Chapter 9 Two Sample t Tests

Sherri Verdugo

October 13, 2014

Contents

Topics for this Week

1. Recap
2. Homework 2 is due
3. Chapter 9 Part One
4. Examples of z scores are on the class website
5. Feedback will be on the class portal for each assignment
6. Winners will be announced on Wednesday
7. Wed. we will be doing a mock research project from start to finish
8. Chapter 10 (introduction to 10)

Chapter 9 Two-Sample t Tests

Prologue and Introduction

We are studying a family of tests in statistics named, the t-tests. This time, we can look at introducing more than one group of individuals to study. We are learning how to look at groups statistically. This chapter builds from chapters 6, 7, and 8.

Remember when I mentioned that we should be able to remember what equation we need to use for the data that we have. The equation we use is always decided by the data!

Outline

1. Independent Samples Versus Dependent Samples
2. The Two-Sample t Test (INDEPENDENT DRAWN SAMPLES)
3. Adjustments for Sigma-Hat Squared $\hat{\sigma}^2$
4. Interpreting a Computer-Generated t Test
5. Computer Applications: Independent Samples t Tests
6. The Two-Sample t Test for Dependent Samples
7. Computer Applications: Dependent Samples t Tests
8. Statistical Significance versus Research Significance
9. Statistical Power
10. Conclusions

Independent Samples Versus Dependent Samples

- Independent Samples (Independent Drawn Samples) [PAGE 273]
  - The composition of one sample is in no way matched or paired to the composition of the other sample. Thus, the two samples reflect two *SEPARATE POPULATIONS*
- Dependent Samples: Matched Pairs [PAGE 275]
  - Situation in which members of one sample are not selected independently, but are instead determined by the makeup of the other sample. (they reflect the other group but are NOT THE SAME GROUP.)
- Dependent Samples: Matched Pairs (paired difference) t test [PAGE 275]
  - The t test used when the two samples are dependent samples.

INDEPENDENT SAMPLES t-Test

- The first introductory t-Test from previous chapters assumes *INDEPENDENT SAMPLES*
- CHARACTERISTICS
  - Two separate groups that are NOT matched or paired on any criteria
  - Two Samples REFLECT two DIFFERENT POPULATIONS
  - Group 1 and 2 has it's own:
    * sample size n
    * mean $\bar{x}$
    * variance σ^2 or s^2 (for now we are using s^2 versus σ^2)
    * standard deviation (the $\sqrt{variance}$)
  - Hypotheses

* NON-DIRECTIONAL: $H_0 : \mu_1 = \mu_2$ and $H_1 : \mu_1 \neq \mu_2$
* DIRECTIONAL (IF PRIOR KNOWLEDGE INDICATES THIS):
  - $H_0 : \mu_1 > \mu_2$ and $H_1 : \mu_1 < \mu_2$
  - $H_0 : \mu_1 < \mu_2$ and $H_1 : \mu_1 > \mu_2$

Relationship between Independent Samples and Population

SAMPLE 1	SAMPLE 2
Size n_1	Size n_2
Mean μ_1	Mean μ_2
Variance s_1^2	Variance s_2^2
$\uparrow$	$\uparrow$
POPULATION 1	POPULATION 2
μ_1 is unknown	μ_2 is unknown
σ_1^2 is unknown	σ_2^2 is unknown

Examples of Independent Samples and Populations

** THESE ARE NOT MATCHED AT ALL!!!!**

- Participants randomly selected to receive a placebo and medication
  - Group 1 = Placebo
  - Group 2 = Medication

PLACEBO	MEDICATION
Size $n_1 = 6$	Size $n_2 = 5$
Mean $\bar{x_1}$	Mean $\bar{x_2}$
Variance s_1^2	Variance s_2^2
$\uparrow$	$\uparrow$
POPULATION 1	POPULATION 2
μ_1 is unknown	μ_2 is unknown
σ_1^2 is unknown	σ_2^2 is unknown

- Participants randomly selected to receive a training and non-training
  - Group 1 = Training
  - Group 2 = Non-Training
- Participants randomly selected to view advertising and not view advertising
  - Group 1 = Advertising

- Group 2 = Non-Advertising

The Two-Sample t Test (INDEPENDENT DRAWN SAMPLES)

- Pooled Estimate of Common Variance [PAGE 276]
  - Estimate based on a weighted average of two sample variances being used to estimate the population variance in finding the standard error.
- F test for Homogeneity of Variance [PAGE 276]
  - A test, based on the sample variances, used to determine the most appropriate t test formula to use.

