the column with the most entries first. StatCrunch does not handle spaces well.)
2. Go to **Stat** > **t-Statistics** > **Two Sample**, then **with data** or **with summary**.
3. If you chose *with data*, select the columns containing the 1st and 2nd samples. Otherwise, enter all the sample statistics.
4. Uncheck "Pool variances" and press **Next**.
5. Set the null mean difference and the alternative hypothesis.
6. Click on **Calculate**.

The results should appear.

A note about the difference between the means: As with the previous section, it's often difficult for students to choose which mean to place first. Again - it doesn't matter! The important thing is to note clearly in your work what the order is, and then to construct your alternative hypothesis accordingly.

Example 1

Problem: Suppose we wish to test whether there is a difference in the performances of men and women in mathematics. An ECC instructor collects exam scores from 2 semesters worth of Beginning Algebra students, shown below by gender.

| Women	Men
94 90 96 73 71 75 52	77 57 79 65 68 65 55
86 93 86 50 30 46 36	72 66 29 69 85 82 43
82 55 75 47 80 92 56	60 64 82 91 51 60 76
43 77 63 76 67 99 93	43 82 87 32 71 77 77
75 90 92 88 76 61 77	79 97 56

49 67 81 89 88 42 51	
58 98 46 96 90 46 50	
67 83 85	

Is there enough evidence at the 5% level of significance to support the claim that men and women perform differently in this class?

Solution:

First, we need to make sure that neither sample contains outliers. (We do have sample sizes of at least 30, so we don't need to check to see if they come from normally distributed populations.)

Exam Scores by Gender
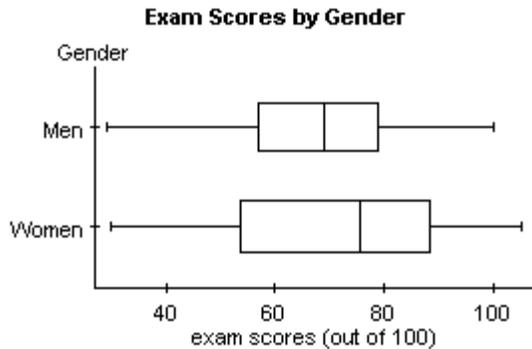
We can see that neither sample contains outliers, so we are free to continue.

Step 1:
In this case, since we're only testing whether the means are *different*, the order doesn't matter at all. It's easiest to simply take the two in the order we receive them, so
$\mu_1 = \mu_W$ (women), and $\mu_2 = \mu_M$ (men).

H_0: $\mu_W - \mu_M = 0$
H_1: $\mu_W - \mu_M \neq 0$

Step 2: $\alpha = 0.05$ (given)

Step 3: (we'll use StatCrunch)

Step 4: Using StatCrunch:

Hypothesis test results:
μ_1 : mean of Women
μ_2 : mean of Men
$\mu_1 - \mu_2$: mean difference
$H_0 : \mu_1 - \mu_2 = 0$
$H_A : \mu_1 - \mu_2 \neq 0$
(without pooled variances)

Difference	Sample Mean	Std. Err.	DF	T-Stat	P-value
$\mu_1 - \mu_2$	4.1426797	4.0103817	70.54798	1.0329889	0.3051

Step 5: Since the *P*-value $> \alpha$, we do not reject the null hypothesis.

Step 6: Based on these results, it would appear that there is not enough

evidence (very little, in fact) to support the claim that men and women perform differently in Beginning Algebra.

It should be noted that larger sample sizes will most likely show a statistically significant difference, though the difference may not have any practical meaning.

One final note, you may wonder what the "pooled variances" is referring to, and why we don't use it. By not pooling the variances, we are assuming that the population variances are unequal. There is a test for equal variances (we'll cover it in Section 11.4), but like the earlier tests concerning standard deviations, the distributions must be normal. Your textbook makes this final comment:

> Because the formal *F*-test for testing the equality of variances is so volatile, we are content to use Welch's *t*. This 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 in Welch's *t* than in the pooled *t* when the population variances are equal. However, the two tests typically provide the same conclusion, even if the assumption of equal population standard deviations seems reasonable.
>
> Source: Statistics; Informed Decisions Using Data
> Autor: Michael Sullivan, III

Confidence Intervals about the Difference Between Two Means

Since the distribution of $\bar{x}_1 - \bar{x}_2$ follows the *t*-distribution, we can also create a confidence interval for the difference between two population means.

