Chapter 9: Inferences from Two Samples

§9.1 Introduction

We will be continuing the discussions of confidence intervals and hypothesis testing that we covered in Chapters 7 and 8 for two samples. There are many times that we want to compare subpopulations within a single population and confidence intervals and hypothesis testing on for two samples will give us the means by which to do this.

Creating confidence intervals for two proportions is pretty straight-forward and will be a fairly direct correlation from one sample proportions. Hypothesis testing requires a mild tweak in our process, but again the entire process will fall in place from what we learned about hypothesis testing for one-sample proportions.

Where we will run into a little complication is with the difference between independent and dependent samples for testing claims about means. Independent samples means that we can't make pairing from one data point to another. Examples of dependent data are time series data (same object is measured twice) or data gathered for related experimental units (such as data gathered for twins). Most data will actually be from independent samples but there are times when dependent samples are used.

§9.2 Inferences About Two Proportions

A confidence interval for two proportions is very straight-forward. Once we have identified the fact that we have two proportions then we only tweaking the margin of error slightly. The margin of error must take into account the two sample proportions. It is "kind of like" averaging the two standard errors for the two samples under the radical.

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

Interval: $(\hat{p}_1 - \hat{p}_2) \pm E$

Example: #14 p. 381 in Brase/Brase's 9th edition, *Understandable Statistics* Most married couples have two or three personality preferences in common (see the reference to the Myers-Briggs Test in #13). Myers used a random sample of 375 married couples and found that 132 has three preferences in common. Another random sample of 571 couples showed that 217 had two personality preferences in common. Let p_1 be the population proportions of all married couples with 3 preferences in common and p_2 be the population proportion of all married couples with 2 preferences in common.

a) Find p-hat_1

b) Find p-hat_2

c) Compute the margin of error for the difference between population 1 & 2 at the 90% confidence level.

d) Use your TI-83/84 to calculate the 90% CI for the difference between population proportions for pop 1 & pop 2.

Your Turn
Example: A football team passed the ball 247 times in the 2000 season with 83 completed passes. In the 2001 season the same team passed the ball 258 times with 89 completed passes. Find a 95% CI for the difference in the proportions of completed passes for the 2000 and 2001 seasons. Compute by hand and then check using your calculator.

Differences between $p_1 - p_2$

ASSUMPTIONS

1) Independent samples
2) Normal assumptions: $np \geq 5$ and $nq \geq 5$

INFORMATION NEEDED

$\hat{p}_1$, $\hat{q}_1$, x_1 , n_1 **Note**: x_1 is # successes, equivalent to p-hat times n

$\hat{p}_2$, $\hat{q}_2$, x_2 , n_2

p_1 & p_2 known, but since they usually aren't known we need something called the pooled best estimate for p, or p-bar.

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2} \qquad \bar{q} = 1 - \bar{p}$$

HYPOTHESES

The hypotheses are based upon the differences being positive, negative or zero

H_0: $p_1 - p_2 \geq 0$ or H_0: $p_1 \geq p_2$

$H_{A \text{ or } 1}$: $p_1 - p_2 < 0$ $H_{A \text{ or } 1}$: $p_1 < p_2$

H_0: $p_1 - p_2 \leq 0$ or H_0: $p_1 \leq p_2$

$H_{A \text{ or } 1}$: $p_1 - p_2 > 0$ $H_{A \text{ or } 1}$: $p_1 > p_2$

H_0: $p_1 - p_2 = 0$ or H_0: $p_1 = p_2$

$H_{A \text{ or } 1}$: $p_1 - p_2 \neq 0$ $H_{A \text{ or } 1}$: $p_1 \neq p_2$

CRITICAL VALUE

$z_{\alpha/2}$

TEST STATISTIC

$$z = \frac{\hat{p}_1 - \hat{p}_2 - 0}{\sqrt{\bar{p}\bar{q}/n_1 + \bar{p}\bar{q}/n_2}}$$

USING YOUR CALCULATOR

1) Perform a 2-Prop ZTest

Input: x_1, n_1, x_2, n_2, Alternate hypothesis

Output: alt. hyp. , z-test stat, p-value of test stat (recall if two tail it is twice p-value), p-hat1, p-hat2 & p-hat which is the pooled p, s's and n's

**Recall that the CI method is not valid for conducting a hypothesis tests for proportions.

Example: Using the problem from above. Myers used a random sample of 375 married couples and found that 132 have three preferences in common. Another random sample of 571 couples showed that 217 had two personality preferences in common. Let p_1 be the population proportions of all married couples with 3 preferences in common and p_2 be the population proportion of all married couples with 2 preferences in common. Test the claim that the proportion with 3 preferences is less than those with 2 preferences in common at the 95% confidence level. Do the problem by hand and then check with your calculator.