Steps for Two-Sample t Test (INDEPENDENT DRAWN SAMPLES)

1. Write out Hypotheses for the original problem (comparison of the two sample means)
   - The Null H_0
   - The Alternative H_1
2. For EACH SAMPLE
   - n_1 (sample size) and n_2 (sample size)
   - $\bar{x_1}$ (mean) and $\bar{x_2}$ (mean)
   - s_1^2 (sample variance) and s_2^2 (sample variance)
3. HOW TO DETERMINE WHICH t formula to use
   - for EQUAL OR UNEQUAL VARIANCES use the F test for Homogeneity of Variances
     - Write out the H_0 and H_1 for the F Test
     - Calculate F and it's two degrees of freedom
     - Compare the obtained F to $F_{critical}$, .05 level (from the F table pp. 566-568)
     - Decisions for F
       * If $F_{obtained} \geq F_{critical}$ ASSUME UNEQUAL POPULATION VARIANCES
       * If $F_{obtained} \leq F_{critical}$ ASSUME EQUAL POPULATION VARIANCES
4. Perform the appropriate t test as determined by the F test

t-test Solving Example 1 Page 277: EQUAL VARIANCES

- Scenario: T.V. Commercial.
- A group of 6 people are assigned to the experimental group and will see a commercial. Those 6 people in the experimental group will then evaluate the product's favor ability. Next, 5 people will be assigned to the control group and then evaluate the product's favor ability without seeing the commercial. In addition, the marketing department has indicated that the commercial has been rated as favorable in the past.

1. Write out H_0 and H_1
   - WE USE DIRECTIONAL BECAUSE OF PREVIOUS INFORMATION FROM THE MARKETING DEPARTMENT.
   - Experimental group saw the commercial
   - Control group did not see the commercial
   - $H_0 : \mu_{experimental} = \mu_{control}$
   - $H_1 : \mu_{experimental} > \mu_{control}$
2. Determine n, $\bar{x}$, and s^2 for each group

Experimental Group	Control Group
$x_{experimental}$	$x_{control}$
10	NA
6	8
8	3
7	5
9	6
7	7

- $n_{experimental} = 6$ and $n_{control} = 5$
- $\bar{x}_{experimental} = 47$ and $\bar{x}_{control} = 29$
- Means for our groups
  - $\bar{x}_{experimental} = \frac{\Sigma x_{experimental}}{n_{experimental}} = \frac{47}{6} = 7.83$
  - $\bar{x}_{control} = \frac{\Sigma x_{control}}{n_{control}} = \frac{29}{5} = 5.80$

3. Calculate the Variance (from scratch, these may or may not be provided to you in an exam or homework… know this skill even if you don't need it for a particular problem.)
   - Square the raw scores

Experimental Group	Control Group
$x_{experimental}$	$x_{control}$
100	NA
36	64
64	9
49	25
81	36
49	49

```
* For Experimental Variance
```

$\Sigma x_1^2 = 379$

$s_1^2 = \frac{\Sigma x_1^2 - \frac{(\Sigma x_1)^2}{n_1}}{n_1}$

$s_1^2 = \frac{379 - \frac{(47)^2}{6}}{6}$

$s_1^2 = \frac{379 - \frac{2206}{6}}{6}$

$s_1^2 = \frac{379 - 368.17}{6}$

$s_1^2 = 1.81$

```
* For Control Variance
```

$\Sigma x_2^2 = 183$

$s_2^2 = \frac{\Sigma x_2^2 - \frac{(\Sigma x_2)^2}{n_2}}{n_2}$

$s_2^2 = \frac{183 - \frac{(29)^2}{5}}{5}$

$s_2^2 = \frac{183 - \frac{841}{5}}{5}$

$s_2^2 = \frac{183 - 168.2}{5}$

$s_2^2 = 2.96$

Summarizing the results So Far

Experimental	Control
$n_1 = 6$	$n_2 = 5$
$\bar{x_1} = 7.83$	$\bar{x_2} = 5.80$
$s_1^2 = 1.81$	$s_2^2 = 2.96$

4. Perform the F Test for the Variances

- F test Hypothesis

$H_0 : \sigma_1^1 = \sigma_2^1$ and $H_1 : \sigma_1^1 \neq \sigma_2^1$

- F test Statistic

$F = \frac{s_{larger}^2}{s_{smaller}^2} = F = \frac{2.96}{1.81} = F = 1.64$ with

$df_{numerator}$ and $df_{denominator}$ ==> This is 2 steps!

