Teaching Psychometrics and Analysing Educational Tests with ShinyItemAnalysis

Patricia Martinkova et al.
useR!2017, Brussels, July 7

Department of Statistical Modelling, Institute of Computer Science
Czech Academy of Sciences
Joint work...

Patricia Martinkova

Adela Drabinova

Jakub Houdek

Ondrej Leder

Lubomir Stepanek

1Department of Statistical Modelling, Institute of Computer Science, Czech Academy of Sciences
2Department of Probability and Mathematical Statistics, Charles University, Prague
3Faculty of Informatics and Statistics, University of Economics, Prague
4First Faculty of Medicine, Charles University, Prague

Acknowledgement: Czech Science Foundation grant GJ15-15856Y
1. Introduction

2. ShinyItemAnalysis for TEACHING

3. ShinyItemAnalysis for RESEARCH

4. ShinyItemAnalysis for PRODUCTION

5. Conclusion
Introduction
Motivation

- Admission tests development
  - to select the best applicants
  - need for valid, reliable and fair tests, well functioning items
- Development of tests for classroom use
  - to keep students interested in the subject
  - to avoid discrimination of minorities
- To teach psychometric models and concepts
- To promote our own psychometric research

Need for user-friendly and freely available tool
ShinyItemAnalysis

Interactive (and step by step) analysis of educational tests and their items

Available as:

- R package
  - Version 1.2.0 now on CRAN
  - Newest version on GitHub
- Online shiny application
  - ICS server in Prague, CZ:
    - https://shiny.cs.cas.cz/ShinyItemAnalysis/
  - shinyapps.io:
    - https://patriciamar.shinyapps.io/ShinyItemAnalysis/
ShinyItemAnalysis Application

Description

ShinyItemAnalysis provides analysis of educational tests (such as admission tests) and their items including:

- Exploration of total and standard scores on Summary page.
- Correlation structure and predictive validity analysis on Validity page.
- Item and distractor analysis on Item analysis page.
- Item analysis by logistic models on Regression page.
- Item analysis by item response theory models on IRT models page.
- Differential item functioning (DIF) and differential distractor functioning (DDF) methods on DIF/Fairness page.

This application is based on the R statistical software R and its shiny package.

For all graphical outputs a download button is provided. Moreover, on Reports page HTML or PDF report can be created. Additionally, all application outputs are complemented by selected R code hence the similar analysis can be run and modified in R.

Data

For demonstration purposes, by default, 20-item dataset *person* from R *distr* package is used. Other three datasets are available: *gswiz* and *medical 20* from *distr* package and *medical 100* from *shinyItemAnalysis* package. You can change the dataset (and try your own one) on page Data.

Availability

Application can be downloaded as R package from CRAN. It is also available online at Czech Academy of Sciences in case of busy server you can try other mirrors: Charles University or shinyapps.io.

Version

Current version of ShinyItemAnalysis available on CRAN is 1.2.0. Version available online is 1.2.0. The newest development version available on GitHub is 1.2.0.

See also older versions: 0.1.0, 0.2.0, 1.0.0, 1.1.0.

Authors and contributors

Patricia Martinkova
Adela Dražinová
Ondřej Leder
Jakub Houdek
Libor Stepanek
ShinyItemAnalysis for TEACHING
ShinyItemAnalysis for Teaching

Who do we teach:

- students of educational measurement
- faculties, university stakeholders

Some helpful features:

- Interactive plots
- Example datasets, allows to upload own data
- Shows model equations, provides interpretation of results
- Allows to download plots, generate extensive reports
- Provides sample R code
Datasets

- Four toy datasets are available
- Allows to upload one’s own dataset
Summary of Total Scores

- Summary statistics
- Interactive histogram
Traditional Item Analysis

- Difficulty, discrimination
- Cronbach’s alpha w/o item, index RIT, RIR, etc.
Distractor Analysis

- Displays option selection percentage by total score group
- Number of groups can be changed
Logistic Regression

- Displays probability of correct answer by total score
- Parameterization can be changed (Z scores, IRT parameterization)
Nonlinear Regression

- Allows for guessing (and inattention)
IRT Models

- Conceptualized as nonlinear mixed effect models
- More precise ability estimation
Selected R Code

- Sample R code may be run in separate R session
IRT Models - construct your own item

- Plots Item Characteristic and Information Curves (ICC and IIC) based on selected parameters
ShinyItemAnalysis for RESEARCH
ShinyItemAnalysis for Research

To widespread novel methodology:

- write scientific paper
- ... and provide R code
- ... and provide dataset
- ... and write accompanying R package
- ... and prepare shiny application
DIF: Students from two groups and *with the same underlying latent ability* have different probability of answering the item correctly.
Differential Distractor Functioning (DDF)

**DDF**: Students from two groups and *with the same underlying latent ability* have different probability of selecting given options.

To learn more, go to
Adela Drabinova: *difNLR*: Detection of potential gender/minority bias with extensions of logistic regression (Thursday, July 6)
Why DIF Analysis Should Be Analyzed Routinely?

- Simulated GMAT data: total scores may have exactly the same distribution, yet there may be DIF present in some items!

Martinkova et al. (2017): Checking Equity: Why DIF Analysis should be a Routine Part of Developing Conceptual Assessments. CBE-Life Sciences Education, 16(2), rm2.  

Online
ShinyItemAnalysis for PRODUCTION
Report Generation

- Upload your own data
- Generate PDF/HTML report
PDF Reports

Item Analysis Report

Summary of total scores

Traditional item analysis report

Summary of total scores

Correlation structure

Item analysis report

Summary of total scores

Traditional item analysis report

Summary of total scores

Correlation structure
Report Generation

- shiny provides a user interface.
- rmarkdown for creating templates for PDF/HTML report generation.
- knitr for compiling R markdown syntax into HTML/PDF
- TEX for creating PDF reports (latest distribution of TEX is needed).

To learn more see UseR!2017 poster by Jakub Houdek et al.
Conclusion
ShinyItemAnalysis is a shiny application for interactive and step-by-step analysis of educational tests. It is useful for:

- **TEACHING** of psychometrics and educational measurement
  - offers example datasets, upload of new datasets
  - visualization, interpretation of results
  - sample R Code

- **PRODUCTION**
  - generates extensive reports for supplied data

ShinyItemAnalysis also promotes our RESEARCH in DIF/DDF detection

https://shiny.cs.cas.cz/ShinyItemAnalysis/

martinkova@cs.cas.cz
Thank you for your attention!

www.cs.cas.cz/martinkova

martinkova@cs.cas.cz

This research was supported by Czech Science Foundation grant GJ15-15856Y
• Martinkova, Drabinova, Leder & Houdek (2017): ShinyItemAnalysis: Test and Item Analysis with Shiny. 
https://shiny.cs.cas.cz/ShinyItemAnalysis/ 
https://CRAN.R-project.org/package=ShinyItemAnalysis

https://CRAN.R-project.org/package=difNLR

• Drabinova & Martinkova (under review): Detection of DIF Based with Non-Linear Regression: Non-IRT Approach Accounting for Guessing.

• Martinkova, Drabinova, Liaw, Sanders, McFarland & Price (2017): Checking Equity: Why DIF Analysis should be a Routine Part of Developing Conceptual Assessments. CBE-Life Sciences Education, 16(2), rm2. www.lifescied.org/content/16/2/rm2

• Martinkova, Drabinova & Houdek (2017): ShinyItemAnalysis: Analysis of admission and other educational and psychological tests. Testforum, Accepted.