

ИСПОЛЬЗОВАНИЕ ОДНОМЕРНЫХ И МНОГОМЕРНЫХ АДАПТИВНЫХ ТЕСТОВ В ДИСТАНЦИОННОМ ОБРАЗОВАНИИ

***Аннотация.** Рассмотрены основные принципы компьютерного адаптивного тестирования. Описаны различия между одномерными и многомерными адаптивными тестами. Для реализации адаптивного теста предлагается использовать среду статистической обработки данных R.*

Сегодня вопросам адаптивного тестирования уделяется много внимания, что обусловлено высокими темпами внедрения в образовательный процесс дистанционных и массовых онлайн–курсов. Методы компьютерного адаптивного тестирования (КАТ), которые применяются для проверки усвоения учебного материала, являются не только эффективными механизмами контроля процесса обучения, но и выполняют мотивационную функцию, позволяя создавать тесты, соответствующие по сложности уровню подготовки учащегося. Основные преимущества КАТ:

- неинформативные вопросы постепенно исключаются из теста (например, хорошо подготовленный ученик не должен получать только легкие вопросы). Благодаря чему тест становится более эффективным и точным;
- позволяет получать более точные оценки уровня подготовленности для тех участников, которые находятся в крайних точках распределения (с очень низкими или с очень высокими баллами);
- позволяет избежать эффекта угадывания;
- повторные тесты могут быть совершенно различными.

Главное преимущество адаптивного тестирования – каждый исследуемый получает свой собственный набор задач, различающихся содержанием и длиной теста. Однако, процесс реализации КАТ, является достаточно сложным. Как правило, компонентами КАТ являются:

- откалиброванный банк заданий (для этого используют математический аппарат IRT);
- задаются начальные правила (определяется начальный уровень подготовки тестируемого с помощью предварительного КАТ и т.д.);
- назначается алгоритм подбора вопросов;
- определяется алгоритм пересчета значений уровня подготовки тестируемого;
- выбирается критерий остановки теста.

Выделяют два вида КАТ – одномерное и многомерное. Основная идея для первого случая заключается в том, что полученное значение $\hat{\theta}_n$ (рассчитанный уровень подготовки тестируемого) учитывается в следующем задании таким образом, чтобы отклонение было минимальным $SE(\hat{\theta}_{n+1}) \rightarrow \min$. Дополнительные критерии тоже существуют, но используются реже. Многомерные модели несколько сложнее. Предполагается, что один вопрос может измерять сразу несколько разных способностей. Если скрытые признаки сильно связаны между собой, то появляется избыточность информации. Некоторые из них могут быть более важными для измерения, чем другие. Если имеет место какой-либо из выше перечисленных критериев, то многомерное КАТ считается более предпочтительным, чем несколько одномерных КАТ направленных на оценку разных способностей. Сам процесс подбора вопросов включает в себе пересчет информационной матрицы ($F(\theta)$) для всех оставшихся вопросов в банке заданий. Например, для модели с двумя скрытыми факторами (M2PL) эта матрица выглядит так: $F(\theta) = \begin{bmatrix} a1^2 & a1a2 \\ a1a2 & a2^2 \end{bmatrix} P(\theta)(1 - P(\theta))$ [1]. Для сравнения матриц вопросов используются различные критерии (например, правило максимального определителя). После того как вопрос был назначен и получен ответ на него, значение θ пересчитывается одним из заранее установленным методом (например, взвешенной оценкой максимального правдоподобия). Затем проверяется возможность остановки теста: количество пройденных вопросов отвечает заранее установленному числу, ошибка полученного θ меньше, чем предопределенное значение, закончилось время тестирования, разница между $\hat{\theta}_n$ и $\hat{\theta}_{n+1}$ меньше некоторого δ и т.д.

Первые два пункта компонентов КАТ довольно просты в применении, а остальной процесс более трудоемкий и сложный. В связи с этим использование специализированных пакетов для реализации КАТ в образовательном процессе авторы считают более предпочтительным.

Сегодня для статистических расчетов и графического анализа широко используется среда RStudio. Это бесплатная программа, в которой реализовано огромное количество методов анализа данных, визуализации их результатов, включая взаимодействие с веб приложениями. Пакет `mirtCat` является на сегодня единственным пакетом для работы с многомерными КАТ в котором предусмотрено использование настраиваемого интерфейса пользователя, обработка ответов с помощью M4PL модели, а также обработка противоречивых типов вопросов. С помощью данного пакета КАТ можно запустить локально или удаленно. Можно выделить три основные функции пакета: `mirtCAT()` – используется как для конструирования КАТ так и для его проведения в реальном времени. Включает инструмент для создания HTML-интерфейса для реализации одномерного и многомерного КАТ с использованием математической теории IRT; `generate_pattern()` – создает шаблон КАТ для моделирования по различным параметрам; `generate.mirt_object()` – функция, с помощью которой рассчитывают параметры модели IRT. На рисунках 1–2 показаны экранные формы теста и график изменения значений θ в зависимости от назначаемого вопроса. Всего в банке заданий было 25 вопросов. Из графика вид-

но, что для оценки уровня подготовки данного студента было достаточно семи вопросов. Была допущена всего одна ошибка на втором вопросе.

NUMERICAL METHODS

Choose the correct sentence

Authors:
National Aerospace
University KhAI

This test is developed
at the Informatics
Department

Click 'Next' to advance

Next

The Secant Method does not exist in numerical methods

The Secant Method is the second name of the Newton Method.

The Secant Method is the most popular of the many variants of the Newton Method.

The Secant Method has nothing in common with the Newton Method.

Рисунок 1 – Экранная форма опросника с выпадающим списком ответов

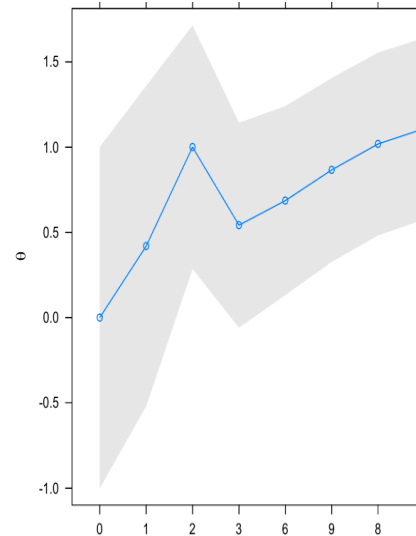


Рисунок 2 – График изменения вычисленного уровня подготовки студента в зависимости от пройденного вопроса

Использование прикладных пакетов среды RStudio для реализации КАТ в дистанционном образовании позволит организовывать гибкий процесс обучения, ориентируясь на индивидуальные способности учащихся.

Список литературы:

1. Chalmers, R. Philip Generating adaptive and non-adaptive test interfaces for multidimensional item response theory applications / R. Philip Chalmers // Journal of Statistical software. / 2016. – V.71. – Issue 5.