$df_{numerator} = n_1 - 1 = 5 - 1 = 4$ and $df_{denominator} = n_2 - 1 = 6 - 1 = 5$

Result

$F_{OBTAINED} = 1.64$ and $F_{CRITICAL} = 5.19$ with $df = (4, 5)$ and our final decision for the equation to use is:

$F_{OBTAINED} = 1.64 < F_{CRITICAL0.5} = 5.19(df = 4and5)$

We do NOT REJECT THE Null and conclude that we use the

t test for EQUAL Population Variances

rarely do we have to change our test statistic because F is calculated from the sample or the population

5. Calculate the test statistic: t Test for EQUAL POPULATION VARIANCES

$$t = \frac{\bar{x_1} - \bar{x_2}}{\sqrt{\left(\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

- Where $df = n_1 + n_2 = 2$

- PRO TIP: "Solve this big equation by finding the components and then plugging them into the equation at the end."

$\bar{x}_1 - \bar{x}_2 = 7.83 - 5.80 = 2.03$

$\frac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2} = \frac{6(1.81)+5(2.96)}{6+5-2} = \frac{10.86+14.80}{11-2} = \frac{25.66}{9} = 2.85$

$\frac{1}{n_1} + \frac{1}{n_2} = \frac{1}{6} + \frac{1}{5} = 0.17 + 0.20 = 0.37$

$\sqrt{\left(\frac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{(2.85)(0.37)} = \sqrt{1.0545} = 1.0268$

- This is where we calculate t:

$t = \frac{\bar{x_1} - \bar{x_2}}{\sqrt{\left(\frac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{2.03}{1.0268} = 1.9708 = t_{obtained} = 1.971$

Degrees of Freedom $df = n_1 + n_2 - 2 = 6 + 5 - 2 = 11 - 2 = 9$

COMPARE t test calculate to critical for decision on rejecting the null hypothesis

- $t_{critical}$ is for a DIRECTIONAL H_1 at $\alpha = 0.5$

Decisions: remember we are using DIRECTIONAL here

- One Tail: At $\alpha = 0.025$ level, $t_{critical}(df = 9) = 2.262 > 1.971$ SINCE $p < 0.05$
- Two Tail: At $\alpha = 0.05$ level, $t_{critical}(df = 9) = 1.833 < 1.971$

We use the One Tail ONLY FOR THIS LONG EXAMPLE. We conclude the following:

We reject H_0 in favor of the alternative H_1 such that $\mu_1 > \mu_2$. This particular commercial does increase favor ability ratings of the product! We tell the marketing department that if the assumption that the entire consumer population had viewed the commercial then their mean support score would increase.

t-test Solving Example 1 Page 283: UNEQUAL VARIANCES

What if our example had different scores? Let's revisit the problem with some new scores and see what happens when we have UNEQUAL VARIANCES.

Experimental Group	Control Group
$x_{experimental}$	$x_{control}$
10	NA
6	10
8	1
7	3
9	6
7	9

- $n_{experimental} = 6$ and $n_{control} = 5$
- $\bar{x}_{experimental} = 47$ and $\bar{x}_{control} = 29$
- Means for our groups
  - $\bar{x}_{experimental} = \frac{\Sigma x_{experimental}}{n_{experimental}} = \frac{47}{6} = 7.83$
  - $\bar{x}_{control} = \frac{\Sigma x_{control}}{n_{control}} = \frac{29}{5} = 5.80$

Calculate the Variance (from scratch, these may or may not be provided to you in an exam or homework… know this skill even if you don't need it for a particular problem.)

