

Е. Ю. Карданова, В. Б. Карпинский (Великий Новгород, НовГУ). **Обнаружение искажений при тестировании с использованием математической модели Г. Раша.**

Главная цель педагогического тестирования — измерение уровня подготовленности учащихся в соответствующей области знаний. Под искажением в процессе измерения будем понимать любое отклонение от достижения этой цели. Наиболее важными искажениями при массовых тестированиях являются искажения в результате нарушения процедуры тестирования (например, списывание или подлог).

В работе, представленной данным сообщением, измерение осуществляется в рамках вероятностных моделей Г. Раша [1, 2]. Применение моделей Раша позволяет получить инвариантные относительно друг друга оценки уровней подготовленности испытуемых и трудности заданий теста, расположенные на единой метрической шкале и сопровождаемые характеристиками точности оценивания.

Для анализа экспериментальных данных и принятия решения о присутствии искажений в измерениях авторы используют различные статистические критерии, как известные из литературы [1, 3], так и специально разработанные. Критерии имеют различную мощность и чувствительны к разным по природе причинам искажений. Исследуются особенности распределений анализируемых критериев и выбора их критических значений, а также эффективность этих критериев. Исследование эффективности статистических критериев проводится путем вычислительного эксперимента на специально сконструированных модельных данных. При этом основная часть виртуальных испытуемых моделируется в соответствии с моделью Раша и добавляются данные с моделированием списывания/подлога. Для моделирования матриц ответов разработано специальное программное обеспечение.

По результатам исследования эффективности различных статистических критериев разработан сводный критерий, эффективность которого выше, чем у всех исследованных базовых критериев. Также показана возможность повышения эффективности критериев путем подбора критического значения, оптимизирующего потери от возможных ошибок первого и второго рода. При этом показано, что оптимальное критическое значение слабо зависит от количества случаев искажений в выборке.

Таким образом, разработаны алгоритмы исследования экспериментальных данных с целью выявления случаев нарушения процедуры тестирования; для их реализации разработано программное обеспечение, что позволяет исследовать большие объемы реальных данных тестирования на предмет наличия искажений в измерениях испытуемых и принятия решения об их достоверности.

СПИСОК ЛИТЕРАТУРЫ

1. *Wright B. D., Masters G. N.* Rating Scale Analysis: Rasch Measurement. Chicago: Mesa Press, 1982.
2. *Карданова Е. Ю., Нейман Ю. М.* Основные модели современной теории тестирования. — Вопросы тестирования в образовании, 2003, № 7, с. 26.
3. *Smith R. M.* Fit analysis in latent trait measurement models. — J. Appl. Meas., 2000, v. 1, № 2, p. 199–218.