**SAN JOSÉ STATE UNIVERSITY**

**College of Social Work**
**S. W. 242**
**Spring 2009**
**Edward Cohen**

**Next week:**
***Meet at ~~lab~~ classroom for *t*-test lab assignments (#2a and 2b)**
**~~*We'll review material for exam, then distribute take home exam~~**
***Introduction to ANOVA**

**Week 7: March 6, 2008**

- **Confidence Intervals**
- **Lab #1 (Chi-square)**
- **Comparing means of two groups: *t*-tests**
- **Final Paper preparation**
- **Take home exam due at beginning of class March 13**

Concepts you should know:

- $H_O$ -- The null hypothesis
- $H_A$ -- The alternative hypothesis
- Confidence interval of the mean
- Standard error of the mean
- Dichotomous variable
- Assumptions for using *t*-test
- Equality of variances
- Independent groups *t*-test
- Dependent group (or paired sample) *t*-test
- Degrees of freedom (and the *t*-test)
- Confidence interval of the difference between means

I. Review:

| **If…** | **The independent variable is :** | **And the dependent variable is:** | **Then the statistical test is:** |
|---|---|---|---|
| | Categorical | Categorical | Chi square |
| | Nominal (dichotomous) | Continuous | *t* - test |

II. *t*–test: comparing means of 2 groups; independent variable is nominal (dichotomous—the 2 groups), and dependent variable is continuous

**Two types:** Independent groups *t*-test
Dependent groups (or paired samples) *t*-test

A. **Independent sample *t*-test** Comparing the means of two independently chosen samples

**Example:**
$\mathbf{H_A}$: SJSU students receiving post-grad test preparation will score higher in the LCSW exam compared to those SJSU students who did not receive post-grad test preparation

**Eight steps to testing hypothesis:**

1. What is the independent variable, and its level of measurement? Group (test prep vs. no test prep); nominal (dichotomous)
2. What is the dependent variable , and its level of measurement? LCSW test score; continuous
3. What's the Null Hypothesis? There is no difference in LCSW test scores between the two groups.
4. Is the alternative hypothesis stated above one-tailed or two-tailed? There is a difference in the test scores between the two groups (two-tailed); OR, The group receiving test preparation will score higher in the LCSW exam than the group not receiving test prep (one-tailed)
5. Statistical test & alpha level: Independent sample *t*-test, alpha $\leq$ .05
6. Review SPSS table of results
7. Describe results & decision to accept or reject Null –using standard reporting of statistic: (*t* = __, *df* = ___, *p* = ___)
8. Discuss results

B. **Dependent (or Paired) sample *t*–test** Comparing the means of two *related* samples, or more typically, the *same sample measured twice*

**Example:**
**$H_A$**: LCSW test preparation will result in reduced anxiety as measured by a ratio level anxiety scale score, comparing scores pre- and post- test preparation

**Eight steps to testing hypothesis:**

1. What is the independent variable, and its level of measurement? One group (pretest, posttest)
2. What is the dependent variable , and its level of measurement? Anxiety scale score; continuous
3. What's the Null Hypothesis? There is no difference in pretest and posttest anxiety scale scores.
4. Is the alternative hypothesis stated above one-tailed or two-tailed? Pretest and posttest anxiety scores are significantly different.
5. Statistical test & alpha level: Independent sample *t*-test, alpha $\leq$ .05
6. Review SPSS table of results
7. Describe results & decision to accept or reject Null –using standard reporting of statistic: (*t* = __, *df* = ___, *p* = ___)
8. Discuss results

C. Which *t*-test?

1. In a one-group pre- and posttest design, anxiety scale scores are measured before and after a 10-week support group. Dependent groups *t*-test

2. In a quasi-experimental study of an innovative infant parent training curriculum, the change in parenting skills in the new curriculum group is compared to that of a matched comparison group provided with standard parenting resource information. Independent groups *t*-test

3. In a randomized treatment control study, the change in anxiety from pre- to post treatment is compared in both groups. Independent groups or paired-samples *t*-

test

4. You suspect that in a San Jose neighborhood African American residents will have higher scores on a ratio level Empowerment Scale than Latino residents. Independent groups *t*-test

5. You suspect that for new Latino immigrants their scores on a ratio level Empowerment Scale will increase one year after their arrival. Dependent groups *t*-test

D. Building block for the t-test: **Confidence interval –** a range into which we would estimate a population parameter to fall based on a sample statistic. The confidence interval of the mean is based on

1. Sample mean

2. Confidence level (e.g. 95%, or 90%, or 99%). 95% is most typical.

3. Amount of variability (standard error of the mean)

   a) **Standard error –** while the standard deviation measures dispersion of the sample data from the mean, the **standard error** is a measure of the precision by which the population mean is estimated from the sample (or, the dispersion of means). It assumes a hypothetical distribution of an infinite number of samples (and their means) drawn from the population.

   b) The formula for the standard error is the *sample's standard deviation divided by the square root of the sample size*. As the sample size (denominator) increases, the standard error becomes smaller. As the dispersion (SD) increases, the standard error increases.

