

ОПРЕДЕЛЕНИЕ ТРУДОЕМКОСТИ ТЕСТОВЫХ ВОПРОСОВ С ПОМОЩЬЮ ITEM RESPONSE THEORY

Щербина Д.Н.

ФГАОУ ВО «Южный федеральный университет»,

НИТЦ Нейротехнологий,

ФГБОУ ВО «Донской государственный технический университет»,

факультет биоинженерии и ветеринарной медицины,

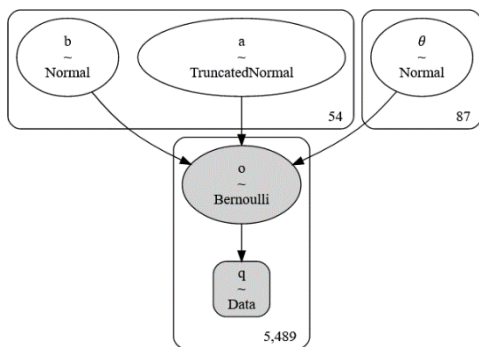
г. Ростов-на-Дону

E-mail: dnsherbina@sfnedu.ru

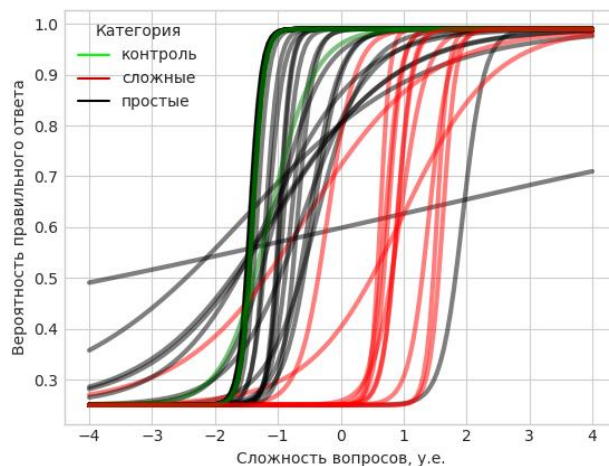
Какое оптимальное время для решения контрольных задач? С одной стороны, подготовленные студенты решают тестовые задания быстрее неподготовленных. С другой стороны – сложные задания требуют большего времени для решения. Если вопросы теста слишком сложные, такие, что у многих в тестовой группе возникают ошибки, то это свидетельствует о том, что времени на решение затрачено недостаточно – по причине спешки или потому что не хватает знаний и оттягивание принятия решения не поможет. Таким образом, при проверке знаний студенты находят компромисс скорости и точности для решения сложных задач. Теория ответов на вопросы (Item Response Theory) утверждает, что существуют латентные/скрытые параметры личности (которые недоступны для непосредственного наблюдения) (Sen, Cohen, 2019). Например, в тесте знаний по предмету — это уровень подготовленности испытуемого и уровень трудности задания.

Мы оценили сложность 54 заданий составного теста, выполненного группой студентов 87 человек, в котором были смешаны простые и сложные вопросы, а также четыре контрольных, которые подразумевали простой готовый ответ. Все задания множественного выбора заключались в выборе одного варианта из четырех предложенных. Параметры угадывания и случайной ошибки фиксировались. Оценку 2PL модели (Рис. 1А) производили в пакете rums3 методом Монте-Карло. Для каждого вопроса были получены параметр сложности **b** и дискриминационный коэффициент **a**.

Два вопроса оказались с низкой дискриминационной способностью из-за того, что формулировка вопроса многими воспринималась некорректно (Рис. 1Б).



А



Б

Рис. 1. Функции ответа на вопросы для получения оценок сложности в рамках Item Response Theory. А. Графическое отображение 2PL-модели. Б. Характеристические функции для трёх категорий вопросов

Эти два вопроса оказались вне общей зависимости повышения среднего времени выполнения задания с ростом сложности. Чтобы избежать влияния нестандартных вопросов, мы использовали пакет statsmodels для оценки робастной регрессии длительности от сложности (Рис. 2). Полученное уравнение средней трудоёмкости задания:

$$d = 27.8 + 11.8b + N(0, 62)$$

где b – безразмерный параметр сложности вопроса.

