Contents

Week 2:

Statistics- logical, principled way of understanding the world
Signal/ noise ratio = effect/error
Standard error depends on the variability in the sample and the number of people.
Therefore, we can reduce error by increasing the number of people or decreasing the variability
We need to conduct research in order to acquire accurate knowledge about why people think, feel and behaviour
Psychology topics are largely subjective, but it is still important to engage in objective research- need to engage in evidence based practice
It is important that people interpreting results also have an understanding of sources of bias
Need to get it right from the start- hard to remove confounds after you run the analysis

Research process

1. Find a research idea- topic and hypothesis
    - You need to be able to actually measure what the variable is
    - The overall question drives everything
2. Define and measure variables
    - Understand how the measures you are using work and how they will help you answer your research question
3. Identify participants/subjects
4. Select a research strategy
5. Select a research design
    - Need to choose a study design with the future analysis in mind
    - Do not alter standard scales without clearly defined reason and justification

- Easy to collapse across continuous variables to create categorical data, but you cannot go the other way
- Therefore, it is better to begin with broad data so that any number of analyses are possible- this usually means that you would use interval/ratio measures

6. Conduct the study
    - Participate actively in the data collection
    - Do the tests yourself so you can understand the strengths and limitations
7. Evaluate data
    - Coding and analysis
        - Check the scoring key against the survey instrument to identify item weights and reverse scoring
        - Assign unique numbers to missing values to distinguish them from missing values, as you should never delete data
        - Check the order of the data input
        - Make sure the data sets are comparable before combining them to make a larger sample
        - All data should have clear and permanent labels
        - Do not discard raw data
    - the statistical tests used must be appropriate to the type of data
    - Things to consider:
        - How many outcome variables are there?
        - What types of outcomes are they? (Categorical or continuous variables)
        - How many predictor variables?
        - What type of predictor? (independent or repeated measures)
        - How many categories does the predictor variable have?
        - Are the assumptions of the linear model met?
    - Two approaches to analysis:
        - Null hypothesis significance testing
            - Using p values
            - We are hoping to reject the null hypothesis
            - Develop the alternative hypothesis and identify the null hypothesis
            - If the data are unlikely under the null hypothesis then the H1 has survived another day- never prove or accept, it is the best explanation
            - Otherwise, the null hypothesis is a good explanation for the data- better than the alternative
            - The p value tells us the probablility that we would obtain the found effect if the null hypothesis was true

    - Therefore, when we accept the significance level at .05, we are saying that there is a 5% chance that we are determining there is an effect when there really isn't.
  - Parameter estimation with CIs
    - If the null hypothesis value is outside of the 95% CI (the region of plausible values) then we can reject it
    - Gives an idea of how precise the study is
    - narrow is good, broad is bad
8. Report results
9. Refine and reformulate your research idea

Bias- someone isn't evaluating the evidence in an objective way

Linear hypothesis model- predict an outcome from some kind of model.

Three contests of bias:

- Things that bias the parameter estimates
- Things that bias standard errors and confidence intervals
- Things that bias test statistics and p-values
- All related

Assumptions

- Parametric- we are relying on the data being a certain way
- Most sources of bias come from the violation of assumptions
- If we use tests on data that violate the assumptions we are contradicting ourselves
- All tests basically have the same assumptions
- Main ones are:
  - Additivity and linearity
  - Normality
  - Homoscedasticity or homogeneity of variance
  - Independence

Additivity and linearity

- The outcome variable is linearly related to predictors
- This is important for the model to be described correctly

Assumption of normality

  - Not necessarily that the data follows a normal distribution
  - We are talking about normality within the groups, not between them. So if we have two groups in our experiment, the assumption of normality applies within these two groups, not to the overall sample.
  - Means different things in different contexts:
    - For significance tests the sampling distribution of what is being tested must be normally distributed. Eg. Sampling distribution of means needs to be normal to test whether two means are different
      - sample distribution- the frequency distribution of the data in our sample- this is not generally assumed to be normal