Your Turn

Example: A football team passed the ball 247 times in the 2000 season with 83 completed passes. In the 2001 season the same team passed the ball 258 times with 89 completed passes. The coach claims that there is no difference in pass completion between the two seasons. Test the coach's claim at the 90% confidence level.

§9.3 Inferences About Independent Means

Confidence intervals for the difference of means look very similar to confidence intervals for one mean – the only difference is the standard error. Again, much like the proportions it looks like "an average of the standard errors" of each of the populations. Do note that the **variance** is used since everything is under the radical. Just as in one populations, you must decide if the standard deviation of the population is known or unknown to decide whether the distribution of the difference of the sample means follows a standard normal or a t-distribution! With the t-distribution we have an additional complication – that of a pooled variance. If there is an assumption made that the standard deviations of the population are known to be equal, even though they are unknown then a pooled variance is used. When I test on this information, I do not test on pooled variances – I believe you already have enough to worry about in decision making, but I do believe that you should be aware of all the possibilities out there.

Means with σ known

$$E = z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Interval: $(\bar{x}_1 - \bar{x}_2) \pm E$

Means with σ unknown – not pooled (*meaning no assumption about σ_1 & σ_2 are made***)**

$$E = t_{nsmall-1,\, \alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Interval: $(\bar{x}_1 - \bar{x}_2) \pm E$

Note: *There is also another more complicated computation for the degrees of freedom, one that our calculators use instead of the simplified smaller degrees of freedom.*

Means with σ unknown, but assumed equal – pooled

$$E = t_{n1+n2-2,\, \alpha/2} \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Interval: $(\bar{x}_1 - \bar{x}_2) \pm E$

Without further explanation, let's do some examples.

Example 1: #8 p. 378 from Brase/Brase's 9th edition, *Understandable Statistics*
Inorganic phosphorous is a naturally occurring element in all plants and animals, with concentrations increasing progressively up the food chain. Geochemical surveys take soil samples to determine phosphorous content (in ppm). A high phosphorous content may or may not indicate an ancient burial site, food storage site, or even a garbage dump. Independent random samples from two regions gave the following phosphorous measurements (in ppm). Assume the distribution of phosphorous is mound-shaped and symmetric for these two regions.

Region 1:	855	1550	1230	875	1080	2330	1850	1860	2340
	1080	910	1130	1450	1260	1010			
Region 2:	540	810	790	1230	1770	960	1650	860	890
	640	1180	1160	1050	1020				

a) Which should be used, a z-interval or a t-interval? Why?

b) What are the degrees of freedom?

c) Find the critical value for an 80% CI using your calculator.

d) Write out the margin of error calculations for a confidence level of 80% for the difference between the mean in R1 & R2.

e) Using your TI-83/84 find an 80% CI for $\mu_1 - \mu_2$

Note: *We can use the data to compute the CI without even finding the mean and standard deviation with the TI 83/84 calculators*

f) Explain what the confidence interval means in context of this problem. Use information about the interval containing all positive, all negative or both positive and negative.

g) At the 80% level of confidence, is one region more interesting than the other from a geochemical perspective? Use the information stated in f).

Example 2: #18 p. 381 in Brase/Brase's 9th edition, *Understandable Statistics* "Parental Sensitivity to Infant Cues: Similarities and Differences Between Mothers and Fathers," by MV Graham (*Journal of Pediatric Nursing,* Vol. 8, No. 6), reports a study of parental empathy for sensitivity cues and baby temperament (higher mean scores means more empathy). Let x_1 be a random variable that represent the score of a mother on an empathy test (in regards to her baby). Let x_2 be the empathy score of a father. A random sample of 32 mothers gave a sample mean of 69.44. Another random sample of 32 fathers gave a mean of 59. Assume the population standard deviation is 11.69 for mothers and 11.60 for fathers.

a) Which should be used, a z-interval or a t-interval? Why?

b) What is the correct critical value for a 99% CI?

c) Compute the margin of error for the difference between mother's mean empathy score and father's mean empathy score.

d) Use your calculator to compute the 99% CI for the difference in mother's and father's mean empathy scores.

e) What does the confidence interval tell about the relationship between empathy scores for mothers and fathers at the 99% confidence level?

f) Do you think we could draw the same conclusions as in e) for the 95% confidence level? Why or why not?