Данная модель трудоёмкости может быть уточнена с введением в модель индивидуальной скорости выполнения операций, а также учета таких концепций как уверенность и эмоциональная вовлеченность. Для учёта уверенности необходимо внедрять методы гранулярной сегментации времени решения задачи с фиксацией непродуктивных затрат времени на предварительное знакомство с вариантами ответа и проверку (Щербина, 2016). Переход от интеллектуальной системы обучения к аффективной (Hasan et al., 2020) может полностью перевернуть стремление к ускорению образовательного процесса: если целью задания будет увлечь студента на длительное время независимо от корректности ответа, то сложность задания перестанет быть определяющим фактором трудоёмкости, уступив эту роль социально-игровым и физиологическим аспектам.

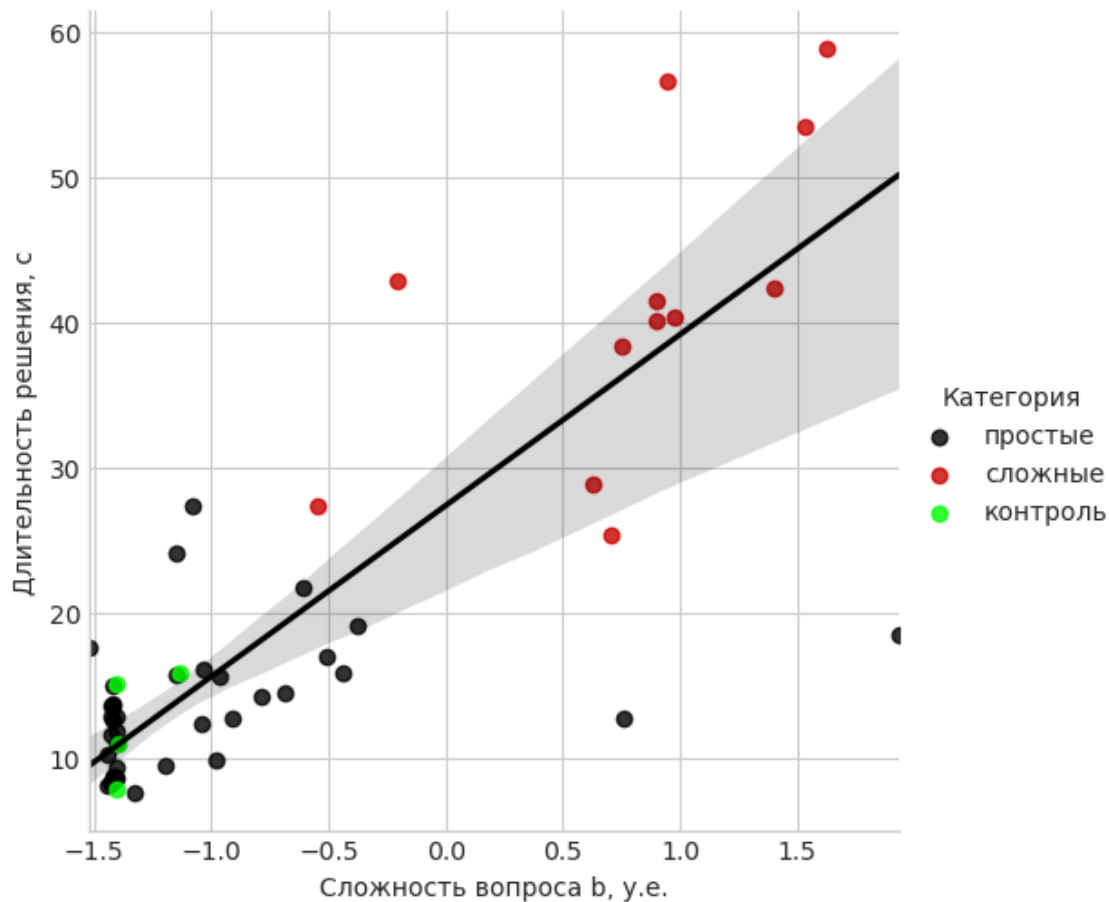


Рис. 2. Зависимость длительности решения заданий от сложности

Литература

1. Sen S., Cohen A.S. Applications of mixture IRT models: a literature review//Measurement: Interdisciplinary Research and Perspectives, 2019, Т. 17, Applications of mixture IRT models, N 4, С. 177-191.
2. Щербина Д.Н. Повышение эффективности контроля знаний студентов на основании анализа последовательности решения тестовых заданий//Образовательные технологии и общество, 2016, Т. 19, N 4, С. 346-363.
3. Hasan M.A., Noor N.F.M., Rahman S.S.B.A., Rahman M.M. The Transition From Intelligent to Affective Tutoring System: A Review and Open Issues//IEEE Access, 2020, Т. 8, The Transition From Intelligent to Affective Tutoring System, С. 204612-204638.