PSY3062 RESEARCH METHODS AND THEORY

CONTENTS

INTRODUCTION TO PSYCHOLOGICAL STATISTICAL TESTS

Test	How many predictor variables (IV)?	How many response variables (DV)?	What type of data is/are the IV's?	What type of data is/are the DV's?	Example research question
Multiple Regression	Multiple	One	Continuous	Continuous	Do attitude towards politics and empathy predict attitudes towards asylum seekers?
Moderated Regression	Multiple	One	Continuous	Continuous	Is the relationship between depression scores and daytime sleepiness different at varying levels of physical activity?
Hierarchical Regression	Multiple	One	Continuous	Continuous	Controlling for age, does height predict calorie intake?
One-way ANOVA	One	One	Categorical	Continuous	Geographical location vs days of exercise per week
Factorial ANOVA	2 or more	One	Categorical	Continuous	Geographical location and gender vs days of exercise per week

INTRODUCTION TO ANOVA

What is an ANOVA?

- Compares means between groups
- Can test the difference between two or more group means
- Tests the null hypothesis

What is the null hypothesis?

- Hypothesis that the means for all of the groups in the research are identical.
- $H_0 = \mu_1 = \mu_2 = \mu_3$...
- Support for the research/alternate hypothesis must result in rejection of the null hypothesis.

Why not use multiple t-tests?

- We cannot look at more than one IV at a time.
- Inflates our Type I error rate.

- Familywise alpha level (αFW) – probability of making at least one type I error amongst a series of comparisons.
- Decisionwise alpha level (αDW) – alpha level for each comparison.

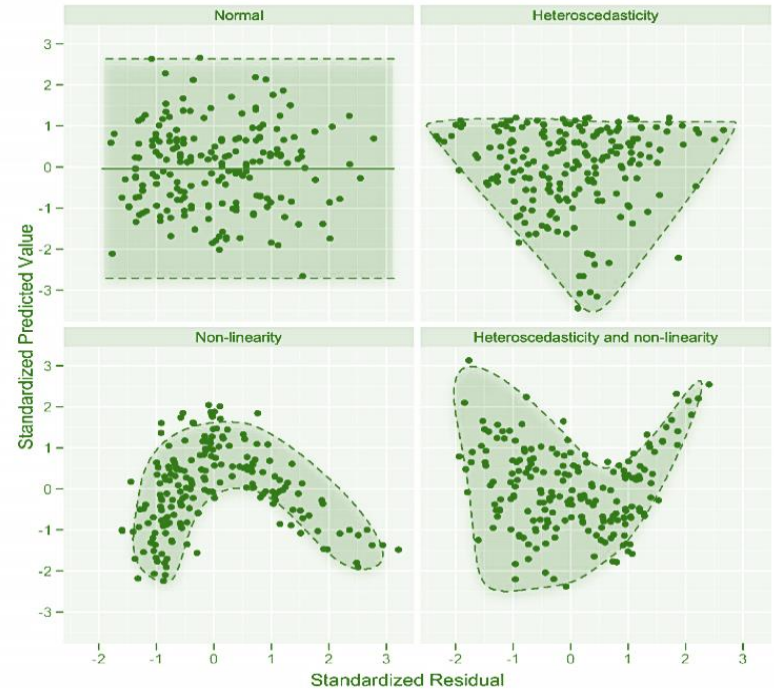
- When assumptions are met, ANOVA is more powerful than t-tests (for 2+ groups).
- ANOVA uses the pooled variance estimate across all groups, and the df draws on the sample size. This means that a larger sample size results in a higher power to detect statistically significant differences.

Underlying theory/definitions

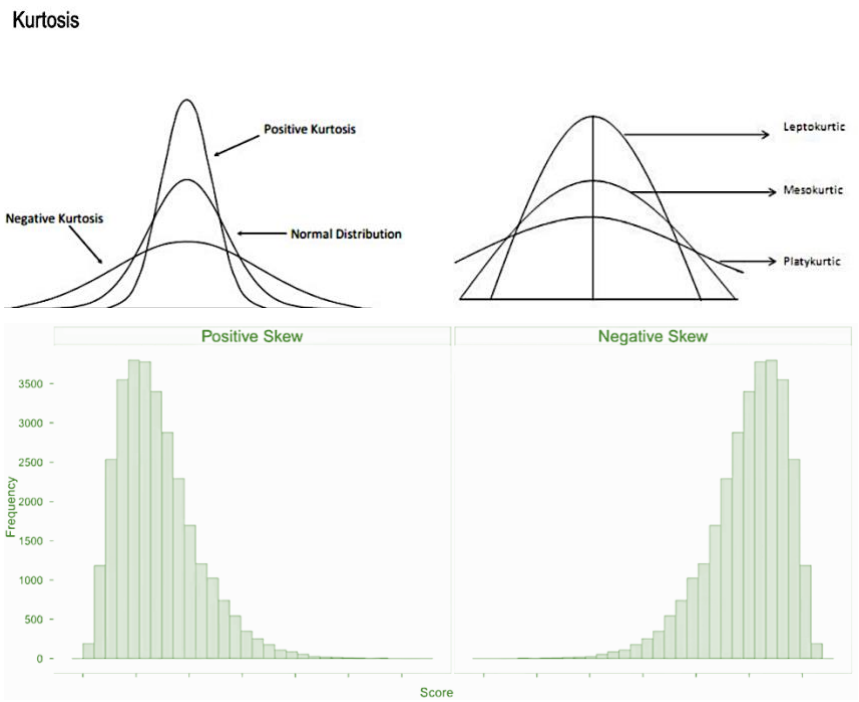
- Total Sum of Squares (SS_T) – total variability between scores
- Model Sum of Squares (SS_M) – variability between group means
  - This is the explained variance (how much variability is accounted for by the IV).
- Residual Sum of Squares (SS_R) – the unexplained variance (due to chance).
- F-ratio – statistic used to determine whether the model explains more variability than the residuals.
  - Large f-ratio → variability attributable to the model is greater than the variability that occurs simply due to chance (eg. error).

