Using a Computerized Adaptive Test for optimizing questionnaires - Application of Item Response Theory

Augusto Sousa da Silva Filho[1], Rodrigo Tomás Nogueira Cardoso[2] & Marcos Antônio da Cunha Santos [3]

[1] Centro Federal de Educação Tecnológica de Minas Gerais - CEFET-MG
[2] Centro Federal de Educação Tecnológica de Minas Gerais - CEFET-MG
[3] Universidade Federal de Minas Gerais - UFMG

Abstract

When a researcher wants to determine the degree or intensity of the symptoms of a patient generally applies a pen and paper questionnaire with a series of pre-defined questions the same happens when a teacher seeks to determine the ability of a student in a particular area of knowledge. In almost all cases, such questionnaires are too strenuous for the respondent and, because of this risk does not represent a significant patient status or level of student's knowledge. An alternative to this scenario is the construction of a Computerized Adaptive Testing - CAT. The CAT seeks to optimize the test for each respondent, seeking to estimate the predominant characteristic of an individual responded to each item. In this poster, CAT developed not use the Classical Theory of Measures, which measures only the proportion of hits and misses on a test. Instead, the Item Response Theory, which has the advantage of making comparisons between latent traits of individuals from different populations when subjected to tests or questionnaires that have some common items and also allows the comparison of individuals will be used same population subjected to totally different tests, this is possible because the TRI has as central items and questionnaire or not the test as a whole. The IRT presents mathematical models that have been widely discussed in the literature, such as the cumulative models, due to its ease of implementation. However, other models do not have the same ease of implementation as models of developments. This poster will propose the use of generalized models of graduate unfolding - GGUM in Computerized Adaptive Testing of construction, since they can be used both as dichotomous responses to polytomous ordinal responses.

Palavras-chave: Computerized Adaptive, Item Response Theory tests.

1. Introduction

In Brazil, some reviews have adapted their evidence to the computerized system where the same are conducted via computer. However, this type of application it is not a Computerized Adaptive Test, since it is only the reproduction of proof "paper and pencil" on the computer. In CAT items are selected according to the examinee who perform the test and its latent trait is measured iteratively. To implement a CAT construction of an item bank with psychometric and pedagogical qualities is indispensable to use models of Item Response Theory is necessary. The TRI will relate the latent trait of the examinee with the response to the item, so that the higher the examinee's latent trait, the greater the probability of getting the item. With the use of IRT calibrated if items. The estimates obtained in this step will feed the CAT, a procedure with which to estimate the abilities of examinees during testing, and it becomes possible to compare the abilities of different test takers at the end of the process, optimizing the number of items (questions) that will be made to the student.

2. Theoretical Framework

The majority of studies using computerized adaptive tests based on item response theory, the authors use the cumulative models 1, 2 and 3 parameters, because they are considered simpler to implement. Araújo, Andrade and Bortolotti, claim that in contrast, the models of split (not

cumulative), including the Model Parella, generalized graded unfolding model and model Hyperbolic Cosine, have not yet been sufficiently exploited in this way, "did not make much progress and the reason is mainly due to the understanding of its mechanism of response and lack of computer programs to estimate the parameters of this model type".

Models of developments are distinguished by being models of cumulative proximity models, where higher categories of responses are more likely (indicative of stronger levels of agreement) when the distance between the parameters of the individual and the position of the item on the scale decreases. That is, the probability of an individual to respond to an item is based on the distance between the parameters of the individual and the position of the item on the scale, and not as a function of the parameter of the individual, as in the cumulative models. The model of generalized graded unfolding is given below:

$$P\left(Z_i = z|\theta_j\right) = \frac{\exp\left[\alpha_i\left(z\left(\theta_j - \delta_i\right) - \sum_{k=0}^{z}\tau_{ik}\right)\right] + \exp\left[\alpha_i\left(\left(M-z\right)\left(\theta_j - \delta_i\right) - \sum_{K=0}^{z}\tau_{ik}\right)\right]}{\sum_{v=0}^{H}\left[\exp\left(\alpha_i\left[v\left(\theta_j - \delta_i\right) - \sum_{k=0}^{v}\tau_{ik}\right]\right) + \exp\left(\alpha_i\left[\left(M-v\right)\left(\theta_j - \delta_i\right) - \sum_{k=0}^{v}\tau_{ik}\right]\right)\right]} \quad (1)$$

Where:

i) Z_i is an observable response to an item of attitude i;

ii) $z = 0, 1, 2, 3, \ldots H; z = 0$ corresponds to the strongest level of disagreement and $z = H$ corresponds to the strongest level of agreement;

iii) H is the number of categories of observable responses with 1 less $M = 2h + 1$;

iv) θ_j parameter is the location of the individual j the scale of the latent trait;

v) δ_i is the parameter of the item's location i the scale of the latent trait;

vi) α_i is the discrimination parameter of item i;

vii) t_{ik} parameter is the position of the threshold category of subjective response k the scale of the latent trait on the position of the item i; corresponds to the value of the distance between θ_j and δ_i determining the point at which the k-th category subjective response is replaced response probability of $(k-1)$*k-th* category of subjective response to the individual j in item i e τ_{i0} is, by definition equal to zero;

viii) M is the number of categories of subjective responses less 1.

According Bortolotti, 2013 models unfolding item response theory are models which are estimated item parameters and the parameters of the individuals in this way, the parameters of individuals and items are placed on the same scale. Therefore, the parameters of the individual are allocated according to their opinion and the items are located on the same scale according to their content. Therefore, these models are also known as proximity models, because the response categories of the highest items are more likely when the distance between the position of the individual parameter and the item on the scale decreases. In an unfolding model, the probability of an individual to respond to an item is based on the distance between the position of the individual parameter and the position of the item on the scale parameter. Soon a model of development for a measure of attitude the individual agrees with a response category of an item in that the sentiment conveyed by the response of the item fits your own review.

3. Development

The overall objective of this poster is to present a systematic approach for the development and implementation of Computerized Adaptive Testing based on Item Response Theory Models

via Generalized Splits Alumni. And show how a CAT can achieve a great result with the least possible number of questions asked to a student.

Specific goals expected for this poster, we can highlight:

- Present the steps necessary for the implementation of a CAT, according to their specific characteristics;
- Present and comment on the algorithms used during the preparation of the poster;
- Implement in Software R the various algorithms of this poster;
- Create the database items and the parameter estimates;
- Make adjustments to the CBAT-2 algorithm for the proposed issue;

4. Results and comments

The union Computerized Adaptive Testing and Item Response Theory is considered a major breakthrough in the field of statistics and computer science. A Computerized Adaptive Test held under the support of Item Response Theory will bring together the advantages of these two tools, producing a pattern with a far superior to conventional tests and providing quality professional education the ideal tool for measuring the degree of student learning with a significantly smaller number of items which will facilitate your decision making and new approaches in order to better absorption of knowledge to the student.

5. Conclusion

This is a work in progress and its main objective is the construction of Computerized Adaptive Testing using Generalized models of prescription developments. This poster is on the dissertation in progress carried out in the Federal Center of Technological Education of Minas Gerais. With this project it will be possible to create a package in R with the goal of uniting the algorithms used in the construction of models of generalized graded unfolding. Studies of this nature has as main objective the optimization of the number of items used in large-scale testing and its use can be considered a trend for years to come.

Bibliography

BORTOLOTTI, S. Aplicação de um modelo de desdobramento graduado generalizado da teoria de resposta ao item - TRI. Dissertação de Mestrado - UFSC, Florianopólis, 2003.

FETZER, M. et al. Computer adaptative testing (cat) in an employment context. Roswell, USA, 2008.