The 6th Baltic-Nordic Conference on Survey Statistics
Helsinki, August 2023

TECHNOLOGIES FOR CREATING AND ANALYZING TESTS IN ADVANCED MATHEMATICS

N. Kruglova[1], O. Dykhovychnyi[2] and M. Poprozhuk[3]
[1] National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine
e-mail: natahak@ukr.net

[2] National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine
e-mail: a.dyx@ukr.net

[3] National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine
e-mail: poprozhuk19@gmail.com

Abstract

Over the past three years, distance education has become one of the primary formats of education for Ukrainian students, initially due to the COVID-19 pandemic and currently because of ongoing Russian missile strikes on civilian targets in Ukraine. It is highly likely that distance education will remain the only viable option for a considerable period of time. Therefore, creating high-quality content to test students' knowledge under these conditions has become an exceedingly critical task.

A team of instructors at National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute" are currently working on mathematics tests design in the two main areas:

A. Test item design. To facilitate this process, we employ Wolfram Mathematica as a practical tool that allows us to perform problems' solutions verification, designing numerical problems with integer answers, modeling high-quality distractors for multiple-choice questions, creating images for problem illustration, and identifying problems that are likely to be solved by students using online resources.

B. Statistical analysis of test quality. The quality of the designed tests is analyzed by testing them on a control group of students, with the aim of improving the tests by removing or reformulating items that are either too easy or too difficult.

Classical Test Theory (CTT) and Item Response Theory (IRT), Crocker (2006) are the primary conventional tools used for this purpose. However, we apply a more advanced framework called Multidimensional Item Response Theory (MIRT), Reckase (2009). This framework permits a more nuanced differentiation of the individual characteristics and abilities of examinees. A key challenge when applying MIRT is to select an appropriate model for analysis, particularly when it comes to choosing the model's dimensionality. The model selection methodology we employ is as follows:

1) Exploratory Factor Analysis (EFA). We conduct EFA to initially select the model's dimensionality based on Auerswald (2019) approach. At this stage, we use methods like Parallel Analysis, Empirical Kaiser Criterion, etc. to determine the number of competencies being tested as well as the number of latent numerical parameters that characterize a student.

2) Estimation of model parameters. For both compensatory and non-compensatory MIRT models, we utilize Confirmatory Factor Analysis (CFA), Chalmers (2012) approach for model parameters estimation employing EM or NH-RM algorithms.

3) Assessing model adequacy. Since different EFA algorithms can yield different dimensionalities, the most adequate model is selected based on the following criteria: M2, RMSEA, TLI, CFI.

We chose the R programming language for statistical analysis of test quality.

Keywords: IRT, MIRT, EFA, CFA

References

Auerswald, A., and Moshagen, M. (2019) How to Determine the Number of Factors to Retain in Exploratory Factor Analysis: A Comparison of Extraction Methods Under Realistic Conditions. *Psychological Methods*. Advance online publication. 1-24.

Chalmers, R. (2012) Multidimensional Item Response Theory Package for the R Environment. *Journal of Statistical Software*, Volume 48, Issue 6, 1-25.

Crocker, L., and Algina, J. (2006) *Introduction to Classical and Modern Test Theor*y, Belmont, CA:Wadsworth.

Reckase, M. D. (2009) *Multidimensional item response theory*, Springer, New York.