

TESTES ADAPTATIVOS COMPUTADORIZADOS APLICADOS A PROVA DE CIÊNCIAS DA NATUREZA E SUAS TECNOLOGIAS

Erica Castilho Rodrigues (1); Edio da Costa Junior (2);

(Universidade Federal de Ouro Preto, ericarodrigues@ufop.edu.br)

(Instituto Federal de Minas Gerais, edio.junior@ifmg.edu.br)

Introdução

Os testes adaptativos computadorizados têm se mostrado muito eficientes para o cálculo das habilidades dos alunos nos mais diferentes tipos de testes. Eles são uma alternativa aos testes sequenciais ou fixos. Os itens são ministrados aos respondentes em uma sequência otimizada, tal que em cada passo o item selecionado é aquele que possui mais informação para estimar a habilidade do aluno naquele momento. Essa escolha é baseada nos itens respondidos anteriormente, as respostas dadas pelo aluno a essas questões e a estimativa atual de sua habilidade. Esse tipo de teste tem muitas vantagens em relação aos testes tradicionais, dentre elas uma redução no número de itens ministrados aos alunos, diminui o risco de fraude, fornece o resultado do teste assim no momento em que o mesmo é finalizado.

Para o caso do ENEM, em particular, existe uma grande queixa por parte dos alunos em relação ao grande número de itens a serem respondidos. Um dos motivos para a necessidade de aplicação tantos itens se deve ao fato da prova ser única para todos os candidatos e, portanto, deve ser capaz de medir habilidades em toda a escala de proficiência. Nos testes adaptativos a seleção dos itens é feita de maneira dinâmica e se adequa a realidade e conhecimentos de cada aluno. Aos invés de ministrar a todos os alunos o mesmo teste, a seleção de itens se adapta ao nível de habilidade de cada um deles. Isso viabiliza a aplicação de testes mais curtos e que estimam as habilidades com uma precisão muito maior. Além disso, os custos com aplicação, armazenamento e correção dos testes são reduzidos, sendo que o resultado apresentado ao aluno assim que o teste é finalizado. A aplicação deste tipo de teste ainda é muito incipiente no Brasil, o que evidencia a necessidade de estudos detalhados sobre quais seriam seus possíveis impactos, benefícios e dificuldades. Um dos aspectos cruciais para aplicação desse tipo de teste se refere à segurança dos itens. Os candidatos podem memorizá-los e dividir com seus colegas. Portanto, métodos para controlar a taxa de exposição dos itens tornam-se necessários.

O objetivo principal deste trabalho é verificar, por meio de estudos de simulação, como os itens da prova Ciências da Natureza e suas Tecnologias, que faz parte do ENEM (Exame Nacional do Ensino Médio), se comportam ao utilizar uma abordagem adaptativa na execução do teste. Espera-se que ao se utilizar a abordagem adaptativa seja possível estimar as habilidades dos alunos com número menor de itens.

Metodologia

Todo o estudo de simulação foi realizado no Software R, com o auxílio do pacote catR (Davis and Raiche (2012)). Ele foi desenvolvido para gerar padrões de respostas para testes adaptativos. Esse pacote disponibiliza diferentes métodos para seleção dos itens, para estimação das habilidades e para o interrupção do teste. Permite ainda gerar diferentes tipos de padrões de resposta para uma dada habilidade. Além disso, ele possui uma ferramenta importante que consiste em controlar a taxa de exposição dos itens do banco.

O teste adaptativo precisa de um modelo de Teoria de Resposta ao Item subjacente. O modelo implementado no pacote catR é o mais genérico de todos, que é modelo de 4 parâmetros proposto por Barton and Lord (1981). Os parâmetros desse modelo são os seguintes: discriminação, dificuldade, acerto ao acaso e o parâmetro de desatenção. Entretanto o modelo utilizado pelo INEP não considera

este último parâmetro e, portanto, possui apenas 3. Esse será, então, o modelo considerado neste trabalho. Para a execução da simulação é necessário ter um conjunto de itens calibrados, porém o INEP (Instituto Nacional de Pesquisas Educacionais) não divulga os valores estimados para os parâmetros dos itens. Dessa maneira, os itens foram recalibrados utilizando o mesmo modelo de Teoria de Resposta ao Item adotado pelo INEP. Essa calibração dos itens foi feita utilizando o pacote Irm proposto por Rizopoulos (2006). O método de estimação implementado no pacote é a Máxima Verossimilhança Marginal. Como o tamanho da base de dados era muito grande, mais de 7 milhões de alunos participantes, foi selecionada uma amostra aleatória dentre eles e foi ajustado o modelo TRI de 3 parâmetros. A amostra selecionada foi de 200000 alunos e as habilidades foram reestimadas e seus valores foram comparados com os valores estimados pelo INEP. A correlação entre os valores estimados após a recalibração dos itens apresentou uma correlação de 0,92 com as habilidades originais da base de dados.

