Assessing Fit of IRT Models

Nina van Gerwen & Dave Hessen

14th of December, 2022

Methodology & Statistics for the Behavioural, Biomedical and Social Sciences

Utrecht University

Outline

Utrecht University

1. Background information

- Item Response Theory
  - Latent constructs
  - Tests/questionnaires
- Assessing model fit
  - Goodness-of-fit tests
  - Fit indices

Utrecht University

What's the issue?

- Few goodness-of-fit tests
  - Multiple issues with current tests [1]
- Scarce studies for TLI[2] and CFI[3]

[1] (Curran, West, & Finch, 1996; Barton & Lord, 1981)
[2] (Tucker & Lewis, 1973)
[3] (Bentler, 1990)

Utrecht University

3. The current study

- New goodness-of-fit test
- TLI and CFI

3.1. Research questions

Utrecht University

3. The current study

- New goodness-of-fit test
- TLI and CFI

3.1. Research questions

1. Required sample size?

Utrecht University

3. The current study

- New goodness-of-fit test
- TLI and CFI

3.1. Research questions

1. Required sample size?
2. Comparing performance

Utrecht University

3. The current study

- New goodness-of-fit test
- TLI and CFI

3.1. Research questions

1. Required sample size?
2. Comparing performance
3. Performance of TLI and CFI?

Utrecht University

How will we research this?

Through a simulation study of course!

Utrecht University

4. Simulation study

4.1. Data generation

- Dichotomous items
- Static item parameters
- Varying four factors

Utrecht University

4.2. Simulation Design

Table: Overview of Simulation Conditions for Each Factor

Factor	Conditions	Description
Test length	5 - 10 - 20	The total number of items that the test will consist of
Sample size	100 - 200 - 500 1000 - 1500	The total number of observations that are available for each item
Model type	2PL - 3PL	The models that we will use as the basis for data generation
Number of groups	2 - 3 - 4	The number of groups that the data gets divided into for the LR Randomisation test calculations

Utrecht University

4.3. Performance metrics

Goodness-of-fit tests

- Power
- Empirical α

Fit indices

- Mean (SE)

Utrecht University

5. Preview of the results

Table 2. Temporary Results for Empirical Alpha

Conditions		Goodness-of-fit test				
I	N	LR2	LR3	LR4	χ^2	P-χ^2
5	100	...	...	...	...	...
	200	...	...	...	...	...
	500	...	...	...	...	...
	1000	...	...	...	...	...
	1500	...	...	...	...	...
10	100	...	...	...	...	...
	200	...	...	...	...	...
	500	...	...	...	...	...
	1000	...	...	...	...	...
	1500	...	...	...	...	...
20	100	...	...	...	...	...
	200	...	...	...	...	...
	500	...	...	...	...	...
	1000	...	...	...	...	...
	1500	...	...	...	...	...

Note. Fitted model = two-parameter logistic model; Data generating model = two-parameter logistic model; I = test length; N = sample size; LR2 = LR Randomisation test with $g = 2$; LR3 = LR Randomisation test with $g = 3$; LR4 = LR Randomisation test with $g = 4$; $\chi^2 = \chi^2$-difference test under the three-parameter logistic model with no constraints; P-χ^2 = Pearson's χ^2 test.

Table 4. Fit indices values for correct model specification

Conditions		TLI	CFI
I	N	M (SE)	M (SE)
5	100	...	...
	200	...	...
	500	...	...
	1000	...	...
	1500	...	...
10	100	...	...
	200	...	...
	500	...	...
	1000	...	...
	1500	...	...
20	100	...	...
	200	...	...
	500	...	...
	1000	...	...
	1500	...	...

Note. Fitted model = two-parameter logistic model; Data generating model = two-parameter logistic model; I = test length; N = sample size; TLI = tucker-lewis index; CFI = comparative fit index; M = mean; SE = standard error.

Utrecht University

Thank you for listening!

Utrecht University

References

Barton, M. A., & Lord, F. M. (1981). An upper asymptote for the three-parameter logistic item-response model. *ETS Research Report Series*, *1981*(1), i–8.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological bulletin*, *107*(2), 238.

Curran, P., West, S., & Finch, J. (1996, March). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*(1), 16–29. doi: 10.1037/1082-989X.1.1.16

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*(1), 1–10.

Utrecht University