

Исследование уровней трудности заданий теста по моделям современной теории тестирования

С. А. Евдокимова, email: evdsv@mail.ru¹

М. А. Кащенко, email: marg.kashenko2015@yandex.ru¹

¹ ФГБОУ ВО «Воронежский государственный лесотехнический университет имени Г.Ф. Морозова»

Аннотация. В данной работе проводится анализ теста по дисциплине «Информационные технологии в лесном хозяйстве» на основе современной теории тестирования IRT. Построены характеристические кривые трудности заданий и подготовленности студентов, которые показали задания, требующие доработки.

Ключевые слова: современная теория тестирования, модель Г. Раша, двухпараметрическая модель А. Бирнбаума, характеристические кривые, информационно-обучающие системы.

Введение

В настоящее время очень часто для проверки достижений учащихся при очном обучении или изучении дистанционных курсов используется тестирование. Чтобы не было завышения или занижения оценки обучаемых, разработке тестов надо уделять особое внимание.

Проведем исследование заданий теста по дисциплине «Информационные технологии в лесном хозяйстве» на основе современной теории тестирования (IRT), которая позволяет установить взаимосвязь между подготовленностью испытуемого и уровнем трудности заданий, определить какие задания надо доработать. Тест был выполнен студентами «ВГЛТУ» при изучении указанной дисциплины с использованием информационно-обучающей системы [1, 2].

1. Постановка задачи

Анализируемый тест по разделу «Применение информационных технологий в лесном хозяйстве» охватывает 2 темы и содержит 10 заданий закрытого типа, т.е. в каждом задании предлагаются 4 варианта ответа, из которых правильный только один. Правильный ответ оценивается по формуле:

$$x_{ij} = \begin{cases} 1, & \text{если ответ верный;} \\ 0, & \text{если ответ неверный.} \end{cases} \quad (1)$$

где x_{ij} – результат ответа i -го студента ($i \in [1, n]$) на j -ый вопрос ($j \in [1, m]$);

m – количество вопросов в тесте, $m = 10$;

n – количество студентов, отвечающих на тест, $n = 100$.

Результаты ответов на задания тестов сводятся в матрицу, столбцы которой представляют собой номера заданий, а строки – ответы студентов. При этом исключаются строки и столбцы, состоящие из одних нулей или единиц, потому что в этом случае судить об уровне трудности заданий и уровни подготовленности студентов нельзя [3].

В основе современной теории тестирования лежат модели Г. Раша и А. Бирнбаума [3, 4]. Модель Г. Раша представляет собой функции:

$$P_j(\theta) = \frac{e^{1.7(\theta - \beta_j)}}{1 + e^{1.7(\theta - \beta_j)}}, \quad (2)$$

$$P_i(\beta) = \frac{e^{1.7(\theta_i - \beta)}}{1 + e^{1.7(\theta_i - \beta)}}, \quad (3)$$

где θ_i – уровень подготовленности студента;

β_j – уровень трудности заданий в тесте;

P_{ij} – вероятность правильного ответа i -го студента на j -ое задание.

Важной характеристикой модели Г. Раша является инвариантность полученных оценок трудности заданий относительно уровня подготовки испытуемых и наоборот [3-6].

Двухпараметрическая модель А. Бирнбаума учитывает дифференцирующую способность заданий:

$$P_j(\theta) = \frac{e^{1.7 a_j (\theta - \beta_j)}}{1 + e^{1.7 a_j (\theta - \beta_j)}}, \quad (4)$$

где a_j – дифференцирующая способность задания.

2. Анализ результатов расчетов уровня трудности заданий теста

Алгоритм действий при исследовании результатов тестирования на основе теории ИРТ изложен в работах [3, 4] и ранее рассматривался авторами в работе [6].

При расчете по заданной последовательности действий были получены следующие функциональные зависимости:

$$\theta_i = 1.171 \cdot \theta_i^0 - 0.42, \quad i \in [1, n], \quad (5)$$

$$\beta_j = 1.346 \cdot \beta_j^0 + 0.483, j \in [1, m], \quad (6)$$

где θ_i – уровень подготовленности студента;

θ_i^0 – начальные значения уровни подготовленности студентов;

β_j – уровень трудности заданий в тесте;

β_j^0 – начальные значения уровня трудности заданий в тесте.

