Inferences Based on Two Samples

In the following sections, our goal is to compare two population parameters to each other. We want to know the relationship between the parameters (if they are equal or if one is larger than the other). For example, I may want to compare the entrance exam results for men and women trying to be admitted to a prestigious prep school. In that scenario, we would look at the mean scores men and women and see if there is a difference between them.

Example: Since court liability awards vary over time, an insurance company wants to compare the mean level of current personal liability awards with those from one year earlier. Random samples of cases were selected from each year. The data is summarized below:

Year	Sample Size	Sample Mean	Sample Variance
Current	50	1.32	0.9734
Previous	55	1.04	0.7291

If we want to estimate the *true difference between the average amounts of awards over the two years*, our best point-estimate of that difference is $\bar{X}_1 - \bar{X}_2$ the difference of the sample means. If we also would like to form a confidence interval using the same format as we used in earlier sections, we need to know some properties of the sampling distribution of $\bar{X}_1 - \bar{X}_2$. To understand these properties note that $E(X \pm Y) = \mu_X \pm \mu_Y$ and $Var(X \pm Y) = \sigma_X^2 + \sigma_Y^2 \pm 2Cov(X,Y)$.

Properties of the Sampling Distribution of $\bar{X}_1 - \bar{X}_2$:

1. $\mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2$
2. $\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
3. If the sampled populations are normally distributed then so is the distribution of $\bar{X}_1 - \bar{X}_2$ regardless of the sample size.
4. If the sampled populations are not normal then we will need to have large sample sizes to ensure that we can approximate the distribution of $\bar{X}_1 - \bar{X}_2$ by the normal distribution.

***Note: We are assuming that the samples drawn are independent.

† †Note: If σ_1^2 & σ_2^2 are unknown, we may use their sample estimates as approximations as long as the sample sizes are large (> 30).

Using the properties above and the same structure as we used in previous sections, we can find a formula for the:

($1-\alpha$)100% **Confidence Interval for the True Difference Between the Population Means**

(Point Estimator) $\pm$ (Number of Standard Deviations)(Standard Error)

$$\bar{X}_1 - \bar{X}_2 \pm Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

* assuming that the two samples are independent.

Remember, that if we do not know the population standard deviation, but the sample size is large we can use the sample estimates. In other words,

$$(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\sigma_{(\bar{x}_1-\bar{x}_2)} = (\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

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

Now let's use the above formula we just found on the opening example regarding court awards:

Steps for constructing the **Confidence Interval for the True Difference between the Population Means:**

Step 1 Gather Data from Problem, Calculate $\bar{X}_1 - \bar{X}_2$, and Calculate $\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$.

Step 2 Find $Z_{\alpha/2}$

Step 3 Use the results from steps 2 and 1 to get the margin of error, E = $Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$

Step 4 Form $\left[(\bar{X}_1 - \bar{X}_2) - E, (\bar{X}_1 - \bar{X}_2) + E\right]$

To use the above formula we must have large sample sizes (>30), and the samples must be randomly drawn from independent populations.

Example 131: Since court liability awards vary over time, an insurance company wants to compare the mean level of current personal liability awards with those from one year earlier. Random samples of cases were selected from each year. The data is summarized below:

Year	Sample Size	Sample Mean	Sample Variance
Current	50	1.32	0.9734
Previous	55	1.04	0.7291

Let's do that one more time:

Example 132: Two samples concerning retention rates for first-year students at private and public institutions were obtained from the Department of Education's data base to see if there was a significant difference in the two types of colleges. Construct the 98% confidence interval for the true difference between the retention rates for the two types of colleges.

Type of College	Sample Size	Sample Mean	Sample Standard Dev.
Private	71	78.17	9.55
Public	32	84	9.88

Click here to see the solution:
http://online4.fiu.edu/SA1/Flash/McGuckian/STA3123/Example132/Example132.html

A Large Sample Hypothesis Test for the Difference between Two Population Means

Next, we will look at the method of testing hypotheses of the form: $H_0:(\mu_1-\mu_2)=D_0$ vs. $H_A:(\mu_1-\mu_2)<D_0$ (note: as usual the null hypothesis may have the symbols $\leq$ or $\geq$, and the alternative hypothesis may have > or $\neq$). The D here refers to the specified difference you are looking to detect. Many times we want to test that no difference exists. What will the value of D_0 be in those cases? (Answer: $D_0 = 0$)

Steps to test a hypothesis:

1. Express the original claim symbolically *
2. Identify the Null and Alternative hypothesis*
3. Record the data from the problem
4. Calculate the test statistic*
5. Determine your rejection region
6. Find the initial conclusion
7. Word your final conclusion

