Module 8 Handout: t Test Examples

Jon Starkweather, PhD

October 15, 2010

Jon Starkweather, PhD
jonathan.starkweather@unt.edu
Consultant
Research and **S**tatistical **S**upport

http://www.unt.edu

http://www.unt.edu/rss

The Research and Statistical Support (RSS) office at the University of North Texas hosts a number of "Short Courses".
A list of them is available at:
http://www.unt.edu/rss/Instructional.htm

Contents

1 One Sample t Test Example

Step 1: Define the populations and restate the research question as null and alternative hypotheses.

- Research question: Does this class of four students get **less** sleep than all UNT students ($\mu = 6.08$ hours)?

- Population 1: Students in this class.

- Population 2: All other UNT students.

$$H_0 = \mu_1 \geq \mu_2$$
$$H_1 = \mu_1 < \mu_2$$

- Note, the Alternative hypothesis (H_1) is directional: one-tailed test.

Step 2: Comparison Distribution

Table 2: Comparison dist. (estimate σ with S_M)

X	$\overline{X}$	$X - \overline{X}$	$\left(X - \overline{X}\right)^2$
5	4	1	1
3	4	-1	1
6	4	2	4
2	4	-2	4
16			$SOS = 10$

$$S^2 = \frac{\sum \left(X - \overline{X}\right)^2}{n-1} = \frac{10}{4-1} = 3.33$$

$$S_M^2 = \frac{S^2}{n} = \frac{3.33}{4} = 0.8325$$

$$S_M = \sqrt{S_M^2} = \sqrt{0.8325} = 0.912$$

Step 3: Determine the critical score

- Determine the cutoff sample score.

- Significance level = .05, one-tailed test, $df = 4 - 1 = 3$

- Look in the appropriate column of the t table.

    - One-tailed, significance level = .05

        http://www.math.unb.ca/~knight/utility/t-table.htm

3

- Critical t value is 2.353

- Which means, we need t_{calc} to be greater than |2.353| (absolute value of 2.353) to find a significant difference.

Step 4: Calculate t (t_{calc})

- From previous calculating, we have:

$$\overline{X} = 4.00, \, S_M = 0.912$$

- So, we can plug in the values, including the population mean which was given.

$$t_{calc} = \frac{\overline{X} - \mu}{S_M} = \frac{4.00 - 6.08}{0.912} = \frac{-2.08}{0.912} = -2.28$$

Step 5: Compare and Make a Decision

- Since $t_{calc} = |-2.28| < |2.353| = t_{crit}$ we fail to reject the null hypothesis.

- The interpretation is...Students in this class ($M = 4.00$) do not sleep significantly fewer hours than all other students at UNT ($\mu = 6.08$), $t(3) = -2.28, p > .05$.

- Do not be tempted to say something like; nearly significant, just missed significance, etc.

  - Although the sample mean is numerically smaller, it is not statistically significantly smaller.

  - Remember, we have a *very* small sample size, so we would need a *very* large mean difference to achieve significance (or a larger sample).

Effect Size

- The appropriate effect size measure for the one sample t test is Cohen's d.

- Calculation of d in its general form (as it was with the Z-test) is:

$$d = \frac{\overline{X} - \mu}{\sigma}$$

- However, we do not know the population standard deviation (σ) in the t situation, so we estimate with $S = \sqrt{S^2}$

$$d = \frac{\overline{X} - \mu}{S} = \frac{4 - 6.08}{\sqrt{3.33}} = \frac{-2.08}{1.82} = 1.143$$

- So, although we have a large effect size (standardized difference), we did not achieve statistical significance. However, keep in mind that with a larger sample, this amount of mean difference may have been significant.

  - Statistical significance is **directly** tied to sample size, effect size is not.

Calculating a CI

- In this example, we are calculating a 95% confidence interval (CI_{95}) because our critical value was based on a significance level of .05 (1 - .05 = .95 or 95%).

- $t_{crit} = 2.353$, $S_M = 0.912$, $\overline{X} = 4.00$

$$UL = +2.353 * 0.912 + 4.00 = 6.146$$
$$LL = -2.353 * 0.912 + 4.00 = 1.854$$

- If we were to take an infinite number of samples of students in this class, 95% of those samples' means would be between 6.146 and 1.854 hours of sleep.

