thurstonianIRT: Thurstonian IRT Models in R

Paul-Christian Bürkner

11 August 2019

Summary

In the human sciences, we often aim to measure certain person characteristics which 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-report measures based on rating scales where persons answer how much they agree to an item. This format is easily fakeable for obvious reasons if participants know which answers are desireble. Thus its application in high stakes situations (e.g., in personnel selection) is problematic as participants may have motiviation to not answer honestly (Brown and Maydeu-Olivares 2011).

An an alternative, forced-choice formats have been proposed in which persons are required to make comparative judgments between two or more items. As a result, it is not possible 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 and Maydeu-Olivares (2011). Forced-choice questionnaires and corresponding statistical models 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, and 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 and Maydeu-Olivares (2011) and Bürkner, Schulte, and Holling (2019). For the 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, Schulte, and Holling (2019) but has since been developed further to fit and post-process a more broader 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, Anna, and Alberto Maydeu-Olivares. 2011. "Item Response Modeling of Forced-Choice Questionnaires." *Educational and Psychological Measurement* 71: 460–502. https://doi.org/10.1177/0013164410375112.

Bürkner, Paul-Christian, Niklas Schulte, and Heinz Holling. 2019. "On the Statistical and Practical Limitations of Thurstonian Irt Models." *Educational and Psychological Measurement*, 0013164419832063. https://doi.org/10.1177/0013164419832063.

Carpenter, B., A. Gelman, M. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. A. Brubaker, J. Guo, P. Li, and A. Ridell. 2017. "Stan: A Probabilistic Programming Language." *Journal of Statistical Software*. https://doi.org/10.18637/jss.v076.i01.

Rosseel, Yves. 2012. "lavaan: An R Package for Structural Equation Modeling." *Journal of Statistical Software* 48 (2): 1–36. https://doi.org/10.18637/jss.v048.i02.