These are basically the same seven steps that we used to test hypotheses in previous sections; however, there are some minor changes that occur in the steps with asterisks by them. For example, the hypotheses will look as described above, and we will have a new test statistic, given by:

$$z = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Another important thing to consider are the test assumptions, which are as follows:

1. The samples are randomly selected and independent.

2. The sample sizes are both > 30.

Example 133: Tests of the effectiveness of Echinacea in preventing upper respiratory infections in children were conducted in double blind studies. "Days of Fever" was one criteria used to test the effectiveness of Echinacea. Among 337 children treated with Echinacea, the mean number of days with fever was 0.81, with a standard deviation of 1.5 days. Among 370 children given a placebo, the mean number of days of fever was 0.64 with a standard deviation of 1.16 days (JAMA Vol. 290). Use a 5% significance level to test the claim that Echinacea affects the number of days with fever. What is the p-value for the test?

Example 134: In a case in Ireland, a class action lawsuit was brought against the government alleging age discrimination. The ages of 31 randomly selected unsuccessful applicants for promotion yielded a mean age of 47 and a standard deviation of 7.2 years. The ages of 30 randomly selected successful applicants yielded an average age of 43.9 and a standard deviation of 5.9 years. Using a 5% significance level, test the claim that the government is discriminating based on age.

Click here to see the solution:
http://online4.fiu.edu/SA1/Flash/McGuckian/STA3123/example134/example134.html

Small Samples Scenario

What happens when the sample sizes are not greater than 30? There are two consequences of this:

1. We cannot assume the CLT can give us approximate normality (solution: We must know the samples are normally distributed to start with—actually as long as there are no outliers in the data and there isn't too much of a departure from normality we can apply this method).
2. The sample standard deviations may not be reliable estimates of their population counter parts (solution: We will need to use the t-distribution).

In order to use the t-distribution for this problem, we need to decide between two possible approaches. The first scenario assumes that the population variances are equal. Assuming the variances are equal will allow us to use a random variable that has an exact t-distribution. If we cannot assume this, we may use some other approach such as the **Welch-Satterthwaite method** which allows us to approximate the t-distribution for this kind of problem.

Example 135: Among 28 subjects using the Weight Watchers diet, the mean weight loss after a year was 3.0 pounds with a standard deviation of 4.9 pounds. Among 25 subjects using the Atkins diet, the mean weight loss after one year was 2.1 pounds with a standard deviation of 4.8 pounds. Construct a 95% confidence interval estimate of the difference between the mean weight losses for the two diets (assume weight loss is a normally distributed random variable). Does there appear to be a difference between the effectiveness of the two diets?

To do the above problem using the t-distribution, we will first assume that the two variances are equal. It is then reasonable to pool the two sample variances into one sample estimator of σ^2. We call this estimator the pooled sample estimator of σ^2 (notice it is just a weighted average with the degrees of freedom as the weights):

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

To form our confidence interval we will follow the following set of steps:

Step 1 Gather Data from Problem, Calculate $\overline{X}_1 - \overline{X}_2$, and Calculate $S_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$

Step 2 Find $t_{\alpha/2}$ using $n_1 + n_2 - 2$ as the degrees of freedom

Step 3 Find E = $t_{\alpha/2}\sqrt{\frac{S_p^2}{n_1}+\frac{S_p^2}{n_2}}$

Step 4 Form $\left[(\bar{X}_1-\bar{X}_2)-E,(\bar{X}_1-\bar{X}_2)+E\right]$

Example 136: Let's redo this problem without assuming equal variances. Among 28 subjects using the Weight Watchers diet, the mean weight loss after a year was 3.0 pounds with a standard deviation of 4.9 pounds. Among 25 subjects using the Atkins diet, the mean weight loss after one year was 2.1 pounds with a standard deviation of 4.8 pounds. Construct a 95% confidence interval estimate of the difference between the mean weight losses for the two diets (assume weight loss is a normally distributed random variable). Does there appear to be a difference between the effectiveness of the two diets?

To do the above problem without assuming equal variances, we will need to use an approximation method called the Welch-Satterthwaite method. To do this, let's modify the traditional four steps:

Step 1 Gather Data from Problem, Calculate $\bar{X}_1-\bar{X}_2$, Calculate $A=\frac{s_1^2}{n_1}$, Calculate $B=\frac{s_2^2}{n_2}$, and Calculate $\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$

Step 2 Find $t_{\alpha/2}$ using degrees of freedom = $\frac{(A+B)^2}{\frac{A^2}{n_1-1}+\frac{B^2}{n_2-1}}$ *(truncate to the nearest whole number)

Step 3 Find E = $t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$

Step 4 Form $\left[(\bar{X}_1-\bar{X}_2)-E,(\bar{X}_1-\bar{X}_2)+E\right]$

Note: Do not assume equal variances for the small sample size problems unless I specify. That means you will use the **Welch-Satterthwaite method** unless I say otherwise.

