Lab 1 - Correlations, reliability and consistency

Tony Tan & Jarl Kristensen

University of Oslo

Monday, 7 November 2022

Today

Estimate and interpret:

- Correlations between variables
- Test-retest reliability
- Alternate/parallel test-forms
- Internal consistency
  - Cronbach's alpha
  - Single factor model
  - McDonald's omega

Correlation

$$\rho_{XY} = \frac{\mathrm{Cov}\,(X, Y)}{\sigma_X\,\sigma_Y}$$

- Correlation is a measure of the strength of the linear relationship between two variables.
- The correlation coefficient is a number between -1 and 1, where 0 indicates no linear relationship, and 1 or -1 indicates perfect a linear relationship.
- Correlations are superior to covariances for interpreting interrelationship between two variables thanks to the standardisation procedure.

Task 1: Correlation

Estimate the correlation between a pair of variables in likert_data.rds and comment on the relationship between the variables you chose.

Test reliability

- Test-retest reliability
- Alternate test forms

Use the correlation of the sum scores of the two test forms to estimate the reliability of the test.

Cronbach's α

$$\alpha = \frac{m}{m-1}\left[1 - \frac{\sum_{j=1}^{m} \mathrm{Var}\left(X_j\right)}{\sigma_Y^2}\right]$$

Task 2: Test reliability

- Estimate coefficient alpha for one of the scales in the dich_data.rds dataset.
- How would you describe the estimated value?

Cronbach's α: Pitfalls

- Dependent on the number of items
- Assumes a single factor model with equal loadings for all items
- A lower bound estimate of reliability as long as assumptions are met

Single factor model

- lavaan and other R packages
- This lab focuses on lavaan
- Lecture 6: single factor model

lavaan: Syntax

Operator	Reading	Meaning
$=\sim$	is measured by	define a latent variable
$\sim$	is regressed on	define a regression model
$\sim\sim$	is correlated with	specify covariances

lavaan: Example code

```r
# Load lavaan
library(lavaan)

# Define latent variable
lat_var = " y =~ x1 + x2 + x3 "

# Run a confirmatory factor analysis
cfa <- cfa(lat_var, data = mydata)

# Model evaluation
summary(cfa, fit.measures = TRUE)

# Extract model coefficients
coef(cfa)
```

Task 3: Single factor model

- Estimate the single factor model using the scale you chose in Task 2
- Evaluate the model fit.
- Evaluate if the α you calculated in Task 2 violates the assumptions α relies on.

Coefficient omega

$$\omega = \frac{\sigma_C^2}{\sigma_C^2 + \sigma_U^2} = \frac{\sigma_Y^2}{\sigma_{T_Y}^2 + \sigma_{E_Y}^2} = \frac{\left(\sum_{j=1}^{m} \lambda_j\right)^2}{\left(\sum_{j=1}^{m} \lambda_j\right)^2 + \sum_{j=1}^{m} \Psi_i^2}$$

Coefficient omega is the ratio of the true score variance (common) to total score variance (common + unique).

Task 4: Standardised model

- Use the coefficients from Task 3 to calculate ω.
- Evaluate the reliability of the scale.
- Compare ω (Task 4) and α (Task 3).