In general, a **(1-α)100% confidence interval for μ₁-μ₂** is

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

where $t_{\alpha/2}$ is computed with min{n_1-1, n_2-1} degrees of freedom.

Note: The sample sizes must be large ($n_1, n_2 \geq 30$) with no outliers or the populations must be normally distributed.

Confidence Intervals About μ₁-μ₂ Using StatCrunch

1. Enter the data. (Note: If you're copying from another file, be careful - put the column with the most entries first. StatCrunch does not handle spaces well.)
2. Go to **Stat** > **t-Statistics** > **Two Sample**, then **with data** or **with summary**.
3. If you chose *with data*, select the columns containing the 1st and 2nd samples. Otherwise, enter all the sample statistics.
4. Uncheck "Pool variances" and press **Next**.
5. Select "Confidence Interval" and select the confidence level.
6. Click on **Calculate**.

 The results should appear.

Example 2

Problem: Consider the data comparing men and women from Example 1.

Women	Men
94 90 96 73 71 75 52	77 57 79 65 68 65 55
86 93 86 50 30 46 36	72 66 29 69 85 82 43
82 55 75 47 80 92 56	60 64 82 91 51 60 76
43 77 63 76 67 99 93	43 82 87 32 71 77 77
75 90 92 88 76 61 77	79 97 56
49 67 81 89 88 42 51	
58 98 46 96 90 46 50	
67 83 85	

Find a 90% confidence interval for the population mean difference.

Solution: From Example 1, we know that neither sample contains outliers, so we can find the confidence interval.

Using StatCrunch:

90% confidence interval results:

μ_1 : mean of Women
μ_2 : mean of Men
$\mu_1 - \mu_2$: mean difference
(without pooled variances)

Difference	Sample Mean	Std. Err.	DF	L. Limit	U. Limit
$\mu_1 - \mu_2$	4.1426797	4.0103817	70.54798	-2.541587	10.826947

So we can say that we're 90% confident that the difference between the two means is between -2.5 and 10.8.

Section 11.4: Inference about Two Population Standard Deviations

Objectives

By the end of this lesson, you will be able to...

1. find critical values of the *F*-distribution
2. test hypotheses regarding two population standard deviations

The last parameters we need to compare between two populations are the variance and standard deviation. Before we can develop a hypothesis test comparing two population parameters, we need another distribution.

Fisher's *F*-distribution

Unlike the mean, the standard deviation is extremely susceptible to extreme values, and consequently does a very poor job of measuring spread for distributions that are not symmetric. So before we do any inference regarding population standard deviations, we must first verify that the samples come from normally distributed populations.

Fisher's *F*-distribution

If $\sigma_1^2 = \sigma_2^2$ and s_1^2 and s_2^2 are sample variances from independent simple random samples of size n_1 and n_2, respectively, drawn from normal populations, then

$$F = \frac{s_1^2}{s_2^2}$$

follows the **F-distribution** with n_1-1 degrees of freedom in the numerator and n_2-1 degrees of freedom in the denominator.

Properties of the *F*-distribution

1. Like the X^2 distribution, it is *not* symmetric. It is skewed right
2. The shape depends on the degrees of freedom in the numerator and denominator.
3. $F \geq 0$

Notice here that the samples *must* come from normally distributed populations.

Finding Critical Values

Find critical values in the *F*-distribution using a table is done in a similar manner to the *t* and X^2 tables, though with some differences. The values in the table still represent values with the indicidated α area to the right, but because the *F* distribution has two degreees of freedom rather than one, it requires a separate table for each α.

Before we start the section, you need a copy of the table. You can download a printable copy of this table, or use

Table VII starting on page A-14 in the back of your textbook. That table will look different than the printable version linked above because the publisher did not provide a digital version.