Small-Sample Hypothesis Test of $H_0:(\mu_1-\mu_2)=D_0, H_0:(\mu_1-\mu_2)\le D_0$, or $H_0:(\mu_1-\mu_2)\ge D_0$

We will use the same seven steps as always; however, we will need new t-test statistic. There will be two cases again based on our assumptions.

The equal variance case:

$$t = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - D_0}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}}, \; where \; S_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}, \text{ and d.f.=} n_1 + n_2 - 2$$

The Welch-Satterthwaite method (unequal variance case):

$$t = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}, \textit{ where } \text{t has degrees of freedom} = \frac{(A+B)^2}{\frac{A^2}{n_1 - 1} + \frac{B^2}{n_2 - 1}} \text{ and } A = \frac{s_1^2}{n_1} \text{ and } B = \frac{s_2^2}{n_2}$$

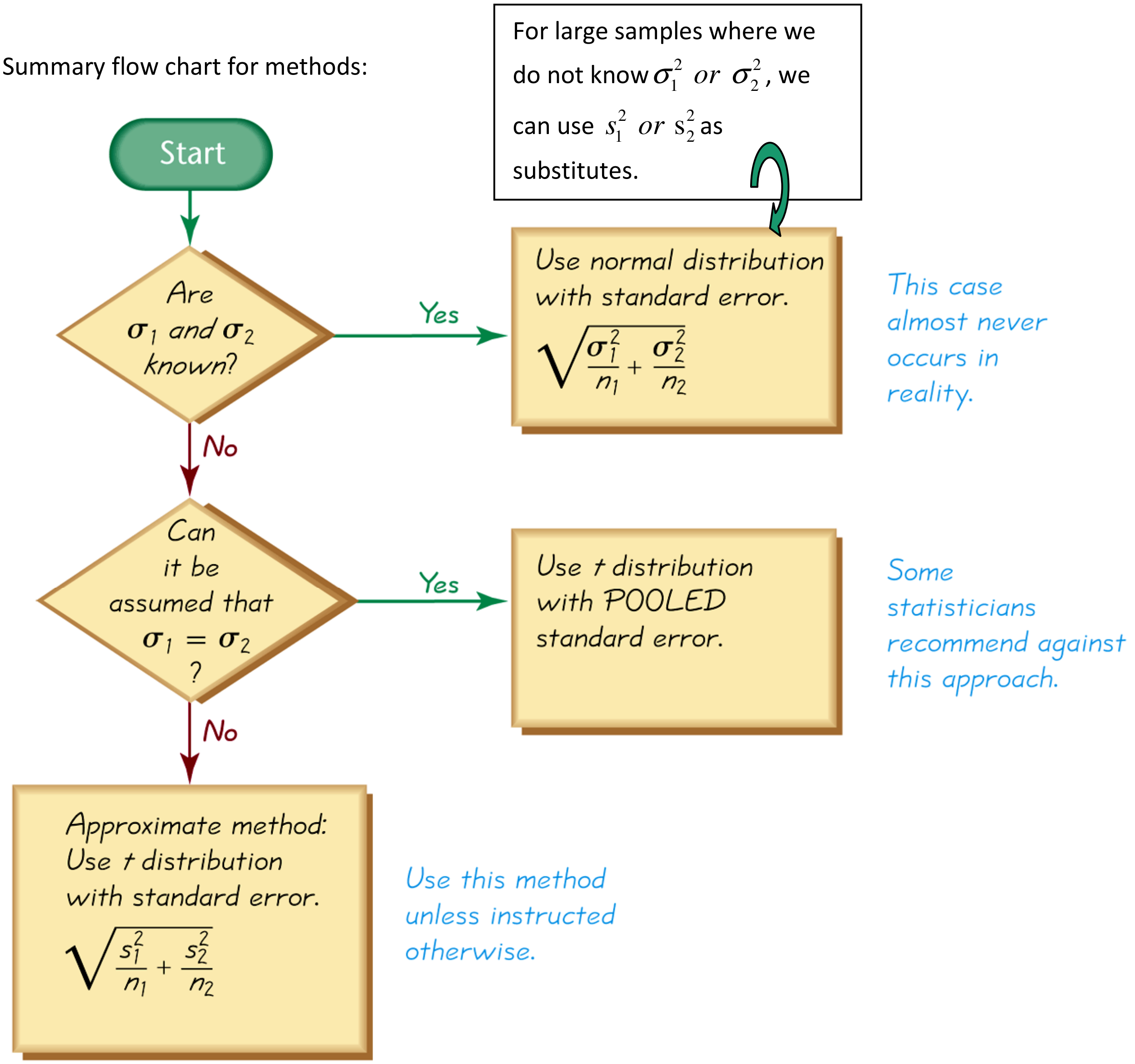
Example 137: When investigating the link between birth weights and IQ scores, researchers found that 28 subjects with birth weights less than 1000g had a mean IQ of 95.5 and a standard deviation of 16.0. For 27 subjects with normal birth weights, the mean IQ was 104.9 with a standard deviation of 14.5. At the 5% significance level, test the claim that low birth weight children have lower IQs than normal birth weight children (assume IQ scores are normally distributed and the variances are equal). Is there a possible problem with the assumption of equal variances here?

