



UMEÅ UNIVERSITET

Extensions and Applications of Item Response Theory

Joakim Wallmark

Akademisk avhandling

som med vederbörligt tillstånd av Rektor vid Umeå universitet för avläggande av filosofie doktorsexamen framläggs till offentligt försvar i Hörsal HUM.D.220 - Hjortronlandet, Humanisthuset, fredagen den 7 februari, kl. 9:00.

Avhandlingen kommer att försvaras på engelska.

Fakultetsopponent: Dr., Patricia Martinkova,
Faculty of Education, Charles University, Prague, Czech Republic
and Department of Statistical Modeling, Institute of Computer
Science of the Czech Academy of Sciences, Prague, Czech Republic

Department of Statistics, USBE

Organization

Umeå University
Department of Statistics
Umeå School of Business,
Economics and Statistics

Document type

Doctoral thesis

Date of publication

17 January 2025

Author

Joakim Wallmark

Title

Extensions and Applications of Item Response Theory

Abstract

This doctoral thesis focuses on Item Response Theory (IRT), a statistical method widely used in fields such as education and psychology to analyze response patterns on tests and surveys. In practice, IRT models are estimated using collected test data, which allows researchers to assess both how effectively each item measures the underlying trait—such as subject knowledge or personality characteristics—that the test aims to evaluate, and to estimate each individual's level of that trait. Unlike traditional methods that simply sum predetermined item scores, IRT accounts for the difficulty of each item and its ability to measure the intended trait.

The thesis consists of four research articles, each addressing different aspects of IRT and its applications. The first article is focused on test equating, ensuring that scores from different versions of a test are comparable. Equating methods with and without IRT are compared using simulations to explore the advantages and disadvantages of incorporating IRT into the kernel equating framework. The second and third articles introduce and compare different types of IRT models. Through simulations and real test data examples, it is demonstrated that more flexible models can better capture the true relationships between test responses and the underlying traits being measured.

Finally, the IRTorch Python package is presented in the fourth study. IRTorch supports various IRT models and estimation methods and can be used to analyze data from different types of tests and surveys. In summary, the thesis demonstrates how IRT-based equating methods can serve as an alternative to traditional equating methods, how more flexible IRT models can improve the precision of test results, and how user-friendly software can make advanced statistical models accessible to a wider audience.

Keywords

Item response theory, Test equating, Statistical software, Educational assessment, Latent variable modelling

Language

English

ISBN

print: 978-91-8070-571-4
pdf: 978-91-8070-572-1

ISSN

1100-8989

Number of pages

25 + 4 papers