F		Degrees of Freedom in the Numerator						
α = 0.1		1	2	3	4	5	6	7
	1	39.86	49.50	53.59	55.83	57.24	58.20	58.91
	2	8.53	9.00	9.16	9.24	9.29	9.33	9.35
	3	5.54	5.46	5.39	5.34	5.31	5.28	5.27
	4	4.54	4.32	4.19	4.11	4.05	4.01	3.98
	5	4.06	3.78	3.62	3.52	3.45	3.40	3.37
	6	3.78	3.46	3.29	3.18	3.11	3.05	3.01
	7	3.59	3.26	3.07	2.96	2.88	2.83	2.78
	8	3.46	3.11	2.92	2.81	2.73	2.67	2.62
	9	3.36	3.01	2.81	2.69	2.61	2.55	2.51
	10	3.29	2.92	2.73	2.61	2.52	2.46	2.41
	12	3.18	2.81	2.61	2.48	2.39	2.33	2.28
	15	3.07	2.70	2.49	2.36	2.27	2.21	2.16
	20	2.97	2.59	2.38	2.25	2.16	2.09	2.04
	25	2.92	2.53	2.32	2.18	2.09	2.02	1.97
	30	2.88	2.49	2.28	2.14	2.05	1.98	1.93
	40	2.84	2.44	2.23	2.09	2.00	1.93	1.87
	50	2.81	2.41	2.20	2.06	1.97	1.90	1.84

(Left axis label: Degress of Freedom in the Denominator)

So the values in the table above are the critical values:

$$F_{\alpha, n_1-1, n_2-1}$$

You may wonder how we find critical values for left-tailed tests. To do that, we use the same table and the following formula:

$$F_{1-\alpha, n_1-1, n_2-1} = \frac{1}{F_{\alpha, n_2-1, n_1-1}}$$

Let's try a couple examples.

Example 1

Find the value of the *F*-distribution that has α=0.05 area to the right, with 10 degrees of freedom in the numerator, and 15 degrees of freedom in the denominator.

[reveal answer]

Example 2

Find the value of the *F*-distribution that has α=0.05 area to the *left*, with 20 degrees of freedom in the numerator, and 8 degrees of freedom in the denominator.

[reveal answer]

Finding Critical Values Using StatCrunch

Click on **Stat** > **Calculators** > **F**

Enter the numerator and denominator degrees of freedom, the direction of the inequality, and the probability (leave X blank). Then press **Compute**.

Example 3

Use the technology of your choice to find the value from the *F*-distribution with $\alpha=0.01$ area to the right if samples of size 15 and 20 are taken.

[reveal answer]

Performing a Hypothesis Test Regarding Two Population Standard Deviations

Step 1: State the null and alternative hypotheses.

Two-Tailed   **Left-Tailed**   **Right-Tailed**

$H_0: \sigma_1^2 = \sigma_2^2$ $H_0: \sigma_1^2 = \sigma_2^2$ $H_0: \sigma_1^2 = \sigma_2^2$

$H_1: \sigma_1^2 \neq \sigma_2^2$ $H_1: \sigma_1^2 < \sigma_2^2$ $H_1: \sigma_1^2 > \sigma_2^2$

Step 2: Decide on a level of significance, α.

Step 3: Compute the test statistic, $F = \dfrac{s_1^2}{s_2^2}$.

Step 4: Determine the *P*-value.

Step 5: Reject the null hypothesis if the *P*-value is less than the level of significance, α.

Step 6: State the conclusion.

Hypothesis Testing Regarding Two Population Standard Deviations Using StatCrunch

1. Go to **Stat** > **Variance** > **Two Sample** > **data/summary**
2. Enter the sample variances or select the appropriate column
3. Select **Next**.
4. Set the null variance ratio (standard is 1) and the alternative hypothesis.
5. Click on **Calculate**.

The results should appear.

Example 4

Problem: In Example 1 in Section 11.2, we compared the average scores of men and women on a Mth096 exam. In that test, we assumed that the standard deviations of the two groups were equal. Test the assumption at the $\alpha=0.1$ level of significance.

Solution:

First, we need to check the conditions. We know from Example 1 that neither sample contains outliers, but we do not know if they come from normally distributed populations. We'll use StatCrunch to perform Q-Q plots.
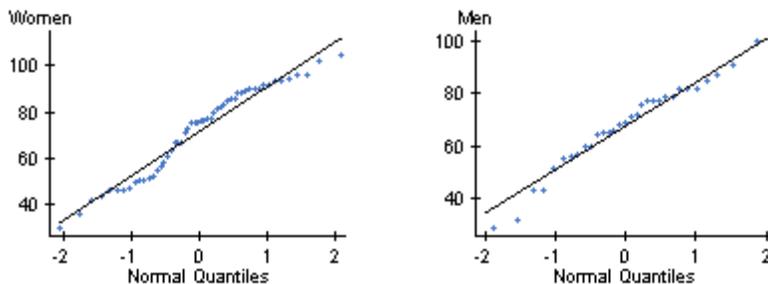
While the two plots aren't exactly linear, it does appear that the samples could come from normally distributed populations, so our conditions are satisfied.