Steps:

1. Express the original claim symbolically *
2. Identify the Null and Alternative hypothesis*
3. Record the data from the problem
4. Calculate the test statistic*
5. Determine your rejection region
6. Find the initial conclusion
7. Word your final conclusion

Example 138: Redo the problem without assuming equal variances: When investigating the link between birth weights and IQ scores, researchers found that 28 subjects with birth weights less than 1000g had a mean IQ of 95.5 and a standard deviation of 16.0. For 27 subjects with normal birth weights, the mean IQ was 104.9 with a standard deviation of 14.5. At the 5% significance level, test the claim that low birth weight children have lower IQs than normal birth weight children (assume IQ scores are normally distributed).

Summary flow chart for methods:



Comparing Two Populations Means: Paired Difference Experiments

Comparing Two Population Means: Matched Pairs

Recall that our goal in the previous section was to be able to detect a difference between two population averages. The example problem below requires the same kind of analysis. We would like to be able to detect if the scores for students taking the FCAT math section improve after completing a series of FCAT prep classes.

Example 139: Below is a table of FCAT SSS developmental scores for a group of students who were struggling with math in the 3rd grade.

FCAT math scores for 8 struggling students		
Student	After Prep	Before Prep
1	290	275
2	275	270
3	380	370
4	260	245
5	340	325
6	270	260
7	280	270
8	215	200

Click here to see the solution: http://online4.fiu.edu/SA1/Flash/McGuckian/STA3123/ex139/ex139.html

We want to test the claim: The prep classes work to improve FCAT math SSS scores.

(In symbolic form: $\mu_{After} - \mu_{Before} > 0$)

Using the method from the previous section, we get a test statistic of: t = 0.467, which is not significant. This means we cannot reject the null ($\mu_{After} - \mu_{Before} \leq 0$), and we conclude that the prep classes are not effective at raising FCAT SSS scores. Does that seem correct when we look at the table of values above? Isn't it true that every student improved their math score after attending the prep classes? Then why did we get this result?

The answer is that the method we used is not valid here. In the previous section, our assumption was that the two samples were independent, but that is not true here. The two samples above were drawn from the same students. We gave them the FCAT, and then we gave them prep classes and retested the same students.

Okay, so we violated the assumptions—so what? Why does that affect our ability to detect the difference between the two FCAT performances? The answer lies in the quantity:

$S_p^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$ = 2,581.03 This is our pooled variance for the t-test we conducted above. What is this measuring in this case? It is measuring the variation of the FCAT scores between the students—not the difference between the scores each individual earned before and after prep. In other words it is **not** looking at the difference between before and after exams, but rather looking at the differences between student abilities. We know there is a lot of difference between individual student ability, but how is that affecting our hypothesis test?

We are trying to look at the distance between the average difference of before and after FCAT scores and zero in terms of the standard deviation (natural variation) between FCAT before and after scores. If the average distance between the before and after FCAT scores and zero is not much larger than (or is small) compared to the natural variation that occurs between retakes of the FCAT, we will conclude there is no significant improvement due to the prep classes.

Consider this simple example: Johnny scores a 170 on his FCAT math section, and Suzy scores a 460 on her FCAT math section. After taking the prep classes, Johnny retakes the FCAT and improves his grade to a 190, while Suzy jumps to a 490.

Suzy's score change = 490 – 460 = +30, Johnny's score change = 190 – 170 = +20

Average score change = + 25

What is the difference in their scores however?

Before difference = Suzy(460) – Johnny(170) = 290

After difference = Suzy(490) – Johnny(190) = 300

Average difference = 295

If you compare these two numbers, it is clear the average score change is quite small compared to the differences between Johnny's and Suzy's FCAT scores, but we do not want to compare these two quantities do we? No, we would want to compare the Average score change against the variation of the individual score changes not the variation of the individual scores.

