Applied Research Methods (MULT30018)

Tutorial 5

Nhan La

W1. Introduction
W2. Descriptive statistics
W3. Crosstabs and correlations
W4. Regression 1
W5. Regression 2
W6. Experiments
W7. Counterfactuals
W8. Mixed methods
W9. Group presentations
W10. Final assignment workshop

- Review basics of regression
- Get to know the assumptions of linear regressions
    o Collinearity problem
- SPSS practice
    o Group activity

- Purpose
- Formulation
- Interpretation
  - Line and slope
  - Coefficients
  - Significance
  - Model fit (R-squared)
    - Perfect model?
  - Degrees of freedom
  - Standard error

Linear regression assumptions:

-   Linearity in parameters between the dependent variable and independent variables
    o   Violation: non-linear relationships
-   Error term has expected value (or mean value) of 0
    o   Violation: intercept is biased
-   Error term has constant value across variables and over time
    o   Violation: model uncertainty
-   Random sampling
    o   Violation: serial correlation/autocorrelation
-   Error term is uncorrelated with independent variables
    o   Violation: omitted variable bias
-   Multicollinearity
    o   Violation: model uncertainty/unstability

- The dependent variable must have 3 or more categories that can be ranked
- Need theory to run meaningful regressions
    - Junk-in, junk-out
- Linear regression only measures linear associations
- Only make predictions within sample range

Satisfaction with your life [V23] = Marital status [Recoded] + Scales of income [V239]

- Identify a dependent variable and two or three independent variables
    - Choose one that need dichotomisation
- Run the regressions, including multicollinearity test
- Discuss the findings in class
    o What percentage of the variance in the outcome variable do your predictors explain as a whole?
    o Which of your predictors was relatively the most important in explaining the dependent variable?
    o Which predictors are positively and which are negatively associated with the dependent variable?
    o To what magnitude does each of the independent variables affect your dependent variable?
    o How many cases (i.e., participants) are included in your regression model?
    o Are there any problems with multicollinearity?
    o Should you include all of the independent variables in your model, or should you drop some of them?

?