In the field of psychometrics, there is increasing interest in the evaluation of fairness in standardized tests. Differential item functioning (DIF) is a phenomenon observed when two groups of respondents have different response patterns. My dissertation work was motivated by an analysis of DIF using a two-tier full-information item factor analysis model. This particular factor analysis model provides deep analysis of the latent variable of each respondent. Since existing methodologies do not provide an analysis of this nature, I developed a new model to divide the latent variable into primary dimensions and specific dimensions.

In my dissertation, I sought to understand the data and direction of research under the supervision of three professors in the UCLA Statistics Department: Peter Bentler, Jan de Leeuw, and Rick Paik Schoenberg. Following the advice of Education Professor Li Cai, I developed a method of analysis using intricate details derived from fundamental theorems of psychometrics, such as Item Response Theory (IRT). Statistics-based methods for estimating parameters and fitting models can be applied in the field of psychometrics.

**Differential Item Functioning**

DIF is an aspect of bias capable of providing statistical evidence than an item is biased; it is slightly different in concept from impact and item bias. From the IRT point of view, DIF is present when different item characteristic curves (ICCs) can be identified from the different subgroups [5]. In contrast, impact measures the difference in performance on an item between two groups, which comes from the difference in average ability of the groups [4]. Therefore, impact measures how the difference in performance is affected by difference in ability, whereas DIF measures how the difference in performance emerges despite the lack of difference in ability. Item bias is also different from DIF. If an item shows DIF, experts can implement an evaluation to see, in the social or content aspect, whether the item favors one group over the other [1]. DIF is a condition that is necessarily present when an item is biased, but the converse is not true.
Two-Tier Full-Information Item Factor Analysis

The two-tier full-information item factor analysis model is a confirmatory item factor model having features that are useful in psychometric research [3]. One of the features is flexibility while handling residual dependence of item responses without using copula functions [2]. Another feature is maximum likelihood estimation, with proven accuracy and efficiency from demonstrations using data. Dimension reduction, yet another feature, can reduce the dimensionality of the latent variable, which gives the model a computational advantage over other models.

Future work

My immediate goal is to publish the results of my dissertation. For further research, I would like to explore some new concepts for detecting DIF among items. One of those concepts is the determination of a set of items with zero DIF, to facilitate obtaining an accurate scale for DIF magnitude. This new method should be more reliable than the method in which the average DIF magnitude is assumed to be zero for all groups. One candidate is a purification method, which is capable of locating a set of DIF free items. I have an interest in participating in projects developing unbiased exams and providing better education.

References


