Psychometrics in R: Rasch Model and beyond

Christophe Lalanne[1,2,*], Martin Duracinsky[2,3], Laurence Vaivre-Douret[1,4], Olivier Chassany[2]

1. INSERM U669, Univ Paris Sud and Univ Paris Descartes, UMR-SO669, Paris, France
2. Hôpital Saint-Louis, Department of Clinical Research, Paris, France
3. Hôpital Bicêtre, Internal Medicine and Infectious Diseases Department, Paris, France
4. Univ Paris 10, Nanterre, France

* Contact author: ch.lalanne@gmail.com

Keywords: psychometrics, item-response models, Rasch, HRQoL

The analysis of a set of items responses collected on a sample of individuals usually relies on classical test theory and item response theory (IRT), whereby test or composite scores are modelled under either a linear or generalized linear model. The well-known Rasch Model shows how it is possible to introduce probability considerations into subjects' responses, and statistical analysis focuses on several form of validity and reliability of test scores. Furthermore, it can be shown that most of IRT models can be expressed as mixed-effects models (e.g. De Boeck and Wilson, 2004), which facilitates exploration of complex effects like Differential Item Functionning (DIF).

Though several compiled packages have been made available for psychometrics in educational assessmement or biomedical investigation, only recently have languages such as SAS and Stata incorporated IRT modelling capabilities (e.g. Rabe-Hesketh and Skrondal, 2008; Hardouin and Mesbah, 2007, but see the Free IRT Project, `http://freeirt.anaqol.org/`, for an overview). The open source R software also provides a well integrated statistical framework for psychometrics, ranging from classical analysis to modern techniques based on IRT and derived methods (but see *Journal of Statistical Software*, volume 20, 2007). For instance, two packages (`eRm` and `ltm`) available on CRAN website allow to estimate IRT models under conditional or marginal likelihood approach.

We show how R core functionnalities may be used in applied biomedical research, in particular for the analysis of Patient Reported Outcomes (PRO). It is now well recognized that PRO and Health related Quality of Life have to be taken into account in the evaluation of therapeutic strategies. Following standard guidelines, a stepwise analytic approach must demonstrate that a given HRQoL questionnaire has all of the desirable characteristics of a valid and reliable measurement instrument. After having analyzed the inter-items correlation matrix and responses distribution, a factorial analysis of polychoric or linear correlation matrix aims at determining a minimal set of underlying latent factors which explains the maximum of scores variance. Scores gathered on other questionnaires allow to study convergent and discriminant validities, together with Multi-Trait scaling analysis. Finally, studying composite scores in relation to biomedical indicators help to highlight sensibility or responsiveness of the questionnaire to patients' relative disease. Likewise, scores on unidimensional domains, like physical or mental states, may vary according to cultural-based factors, whether it be DIF or not, and this can be studied using explanatory IRT models.

In summary, R now offers a flexible and reliable way to carry out Exploratory Factor Analysis, Multi-Trait Scaling, Reliability assessment and IRT modelling. Obviously, this avoids the need to switch between different dedicated software and facilitates a seemly integration of statistical analysis and project management.

References

De Boeck, P. and Wilson, M. (2004). *Explanatory Item Response Models. A Generalized Linear and Nonlinear Approach.* Springer.

Hardouin, J.-B. and Mesbah, M. (2007). The SAS macro-program `\%AnaQol` to estimate the parameters of IRT models. *Communications in Statistics: Simulation and Computation*, *36(2)*, 437–453.

Rabe-Hesketh, S. and Skrondal, A. (2008). *Multilevel and Longitudinal Modeling Using Stata*, 2nd ed. Stat Press