  - Remember, the population mean is fixed (but unknown); while each sample has its own mean (sample means fluctuate).

CI_{95} Considerations and Interpretations
The current example resulted in $CI_{95} = 6.146 : 1.854$.

- Notice the population 2 mean ($\mu_2 = 6.08$), representing all UNT students, falls **inside** our interval.

  - Recall, we did not reject the null; meaning our sample *did* come from population 2, not a distinct population (i.e. population 1).

- If μ_2 were greater than 6.146, then we would have rejected the null hypothesis and inferred that our sample came from population 1; meaning, population 1 would have been significantly different from population 2.

- Important: the interval is **not** interpreted as "we are 95% confident that population 1's mean is between 6.146 and 1.854."

  - We are dealing with a sample; we do not know what μ_1 is; so, we can not know what that interval would be.

2 Dependent Samples t Test Example

Step 1: State the Null and Alternative Hypotheses
Define the populations: Relationship Satisfaction = RS

- Population 1: RS prior to college separation.

- Population 2: RS after the first semester.

State the Hypotheses:

- The null hypothesis is: there will be no difference between the means of the ratings (before college vs. after the first semester).

$$H_0 : \mu_1 - \mu_2 \leq 0 \qquad \text{or} \qquad \mu_D \leq 0$$

- The alternative hypothesis is: there will be a significant **decrease** in the ratings.

$$H_1 : \mu_1 - \mu_2 > 0 \qquad \text{or} \qquad \mu_D > 0$$

Note the directional alternative hypothesis; pay careful attention to how we specified the hypotheses in symbols.

- If the alternative hypothesis expected an increase,
  then $H_1 : \mu_1 - \mu_2 < 0$ or $\mu_D < 0$

Step 2: Comparison Distribution (estimate σ with S_M).

Pair	Before	After	D	$\overline{D}$	$(D - \overline{D})$	$\left(D - \overline{D}\right)^2$
1	40	32	8	8	0	0
2	38	31	7	8	-1	1
3	36	30	6	8	-2	4
4	42	31	11	8	3	9
4			32			$SOS = 14$

$$\overline{D} = \sum(D)/n_D = 32/4 = 8$$

$$S^2 = SOS/df = 14/n_D - 1 = 14/3 = 4.67$$

$$S_M^2 = \frac{S^2}{n_D} = \frac{4.67}{4} = 1.1675$$

$$S_M = \sqrt{S_M^2} = \sqrt{1.1675} = 1.08$$

Step 3: Determine the critical value

- Deja-vu...?

- Determine the critical t value.

- Significance level = .05, one-tailed test, $df = n_D - 1 = 3$.

- Look in the appropriate column of the t table.

  – One-tailed, significance level = .05

  http://www.math.unb.ca/~knight/utility/t-table.htm

- Critical t value is 2.353

- Which means, we need t_{calc} to be greater than |2.353| (absolute value of 2.353) to find a significant difference.

Same critical value we used with the One Sample t Test.

Step 4: Calculate t

- The general t formula:

$$t = \frac{\overline{X} - \mu}{S_M}$$

- Remember we are dealing with the *difference scores (D)* but the t formula is essentially the same as above.

- From previous calculations, we have: $\overline{D} = 8, \mu_D = 0, S_M = 1.08$

$$t = \frac{\overline{D} - \mu_D}{S_M} = \frac{8-0}{1.08} = 7.41$$

- So, our $t_{calc} = 7.41$ which is fairly large.

Step 5: Compare and make a decision.

- Since $t_{calc} = 7.41 > 2.353 = t_{crit}$ we reject the null hypothesis.

- The initial interpretation is... There was a significant decrease in ratings of relationship satisfaction from before college separation ($M = 39, SD = 2.58$) to after the first semester of college separation ($M = 31, SD = 0.82$), $t(3) = 7.41, p = .005$ (one-tailed).

  - One could also state the results as; $t(3) = 7.41, p < .05$ because .05 was our significance level.
  - The 3 above is the degrees of freedom (*df*).