4. So, a 95% CI formula is: Mean $\pm$ 1.96*Standard error

   a) Where does the 1.96 come from? 1.96 is the *z* score for the tails of the normal curve—in between -2 and +2 standard deviations (or -1.96 and +1.96 in *z* values) are 95% of the theoretical observations

   b) For a 99% confidence interval the *z* would be 2.58, yielding a wider confidence interval

5. Example: Our sample of client satisfactions scores have a mean of 40 and standard deviation of 5. We want to be 95% sure of the *true* population mean (meaning we need to know the potential *range* of means).

   a) First we need the standard error: SE = SD/Sq root of sample size

SE = 5/sq root of 100 = 5/10 = 0.5

   b) Then the 95% Confidence Interval:

95% CI = Mean $\pm$ 1.96*Standard error = 40 $\pm$ 1.96*.5
The *lower* limit of the range is 40 - (1.96*.5) = 40 - .98 = 39.02

The *upper* limit of the range is 40 + (1.96*.5) = 40 + .98 = 40.98

**c)** So we can say that we are 95% confident that any sample we draw will have a mean between 39.02 and 40.98 (which is our best estimate of the population mean) **Repeat this aloud several times.**

6. *Margin of error* is a one-number indicator of range, or the "radius" of the distance between the upper and lower limit, usually derived from the 95% confidence interval. (In election polls the question is usually "Whom would you vote for?" which results in nominal data summarized as proportions, not ratio data is in above examples. You can also have confidence intervals of proportions.) So when 40% of registered Democrats surveyed in Pennsylvania said they would vote for Hillary vs. 39% for Barack, 'with a margin of error of 2' statistically they were tied!

III. Data examples of Paired Samples and Independent Groups *t*-test with randomly assigned groups

**Dependent Groups (or Paired Samples) *t*-test.** These are data from a pilot intervention program designed to treat anxiety disorders. Fifteen participants are pre- and post-tested with a ratio-level anxiety scale, where higher scores indicate higher anxiety. We want to know if anxiety scores improve (are lowered) over time for the sample.

| Participant_ID | Pre_anxiety | Post_anxiety |
|---|---|---|
| 1 | 10 | 9 |
| 2 | 12 | 9 |
| 3 | 8 | 11 |
| 4 | 8 | 6 |
| 5 | 12 | 8 |
| 6 | 12 | 12 |
| 7 | 11 | 9 |
| 8 | 10 | 8 |
| 9 | 5 | 6 |
| 10 | 13 | 9 |
| 11 | 14 | 8 |
| 12 | 8 | 8 |
| 13 | 8 | 8 |
| 14 | 13 | 11 |
| 15 | 14 | 12 |

Using the eight steps for hypothesis testing:

1. **What is the independent variable and its level of measurement?**

The independent variable is "treatment" which is a nominal/dichotomous variable (but see next sentence).

In the case of a paired sample, it is customary to say that the independent variable is "treatment," *although here "treatment" is a constant*—everyone in this group gets the experimental treatment. (In terms of research design this scenario is a pre-experimental one group pretest, posttest design.)

2. **What is the dependent variable and its level of measurement?**

The dependent variable is "anxiety score" which is a continuous variable.

3. **State the null hypothesis.**

The null hypothesis (or ($H_O$) is "There is no difference between the pretest and posttest score." OR, "The experimental treatment does not result in improved anxiety scores." *The first hypothesis is much preferred. Remember, the pre-experimental research design does not lend itself to determining causality.*

4. **State the alternative hypothesis.**

$H_A$ : "The posttest anxiety score will be lower than the pretest score."

5. **Identify specific statistical test and alpha level.**

"Since we are comparing means for paired (dependent) samples, we will use the Paired Samples *t*-test. The alpha level is $p \leq .05$."

The Paired-Samples *t*-test statistical formula to calculate *t* is

$$t = \frac{average\ of\ differences}{\frac{SD}{\sqrt{n}}}$$

(Note: you don't have to write the formula in your exams or papers.)