We need to find a way to ignore these differences between different students' FCAT scores. These differences are not important to us. In other words, we know Suzy scores higher than Johnny already. We want to know if the prep course helps students do better regardless of how they did before. We do not want the differences in student ability to obscure the differences between before and after scores that we are interested in. If we can't block out the differences between students, we will never be able to detect the smaller differences that are occurring between before and after test scores. It would be like trying to hear the footsteps of a mouse running across a concert hall floor while a rock concert is being played in the same hall.

Blocking

We do have a very simple solution to this problem: we will run a one-sample t-test on the differences between before and after scores:

Subject	1	2	3	4	5	6	7	8
After	290	275	380	260	340	270	280	215
Before	275	270	370	245	325	260	270	200

Just subtract each subject's after and before scores…

Subject	1	2	3	4	5	6	7	8
After	290	275	380	260	340	270	280	215
Before	275	270	370	245	325	260	270	200
Difference	**15**	**5**	**10**	**15**	**15**	**10**	**10**	**15**

Then treat the row of differences as a single sample. We can get the average difference $\overline{X_d} = 11.875$, the standard deviation for the differences $S_d = 3.720$, and the number of differences $n_d = 8$. Then we can use the same test statistic we used for a one-sample t-test:

$$t = \frac{\overline{X_d} - \mu_d}{\frac{S_d}{\sqrt{n_d}}} \quad \text{with degrees of freedom = } n_d - 1$$

Where do we get μ_d ? That is the hypothesized value for the true average difference. This leads us to the question, "what will our claims look like?"

We will be conducting our hypothesis test using the following pair of competing claims:

$$H_0 : \mu_d \le D_0$$
$$H_A : \mu_d > D_0 \,*$$

*Of course, all the usual null/alternative pairings are possible

Now let's finish our example properly:

1. Express the original claim symbolically: $\mu_d > 0$ *
2. Identify the Null and Alternative hypothesis: $H_0 : \mu_d \le 0$, $H_A : \mu_d > 0$
3. Record the data from the problem: $\overline{X_d} = 11.875, S_d = 3.720, n_d = 8, \alpha = 0.05$
4. Calculate the test statistic: $t = \frac{\overline{X_d} - \mu_d}{\frac{S_d}{\sqrt{n_d}}} = \frac{11.875}{\frac{3.720}{\sqrt{8}}} \approx 9.029$
5. Determine your rejection region: t > 1.895
6. Find the initial conclusion: Reject the null, support the alternative
7. Word your final conclusion: The sample data support the claim that the prep classes are effective at improving student's FCAT SSS math scores.

*Note: our claim that the prep classes improve scores indicates that theoretically the 'after' exam will be better than the 'before' exam. This means if we form the difference d = After – Before, the differences should be positive, i.e. d > 0.

Some other examples where blocking would be necessary:

- A study of the effect of a Gulf hurricane on gas prices. If we poll stations across the country and look at the price of gas before and after the storm at those stations. Since prices vary from state to state, region to region, we would want to block out those differences by looking at every station twice (once before the storm and once after).

- A study wants to look at the difference between starting salaries of their male and female graduates. Since differences between majors and GPAs could affect salaries we would pair male and females in the study who have similar GPAs and majors. Then we would look at the differences in their starting offers.

- A study wants to compare the absorption rates of two different drugs for pain relief. Since there are so many possible differences between patients it would be best to make sure every study participant takes both drugs one after the other (allowing for the total elimination of the drug that was administered first). This will ensure the differences between the patients do not obscure the differences in the absorption rates between the drugs.

Example 140 The drug Depo-Provera is often used for patients who exhibit hypersexual behavior caused by traumatic brain injury (TMI). The results of a study conducted on 8 patients who exhibited such behavior is included below. Each patient had their testosterone levels taken before treatment with Depo and after treatment. At the 1% significance level, test the claim that the drug is effective at reducing testosterone levels in TMI patients.

Patient	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Pretreatment	849	903	890	1092	362	900	1006	672
After Depo	96	41	31	124	46	53	113	174
Differences	**753**	**862**	**859**	**968**	**316**	**847**	**893**	**498**

Next, we will look at the confidence interval for paired differences data.

Confidence Interval for Paired Differences:

$$\left[\overline{X_d} - t_{\alpha/2}\frac{S_d}{\sqrt{n_d}}, \overline{X_d} + t_{\alpha/2}\frac{S_d}{\sqrt{n_d}}\right]$$

Required assumptions:

1. We must assume that the sample was chosen randomly from the target population.
2. We must assume that the population of differences has a normal distribution.

Example 141 Use the data from the Depo-Provera study to form a 98% confidence interval for μ_d , then interpret the results. Are they consistent with the results found in the hypothesis test we conducted?

