Analyses of Variance (ANOVA's)

A Basic Introduction

What ANOVA's do

- ANOVA's are used to answer a simple question:
  Do different treatment groups differ on a certain measure?
- In other words, did our treatment have an effect?

The Null-Hypothesis

- The Null-hypothesis is that the groups are NOT different
- What we try to do is to show that the Null-hypothesis is FALSE
- DON'T try to PROVE the Null-Hypothesis!

Independent Variable

- 'Treatment' can mean any two (or more) different ways in which we vary an experimental factor
- For example, we might vary whether subjects have a time constraint or not on making a grammaticality judgment.
- This is our INDEPENDENT VARIABLE
  ➔ It has to be CATEGORICAL

Dependent Variable

- What we measure is the DEPENDENT VARIABLE
- An example would be the percentage of how often a certain construction is judged to be grammatical.

  ➔ It has to be CONTINUOUS (or QUANTITATIVE)

Generalizing our Finding

- Just because for the few people we happened to check there was a difference in the different treatment conditions, that doesn't mean that that is true for the entire POPULATION
- ANOVA's tell you whether it is reasonable to generalize your finding

Sources of Variability

- An ANOVA does that by comparing how different sources of VARIABILITY contribute to the overall variation:
  - Some variance is just due to individual differences between people etc.
  - Some variance is due to our varying the independent variable

Within vs. Between Groups

- To measure how much of the variability is due to other factors, we look at how much the people WITHIN ONE TREATMENT group differ.
- To measure how much of the variability is due to our factor, we look at how much of a difference there is BETWEEN GROUPS.

The heart of ANOVA's

- The basic ratio of ANOVA's:

$$F = \frac{\text{BETWEEN GROUP VARIANCE}}{\text{WITHIN GROUP VARIANCE}}$$

What are the Odds? (I)

- The more variance is due to the treatment, and the less is due to other factors, the bigger your F-value
- The bigger your F-value, the less likely it is that the differences you found are due to chance

4 Research Scenarios

You think there is…

Actually, there is…	**A Difference**	**No Difference**
A Difference	Yeah!	Hmm. Try again TYPE II ERROR
No difference	Lucky (sort of) TYPE I ERROR	Wrong Question?

What are the Odds? (II)

You think there is…

Actually, there is…	**A Difference**	**No Difference**
A Difference		TYPE II ERROR ➔ POWER
No difference	TYPE I ERROR ➔ 5 %	

What are the Odds (III)

When 20 of us run the same study, one of us will find a significant difference between groups,
EVEN IF THERE ACTUALLY IS NO DIFFERENCE!

Calculating Sums of Squares

An Example

- Let's calculate a simple example
- For simplicity, we will look at a between subjects design

➔ each subject contributes one data point

An Example

- We want to know whether taking a linguistics class affects your grammaticality judgments.
- 2 groups: students that have taken a class and students that have not
(this is the independent variable)
- We give them 20 grammatical sentences taken from syntax papers. They have to decide whether they are grammatical or not.
Dependent Variable: How many out of the 20 sentences do they find grammatical?
- There are 4 subjects in each treatment group

Total Sums of Squares

	Score (X)	X – overall mean (10)	Squared
Group 1 (mean 5)	3	-7	49
	4	-6	36
	5	-5	25
	8	-2	4
	Group1 total SS		**114**
Group 2 (mean 15)	13	3	9
	13	3	9
	16	6	36
	18	8	64
	Group 2 total SS		**118**
	Total SS		***232***

Variance

- We get (something very close to) the mean of the SS by dividing the SS by the degrees of freedom (usually # of x – 1)
- MS = variance = $\frac{SS}{df} = \frac{232}{7} = 33.1$

Standard deviation

- We get (something very close to) the mean distance of the data points from the overall mean by taking the square root of the variance
- Variance = 33.1
- Standard deviation = square root of variance = 5.8

More Sums of Squares

- We have calculated the overall variance
- How much of this is due to
  - Variability within groups?
  - Variability between groups?
- Let’s calculate Sums of Squares for these

Within Group Sums of Squares

	Score (X)	X - group mean	Squared
Group 1 (mean 5)	3	-2	4
	4	-1	1
	5	0	0
	8	3	9
	Group1 within group SS		**14**
Group 2 (mean 15)	13	-2	4
	13	-2	4
	16	1	1
	18	3	9
	Group 2 within group SS		**18**
	Total Within Group SS		***32***

Between Group Sums of Squares

	Score (X)	group mean	Group mean – overall mean	Squared
Group 1 (mean 5)	3	5	-5	25
	4	5	-5	25
	5	5	-5	25
	8	5	-5	25
	Group1 contribution to between group SS			**100**
Group 2 (mean 15)	13	15	5	25
	13	15	5	25
	16	15	5	25
	18	15	5	25
	Group 2 contribution to between group SS			**100**
	Total Between Group SS			***200***

Relation between SS

- You may have noted that the total SS is the sum of the within group and between group SS:

$$SS_{total} = SS_{within} + SS_{between}$$

The F-ratio

- We almost have everything we need for calculating the F-ratio. We just need to calculate the within group and between group variance

- MS_{within} = 32/6 = 5.33
- $MS_{between}$ = 200/1 = 200

The F-ratio

$$F = \frac{MS_{between}}{MS_{within}} = \frac{200}{5.33} = \mathbf{37.5}$$

What does this tell you?

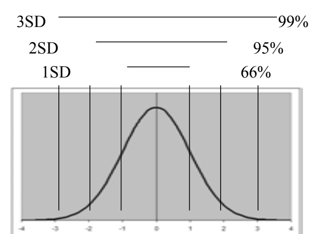
- How likely is it that you would have ended up with this F-value by chance?
- This depends on how many subjects you ran in how many conditions.
- You could look this up in an F-table, but any program will give you the desired p-value.

What are the odds, really?

- Imagine your study would have been run a large number of times
- Even if there was no difference, every once in a while you'd get a high F-value by chance
- F-tables tell you how likely it is that this happened

What are the odds, really?

- The F-values one would get for running a study over and over would form a normal curve.



- The p-value tells you the probability of having found a difference even though there isn't one

The p-value

- The typical cutoff accepted in the social sciences is $p = .05$
- That is, if we are 95% sure that our result did not come about by chance, we accept it

(This is, of course, somewhat random)

Appendix I - Formulas

- MS = SS/ df
- SD = square root of MS
- $F = MS_{between} / Ms_{within}$

Appendix II – degrees of Freedom

- n: # of subjects per treatment group
- a: # of treatment groups

- $df_{total} = (a * n) - 1$
- $df_{between} = a - 1$
- $df_{within} = a (n-1)$

- Note: $df_{total} = df_{between} + df_{within}$

Designing your study

- We usually use WITHIN SUBJECT designs I.e. every subject sees every condition
- A very typical design:
  - 2/4 conditions
  - 6 items per condition
  - ~ 24 subjects
  - Counterbalance items/conditions

	S1	S2
I1	6 x A	6 x B
I2	6 x B	6 x A