This involves subtracting each post-test score from the corresponding pretest score, then averaging the total differences. Note the denominator. The standard deviation of the difference scores is "averaged" by dividing it with the square root of the sample size. This is called the *standard error of the difference*, an approximation of the average population standard deviation you would get if you drew an infinite number of samples from the larger population and calculated their anxiety difference scores. So, the formula adjusts the average difference of the scores by the standard error of the difference. This resulting *t* statistic can then be located (by SPSS) in the area under the normal curve, the theorized distribution of all possible difference scores in the population.

6. **Review SPSS output.**

**Paired Samples Statistics**

| | | Mean | N | Std. Deviation | Std. Error Mean |
|---|---|---|---|---|---|
| Pair 1 | Posttreatment anxiety | 8.9333 | 15 | 1.86956 | .48272 |
| | Pretreatment anxiety | 10.5333 | 15 | 2.66905 | .68914 |

This is a simple report about the means, standard deviations and standard errors of both scores. It's handy to have when you report results. You can see the means confirm $H_A$ but is the difference significant?

**Paired Samples Correlations**

| | | N | Correlation | Sig. |
|---|---|---|---|---|
| Pair 1 | Posttreatment anxiety & Pretreatment anxiety | 15 | .566 | .028 |

For now this table is not very relevant. It shows you the Pearson Correlation Coefficient between the pre- and posttest scores. The correlation of .57 is relatively high, which makes sense since low pretest scores will be associated with low posttest scores.

**Paired Samples Test**

| | | Paired Differences | | | | | t | df | Sig. (2-tailed) |
|---|---|---|---|---|---|---|---|---|---|
| | | Mean | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Difference | | | | |
| | | | | | Lower | Upper | | | |
| Pair 1 | Posttreatment anxiety - Pretreatment anxiety | -1.60000 | 2.22967 | .57570 | -2.83475 | -.36525 | -2.779 | 14 | .015 |

This table contains what you'll need to report the results of your main research question.

The mean difference is -1.60. Anxiety dropped an average of 1.6 units in score from pretest to posttest.

The *t* value is -2.779. (Divide -1.60 by the std. error of the mean, .5757 and you'll get the same result.) "Degrees of freedom" is an indicator of sample size and is required to locate the *t* statistic in the table of critical values (those calculations are programmed into SPSS). The *p* value is *p* = .015. Since this is less than our alpha of .05, we can *reject the null in favor of our alternative hypothesis*. (While there is no practical direct interpretation of the *t* value, we can at least say that the probability of obtaining the *t* statistic of -.2779 by chance alone is only .015.)

The 95% Confidence Interval of the Difference shows the range of possible *t* statistics that could exist in several samples drawn from the population of average difference scores. So, we can be 95% sure that the any given "anxiety difference score" (from the same instrument) drawn from the population would fall within this range.

7. **Describe results and decision to reject or not reject Null (this is what you would write in your Results Section)**

Anxiety was significantly lowered by -1.60 units of the Anxiety Scale (*t* = -2.779, df = 14, *p* =

.015). Since the *p* value is less than the alpha of .05, the null hypothesis can be rejected in favor of the alternative hypothesis that there is a significant difference between pre- and posttest scores.

8. **Discuss results and implications (this would be part of a Discussion section of a final research report)**

This study showed that anxiety scores decreased significantly from pre- to posttest indicating that the pilot experimental treatment can potentially result in lowered anxiety. However, the pre-experimental research design cannot confirm a causal relationship between the treatment and outcome; rival hypotheses such as history and maturation may have resulted in lowered anxiety, independent of the treatment. Further research with randomized treatment and control groups would confirm the effectiveness of this treatment. [Then, depending on the sample, more information about transcultural implications, implications for social work practice…etc.]

**Independent Groups *t*-test**

Now, let's expand this example to include a randomized treatment and control group. Treatment" is a variable indicating assignment to experimental treatment (coded "1") or treatment-as-usual (coded "2").

"Pre-test anxiety score" and "Post-test anxiety score" are from a ratio-level anxiety scale, where lower scores indicate lower anxiety. **We are interested in the change in anxiety as a result of treatment, from pre- to post-. We will be comparing both groups on a "Change score" shown in the fifth column which we have to create in SPSS (the independent *t* test won't do that for us, as it did with the paired samples *t* test).**

