Confirmatory factor analysis

Wan Nor Arifin

Unit of Biostatistics and Research Methodology, Universiti Sains Malaysia.

email: wnarifin@usm.my

December 15, 2018

1. Introduction
2. Confirmatory vs Exploratory factor analysis
3. Confirmatory factor analysis (CFA)

Learning outcomes

- understand the basic concepts and application of CFA.
- able to perform CFA using R.

Introduction

Introduction

Structural equation modeling (SEM):

- measurement model – CFA.
- structural model – path analysis.

Introduction

- confirmatory method – e.g. at final stage of questionnaire development.
- theory confirmation.
- based on common factor model – similar to EFA.
- accounts for measurement errors.
- analysis done on variance-covariance matrix.
- allows assessment of model fit.

Confirmatory vs Exploratory factor analysis

CFA vs EFA

CFA items:

- I love fast food
- I hate vegetable
- I hate eating fruits
- I hate exercise

$\rightarrow$ Obesity

CFA vs EFA

EFA items:

- I love cat
- I hate snake
- I love statistics
- I love snorkelling
- I support Harimau Malaya team
- I love driving car
- I love computer game
- I like to have everything normally distributed
- I think of independent t-test everyday
- My favourite food is nasi ayam
- I used to eat a lot of pisang goreng
- I spend most of my time in front of computer
- I love SPSS

$\rightarrow$ What factors???

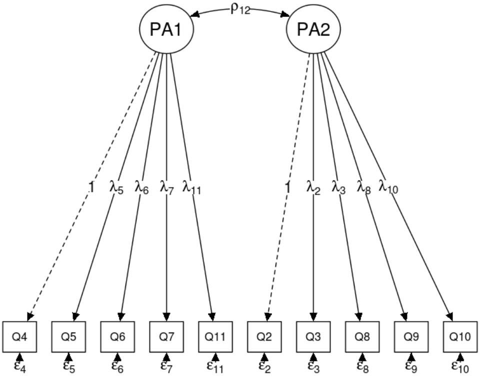
CFA vs EFA

EFA	CFA
Exploratory procedure.	Confirmatory procedure.
No pre-requisite to specify theoretical factors for a collections of items.	Pre-specified theoretical factors.
Aims to explore the items and extract common ideas. Theory generating based on empirical findings.	Strong theory. Just want to confirm.
Items free loading and not fixed to factors.	Items are fixed to pre-specified factors.
Rotation of factors is used to allow simpler solution.	Rotation not used.
Explicit hypothesis is not tested.	Explicit hypothesis testing. Allows assessment of model fit (χ^2 GOF, Fit indices).

Confirmatory factor analysis (CFA)

CFA

Path diagram representation of our previous EFA model:



CFA

Research question:

- Does our theoretical measurement model (factor + item + correlation) fit the data?
- Strength of relationship between items and the factors?
- Factor correlations?

$\rightarrow$ confirm the theory.

CFA

Applications:

- confirm the measurement model of newly developed questionnaires in target populations.
- confirm the measurement model of existing questionnaires in new populations.
- confirm the measurement model translated questionnaires in new populations.

$\rightarrow$ confirmatory role of the analysis.

CFA

Estimations methods:

- numerical scale + multivariate normal data → maximum likelihood.
- numerical scale + non-multivariate normal data → robust maximum likelihood.
- categorical scale → weighted least squares.

Many more ... http://lavaan.ugent.be/tutorial/est.html

CFA

Results to focus on,

1. Overall model fit – by fit indices.
2. Parameter estimates
    - Factor loadings.
    - Factor correlations.
3. Localized areas of misfit
    - Modification indices.
    - Residuals.

1. Fit indices.

The following are a number of selected fit indices and the recommended cut-off values (Brown, 2015; Schreiber, Nora, Stage, Barlow, & King, 2006):

Category	Fit index	Cut-off
Absolute fit	χ^2	$P>0.05$
	Standardized root mean square (SRMR)	≤ 0.08
Parsimony correction	Root mean square error of approximation (RMSEA)	and its 90% CI ≤ 0.08, CFit $P>0.05$
Comparative fit	Comparative fit index (CFI)	≥ 0.95
	Tucker-Lewis index (TLI)	

CFA

2. Parameter estimates: Factor loadings (FLs).

- The guideline for EFA is applicable also to CFA (FL > 0.3).
- In addition, the *P*-values of the FLs must be significant (at $\alpha = 0.05$).
- Also look for out-of-range values – FLs should be in range of 0 to 1 (absolute values).

2. Parameter estimates: Factor correlations.

- Similar to EFA, a factor correlation must be < 0.85, which indicates that the factors are distinct.
- A correlation > 0.85 indicates multicollinearity problem (Brown, 2015).
- Also look for out-of-range values – factor correlations should be in range of 0 to 1 (absolute values).

CFA

2. Parameter estimates

- Out-of-range values are called *Heywood cases* or *offending estimates* (Brown, 2015).
- When a model has Heywood cases, the solution is not acceptable.
- The variance-covariance matrix (of our data) could be *non-positive definite* i.e. the matrix is not invertible for the analysis.

CFA

3. Localized areas of misfit: Residuals (Brown, 2015):

   - Residuals are the difference between the values in the sample and model-implied variance-covariance matrices.
   - Standardized residuals (SRs) $> |2.58|$ indicate the standardized discrepancy between the matrices.

CFA

3. Localized areas of misfit: Modification indices (MIs) (Brown, 2015):
   - A modification index indicates the expected parameter change if we include a particular specification in the model.
   - e.g. by correlating between errors of Q1 and Q2.
   - Specifications with MIs > 3.84 should be investigated.

CFA

Model revision – to improve model fit. Causes of poor fit:

- Item – low FL, wrong factor.
- Factor – multicollinearity.
- Correlated error – items with similar wording/meaning.
- Model not supported by data.

Model-to-model comparison – to choose the best model.

1. Compare models that do not involve item removal – Nested model
   - χ^2 difference.
   - AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion).
   - Better model = Smaller AIC/BIC (Brown, 2015).
2. Compare models that involves item removal – Non-nested model
   - AIC and BIC.

Practical

cfa_short.R

Bibliography

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. New York: The Guilford Press.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis*. New Jersey: Prentice Hall.

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research*, *99*(6), 323–338.