- At this point in the course, you should be able to calculate the means and standard deviations of each group of scores; however the exact p value was obtained by verifying the results above in SPSS (see below).

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	before	39.0000	4	2.58199	1.29099
	after	31.0000	4	.81650	.40825

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	before & after	4	.632	.368

Paired Samples Test

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	before - after	8.00000	2.16025	1.08012	4.56257	11.43743	7.407	3	.005

Effect Size

- Recall, Cohen's d:

$$d = \frac{\overline{X} - \mu}{\sigma}$$

7

- Here, we do not know σ so we use S to estimate (not S_M because it is influenced heavily by sample size).

  - Effect sizes are not influenced by sample size.

- Also, because we are dealing with difference scores, we use slightly different symbols in the formula:

$$d = \frac{\overline{D} - \mu_D}{S}$$

Effect Size continued

- Cohen's d for difference scores:

$$d = \frac{\overline{D} - \mu_D}{S}$$

- We have $S^2 = 4.67$ from above, which leads to an estimate of σ:

$$S = \sqrt{4.67} = 2.16$$

- Which in turn, leads to d:

$$d = \frac{\overline{D} - \mu_D}{S} = \frac{8-0}{2.16} = 3.70$$

- A large effect size.

Calculating a CI

Calculating a Confidence Interval with $\overline{D}$

- Using essentially the same procedures we used with the one sample t test, we can calculate the lower limit (LL) and upper limit (UL).

- Recall, the general formulas for a confidence interval are: LL = (-crit)*(SE) + mean and UL = (+crit)*(SE) + mean

- When in the Dependent Samples situation, we simply use the difference score mean.

$$LL = -t_{crit} * S_M + \overline{D} = -2.353 * 1.08 + 8 = -2.541 + 8 = 5.459$$

$$UL = +t_{crit} * S_M + \overline{D} = +2.353 * 1.08 + 8 = +2.541 + 8 = 10.541$$

Interpretation of CI_95

- In this example, we calculated a 95% confidence interval (CI_{95}) because our critical value was based on a significance level of .05.

$$LL = -2.353 * 1.08 + 8 = 5.459$$
$$UL = +2.353 * 1.08 + 8 = 10.541$$

- If we drew an infinite number of samples of young adults' relationship satisfaction ratings, 95% of those samples' difference score means would be between 5.459 and 10.541.

  - Remember, the population difference score mean is fixed (but unknown); while each sample has its own difference score mean (samples fluctuate).

3 Independent Samples t Test Example

Step 1

- Define the populations and restate the research question as null and alternative hypotheses.

- Population 1: Americans who watch *The Daily Show*.

- Population 2: Americans who watch *The O'Reilly Factor*.

$$H_0 : \mu_1 \leq \mu_2 \text{ or } H_0 : \mu_1 - \mu_2 \leq 0$$

$$H_1 : \mu_1 > \mu_2 \text{ or } H_1 : \mu_1 - \mu_2 > 0$$

- In terms of knowledge about current events.

- Notice the directional alternative hypothesis (H_1) which indicates a one-tailed test.

The Daily show group's data

Table 3: Daily Show Group

X_1	$\overline{X}_1$	$X_1 - \overline{X}_1$	$\left(X_1 - \overline{X}_1\right)^2$
6	6.75	-0.750	0.563
6	6.75	-0.750	0.563
9	6.75	2.250	5.063
8	6.75	1.250	1.563
4	6.75	-2.275	7.563
6	6.75	-0.750	0.563
7	6.75	0.250	0.063
8	6.75	1.250	1.563
$54 = \sum X_1$			$SOS_1 = 17.50$
$8 = n_1$			

$$S_1^2 = \frac{\sum \left(X_1 - \overline{X}_1\right)^2}{n_1 - 1} = \frac{SOS_1}{df_1} = \frac{17.50}{8-1} = \frac{17.50}{7} = 2.50$$