- Square the raw scores

Experimental Group	Control Group
$x_{experimental}$	$x_{control}$
100	NA
36	100
64	1
49	9
81	36
49	81

```
* For Experimental Variance
```

$\Sigma x_1^2 = 379$

$s_1^2 = \frac{\Sigma x_1^2 - \frac{(\Sigma x_1)^2}{n_1}}{n_1}$

$s_1^2 = \frac{379 - \frac{(47)^2}{6}}{6}$

$s_1^2 = \frac{379 - \frac{2206}{6}}{6}$

$s_1^2 = \frac{379 - 368.17}{6}$

$s_1^2 = 1.81$

```
* For Control Variance
```

$\Sigma x_2^2 = 227$

$s_2^2 = \frac{\Sigma x_2^2 - \frac{(\Sigma x_2)^2}{n_2}}{n_2}$

$s_2^2 = \frac{227 - \frac{(29)^2}{5}}{5}$

$s_2^2 = \frac{227 - \frac{841}{5}}{5}$

$s_2^2 = \frac{227 - 168.2}{5} = \frac{58.8}{5} = 11.76$

Summarizing the results So Far

Experimental	Control
$n_1 = 6$	$n_2 = 5$
$\bar{x_1} = 7.83$	$\bar{x_2} = 5.80$
$s_1^2 = 1.81$	$s_2^2 = 11.76$

Next, Perform the F Test for the Variances

- F test Hypothesis

$H_0 : \sigma_1^1 = \sigma_2^1$ and $H_1 : \sigma_1^1 \neq \sigma_2^1$

- F test Statistic

$F = \frac{s_{larger}^2}{s_{smaller}^2} = F = \frac{11.76}{1.81} = F = 6.497$ with

$df_{numerator}$ and $df_{denominator}$ ==> This is 2 steps!

$df_{numerator} = n_1 - 1 = 5 - 1 = 4$ and $df_{denominator} = n_2 - 1 = 6 - 1 = 5$

Result

$F_{OBTAINED} = 6.497$ and $F_{CRITICAL} = 5.19$ with $df = (4, 5)$ and our final decision for the equation to use is:

$F_{OBTAINED} = 6.497 > F_{CRITICAL0.5} = 5.19(df = 4and5)$

We REJECT THE Null and conclude that we use the

t test for UNEQUAL Population Variances

rarely do we have to change our test statistic because F is calculated from the sample or the population

Again, we calculate the test statistic: t Test BUT THIS TIME WE USE THE UNEQUAL POPULATION VARIANCES

$t = \frac{\bar{x_1} - \bar{x_2}}{\sqrt{\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 1}}}$

- df Estimate: whichever n is smaller
- df Exact:

$$df_{exact} = \frac{\left(\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 2}\right)^2}{\left[\frac{\left(\frac{s_1^2}{n_1 - 1}\right)^2}{\left(n_1 - 1\right)}\right] + \left[\frac{\left(\frac{s_2^2}{n_2 - 1}\right)^2}{\left(n_2 - 1\right)}\right]}$$

- PRO TIP: “Solve this big equation by finding the components and then plugging them into the equation at the end.”

$\bar{x}_1 - \bar{x}_2 = 7.83 - 5.80 = 2.03$

$\frac{s_1^2}{n_1 - 1} = \frac{1.81}{6 - 1} = \frac{1.81}{5} = 0.36$ $\frac{s_2^2}{n_2 - 1} = \frac{11.76}{5 - 1} = \frac{11.76}{4} = 2.94$

$\sqrt{\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 1}} = \sqrt{0.36 + 2.94} = \sqrt{3.3} = 1.82$

- This is where we calculate t:

$t = \frac{\bar{x_1} - \bar{x_2}}{\sqrt{\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 1}}}$

$t = \frac{2.03}{1.82} = 1.115$

Decisions:

$df = n_2 = 5$ or you can use the exact df equation.

COMPARE t test calculate to critical for decision on rejecting the null hypothesis

- $t_{obtained=1.115}$
- $t_{critical}$ is for a DIRECTIONAL H_1 at $\alpha = 0.5$

Decisions: remember we are using DIRECTIONAL here

- One Tail: At $\alpha = 0.025$ level, $t_{critical}(df = 9) = 2.262 > 1.971$ SINCE $p < 0.05$
- Two Tail: At $\alpha = 0.05$ level, $t_{critical}(df = 9) = 1.833 < 1.971$

We use the One Tail ONLY FOR THIS LONG EXAMPLE. We conclude the following:

We CANNOT reject H_0 in favor of the alternative H_1 such that $\mu_1 > \mu_2$. This particular commercial does NOT increase favor ability ratings of the product! We tell the marketing department that if the assumption that the entire consumer population had viewed the commercial then their mean support score would NOT increase.

