The Journal of Open Source Software

thurstonianIRT: Thurstonian IRT Models in R

Paul-Christian Bürkner[1]

1 Aalto University, Department of Computer Science

DOI: 10.21105/joss.01662

Software

- Review
- Repository
- Archive

Submitted: 11 August 2019
Published: 23 October 2019

Summary

In the human sciences, we often aim to measure certain person characteristics that are latent, that is, not directly observable. Examples for these latents characteristics are personality traits such as extraversion or emotional stability as well as performance related traits such as intelligence or creativity. When measuring personality traits, we mostly rely on self-reported measures based on rating scales where people answer how much they agree on an item. This format is easily manipulatable, for example, if participants know which answers are desirable. Thus its application in high stakes situations (e.g., in personnel selection) is problematic, as participants may be motivated to answer dishonestly (Brown & Maydeu-Olivares, 2011).

An an alternative, forced-choice formats have been proposed in which people are required to make comparative judgments between two or more items. In this case, they would not be able to endorse all items at the same time. Analysing data obtained from forced-choice questionnaires requires specialized statistical models. One of these models is the Thurstonian Item Response Theory (IRT) model, which was originally proposed by Brown & Maydeu-Olivares (2011). IRT in general comes with several advantages over classical test theory, for instance, the ability to model varying item difficulties as well as item factor loadings on the participants' traits they are supposed to measure. Moreover, if multiple traits are modeled at the same time, their correlation can be incorporated into an IRT model to improve the overall estimation accuracy (Brown & Maydeu-Olivares, 2011). In addition to those general advantages, forced-choice questionnaires and corresponding IRT models specifically come with the hope of providing more valid inference in situations where participants have motivation to fake. Whether they live up to this hope remains a topic of debate (e.g., see Bürkner, Schulte, & Holling, 2019) but it is in any case necessary to provide software for fitting these statistical models both for research and practical purposes.

The R package *thurstonianIRT* has been developed to fit various IRT models for forced-choice data, in particular the Thurstonian IRT model. In the original formulation, the Thurstonian IRT model assumes responses on dichotomous pairwise comparisons and models the probability of endorsing one versus the other item. This probability depends on parameters related to the items under comparison as well as on parameters related to the participants' latent traits, which are assumed to be measured by the items. For more details see Brown & Maydeu-Olivares (2011) and Bürkner et al. (2019). For model estimation, thurstonianIRT offers multiple backends, most notably the open source packages Stan (Carpenter et al., 2017) and lavaan (Rosseel, 2012). The thurstonianIRT package was originally developed as part of a project that led to the publication of Bürkner et al. (2019) but has since been developed further to fit and postprocess a more broad set of models for analysing forced-choice data. For instance, the formulation of the Thurstonian IRT model may be extended to ordinal or continuous comparative judgements, which are an active area of research facilitated by the thurstonianIRT package.

The source code of the package is available on GitHub (https://github.com/paul-buerkner/thurstonianIRT).

References

Brown, A., & Maydeu-Olivares, A. (2011). Item response modeling of forced-choice questionnaires. *Educational and Psychological Measurement*, *71*(3), 460–502. doi:10.1177/0013164410375112

Bürkner, P.-C., Schulte, N., & Holling, H. (2019). On the statistical and practical limitations of Thurstonian IRT models. *Educational and Psychological Measurement*, *79*(5), 827–854. doi:10.1177/0013164419832063

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., et al. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1), 1–32. doi:10.18637/jss.v076.i01

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. doi:10.18637/jss.v048.i02