Example 142 Use the data from the SAT problem that opened this section to form a 90% confidence interval for μ_d the true average difference between before and after SAT scores and interpret the results.

Click here for the solution: http://online4.fiu.edu/SA1/Flash/McGuckian/STA3123/ex142/ex142.html

Comparing Two Population Proportions Independent Sampling

Example 143: In a randomized controlled trial in Kenya, insecticide treated bed-nets were tested as a way to reduce malaria. Among 343 infants who used the bed-nets, 15 developed malaria. Among 294 infants not using bed-nets, 27 developed malaria (JAMA vol. 291, No. 21). Use a 0.01 significance level to test the claim that the incidence of malaria is lower for infants who use the bed-nets. Do the bed-nets seem to work?

Click here to see the solution: http://online4.fiu.edu/SA1/Flash/McGuckian/STA3123/ex143/ex143.html

To answer this question we need to know what quantities to compare. Let's look at what we have here:

	Bed-Nets	No Nets
X	15	27
N	343	294
$\hat{p}$	0.044	0.092

Clearly, for this sample, bed-nets resulted in a lower infection rate, but is this difference just a coincidence? Maybe this could have happened by chance—after all the number of mosquitoes each group was exposed to was not controlled.

What we can do is express the difference between these two sample proportions $\left(\hat{p}_1 - \hat{p}_2\right)$ in terms of the standard deviation of the sampling distribution of their difference. Then if the sample sizes are large enough*, we can express the probability that we would observe this difference $\left(\hat{p}_1 - \hat{p}_2\right)$ by random chance given that they are really the same (that is the assumption under the null hypothesis).

*Sample sizes are large enough when $\hat{p} \pm 3\sqrt{\frac{\hat{p}\hat{q}}{n}}$ is entirely captured inside [0, 1]. Another rule of thumb that is used is: Samples sizes are large enough if

$$\begin{array}{l} n_1\hat{p}_1 \geq 15 \text{ and } n_1\hat{q}_1 \geq 15 \\ n_2\hat{p}_2 \geq 15 \text{ and } n_2\hat{q}_2 \geq 15 \end{array}$$

So our point estimator will be: $\left(\hat{p}_1 - \hat{p}_2\right)$

The standard error of its sampling distribution will be:

$$\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}} \approx \sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}, where\ \hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

Our competing hypotheses will be: $H_0 : \left(p_1 - p_2\right) = 0$ Vs. $H_A : \left(p_1 - p_2\right) \neq 0$ (<, > are possible also)

Our test statistic will be: $z \approx \dfrac{\left(\hat{p}_{nets} - \hat{p}_{no-nets}\right)}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}, where\ \hat{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$

Now let us determine if we can support the claim that the bed-nets are effective in preventing malaria infection.

1. Express the original claim symbolically: $p_{nets} < p_{no-nets}$
2. Identify the Null and Alternative hypothesis: $H_0 : (p_{nets} - p_{no-nets}) = 0$ $H_A : (p_{nets} - p_{no-nets}) < 0$
3. Record the data from the problem: $p_{nets} = 0.044, p_{no-nets} = 0.092, \alpha = 0.01$
4. Calculate the test statistic:

$$z \approx \frac{(\hat{p}_{nets} - \hat{p}_{no-nets})}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{-0.048}{\sqrt{0.066(0.934)\left(\frac{1}{343} + \frac{1}{294}\right)}} \approx -2.434$$

5. Determine your rejection region: Z < -2.326
6. Find the initial conclusion: Reject the null, support the alternative
7. Word your final conclusion: The sample data support the claim that the bed-nets are effective at preventing malaria.

Example 144 In 1995, the American Cancer Society randomly sampled 1500 adults of which 555 smoked. In 2005, they surveyed 1750 adults and found that 578 of them smoked. At the 5% level of significance test the claim that the proportion of smokers in the population decreased over the ten year period. What is the p-value for the test?

Let's now look at the confidence interval for the difference between two proportions:

Confidence Interval for $(p_1 - p_2)$ **:**

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}} \approx (\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

Example 145 Randy Stinchfield of the University of Minnesota studied the gambling activities of public school students in 1992 and 1998 (*Journal of Gambling Studies*, Winter 2001). His results are reported below. Form a 99% confidence interval to estimate the true difference between the proportion of public school students who gambled in 1992 and 1998. Does there seem to be a difference?

