The General Linear Model – ANOVA Part 1 (Video 1)

Dr Andrew J. Stewart

E: drandrewjstewart@gmail.com
T: @ajstewart_lang
G: ajstewartlang

Analysis of Variance (ANOVA)

- We're going to have our first look at the Analysis of Variance (ANOVA).

- This week we'll look at ANOVA for within-subjects, between-subjects and factorial designs.

- ANOVA is an important statistical model and (in various forms) is used widely across many areas of biomedical and life sciences.

Why ANOVA vs. lots of t–tests?

- t-tests are fine if we're just comparing two means.

- In the real world of experiments, we often have more than two conditions.

- How could we analyse our data?

- One possibility could be that we do multiple t-tests – but there's a problem with that.

Why ANOVA vs. lots of *t*–tests?

- With one *t*-test, at $p < 0.05$ alpha level there is a 5% chance of falsely rejecting our null hypothesis (Type I error).

- If we have three conditions, then we have three pairs of means to compare (Condition 1 *vs*. Condition 2, Condition 2 *vs*. Condition 3 and Condition 1 *vs*. Condition 3).

The Familywise Error Rate

- For each test, there is 0.95 probability of not having a Type I error.
- But when we do three tests the probability is 0.95 x 0.95 x 0.95 (= 0.857)
- So that means there is a 14.3% chance of us (falsely) rejecting the null hypothesis.
- This is known as the **familywise error rate**.
  - familywise error = $1 - (0.95)^n$
- If we had 5 conditions, (10 *t*-tests), our error rate would be 0.4 − a 40% chance of having made at least one Type I error (i.e., thinking we have an effect when none is present).

Similarities between t–tests and ANOVA

- t-tests tell us whether or not two samples have statistically indistinguishable means.

- ANOVA tells us whether two or more samples have statistically indistinguishable means.

- As the t-test produces the t-statistic, the ANOVA gives us an F-statistic or F-ratio which compares the amount of systematic variance in our data with the amount of unsystematic variance.

Similarities between *t*-tests and ANOVA

- ANOVA can tell us that there is a difference between means – but it doesn't tell us where the difference is.

- For three Conditions, it doesn't tell us whether the mean of Condition 1 differs from both the means of Conditions 2 and 3, or whether the mean of Condition 2 differs from the mean of Condition 3 but not from the mean of Condition 2 etc.

ANOVA – an example

- Imagine we're interested in the impact of caffeine consumption on an individual's motor performance.

- It's a between-subjects design with 3 conditions:
  - A low amount of caffeine (single espresso)
  - A large amount of caffeine (double espresso)
  - A placebo group (water)

ANOVA – an example

- We conduct an ANOVA and find a significant F-ratio.
- What does it mean?
- The single espresso people could have performed better than the double espresso and water group.
- Or maybe they performed the same as the water group but better than the double espresso group.
- Or maybe (unexpectedly) they performed worse than both the double espresso and water groups.
- To know what is the case we need to do planned contrasts (similar to 1-tailed tests) or post hoc tests (similar to 2-tailed tests).

ANOVA – an example

- We know that at least one of our means differs from at least one of our other means but (so far) we don't know where that difference lies…

- Luckily things easy for us as we can conduct what are known as post hoc tests.

- These will tell us which means differ from which other means (and allow us to begin to tell a story.)

Post hoc tests

- These work by doing pairwise comparisons on all the different combinations of experimental groups.

- They control for the familywise error rate to get round the problem of multiple comparisons.

- Bonferroni method divides our critical p-value (0.05) by the number of tests. If we are conducting ten tests, then for each test the critical p is 0.005 – but this increases our chances of a Type II error – missing an effect when it's there.

Post hoc tests

- When deciding which post hoc test to use :
  - Does it control the Type I error rate ?
  - Does it control the Type II error rate ?
  - Is it reliable when ANOVA assumptions have been violated ?

LSD, Bonferroni, and Tukey tests

- The least significant differences test (LSD) doesn't control the Type I error and is like doing multiple t-tests on the data (but only if the ANOVA itself is significant).

- Bonferroni and Tukey both control for Type I errors but are conservative. Bonferroni works by dividing the critical alpha level by the number of tests conducted.

- Tukey is less conservative than Bonferroni.

The next video...

In the next video, we'll look at how we conduct a between participants ANOVA in R...