The DF long way for UNEQUAL VARIANCES

The equation for this procedure is:

$df_{exact} = \frac{\left(\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 2}\right)^2}{\left[\frac{\left(\frac{s_1^2}{n_1 - 1}\right)^2}{\left(n_1 - 1\right)}\right] + \left[\frac{\left(\frac{s_2^2}{n_2 - 1}\right)^2}{\left(n_2 - 1\right)}\right]}$

To solve it, using the previous example, we would find the components piece by piece and then plug them in to solve the equation.

Steps for the df calculation the long way:

- Step One df: $\frac{s_1^2}{n_1 - 1} = \frac{1.81}{6 - 1} = \frac{1.81}{5} = 0.36$
- Step Two df: $\left(\frac{s_1^2}{n_1 - 1}\right)^2 = (.36)^2 = 0.13$
- Step Three df: $\frac{s_2^2}{n_2 - 1} = \frac{11.76}{5 - 1} = \frac{11.76}{4} = 2.94$
- Step Four df: $\left(\frac{s_2^2}{n_2 - 1}\right)^2 = (2.94)^2 = 8.64$
- Step Five df: $\left(\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 1}\right)^2 = (0.36 + 2.94)^2 = (3.30)^2 = 10.89$

- Step Six df: $\frac{\left(s_1^2\right)^2}{(n_1-1)} = \frac{0.13}{6-1} = \frac{0.13}{5} = 0.26 \approx 0.3$
- Step Seven df: $\frac{\left(s_2^2\right)^2}{(n_2-1)} = \frac{8.64}{5-1} = \frac{8.64}{4} = 2.16$
- Step Eight df FINAL STEP!:
  $df_{exact} = \frac{\left(\frac{s_1^2}{n_1-1}+\frac{s_2^2}{n_2-2}\right)^2}{\left[\frac{\left(\frac{s_1^2}{n_1-1}\right)^2}{\left(n_1-1\right)}\right]+\left[\frac{\left(\frac{s_2^2}{n_2-1}\right)^2}{\left(n_2-1\right)}\right]} = \frac{10.89}{0.3+2.16} = 4.97$
- The computer would list this as df = 4.97
- DO NOT ROUND UP.
- The final df = 4
- We still have the same result, but sometimes this is a slight issue. Go with the computer generated values for most of the time. Rarely, do we ever do this by hand again. Well, unless our computers fail!
- Result: $t_{critical}$, one-tailed, 0.05α level, with $df = 4$
- $2.32 > 1.15$ and we CAN NOT REJECT THE H_0 IN FAVOR OF THE H_1

Warnings about directional hypotheses.

- Always use the critical value for the α that is given. Usually 0.05.
- Decide on one or two tail and stick with that decision throughout the analysis. Do not mix up the critical values for one tail or two tail for comparing your obtained statistic. If you pick one tail....STICK TO ONE TAIL!
- One tail requires prior knowledge to make that "leap" of statistical faith.

END OF LECTURE FOR TODAY... ON WEDNESDAY... OR next time

- Adjustments for Sigma-Hat Squared $\hat{\sigma}^2$
- Interpreting a Computer-Generated t Test
- Computer Applications: Independent Samples t Tests
- The Two-Sample t Test for Dependent Samples
- Computer Applications: Dependent Samples t Tests
- Statistical Significance versus Research Significance
- Statistical Power
- Conclusions/RECAP

Key Concepts

- Independent Samples (Independent Drawn Samples) [PAGE 273]
  - The composition of one sample is in no way matched or paired to the composition of the other sample. Thus, the two samples reflect two *SEPARATE POPULATIONS*
- Dependent Samples: Matched Pairs [PAGE 275]

- Situation in which members of one sample are not selected independently, but are instead determined by the makeup of the other sample. (They reflect the other group but are NOT THE SAME GROUP.)