	1992	1998
Survey *n*	21,484	23,199
Number who gambled	4,684	5,313
Proportion who gambled	.218	.229

Click here to see the solution: http://online4.fiu.edu/SA1/Flash/McGuckian/STA3123/ex145/ex145.html

Example 146 In 1995, the American Cancer Society randomly sampled 1500 adults of which 555 smoked. In 2005, they surveyed 1750 adults and found that 578 of them smoked. Form a 90% confidence interval for the difference between the proportion of smokers in 1995 and 2005. Has the proportion decreased over the ten year period?

Before we begin the next topic, let's learn how to use the F table found in the back of you textbook to find critical values. The F-distribution is the ratio of two chi-square distributions with degrees of freedom ν_1 and ν_2, respectively, where each chi-square has first been divided by its degrees of freedom. When we use the F tables, the degrees of freedom will be just $n_1 - 1$ and $n_2 - 1$ for the two sample sizes in the problem. We will consider one of these to be the numerator degrees of freedom and one to be the denominator degrees of freedom. This will depend how we set up the ratio in the given problem. For now let's just assume $n_1 - 1$ is the numerator degree of freedom and $n_2 - 1$ is the denominator's degree of freedom. We will also need an alpha value in each problem. This α value will be the amount of area that will be found in the upper tail of our F-distribution beyond our critical value.

Example 147 Find the critical value $f_{n_1-1,n_2-1,\alpha} = f_{30,40,0.01}$

Example 148 Find the critical value $f_{n_1-1,n_2-1,\alpha} = f_{6,24,0.025}$

Comparing Two Population Variances: Independent Sampling

There are a lot of situations where we want to know if the populations have the same variances. In fact, we just studied hypothesis testing methods for comparing two means from independent populations. In that section when the sample sizes were small we needed to assume either that the population variances were equal or that they weren't, but with the test we learn here today, we won't have to assume. At other times, we will actually want to compare the variances of two groups instead of just performing the test to know if we can go forward with a test of the means. The following is an example of the latter type of problem:

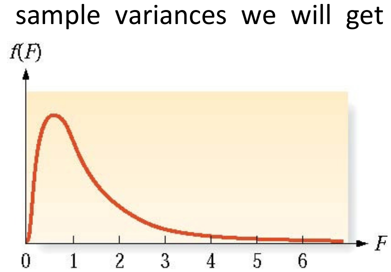
Example 149: Disney is comparing two methods of receiving customers at its City Halls in the Magic Kingdom and Disney Land. They are deciding if using one long line (used in Disney Land) is better than allowing people to line up in separate lines for each teller (used in Magic Kingdom). They collected waiting time data (measured in minutes) for both locations:

$$\text{Disney Land: } n = 41, \overline{X} = 5.15, S = 0.48$$
$$\text{Magic Kingdom: } n = 61, \overline{X} = 5.15, S = 1.23$$

Test the claim at the 5% significance level that Disney Land "City Hall" lines have a smaller variance than the lines at "City Hall" in Magic Kingdom.

To compare variances we will use an F-test. The F-distribution is the ratio of two chi-square distributions with degrees of freedom ν_1 and ν_2, respectively, where each chi-square has first been divided by its degrees of freedom. It turns out that if a random variable is normally distributed and has population variance σ^2 the quantity $\frac{S^2(n-1)}{\sigma^2} \sim \chi^2_{n-1}$ (is chi-squared with degree of freedom n-1). This means that if we divide our two sample variances we will get a random variable that has an F-distribution:

$$\frac{S_1^2}{S_2^2} \sim F$$



Note, actually we are forming the ratio of two chi-squared random variables divided by their degrees of feedom: $\dfrac{\frac{S_1^2(n_1-1)}{\sigma^2}/(n_1-1)}{\frac{S_2^2(n_2-1)}{\sigma^2}/(n_2-1)} = \dfrac{\frac{S_1^2}{\sigma^2}}{\frac{S_2^2}{\sigma^2}} = \dfrac{S_1^2}{S_2^2} \sim F$ (Why isn't it σ_1^2 & σ_2^2 in the denominators? Answer: Our null hypothesis will include the assumption that $\sigma_1^2 = \sigma_2^2$).

To conduct our test we will use the ratio of our two sample variances: $\dfrac{S_1^2}{S_2^2} \sim F_{n_1-1, n_2-1}$ *

*Because of the tables we have in our textbook, we will always put the larger sample variance on top. One way to do this is to always label the sample with the larger sample variance as representing population 1.

Our competing hypotheses will be: $H_0: \dfrac{\sigma_1^2}{\sigma_2^2} \le 1$ Vs. $H_A: \dfrac{\sigma_1^2}{\sigma_2^2} > 1$ (= vs. $\neq$ are also possible).