Результаты расчета уровней трудности заданий представлены в таблице.

Таблица

Расчетные параметры уровня трудности и дифференцирующей способности заданий

Номер задания в тесте, j	Уровень трудности заданий в тесте, β_j	Дифференцирующая способность заданий в тесте, a_j
1	-1,143	0,738
2	-1,651	0,857
3	-0,531	1,015
4	0,538	2,844
5	-0,469	0,518
6	-0,291	0,889
7	1,561	0,932
8	0,808	0,807
9	1,086	11,208
10	-0,722	0,629

Сумма рассчитанных значений уровня трудностей заданий β_j будет равна

$$\sum \beta = \sum_{j=1}^m \beta_j = -0.814. \quad (7)$$

Отрицательное значение суммы (4) показывает, что тест легкий. В хорошем тесте значение суммы (7) должно стремиться к 0.

На основе формул (2) и (3) построены характеристические кривые заданий теста и уровня подготовленности студентов (рис. 1 и 2 соответственно). Проанализировав рис. 1 и 2, можно сделать вывод, что уровень подготовленности студентов θ лежит в пределах от -1.5 до +1.5

логит, уровень трудности заданий – в пределах от -3 до 2 логит, 6 заданий из 10 легкие (их характеристические функции принимают значение $P_j(\theta) = 0.5$ при отрицательном значении θ), трудность заданий 3 и 5 одинакова (их графики практически совпадают).

Так как рассматриваемый тест включает задания по двум темам, каждая из которых состоит из 5 вопросов, то построим отдельно характеристические графики по каждой теме (рис. 3).

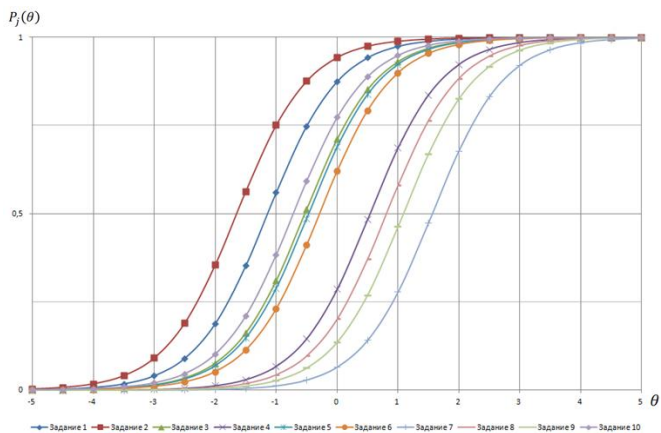


Рис. 1. Характеристические кривые заданий теста

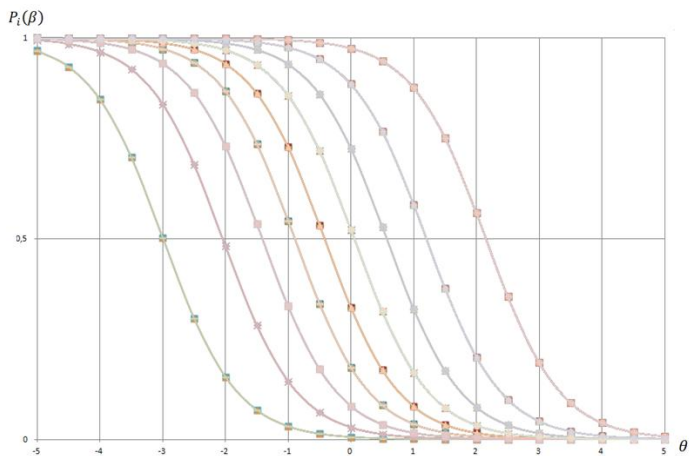
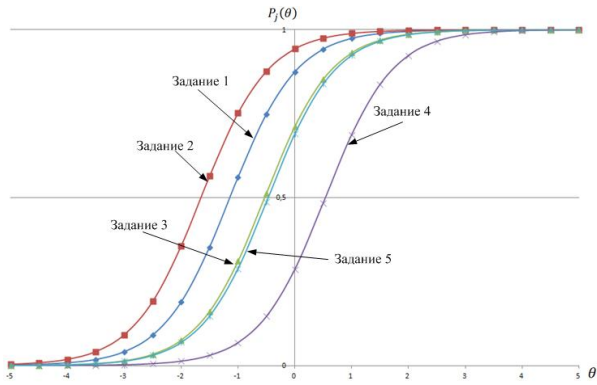
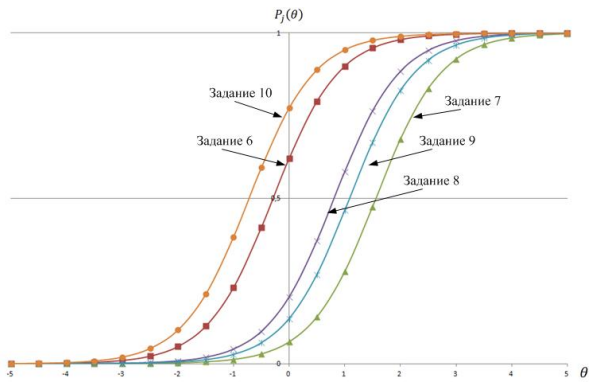


