Week of March 9

1. Measurement: Definitions
2. Classical test theory
3. Methods of assessing reliability
4. Assessing reliability using SPSS

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment.* Thousand Oaks, CA: Sage.

Measurement: Definitions

- Measurement: The assignment of numbers to events, objects, or constructs
  - Measure (n.): A specific method of assigning numbers to represent a specific event, object, or construct
- Reliability: The extent to which a measure yields the same results on repeated trials
- Validity: The extent to which scores on a measure indicate what they are meant to indicate

Classical Test Theory: True Scores and Error

$$X_{ij} = t_i + e_{ij}$$

- *X*: A person's observed score on a measure
- *t*: A person's true score on the construct being measured (i.e., the average score that would be obtained if the person were measured an infinite number of times)
- *e*: A person's error score (i.e., the deviation of their observed score from their true score)
  - Assumed to be normally distributed with a mean of 0
  - Assumed to be uncorrelated with *t* (i.e., random)

Classical Test Theory: Reliability

- Equation: $\rho_{XX} = \frac{\sigma_t^2}{\sigma_X^2} = \frac{\sigma_t^2}{\sigma_t^2 + \sigma_e^2}$
- Words: Reliability is the proportion of variability in a set of observed scores that reflects variability in true scores
  - Reliability is specific to a particular measure used with a particular population or sample
  - In social science research, the conventional standard for acceptable reliability is .70 (Nunnally, 1978)

Why Reliability Matters

- Observed associations are attenuated (made smaller) by measurement error
  - Attenuation due to unreliability:

    $r_{XY} = \rho_{XY}\sqrt{\rho_{XX}\rho_{YY}}$
  - Correction for attenuation:

    $\rho_{XY} = r_{XY} \Big/ \sqrt{\rho_{XX}\rho_{YY}}$

Assessing Reliability: Internal Consistency Method

- Estimates the reliability of a measure from the correlations among the test items
- Cronbach's alpha: The most common internal-consistency estimate; depends on two quantities:
  - r_{ij}: The mean interitem correlation
  - *N*: The number of test items

$$\alpha = N\overline{r}_{ij} \Big/ \left[1 + \overline{r}_{ij}\left(N - 1\right)\right]$$

Cronbach's Alpha: Pros and Cons

- Advantages
  - Only requires data from a single test at a single time point
  - Does not require an arbitrary split of the items
- Disadvantages
  - Cannot estimate the reliability of single-item tests
  - Often underestimates the reliability of tests with non-normally distributed items (e.g., dichotomous items)

Assessing Reliability: Other Methods

- Retest Method: Estimates the reliability of a measure as the correlation between a sample's scores at time 1 and the same sample's scores at time 2
- Parallel Forms Method: Estimates the reliability of a measure as the correlation between a sample's scores on two parallel forms
- Split-Halves Method: Uses the Spearman-Brown prophecy formula to estimate the reliability of a measure from the correlation between two halves of the test items

Assessing Reliability: Using SPSS

1. Make sure that all items are scored in the same direction (i.e., false-keyed items have been reverse-scored)
2. Use Scale → Reliability Analysis
3. Define the Items
4. Choose the reliability Model (alpha or split-half)
5. Recommended: In Statistics, ask for (a) Descriptives for Scale if Item Deleted and (b) Summaries / Correlations
6. Click OK to run the analysis