Now let's work our example:

1. Express the original claim symbolically: $\dfrac{\sigma^2_{MagicKingdom}}{\sigma^2_{DisneyLand}} > 1$ (because we claimed Disney Land's variance was the smaller of the two)
2. Identify the Null and Alternative hypothesis: $H_0: \dfrac{\sigma^2_{MagicKingdom}}{\sigma^2_{DisneyLand}} = 1$ $H_A: \dfrac{\sigma^2_{MagicKingdom}}{\sigma^2_{DisneyLand}} > 1$
3. Record the data from the problem: Disney Land: $n = 41, \overline{X} = 5.15, S = 0.48$ Magic Kingdom: $n = 61, \overline{X} = 5.15, S = 1.23$
4. Calculate the test statistic: $\dfrac{(1.23)^2}{(0.48)^2} = 6.566$
5. Determine your critical value and rejection region: F > 1.64 (see the F-Tables in your book)

Steps to determine the critical value for an F-test:

a. Determine the number of tails and alpha (divide alpha in half if two-tails)
b. Determine the table to use based on step a.
c. Use the numerator degree of freedom for the top row of table and the denominator degree of freedom for the left column of the table.

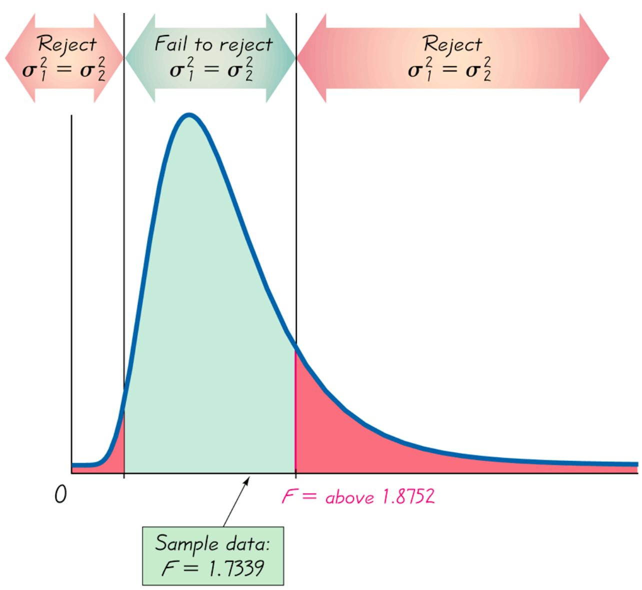
6. Find the initial conclusion: Reject the null, support the alternative
7. Word your final conclusion: The sample data support the claim that the waiting times at Disney Land have less variance than the waiting times at Magic Kingdom.

Assumptions for the above test:

1. The samples are random and independent
2. Both populations are normally distributed (the test is very sensitive to violations of this assumption)

Example 150 Coke Versus Pepsi, the weights (in pounds) of samples of regular Coke and regular Pepsi have been summarized below. Sample statistics are shown. Use the 0.05 significance level to test the claim that the weights of regular Coke and the weights of regular Pepsi have the same standard deviation.

	Regular Coke	**Regular Pepsi**
Sample size	**36**	**36**
Mean	**0.81682**	**0.82410**
Standard deviation	**0.007507**	**0.005701**



Confidence Interval for $\frac{\sigma_1^2}{\sigma_2^2}$: We won't form this interval in the course, but it is provided here for your reference.

A $(1-\alpha)\times 100\%$ Confidence Interval for σ_1^2 / σ_2^2

$$\left(\frac{s_1^2}{s_2^2}\right)\left(\frac{1}{F_{U,\alpha/2}}\right) < \left(\frac{\sigma_1^2}{\sigma_2^2}\right) < \left(\frac{s_1^2}{s_2^2}\right)\left(\frac{1}{F_{L,\alpha/2}}\right)$$

where $F_{L,\alpha/2}$ leaves $(\alpha/2)\%$ of the distribution in a lower tail
and $F_{U,\alpha/2}$ leaves $(\alpha/2)\%$ of the distribution in an upper tail.
Degrees of freedom are $n_1 - 1$ in the numerator and
$n_2 - 1$ in the denominator.

Another way to write the above is:

$\left[\frac{S_1^2}{S_2^2}\left(\frac{1}{F_{L,\alpha/2}}\right), \frac{S_1^2}{S_2^2} F_{U,\alpha/2}\right]$ Where $F_{L,\alpha/2}$ places alpha/2 area in the upper tail of a F_{n_1-1,n_2-1} distribution, and $F_{U,\alpha/2}$ places alpha/2 area in the upper tail of a F_{n_2-1,n_1-1} distribution. Note: this comes from the following relationship, $\frac{1}{f_{n_1-1,n_2-1,1-\alpha/2}} = f_{n_2-1,n_1-1,\alpha/2}$