Рис. 2. Характеристические кривые подготовленности студентов



а



б

а – тема 1, б – тема 2

Рис. 3. Характеристические кривые наборов заданий по темам теста

Как видно из рис. 3, в заданиях по первой теме присутствуют два задания 3 и 5 одинаковой трудности, при этом все вопросы в данной теме легкие, только задание 4 имеет повышенную сложность. Во второй теме есть задания и легкие, и сложные, но в обеих темах задания расположены не по возрастанию их трудности. Это не обязательно, но есть смысл поменять порядок заданий в тесте.

Рассчитанные значения дифференцирующей способности заданий a_j для двухпараметрической модели А. Бирнбаума (4) представлены в

таблице. Анализируя эти данные, мы видим, что отрицательных значений и значений, стремящихся к нулю нет. Такие значения заданий a_j указывают на ситуации, когда уровень подготовленности испытуемого не зависит от трудности заданий, и такие задания надо исключать из теста [3]. На практике рекомендуется использовать в тесте задания со значениями из интервала (0.5; 2.5). В этот интервал не попадают значения заданий 4 и 9, поэтому их надо переработать.

Заключение

На основе моделей современной теории тестирования Г. Раша и А. Бирнбаума проанализирован тест «Применение информационных технологий в лесном хозяйстве», который показал, что тест легкий, задания в нем надо усложнить. Одно из заданий 3 или 5 надо убрать, т.к. они одинаковой трудности, а задания 4 и 9 переработать из-за их завышенной дифференцирующей способности.

Список литературы

1. Евдокимова, С.А. Информационно-обучающая система по дисциплине «Информационные технологии в лесном хозяйстве» / С.А. Евдокимова, А.Г. Шматова // Моделирование систем и процессов. – 2017. – Т. 10, № 3. – С. 4-9.
2. Свидетельство об официальной регистрации. Информационно-обучающая система по дисциплине «Информационные технологии в лесном хозяйстве» / С.А. Евдокимова, А.Г. Шматова. – №2017663067 от 23.11.2017.
3. Чельшкова, М.Б. Теория и практика конструирования педагогических тестов : учебное пособие / М.Б. Чельшкова. – М. : Логос, 2002. – 432 с.
4. Ким, В.С. Тестирование учебных достижений : монография. – Усурийск : Издательство УГПИ, 2007. – 214 с.
5. Евдокимова, С.А. Математико-статистическая оценка результатов теста на основе IRT / С.А. Евдокимова, М.А. Кашенко // Моделирование систем и процессов. – 2020. – Т. 13, № 3. – С. 16-22.
6. Муратова, Л.А. Анализ теста «Теория функций комплексной переменной» с привлечением моделей Раша и Бирнбаума / Л.А. Муратова // Вестник Самарского государственного технического университета. Серия: Психолого-педагогические науки. – 2019. – № 1 (41). – С. 111-126.