***Why would it be better to use the *change score* rather than just using the posttest score as a dependent variable?

| Participant_ID | Treatment | Pre_anxiety | Post_anxiety | Change |
|---|---|---|---|---|
| 1 | 1 | 10 | 9 | -1 |
| 2 | 1 | 12 | 9 | -3 |
| 3 | 1 | 8 | 11 | 3 |
| 4 | 1 | 8 | 6 | -2 |
| 5 | 1 | 12 | 8 | -4 |
| 6 | 1 | 12 | 12 | 0 |
| 7 | 1 | 11 | 9 | -2 |
| 8 | 1 | 10 | 8 | -2 |
| 9 | 1 | 5 | 6 | 1 |
| 10 | 1 | 13 | 9 | -4 |
| 11 | 1 | 14 | 8 | -6 |
| 12 | 1 | 8 | 8 | 0 |
| 13 | 1 | 8 | 8 | 0 |
| 14 | 1 | 13 | 11 | -2 |
| 15 | 1 | 14 | 12 | -2 |
| 16 | 2 | 10 | 11 | 1 |
| 17 | 2 | 12 | 10 | -2 |
| 18 | 2 | 8 | 10 | 2 |
| 19 | 2 | 8 | 8 | 0 |
| 20 | 2 | 12 | 11 | -1 |
| 21 | 2 | 12 | 12 | 0 |
| 22 | 2 | 11 | 12 | 1 |
| 23 | 2 | 10 | 10 | 0 |
| 24 | 2 | 5 | 6 | 1 |
| 25 | 2 | 13 | 12 | -1 |
| 26 | 2 | 14 | 14 | 0 |
| 27 | 2 | 8 | 8 | 0 |
| 28 | 2 | 8 | 8 | 0 |
| 29 | 2 | 13 | 12 | -1 |
| 30 | 2 | 14 | 13 | -1 |

Now, let's go through the 8 Steps to Hypothesis Testing with the Independent Groups *t*-test:

1. What's the independent variable and level of measurement? **Treatment (nominal level of measurement)**
2. What's the dependent variable? **Anxiety score (continuous level of measurement)**
3. State the Null Hyphothesis ($H_O$): **There is no difference in the change scores of the two groups.**
4. State the Alternative Hypothesis ($H_A$): **The treatment will result in decreased anxiety for the experimental group compared to the treatment group. (One-tailed hypothesis).**
5. Select the appropriate statistical test and alpha level: **Independent *t*-test, alpha $\leq$ .05**.
6. Review SPSS table of results

**Group Statistics**

| | Treatment group assignment | N | Mean | Std. Deviation | Std. Error Mean |
|---|---|---|---|---|---|
| Change from pre to posttest | Experiment | 15 | -1.6000 | 2.22967 | .57570 |
| | Treatment as usual | 15 | -.0667 | 1.03280 | .26667 |

The first table shows each group's N, mean, standard deviation and standard error of the mean for the variable "change". The experiment group mean is lower (-1.60) than the control group, *in the direction we hypothesized* (lower scores indicate lower anxiety). *Note: it's very important to compare the means visually—to make sure the direction of change is what you hypothesized.*

**Independent Samples Test**

| | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | 95% Confidence Interval of the Difference | |
| | | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | Lower | Upper |
| Change from pre to posttest | Equal variances assumed | 5.336 | .028 | -2.417 | 28 | .022 | -1.53333 | .63446 | -2.83297 | -.23370 |
| | Equal variances not assumed | | | -2.417 | 19.743 | .025 | -1.53333 | .63446 | -2.85790 | -.20877 |

The next table shows the test results. What is *"Equal Variances Assumed"?* The table gives two different options for *t* statistics and *p* values. This has to do with whether or not the variances of the two groups are the same. "There are different independent-groups *t*-test formulas for both scenarios. Levene's Test is a separate test of the null hypothesis that the variance of the two groups is not different, with its own statistic and *p* value. For now let's stick with the first set of results, "Equal variances assumed."

The *t* value is -2.417. The *p* value is *p* = .022. Since this is less than our alpha of .05, we can *reject the null* that states the groups' change scores are the same. They are indeed different! Note: the *p* value is shown as "Sig. (2-tailed)." Even though our alternative hypothesis is one-tailed, the *t* test tests the Null of "no difference", and since the means go in the direction of our one-tailed alternative hypothesis, we can still reject the Null with these results. If the means were totally reversed in the opposite direction with a significant *p* value, we would still reject the Null, but our alternative hypothesis would not have been supported!

7. **Describe results and decision to accept or reject Null**

As hypothesized, the mean anxiety change scores for the experimental group (-1.60) are lower than those of the control group (-.067), and this difference is statistically significant. Since our *p* value is lower than our alpha, we can reject the null hypothesis in favor of the alternative hypothesis that the treatment resulted in lowered anxiety (*t* = -2.417, *df* = 28, *p* = .02).

8. **Discuss results, implications**

The experimental treatment has been shown to be effective in a randomized trial as shown by lower anxiety scores in the experimental group. The implication for practice is that the experimental treatment may be effective in community settings, however, further implementation research should be conducted with the types of diverse populations seen in those settings. In addition, further research should be conducted with larger samples controlling for ethnicity and other important independent variables.