Differences between $\mu_1 - \mu_2$ when σ is known

ASSUMPTIONS

1) If normally distributed, then no assumptions are necessary
2) If not able to make a normal assumption, then $n_1 \geq 30$ & $n_2 \geq 30$
3) Independent samples

INFORMATION NEEDED

$\overline{x}_1$

$\overline{x}_2$

σ_1

σ_2

HYPOTHESES

The hypotheses are based upon the differences being positive, negative or zero

H_0: $\mu_1 - \mu_2 \geq 0$ or H_0: $\mu_1 \geq \mu_2$

$H_{A \text{ or } 1}$: $\mu_1 - \mu_2 < 0$ $H_{A \text{ or } 1}$: $\mu_1 < \mu_2$

H_0: $\mu_1 - \mu_2 \leq 0$ or H_0: $\mu_1 \leq \mu_2$

$H_{A \text{ or } 1}$: $\mu_1 - \mu_2 > 0$ $H_{A \text{ or } 1}$: $\mu_1 > \mu_2$

H_0: $\mu_1 - \mu_2 = 0$ or H_0: $\mu_1 = \mu_2$

$H_{A \text{ or } 1}$: $\mu_1 - \mu_2 \neq 0$ $H_{A \text{ or } 1}$: $\mu_1 \neq \mu_2$

CRITICAL VALUE

$z_{\alpha/2}$

TEST STATISTIC

$$z = \frac{\overline{x}_1 - \overline{x}_2 - 0}{\sqrt{\sigma^2_1 / n_1 + \sigma^2_2 / n_2}}$$

USING YOUR CALCULATOR

1) Enter sample 1 into register 1 & sample 2 into register 2
2) Perform a 2-Sample ZTest on the differences in L3
   Input: Use Data(Stats you'll need x-bars and n's), σ1,σ2, L1, L2, Freq are 1, and choose the correct alternative hypothesis
   Output: alt. hyp. , z-test stat, p-value of test stat (recall if two tail it is twice p-value),x-bars, s's and n's

Example: Let's do an example from your book, Brase's *Understandable Statistics*, 9^{th} edition p. 471 #10.

Based on information from *Rocky Mountain News*, in a random sample of 12 winter days in Denver the mean pollution index was 43. Previous studies have shown that σ is 21. A similar study in a suburb of Denver, called Englewood, a sample of 14 winter days showed a mean pollution index of 36 and previous studies have shown that σ is 15. If the pollution indexes for both Denver and Englewood are assumed to be normally distributed, at the 99% confidence level test the claim that there is no difference between the mean pollution index in Denver and Englewood.

Differences between $\mu_1 - \mu_2$ when σ is unknown

ASSUMPTIONS

1) If normally distributed, then no assumptions are necessary
2) If not able to make a normal assumption, then $n_1 \geq 30$ & $n_2 \geq 30$
3) Independent samples
4) σ is unknown

INFORMATION NEEDED

$\overline{x_1}$, s_1 , n_1

$\overline{x_2}$, s_2 , n_2

HYPOTHESES

The hypotheses are based upon the differences being positive, negative or zero

H_0: $\mu_1 - \mu_2 \geq 0$ or H_0: $\mu_1 \geq \mu_2$

$H_{A \text{ or } 1}$: $\mu_1 - \mu_2 < 0$ $H_{A \text{ or } 1}$: $\mu_1 < \mu_2$

H_0: $\mu_1 - \mu_2 \leq 0$ or H_0: $\mu_1 \leq \mu_2$

$H_{A \text{ or } 1}$: $\mu_1 - \mu_2 > 0$ $H_{A \text{ or } 1}$: $\mu_1 > \mu_2$

H_0: $\mu_1 - \mu_2 = 0$ or H_0: $\mu_1 = \mu_2$

$H_{A \text{ or } 1}$: $\mu_1 - \mu_2 \neq 0$ $H_{A \text{ or } 1}$: $\mu_1 \neq \mu_2$

CRITICAL VALUE

$t_{\alpha,\, n1-1 \text{ or } n2-1}$ the df is the small of $n_1 - 1$ or $n_2 - 1$

***Note**: Statistical software and your TI-Calculator have a more complicated method of computing degrees of freedom. See exercise #21 for details. Also note that the degrees of freedom for a pooled test is different as well. DF for pooled is $n_1 + n_2 - 2$

TEST STATISTIC

Non-Pooled when no assumptions are made about σ_1 & σ_2

$$t = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{s^2_1 / n_1 + s^2_2 / n_2}}$$

***Note:** This is assuming that the population variances are not the same. If the population variances can be assumed to be the same the standard error (denominator) looks different. The following is for that pooled instance.