Assumptions of ANOVA tests

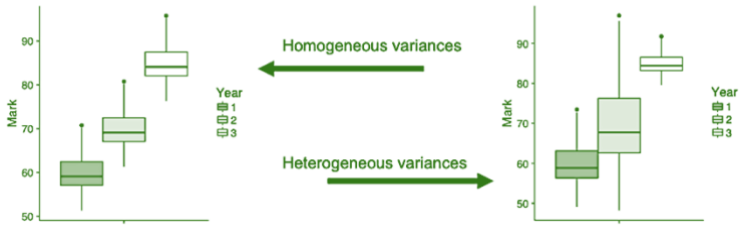
- Independence of observations
  - Participant's results are not impacted by one another in any way.
  - Eg. copying other participants, repeated-measures design.
  - If this is violated there will be a downward bias of standard errors, and incorrect statistical inference (p-values will be wrong).
  - If this is violated do not run the analysis.
- DV should be continuous (interval/ratio data), and IV(s) should be categorical. If this is violated you cannot continue with the ANOVA, must use another test.
- No significant outliers
  - Violated?
    - Check data for significant outliers (eg. boxplot analysis, z-scores outside of +/- 3.29).
    - Remove the participant or trim the data
    - Transform the data;
      - Log (x) – reduce positive skew and stabilise variance, only for positive values.
      - Square root – reduce positive skew and stabilise variance, used for 0 and positive values.
      - Reciprocal (1/x) – reduce the impact of large scores and stabilise variance.
    - Winsorize the scores
      - Change it to the next highest value plus some small value.
      - Convert score to that expected of a z-score of +/- 3.29
      - Convert the score to the mean plus two or three standard deviations.
      - Concert the score to a percentile of the distribution.
- Normality - The scores must be normally distributed around the population mean (μ).
  - Violated?
    - If sample sizes are equal across groups and large, ANOVA should be okay to continue. (Central Limit Theorem)
    - Check Q-Q plot and histograms to assess visual display of normality
      - Q-Q plot will form a straight line if data is normal
    - Check the Shapiro-Wilk Test to see if data is normal (statistically significant = violated)
    - Otherwise, consider transforming the data or using the relevant non-parametric test.



- Kurtosis = heaviness of the tails
  - Leptokurtic = heavy tails
  - Platykurtic = light tails (not desirable)



- Homogeneity of variance - The variance of each group must be approximately equal
  - Violated?
    - If sample sizes are equal between groups, ANOVA should be okay to continue.
    - Otherwise, Levene's statistic in the statistical output should be non-significant.
    - If this statistical output is significant, do not use the F-ratio, must use an alternative statistic suited to the type of ANOVA.



Post Hoc Tests

- Determines where the differences between the means lie.
- In what ways can we do this?
  - Multiple t-tests – inflates our chances of making an error.
  - Orthogonal contrasts/comparisons – use if you want to only assess specific groups based on your hypothesis
  - Trend analysis – use to analyse specific patterns within the data.
  - Post hoc tests – use to compare all groups
- Bonferroni Method
  - Uses a stricter αDW to accept effects as significant.
  - Use the new value to be more careful with our data and decrease our chance of errors.

Effect Size

- Null hypothesis significance testing is either a significant result or non-significant result, however real life is not like this.
  - Statistical significance does not indicate real-world significance.
- Significance testing rules out chance, but p-values do not tell us about the size of effects.
- Different effect sizes:
  - r^2 – regressions and correlations
  - η^2 – ANOVA
    - Biased effect size estimate
    - Overestimated proportion of variability that is accounted for
    - Small effect = .01, medium = .09, large =.25
  - ω^2 – ANOVA
    - Unbiased effect size estimate as it takes more information from the data to give a better estimate
    - Best option if the sample size is not large
    - Small effect = .01, medium = .06, large = .14
  - Cohen's d – t-test or ANOVA
    - Degree of separation between two distributions
    - Small = .2, medium = .5, large = .8

Contrasts – planned a priori

- Planned comparisons based on research hypothesis.
- Trying to figure out where the differences lie across the IV's
  - Test each specific hypothesis to see which groups differ
  - If hypotheses are independent of each other, the familywise type I error will be controlled (it will not be inflated)