The Factor show group's data

Table 4: Factor Show Group

9

X_2	$\overline{X}_2$	$X_2 - \overline{X}_2$	$\left(X_2 - \overline{X}_2\right)^2$
5	3.875	1.125	1.266
4	3.875	0.125	0.016
3	3.875	-0.875	0.766
1	3.875	-2.875	8.266
5	3.875	1.125	1.266
6	3.875	2.125	4.516
3	3.875	-0.875	0.766
4	3.875	0.125	0.016
$31 = \sum X_2$			$SOS_2 = 16.875$
$8 = n_2$			

$$S_2^2 = \frac{\sum \left(X_2 - \overline{X}_2\right)^2}{n_2 - 1} = \frac{SOS_2}{df_2} = \frac{16.875}{8-1} = \frac{16.875}{7} = 2.411$$

Step 2(a)

- 2. Determine the characteristics of the comparison distribution.

- $df_t = df_1 + df_2 = n_1 + n_2 - 2 = 8 + 8 - 2 = 14$

- And from above, $S_1^2 = 2.500$ and $S_2^2 = 2.411$

- (a) Calculate the pooled estimate of the population variance.

$$S_p^2 = (S_1^2) * \frac{df_1}{df_t} + (S_2^2) * \frac{df_2}{df_t} = (2.500) * \frac{7}{14} + (2.411) * \frac{7}{14} = 2.4555$$

- So, $S_p^2 = 2.46$

Step 2(b)

- (b) Calculate the variance of each distribution of means:

$$S_{M1}^2 = \frac{S_p^2}{n_1} = \frac{2.46}{8} = 0.3075$$

$$S_{M2}^2 = \frac{S_p^2}{n_2} = \frac{2.46}{8} = 0.3075$$

- Please note; if the groups were different sizes, the variances of each distribution of means would be different.

Step 2(c) and Step 2(d)

- (c) Calculate the variance of the distribution of differences between means:

$$S_{dif}^2 = S_{M1}^2 + S_{M2}^2 = 0.3075 + 0.3075 = 0.615$$

- So, $S_{dif}^2 = 0.62$

- (d) Calculate the standard deviation of the distribution of differences between means:

$$S_{dif} = \sqrt{S_{dif}^2} = \sqrt{.62} = 0.78$$

- So, $S_{dif} = 0.78$

Step 3

- 3. Determine the critical sample score on the comparison distribution at which the null hypothesis should be rejected.

  - Significance level = .05
  - Two-tailed test (based on H_1).
  - $df_t = n_1 + n_2 - 2 = 7 + 7 - 2 = 14$

    http://www.math.unb.ca/~knight/utility/t-table.htm

- $t_{crit} = 1.761$

Step 4

- 4. Determine the sample's score on the comparison distribution:

  - Compute t_{calc}

$$t = \frac{\overline{X}_1 - \overline{X}_2}{S_{dif}} = \frac{6.75 - 3.875}{0.78} = 3.69$$

- So, $t_{calc} = 3.69$

Step 5

- 5. Compare the scores from Step 3 and Step 4, and make a decision to reject the null hypothesis or fail to reject the null hypothesis.

- Because; $t_{calc} = 3.69 > 1.761 = t_{crit}$ we reject the null hypothesis and conclude there was a statistically significant difference between the two show groups.

  - But, you should know by now, that's not the whole story.

Effect Size
Calculating Effect Size for two Independent Groups

- Recall, the general formula for Cohen's *d*.

$$d = \frac{\mu_1 - \mu_2}{\sigma}$$

- In the current (independent groups) situation, we have:

11

$$d = \frac{\overline{X}_1 - \overline{X}_2}{S_p} = \frac{6.75 - 3.875}{\sqrt{2.46}} = \frac{2.875}{1.57} = 1.83$$

- So, the effect size is fairly large; $d = 1.83$

Calculating a CI

Calculating a Confidence Interval

- Note, we are calculating the interval on the difference between means.

- Recall there are two parts of a confidence interval, the upper limit (UL) and the lower limit (LL).

