Psyc 250 – Statistics & Experimental Design

Single & Paired Samples *t*-tests

Part 1 – Data Entry

For any statistical analysis with any computer program, it is always important that data are entered correctly in a way that lends itself to the respective analysis. This may sound elementary, but the issue becomes particularly important with different types of t-tests. Therefore, we'll start this exercise with data entry.

DATASET: Eating Disorder Example

When creating a database, a few things to consider:

1. **GENERAL RULE OF THUMB:** Each participant should have a designated *row*. All of that participant's data should be on that particular *row*. Data from any one participant should NOT be on the same row as another participant (there are exceptions, but this is a very good general rule)
2. **If participants are in different groups**, create a variable to designate group membership.

The eating disorder dataset presents a unique challenge --- different participants have the same ID number. What should we do?

Enter data for the first 5 Family therapy participants and the first 5 Control group participants. Run "Descriptives" to check the mean and standard deviation with the following:

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
PreTreatment Weight	10	75.60	86.70	82.8100	3.27362
PostTreatment Weight	10	76.40	100.30	87.4100	8.52310
Valid N (listwise)	10				

Now, open the "Stats.Anorexia" dataset that is available on Woodle. Open the file as an SPSS file. You will need to enter the Value Labels for "Group". If you like, you can also enter any Variable Labels.

Single Sample *t*-test

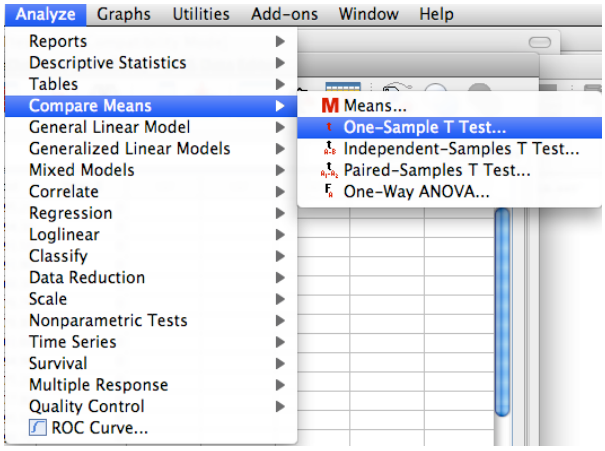
Regardless of group (Control versus Family Therapy), all participants in the dataset are patients diagnosed with anorexia. Thus, we would expect all participants to weight less at pre-treatment than the average weight of the general population of women. In the United States, the average weight for females between 1999 and 2002 was 164.0 lbs (Source: U.S. National Health and Nutrition Examination Survey).

Concept Check:

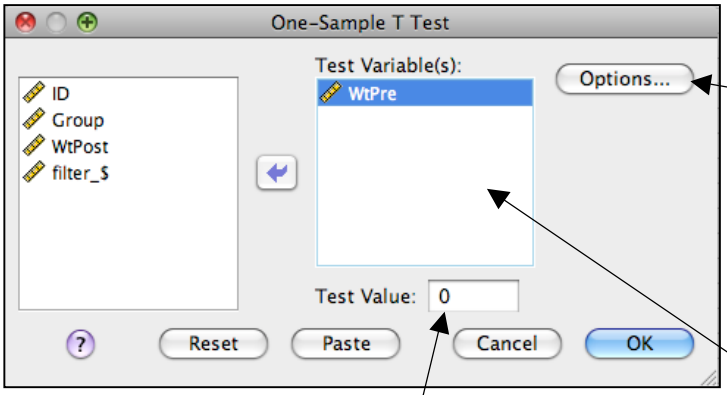
1. What are the research and Null Hypotheses?

To compute a single-sample *t* in SPSS…

[Analyze] → [Compare Means] → [One-Sample T-test]



…which will open a new window…



Options: Nothing too important here for now. However, if you have any missing data in your dataset, you will want to open this box and select "delete listwise".

Test Variable(s): Move variables that you would like to test to this box.

Test Value: Here is where you enter the population mean to which you are comparing the sample mean. In our example, we will enter 135.8.

Then, click on [OK] to run the analysis.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
WtPre	43	83.391	4.8499	.7396

One-Sample Test

	Test Value = 164					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
WtPre	-108.990	42	.000	-80.6093	-82.102	-79.117

Here is the t-value, degrees of freedom, and the significance. In this case, $p < .05$, so we can conclude that there is a significant difference in weight between this sample and the mean population weight for 17 year-old females in the United States.