Step 1:
H_0: $\sigma_1^2 = \sigma_2^2$
H_1: $\sigma_1^2 \neq \sigma_2^2$

Step 2: $\alpha = 0.1$ (given)

Step 3: (we'll use StatCrunch)

Step 4: Using StatCrunch:

Hypothesis test results:

σ_1^2 : variance of Men
σ_2^2 : variance of Women
σ_1^2/σ_2^2 : variance ratio
H_0 : $\sigma_1^2/\sigma_2^2 = 1$
H_A : $\sigma_1^2/\sigma_2^2 \neq 1$

Ratio	n1	n2	Sample Ratio	F-Stat	P-value
σ_1^2/σ_2^2	31	52	0.7472743	0.7472743	0.3959

Step 5: Since the *P*-value > α, we do not reject the null hypothesis.

Step 6: Based on these results, there is no evidence to support the claim that the standard deviations are not equal.

Section 11.5: Putting It Together: Which Method Do I Use?

11.1 Inference about Two Proportions
11.2 Inference about Two Means: Dependent Samples
11.3 Inference about Two Means: Independent Samples
11.4 Inference about Two Standard Deviations
11.5 Putting It Together: Which Method Do I Use?

Objectives

By the end of this lesson, you will be able to...

1. determine the appropriate hypothesis test to perform

Hypothesis Tests Regarding Two Populations

So we now have four new hypothesis tests to add to our arsenal. Here they are again:

Tests Regarding the Difference Between Two Population Means

In order to perform a hypothesis test regarding two population means, the following must be true concerning a sample:

1. the sample is obtained using simple random sampling, and
2. the sample data are matched pairs, and
3. the sample has no outliers and the population from which the sample is drawn is normally distributed, or the sample size is large ($n \geq 30$).

Then the test statistic is

$$t_0 = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}}$$

Tests Regarding the Mean Difference

In order to perform a hypothesis test regarding the mean difference, the following must be true:

1. a simple random sample of size n_1 is taken from a population with unknown mean μ_1 and unknown standard deviation σ_1
2. a simple random sample of size n_2 is taken from a second population with unknown mean μ_2 and unknown standard deviation σ_2
3. the two populations are normally distributed or the sample sizes are sufficiently large (n_1, $n_2 \geq 30$)

Then the test statistic is:

$$t_0 = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Tests Regarding the Difference Between Two Population Proportions

In order to perform a hypothesis test regarding the mean difference, the following must be true:

1. simple random samples size n_1 and n_2 are taken from two populations
2. $n_1 \hat{p}_1 (1 - \hat{p}_1) \geq 10$
3. $n_2 \hat{p}_2 (1 - \hat{p}_2) \geq 10$
4. both sample sizes are less than 5% of their respective populations.

Then the test statistic is:

$$z_0 = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$$

Tests Regarding Two Population Standard Deviations

In order to perform a hypothesis test regarding the mean difference, the following must be true:

1. $\sigma_1^2 = \sigma_2^2$
2. s_1^2 and s_2^2 are sample variances from independent simple random samples of size n_1 and n_2, respectively
3. both populations are normal

Then the test statistic is:

$$F = \frac{s_1^2}{s_2^2}$$

Choosing the Appropriate Hypothesis Test

Now that we've done a (very) quick review of the four various tests, it's helpful to think of a flowchart when deciding which test to apply. Here's a version of the flowchart from your text:
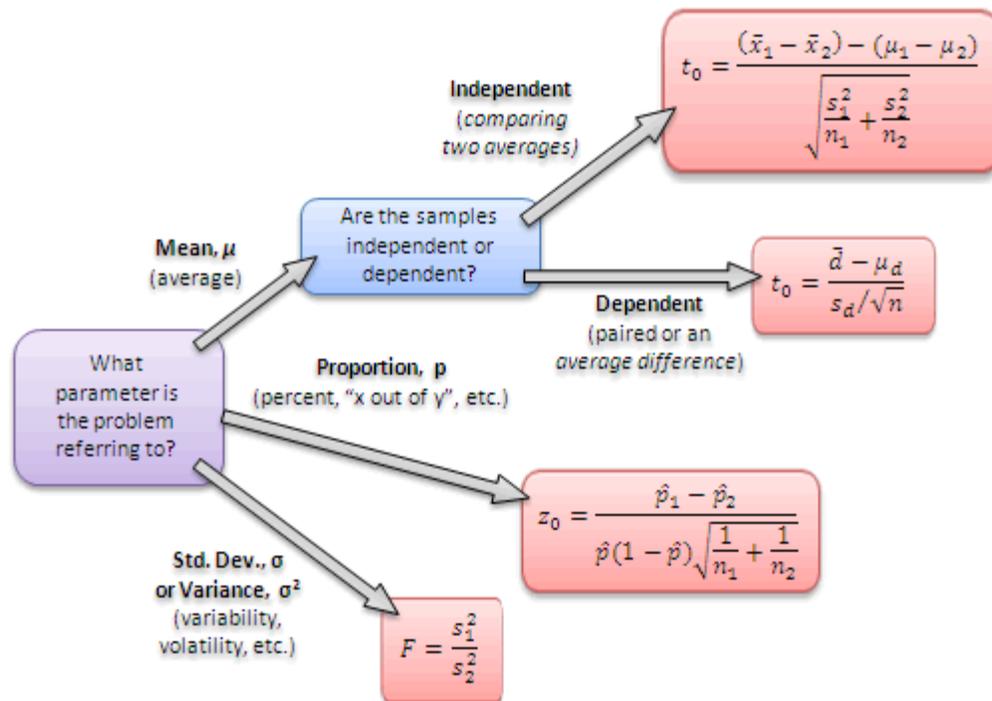
The biggest problems usually occur between the *independent* and *dependent* samples regarding the means. They key is to determine if the samples are somehow paired. A dead giveaway is a problem with *before* and *after*. In that case, they're clearly paired, making them *dependent* samples. If you're comparing two completely different populations (the average mpg for Honda Civics vs. Toyota Camry), then you have *independent* samples.

As we mentioned in the last chapter, there's no quick and easy rule to memorize. You'll need to practice all the

problems on the following page and be sure to do all the assigned homework problems. There's also an extra review for this exam, which also helps you choose which hypothesis test to apply. It's important to practice, practice, practice!

Some Examples

It's time for examples. In each case, don't worry about actually completing the problem. Focus instead on choosing the correct hypothesis test to apply. For more practice, you should look at the Exam 4 Extra Review file, which is available in Desire2Learn.

Example 1

Janis commutes to her Statistics class at ECC. She has two possible routes and would like to determine which is optimal. To help decide, she collects travel times for 60 morning trips, 30 on each route. Her first route has an average travel time of 24.3 minutes, with a standard deviation of 3.8 minutes. The second route has an average travel time of 22.9 minutes, with a standard deviation of 4.4 minutes. Based on these data, does Janis have enough evidence to say that the second route is the optimal one?

[reveal answer]

Example 2

Jay and Sheila are pig farmers in south-western Minnesota. They're changing the feed they use, and they're concerned that one of the new options leads to weights in the pigs that vary too widely. To help determine which choice of feed is more consistent, they take two samples of 100 piglets each. The first sample receives feed from AgraChoice, while the second receives feed from Swine Food. After 6 months, both samples have similar average weights of nearly 200 pounds, but the standard deviations are different. The AgraChoice sample has a standard deviation of 22.1 pounds, while the Swine Food sample has a standard deviation of 24.3 pounds.

Based on these samples, are Jay and Sheila's fears founded? Does the Swine Food yield 6-month-old pigs whose weight varies more than those fed with AgraChoice?

[reveal answer]

Example 3

A statistics professor is interested in the success rates of his students. In particular, anecdotal evidence seems to suggest that those students who are returning to college after an absence seem to be more successful in his courses. He collects data from his and his colleagues' students over one semester, specifically focusing on whether or not students were "returning" (defined for his purposes as those with two or more years away from school) and whether or not they were "successful" (earning a C or better).

He found that out of 184 "returning" students, 132 were successful, and of 429 "traditional" students, 256 were successful. Based on these data, is there evidence to support the professor's anecdotal evidence?

Example 4

A college adminstrator recently learned about a new strategy for encouraging faculty participation in academic committees. She is prepared to implement it, and would like to know if it truly changes faculty participation. She chooses a random sample of 10 faculty and records their current attendance at committee meetings. She then implements the new strategy she learned and records the attendance of these same faculty. The data she collected are as follows (in meetings attended per month):

Faculty:	A	B	C	D	E	F	G	H	I	J
Before	1	2	2	4	3	2	2	3	3	0
After	2	2	2	2	3	1	1	1	2	2

Is there evidence to say the new strategy increased faculty participation in their committees?