- Dependent Samples: Matched Pairs (paired difference) t test[PAGE 275]
  - The t test used when the two samples are dependent samples.
- Pooled Estimate of Common Variance [PAGE 276]
  - Estimate based on a weighted average of two sample variances being used to estimate the population variance in finding the standard error.
- F test for Homogeneity of Variance [PAGE 276]
  - A test, based on the sample variances, used to determine the most appropriate t test formula to use.
- Paired Difference t test [PAGE 298]
  - A one-sample t test applied to the differences in each pairs of scores.
- Research Significance versus Statistical Significance
  - Research Significance [PAGE 306]: Relevance, importance of a particular difference of means or other findings.
  - Statistical Significance [PAGE 306]: The high probability that the difference between the two means or other finding based on a random sample is not the result of sampling error but reflects the characteristics of the population from which the sample was drawn.
- Statistical Power [PAGE 306]
  - The likelihood that our test will reject the null hypothesis when, in fact, H_1, relay is true.
- Type II Error (beta error) β [PAGE 307]
  - The probability that the null hypothesis is really false and H_1 is really true. Our obtained statistic (either z, t, etc.) was too low to enable us to reject the H_0 (null hypothesis) even though it "ought to be" rejected.
- Type I Error (alpha error) α [PAGE 307]
  - The probability of mistakenly rejecting a true null hypothesis (H_1)
- Small Effects [PAGE 308]
  - Hypothesized $.2\sigma$ difference between μ for the experimental group and μ for the control groups.
- Medium Effects [PAGE 307]
  - HYPOTHESIZED $.5\sigma$ difference between the population means.
- Large Effects [PAGE 307]
  - A $.8\sigma$ population mean difference.
  - Only the LARGER MEAN DIFFERENCES are found to be statistically significant for a majority of the time in social sciences.

Equations for this Chapter

Independently Drawn Samples

F Test for Homogeneity of Variables

$F = \frac{s^2_{larger}}{s^2_{smaller}}$

df numerator $= n - 1$ for the sample with the larger variance

df denominator $= n - 1$ for the sample with the smaller variance

The Two-Sample t Test Calculated from S

Equal Population Variances Assumed

$$t = \frac{\bar{x_1} - \bar{x_2}}{\sqrt{\left(\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

- Where $df = n_1 + n_2 = 2$

Unequal Population Variances Assumed

$$t = \frac{\bar{x_1} - \bar{x_2}}{\sqrt{\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 1}}}$$

- df Estimate: whichever n is smaller
- df Exact:

$$df_{exact} = \frac{\left(\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 2}\right)^2}{\left[\frac{\left(\frac{s_1^2}{n_1 - 1}\right)^2}{\left(n_1 - 1\right)}\right] + \left[\frac{\left(\frac{s_2^2}{n_2 - 1}\right)^2}{\left(n_2 - 1\right)}\right]}$$

The Two-Sample t Test Calculated from $\hat{\sigma}^2$

Equal Population Variances Assumed

$$t = \frac{\bar{x_1} - \bar{x_2}}{\sqrt{\left[\frac{(n_1 - 1)\hat{\sigma}_1^2 + (n_2 - 1)\hat{\sigma_2}^2}{n_1 + n_2 - 2}\right]\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}}$$

- Where $df = n_1 + n_2 - 2$

Unequal Population Variances Assumed

$$t = \frac{\bar{x_1} - \bar{x_2}}{\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}}$$

- df Estimate: whichever n is smaller
- df Exact:

$df_{exact} = \frac{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_1}\right)^2}{\left[\frac{\left(\frac{\sigma_1^2}{n_1}\right)^2}{n_1 - 1}\right] + \left[\frac{\left(\frac{\sigma_2^2}{n_2}\right)^2}{n_2 - 1}\right]}$

Dependent Samples

$t = \frac{\bar{D}}{\frac{S_D}{\sqrt{n_p - 1}}}$

- Where $S_D = \sqrt{\frac{\Sigma(D - \bar{D})^2}{n_p}}$

and $df = n_p - 1$

Next Time

1. Don't forget that Homework 2 is due
2. Chapter 10 (introduction to 10)
3. Examples of z scores are on the class website
4. Examples of t will be on the class website
5. Feedback will be on the class portal for each assignment
6. Winners will be announced on Wednesday
7. Wed. we will be doing a mock research project from start to finish
8. Homework 3 will be coming up (11/03)
9. Exam 1 (10/22)
10. Writing Draft (10/29)