Pooled when $\sigma_1 = \sigma_2$ is assumed

$$t = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{(n_1 - 1)s^2_1 + (n_2 - 1)s^2_2 / n_1 + n_2 - 2}}$$

USING YOUR CALCULATOR

1) Enter sample 1 into register 1 & sample 2 into register 2
2) Perform a 2-Sample TTest on the L1 & L2 registers
   Input: Use Data(Stats you'll need x-bars, s's and n's), L1, L2, Freq are 1, and choose the correct alternative hypothesis, for pooled use POOLED Yes & Non-Pooled use No
   Output: alt. hyp. , t-test stat, p-value of test stat (recall if two tail it is twice p-value),df, x-bars, s's and n's

Example: There are two options for treatment for arthritis. One is a magnet and the other is considered a "sham" by the manufacturers of the magnets. To test the claim that the magnets are a better treatment than the "sham," a study is done and measurements are taken. For the magnet a sample of 20 yields a mean of 0.49 and a standard deviation of 0.96 (we aren't told what these are measurements of, just that they are indicators of relief and the higher the number the greater the relief). For the "sham" a sample of 20 yields a mean of 0.44 and a standard deviation 1.4.

§9.4 Inferences About Two Means: Dependent

This section is about testing the differences between paired observations. This is **dependent** data – the data that we mentioned in the discussion in §8.4 but did not discuss further.

Some Examples of Paired Data:

- Height measurements taken for individuals and by the individuals
  - *Paired by individual*
- IQ's of identical twins
  - *Paired by twin set*
- Rainfall in cities across the Bay Area in January and in June
  - *Paired by Bay Area City*
- Death and Birth Rates for Counties in State
  - *Paired by County*
- Percentage of males in a wolf pack in the summer and winter for specific packs
  - *Paired by pack*
- Number of fish caught on shore vs. in a boat during certain months
  - *Paired by month*
- Shoe sizes for the left and right foot
  - *Paired by individual*

There are many applications of this type of paired data analysis. We need to be able to establish a clear *matching link* in order to establish that paired data is to be used. To be paired data there needs to be a **clear, natural way of matching characteristics**.

ASSUMPTIONS

1) If normally distributed, then no assumptions are necessary
2) If not able to make a normal assumption, then $n \geq 30$

HYPOTHESES

The hypotheses are based upon the differences being positive, negative or zero

H_0: $\bar{d}_0 \geq 0$
$H_{A \text{ or } 1}$: $\bar{d}_0 < 0$

H_0: $\bar{d}_0 \leq 0$
$H_{A \text{ or } 1}$: $\bar{d}_0 > 0$

H_0: $\bar{d}_0 = 0$
$H_{A \text{ or } 1}$: $\bar{d}_0 \neq 0$

CRITICAL VALUE

$t_{\alpha,\, n-1}$ where n= # of paired observations

TEST STATISTIC

$$t = \frac{\bar{d} - \bar{d}_0^{\,*}}{s_d / \sqrt{n}}$$

*Note: This value is zero, so the test statistic just looks like: $(\bar{d}\cdot\sqrt{n})/s_d$

d-bar is found by computing the difference between the observed values for sample 1 and sample 2 (or whichever difference is desired) and finding the mean.

s_d is the standard deviation of the differences

USING YOUR CALCULATOR

1) Enter sample 1 into register 1 & sample 2 into register 2
2) Go into L3 heading and key in L1 – L3 to get the differences
3) Perform a t-test on the differences in L3
   Input: $\mu = 0$, Use Data and choose the correct alternative hypothesis
   Output: μ , t-test stat, p-value of test stat (recall if two tail it is twice p-value),d-bar (will be notated as x-bar), s_d (will be notated as s_x)and n

Example: The following are data for the heights of individuals as reported by the individual and as measured by a medical professional. The question is at the 95% Confidence Level is there a difference between the given height of a individual and the measured height of an individual. (Triola, *Elements of Statistics*, 3rd Edition p. 432)

Report	53	64	61	66	64	65	68	63	64	64
Measure	58.1	62.7	61.1	64.8	63.2	66.4	67.6	63.5	66.8	63.9

a) Give the hypotheses

b) Give the critical value

c) Give the test statistic computation

d) Use your calculator to conduct the test & state the conclusion