

Comparing Two Means (Part 1)

Ray Block, Jr.
PS #585 (Research Methods)
Fall 2003

Today's Blueprint

Last Class

- Correlation (Part 2)
 - Measuring Correlation (A Recap)
 - Interpreting Correlation

Today's Class

- Correlation (A Recap)
- Comparing Two Means
 - Means Difference Tests

Correlation (A Recap)

Recap

Testing for Correlation:

- **Step #1:** Arrange the scores of X and Y in a table listing the following:
 - The squares of X and Y (X^2 and Y^2)
 - The products of X and Y (XY)
 - The sums of X and Y (ΣX and ΣY)
 - The sums of squares for X and Y (ΣX^2 and ΣY^2)
 - The sum of products of X and Y (ΣXY)
 - The mean of X and Y ($X\text{-Bar}$ and $Y\text{-Bar}$)
- **Step #2:** Using the sums, calculate Pearson's correlation:

$$r_{XY} = \frac{\sum X_i Y_i - N \bar{X} \bar{Y}}{\sqrt{(\sum X_i^2 - N \bar{X}^2)(\sum Y_i^2 - N \bar{Y}^2)}}$$

- **Step #3:** Find the degrees of freedom
 - The general formula is: $df = N - K$
 - For bivariate correlations: $df = N - 2$
- **Step #4:** Based on the information gained in the above three steps:
 - Compare Pearson's r you obtain with the critical value from Table [see Appendix in most Sstatistics textbooks?]
 - If the obtained value $>$ critical value, reject H_0
 - If the obtained value $<$ critical value, retain H_0

Comparing Two Means

Comparing Two Means

- For this class, we will move from *bivariate correlation* analysis to *bivariate experimental* analysis
- The goal is to examine the relationship between two variables, specifically:
 - How to determine if 2 variables are related
 - (If variables relate), how are they related
 - How to make predictions from one variable to another
- Correlation looks at the *relationship* between 2 variables
 - It tells us whether 2 variables share **something in common** with each other (relatedness)

- If they share something in common, then the variables are correlated (co-related) with one another
- While correlation is about examining *relationships* between variables, experimental analysis is about making *comparisons* between variables
- People use these comparisons to infer *causality*
- ...here's how
- Making comparisons means looking at the *difference* between two variables
 - It tells us whether 2 variables share **nothing in common** with each other
 - If they share nothing in common, then the variables are independent of one another
- The goal is to systematically change (manipulate) the variable to observe its effect on another variable
 - Example: manipulating the independent variable (X) so see its effect on the dependent variable (Y)
- A Simple Experimental Design Includes:
 - One dependent variable
 - One independent variable (manipulated 2 ways)
 - Treatment group
 - Control group

Experimental Analysis:

- *Between-Subjects Design*: Manipulate one of the variables using different subjects
 - Assigning different subjects to different experimental conditions
- *Within-Subjects Design*: Manipulate one of the variables using the same subjects
 - Give the same group of subjects to all the experimental conditions

[Insert Animated Illustration of Experiment Here]

- These examples show that the treatment causes the difference between the treatment and control groups
- However, is this difference significant (meaningful)?
- Use a T-Test to determine whether the difference between groups is significant

What are T-Tests

- They are tests of significance
- They determine whether there is a significant difference between the mean values of two groups
- Two Types (which one you use depends on whether you are testing the difference between related or unrelated groups:
 - *T-Tests for Independent Means* = Tests difference between the means of 2 unrelated groups
 - *T-Tests for Dependent Means* = Tests difference between the means of 2 related groups

T-Tests for Independent Means:

- Asks the question: "Is there a [significant] difference between the means of two unrelated groups?"
 - Unrelated Groups = When there are two experimental conditions and different subjects are assigned to each condition (Between-Subject design)

T-Test for Dependent Means:

- Asks the question: "Is there a difference between the means of two related groups"
 - Related Groups= when there are two experimental conditions and the same subjects took part in both conditions of the experiment (Within-Subject design)

The Logic behind T-Tests:

- Two samples of data are collected (in this case, samples for both groups)
 - You calculate sample means for each group
 - These means might differ a little or a lot

- Two samples of data are collected (in this case, samples for both groups)
 - You calculate sample means for each group
 - These means might differ a little or a lot

- If the samples come from the same population, then we expect their means to be roughly equal (or at least very similar)
 - If this is true, then any differences in the means are due to the treatment

- When testing hypotheses, the null would be that the manipulation had no effect on the subjects
- This means that any differences between means were due to chance and not the treatment
- To conduct this hypothesis test, we would compare the difference between the sample means we collected and the difference between the sample means that we would expect to obtain by chance
- If the difference between the sample means we collect is larger than the difference in sample means that we would expect by chance, then:
 - Either our sample is somehow not representative of the population we were trying to tap
 - Or the two samples we collected come from different populations

- The larger the difference between the obtained sample means and the expected sample means, the more confident we can be in rejecting the null hypothesis
- If we can reject the null, then we are essentially saying that the two sample means differ because of the manipulation

References

- FYI:
 - Levin, Jack and James Alan Fox. 2003. Elementary Statistics in Social Research, 9th Edition. Boston, MA: Pearson Education Group, Inc.
 - Salkind, Neil J. 2003. Exploring Research, 5th Edition. Upper Saddle River, NJ: Prentice Hall.
 - Krantzler, John H. 2003. Statistics for the Terrified, 3rd Edition. Upper Saddle River, NJ: Prentice Hall.