

Introduction to ANOVA



ANOVA

- ANOVA is an acronym that stands for Analysis of Variance
 - Can test for differences between more than two population means
 - Can test multiple factors



One-way versus Two-way

- 1-way ANOVA → testing only one factor
 - Example: the effect of hours spent studying on exam grade
- 2-way ANOVA → testing two factors
 - Example: the effect of hours spent studying and the amount of sleep that the students received the night before the test on their exam grade
 - Note: output will also include a p-value for whether there is significant interaction between the two factors



Assumptions

- The treatment data must be approximately normally distributed
- The variance must be the same for all treatments
- All samples must be randomly selected
- All samples must be independent



Process

- Null hypothesis is always that the means of the different levels of the factor are equal to each other
- If our p-value for a factor is low enough that we reject the null hypothesis we would conclude that the means are not equal to each other
 - We don't know which mean is different – just that there IS a difference



Other Details

- Typically done with computer software
- The test-statistic calculated is an F-statistic
 - It compares the variability between samples with the variability within each sample
 - The formula for the F-statistic for ANOVA is the ratio of the mean square treatments to the mean square error ($F = MST/MSE$)


