catR: An R Package for Computerized Adaptive Testing

David Magis¹ and Gilles Raîche²

Abstract

Computerized adaptive testing (CAT) is an active current research field in psychometrics and educational measurement. However, there is very little software available to handle such adaptive tasks. The R package *catR* was developed to perform adaptive testing with as much flexibility as possible, in an attempt to provide a developmental and testing platform to the interested user. Several item-selection rules and ability estimators are implemented. The item bank can be provided by the user or randomly generated from parent distributions of item parameters. Three stopping rules are available. The output can be graphically displayed.

Keywords

computer program, computerized adaptive testing, estimation, item response theory

Computerized adaptive testing (CAT) is an active current research field in psychometrics and educational measurement. However, there is very little software available to handle such adaptive tasks, the program FireStar (Choi, 2009) and the SAS macro SIMCAT (Raîche & Blais, 2006) being probably the most suitable CAT software to date. The R package *catR* was developed to perform adaptive testing with as much flexibility as possible, in an attempt to provide a developmental and testing platform to the interested user.

The *catR* package permits the user to generate CAT response patterns and holds many design options. First, the item bank can be either provided by the user or randomly generated by specifying the parent distributions of the item parameters. The four-parameter logistic model (Barton & Lord, 1981) is implemented as the baseline item response theory model. The first test item(s) can be selected either by the user or from some prior ability levels, or can be randomly input into the item bank. Four ability estimators can be considered: maximum likelihood, Bayes's modal, expected a posteriori, and weighted likelihood (see, for example, Baker & Kim, 2004; van der Linden & Glas, 2000). Normal prior, uniform prior, and Jeffreys prior are available as prior ability distributions. Seven item-selection criteria are currently implemented: maximum information criterion, maximum likelihood or weighted posterior information, maximum expected information, minimum expected posterior variance, Urry's criterion, and random selection (Choi & Swartz, 2009; van der Linden & Pashley, 2000). The three available stopping rules are based

¹University of Liège and K.U. Leuven, Belgium

²Université du Québec à Montréal, Canada

Corresponding Author:

David Magis, University of Liège, Grande Traverse 12, B-4000 Liège, Belgium

Email: david.magis@ulg.ac.be

either on the length of the test, the precision of the current ability estimate, or whether the ability level can be classified into one group of abilities with some prespecified confidence level. The output of *catR* is organized such that it is easy to read and to interpret, and it contains all relevant information, for example, input options, provisional results of the adaptive process, and final ability estimation and inference. Graphical representations of the estimated ability patterns are also available.

The R package *catR* is available for free from the website <http://www.r-project.org> and works under Windows, Linux, and MacOS platforms. A reference manual (in PDF format) is also available from that website. The current version of *catR* is 1.4. Version 2.12.0 (or later) of the R software should be installed for optimal functioning of *catR*.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The development of the *catR* package was partly funded by a postdoctoral grant “chargé de recherches” of the National Funds for Scientific Research (Belgium), the Research Funds of the K.U. Leuven (Belgium), and a grant from the Social Sciences Research Council of Canada.

References

- Baker, F. B., & Kim, S.-H. (2004). *Item response theory: Parameter estimation techniques* (2nd ed.). New York, NY: Marcel Dekker.
- Barton, M. A., & Lord, F. M. (1981). *An upper asymptote for the three-parameter logistic model* (Technical Report RR-81-20). Princeton, NJ: Educational Testing Service.
- Choi, S. W. (2009). Firestar: Computerized adaptive testing simulation program for polytomous item response theory models. *Applied Psychological Measurement*, 33, 644-645.
- Choi, S. W., & Swartz, R. J. (2009). Comparison of CAT item selection criteria for polytomous items. *Applied Psychological Measurement*, 32, 419-440.
- Raiche, G., & Blais, J.-G. (2006). SIMCAT 1.0-A SAS computer program for simulating computer adaptive testing. *Applied Psychological Measurement*, 30, 60-61.
- van der Linden, W. J., & Glas, C. A. W. (2000). *Computerized adaptive testing: Theory and practice*. Boston, MA: Kluwer.
- van der Linden, W. J., & Pashley, P. J. (2000). Item selection and ability estimation in adaptive testing. In W. J. van der Linden & C. A. W. Glas (Eds.), *Computerized adaptive testing: Theory and practice* (pp. 1-25). Boston, MA: Kluwer.