In recent years, there's been a bit of a push against strict Null Hypothesis Testing. One alternative suggested has been to present the Confidence Interval (CI). Remember, our sample statistic may have some sampling error. The CI tells us that there is a 95% chance that the actual difference between the sample mean and population mean is between -53.9 and -50.9. In this case, the CI does NOT contain zero, which we would interpret as there being a significant difference. Think about it: if the CI contains zero, then there is a good chance that the difference is zero (i.e., no difference).

***When I write the results, it is common to report the most stringent significance level that was reached (.05, .01, or .001). In this case, $p < .001$, because .000 is less than .001.

APA Write-up*:* A single sample *t*-test was conducted to compare the weight of all participants at pretreatment to the average weight of 17-year-old women in the United States. The results was significant, $t(42) = -7.86$, $p < .001$. Women with anorexia weighed significantly less ($M = 83.39$, $SD = 4.85$) than the general population of ($M = 164.0$). The effect size of this difference is $d = 10.81$.

Paired Sample *t*-test

Another question that can be answered with this dataset is: Did participants in the Treatment group gain weight?

Concept Check: What are the research and Null hypotheses?

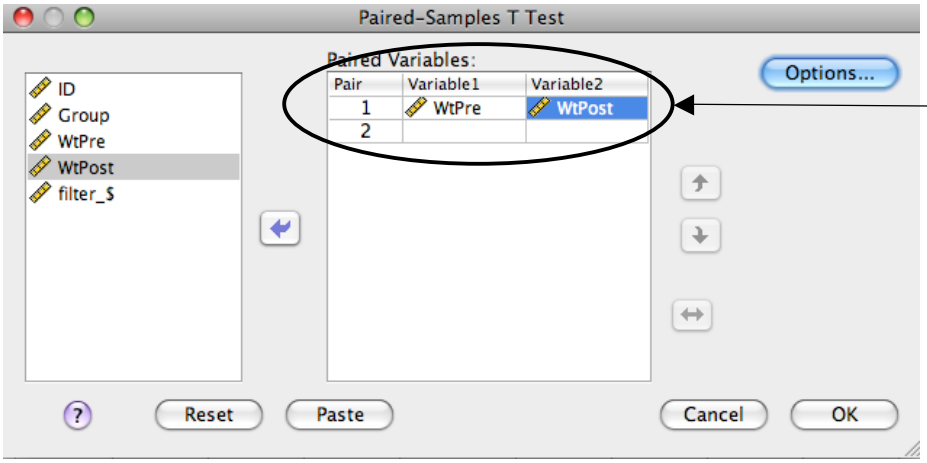
To conduct this analysis, we must first select only participants from the treatment group…

[Data] → [Select Cases] → opens a new window

***Select the variable that you would like to filter, and enter the appropriate equation. In this case, we will want Group to equal 1. (i.e., Filter cases so that only those with group = 1 remain in analyses)

For the paired *t*...

[Analyze] → [Compare Means] → [Paired Samples T-Test]…which will open the following window…



Paired Variables: In a paired *t*, the trickiest part is setting up the dataset appropriately, and then pairing variables the right way. Pairs are denoted by number in the column "Pair". In this example, we are only analyzing 1 pair. The variables for each pair can then be selected and matched under "Variable1" and "Variable2".

Note: "Options" is the same as noted above.

Paired Samples Statistics: This box offers basic descriptive statistics for the variables selected. If a result is significant, you can refer to this box to determine directionality. You will also need this information for the final write-up, because you will need to report Means and SD's of each group.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre-Treatment Weight	83.2294	17	5.01669	1.21673
	Post-Treatment Weight	90.4941	17	8.47507	2.05551

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Pre-Treatment Weight & Post-Treatment Weight	17	.538	.026

Paired Samples Correlations: Recall that a paired samples *t* takes into account the relation (non-independence) between the two samples. This relation is denoted here as a correlation. Not surprisingly, Weights at T1 are associated with Weights at T2 ($p < .05$).

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	Pre-Treatment Weight - Post-Treatment Weight	-7.26471	7.15742	1.73593	-10.94471	-3.58470	-4.185	16	.001

Here is the t-value, degrees of freedom, and the significance. In this case, $p < .05$, so we can conclude that there is a significant difference in weight between Time 1 (pre treatment) and Time 2 (post treatment).