É importante lembrar que para que o teste funcione maneira apropriada o banco de itens deve ser grande, mas isso muitas vezes é inviável por questões de custo ou logística. Pouco se sabe sobre qual o tamanho mínimo necessário para essa base de dados. Weiss e Sain (2015) mostraram que se o banco de itens tem a informação suficiente para estimar as habilidades, o tamanho do banco tem pouca influência sobre a precisão dessa estimativa.

A fim de verificar se essa redução no número de itens é possível no caso do ENEM, foi feita uma simulação para os itens da prova de Ciências da Natureza e suas Tecnologias, aplicada no ano de 2016. Como todas as demais provas, ela é composta por 45 itens pré-calibrados. Os parâmetros desses itens são estimados a partir de um modelo de Teoria de Resposta ao Item de três parâmetros, dificuldade, discriminação e acerto ao acaso. Para mais detalhes sobre esse modelo, consultar Andrade et al. (2000).

Resultados

O estudo de simulação mostrou que é possível, como já esperado, reduzir bastante o número de itens utilizando-se uma abordagem adaptativa. Esse resultado varia de acordo com o critério utilizado para a escolha do próximo item a cada passo do algoritmo. Para o método MFI, que maximiza a função de informação do item, de 20 a 30 questões já se mostraram suficientes para estimar a proficiência do aluno com uma boa precisão. O segundo método utilizado é chamado bOpt e consiste em escolher o item cujo parâmetro de dificuldade mais se aproxime a proficiência atual do candidato. Com essa escolha o número de itens necessários foi ainda menor, 15 a 25. O último método testado é denominado MLW e, assim como o MFI, ele seleciona o item que maximiza a função de informação, porém essa função é ponderada pela verossimilhança. Esse método não apresentou bons resultados e não conseguiu recuperar bem a proficiência dos alunos.

Pretende-se ainda fazer um estudo de simulação mais extenso considerando diferentes faixas de habilidades dos alunos e diferentes critérios de escolha dos itens a fim de verificar quais as melhores configurações

Discussão

Esses são apenas os resultados preliminares do trabalho em desenvolvimento, porém eles já apontam que utilizando a abordagem adaptativa é possível reduzir muito o tamanho do teste.

Uma das desvantagens dos testes adaptativos é que o aluno não pode pular ou revisar itens anteriores, pois isso alteraria e prejudicaria toda a execução da prova. Além disso, eles demandam um investimento financeiro muito grande, tanto no desenvolvimento do banco de itens quanto na aplicação da prova. Entretanto esses custos acabam sendo superados pelas suas vantagens ao longo dos anos.

Referências

DE ANDRADE, Dalton Francisco; TAVARES, Heliton Ribeiro; DA CUNHA VALLE, Raquel. Teoria da Resposta ao Item: conceitos e aplicações. **ABE, Sao Paulo**, 2000.

BARTON, Mark A.; LORD, Frederic M. An upper asymptote for the three- parameter logistic item- response model. **ETS Research Report Series**, v. 1981, n. 1, p. i-8, 1981.

MAGIS, David et al. Random generation of response patterns under computerized adaptive testing with the R package catR. **Journal of Statistical Software**, v. 48, n. 8, p. 1-31, 2012.

SAHIN, Alper; WEISS, David J. Effects of Calibration Sample Size and Item Bank Size on Ability Estimation in Computerized Adaptive Testing. **Educational Sciences: Theory and Practice**, v. 15, n. 6, p. 1585-1595, 2015.

RIZOPOULOS, Dimitris. ltm: An R package for latent variable modeling and item response theory analyses. **Journal of statistical software**, v. 17, n. 5, p. 1-25, 2006.