- The general form of the equations for each limit are:

$$LL = (-crit) * (SE) + mean$$
$$UL = (+crit) * (SE) + mean$$

- In the current situation for the differences between means:

$$LL = (-t_{crit}) * \left(S_{dif}\right) + \left(\overline{X}_1 - \overline{X}_2\right) = -1.761 * 0.78 + (6.75 - 3.875) = -1.374 + 2.875 = 1.501$$
$$UL = (+t_{crit}) * \left(S_{dif}\right) + \left(\overline{X}_1 - \overline{X}_2\right) = +1.761 * 0.78 + (6.75 - 3.875) = +1.374 + 2.875 = 4.249$$
$$LL = 1.50$$
$$UL = 4.25$$

Interpretation of Confidence Interval

- Recall, we had a significance level of .05 (t_{crit}), so we conducted a 95% confidence interval (CI_{95}) on the differences between means.

- The Lower Limit was 1.50 and the Upper Limit was 4.25.

- So, if we drew an infinite number of random samples of viewers of each show, 95% of the differences between means would be between 1.50 and 4.25.

  - Remember, the mean of the population of differences between means is fixed (but unknown); while each sample has its own differences between means (samples fluctuate).

4 Alternative Calculation of Independent Samples t Test

All of what was covered above (i.e., the calculations of each of the three *t* tests) used the definitional formula. Recall, definitional formulas follow the *definition* of the concept. Computational formulas are equivalent (i.e., they produce the same answers), but they are supposed to make calculation easier. Below you will find an example of the computational formula for the Independent Samples *t* Test. The computational formula is applied to the same data used above, therefore, the same answer is produced.

The Daily Show Group is Group 1; meaning all the data below which has a '1' subscript belongs to this group. The O'Reilly Factor Show Group is Group 2; meaning all the data below which has a '2' subscript belongs to this group.

Table 5: Both Groups' Data

X_1	$\overline{X}_1$	$X_1 - \overline{X}_1$	$\left(X_1 - \overline{X}_1\right)^2$	X_2	$\overline{X}_2$	$X_2 - \overline{X}_2$	$\left(X_2 - \overline{X}_2\right)^2$
6	6.75	-0.750	0.563	5	3.875	1.125	1.266
6	6.75	-0.750	0.563	4	3.875	0.125	0.016
9	6.75	2.250	5.063	3	3.875	-0.875	0.766
8	6.75	1.250	1.563	1	3.875	-2.875	8.266
4	6.75	-2.275	7.563	5	3.875	1.125	1.266
6	6.75	-0.750	0.563	6	3.875	2.125	4.516
7	6.75	0.250	0.063	3	3.875	-0.875	0.766
8	6.75	1.250	1.563	4	3.875	0.125	0.016
54	$= \sum X_1$		$SOS_1 = 17.50$	31	$= \sum X_2$		$SOS_2 = 16.875$
8	$= n_1$			8	$= n_2$		

$$S_1^2 = \frac{\sum \left(X_1 - \overline{X}_1\right)^2}{n_1 - 1} = \frac{SOS_1}{df_1} = \frac{17.50}{8-1} = \frac{17.50}{7} = 2.50$$

$$S_2^2 = \frac{\sum \left(X_2 - \overline{X}_2\right)^2}{n_2 - 1} = \frac{SOS_2}{df_2} = \frac{16.875}{8-1} = \frac{16.875}{7} = 2.411$$

$$t_{calc} = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\left[\frac{(n_1-1)*\left(S_1^2\right)+(n_2-1)*\left(S_2^2\right)}{n_1+n_2-2}\right]*\left[\frac{1}{n_1}+\frac{1}{n_2}\right]}}$$

$$t_{calc} = \frac{6.75 - 3.875}{\sqrt{\left[\frac{(8-1)*(2.5)+(8-1)*(2.411)}{8+8-2}\right]*\left[\frac{1}{8}+\frac{1}{8}\right]}}$$

$$t_{calc} = \frac{2.875}{\sqrt{\left[\frac{17.5+16.877}{14}\right]*[.125+.125]}}$$

$$t_{calc} = \frac{2.875}{\sqrt{\left[\frac{34.377}{14}\right]*[.25]}}$$

$$t_{calc} = \frac{2.875}{\sqrt{[2.4555]*[.25]}}$$

$$t_{calc} = \frac{2.875}{\sqrt{.6138}} = \frac{2.875}{.78} = 3.686$$

$$t_{calc} = 3.686$$