

Анализ результатов входного тестирования по технологии IRT

Analysis of entrance testing's results by IRT technology

к.т.н., профессор Хватов Юрий Алексеевич

ФГБОУ ВПО Санкт-Петербургский государственный политехнический университет
+7 (905) 287 5092; chvat@mail.ru
Санкт-Петербург
Российская Федерация

Ph.D., Professor Yuri Alexeevich Khvatov

Saint-Petersburg State Polytechnical University
+7 (905) 287 5092; chvat@mail.ru
Saint-Petersburg
Russian Federation

д.э.н., профессор Счисляева Елена Ростиславовна

ФГБОУ ВПО Санкт-Петербургский государственный политехнический университет
+7 (931) 308 8017, info@igms.info
Санкт-Петербург
Российская Федерация

D. Sc., Professor Elena Rostislavovna Schislaieva

Saint-Petersburg State Polytechnical University
+7 (931) 308 8017, info@igms.info
Saint-Petersburg
Russian Federation

Ключевые слова: входное тестирование по математике, анализ результатов с позиций IRT (Item response theory), валидность заданий теста.

Входное тестирование по математике студентов – первокурсников позволяет преподавателю выяснить реальную подготовленность первокурсников к изучению курса высшей математики.

Ниже проводится сравнительный анализ результатов тестирования первокурсников МВШУ (Международная высшая школа управления), поступившими на бюджетное и платное обучение в 2010-2012 гг. и анализ результатов тестирования, качества заданий теста и качества теста в целом с позиций IRT (Item response theory) (бюджетники 2012 год) Содержание тестовых заданий соответствует программе курса математики для средней общеобразовательной школы. Для выполнения заданий достаточно знаний базового уровня.

Результаты тестирования использовались при формировании групп для дополнительных занятий по математике и для составления программы работы с этими группами, включавшей в себя, помимо текущей тематики, ряд разделов из элементарной математики.

Key words: entrance testing, the analysis of results from positions of IRT (Item response theory), a validity of tasks of test.

Mathematics entrance tests of the first year students allows the teacher to find out the state of the art of their readiness to study Higher Mathematics.

Below there is a comparative analysis of the first year IGMS students testing results (IGMS - International Graduate School of Management). Students enrolled in 2010-2012 to both paid and budgetary places are considered. The analysis of testing results (for budgetary places in 2012), the quality of tasks and tests in general is held from the IRT perspective (IRT - Item Response Theory). The maintenance of test tasks corresponds to the program of a course of mathematics for high comprehensive school. For performance of tasks there are enough knowledge of a basic level.

Results of testing were used when forming groups for additional classes on mathematics and for drawing up the program of work with these groups, including, besides the current subject, a number of sections from elementary mathematics.

1. Введение

Входное тестирование по математике проводится в МВШУ (Международная высшая школа управления) уже более чем 5 лет. Ниже проводится сравнительный анализ результатов тестирования студентов–первокурсников, поступивших на бюджетное и платное обучение в 2010-2012 гг. и анализ результатов тестирования, качества заданий теста и качества теста в целом с позиций IRT (Item response theory).

2. Содержание теста

Основная цель теста – получение информации об уровне подготовленности студентов, только что зачисленных на I курс университета, по элементарной математике. Зачисление студентов происходило по результатам ЕГЭ.

Содержание тестовых заданий соответствует программе курса математики для средней общеобразовательной школы. Для выполнения заданий достаточно знаний базового уровня.

В 2010-2012 гг. первокурсникам предлагался один и тот же тест. Тест содержит 20 заданий открытого типа, расположенных, по мнению составителей теста, в порядке возрастания трудности. Время, отведенное на тест – 90 минут. Вот эти задания:

№1. Вычислите: $\frac{1}{2} + \frac{1}{3} - \frac{11}{6}$.

№2. Стороны квадрата, площадь которого равна 15 см^2 , увеличили в 3 раза. Какова площадь нового квадрата?

№3. Выполните действия: $(a^{-1/2} \cdot a^{1/3})^{1/5} : a^{1/15}$.

№4. Катет прямоугольного треугольника равен 1 см, гипотенуза – $\sqrt{5}$ см. Найдите площадь треугольника.

№5. Найдите корни уравнения $\sin x + \sin 2x = 0$ на промежутке $[0; 2\pi]$:

№6. Найдите решение неравенства $\cos x \leq -1/2$ на промежутке $[0; 2\pi]$

№7. Найдите значение a , при котором число $x = -3$ является корнем многочлена $x^3 - x + a$.

№8. Решите уравнение: $\log_4(3x+1) = 2$.

№9. Решите уравнение: $2^{-x} = \frac{1}{2^{1-x}}$.

№10. Решите уравнение: $\sqrt{5x^2 - 4x + 3} = 2x$.

№ 11. Решите систему уравнений:
$$\begin{cases} 4x - y = 3, \\ 7 - 4x = 5y. \end{cases}$$

№12. Найдите область определения функции: $y = \sqrt[4]{2 - \lg x}$.

№13. Вычислите значение выражения: $\log_5 25 \cdot \log_3 27$.

№ 14. Решите неравенство: $x(x + 3) \leq 0$.

№15. Решите неравенство: $\log_5 x \leq -2$.

№16. Упростите выражение: $(\frac{a^{-2}-a}{a^{-1}-1} - 1) : (a^2 + 1)$.

№17. При каких значениях a уравнение $x^2 + x + 2a = 0$ не имеет решений?

№18. Пусть $\sin \alpha = 0.4$. Чему равен $\cos 2\alpha$?

№19. Высота правильной четырёхугольной пирамиды равна $\frac{1}{2}$, а боковое ребро равно $\frac{5}{2}$.
 Найдите объём пирамиды.

№20. Решите неравенство: $9^x - 6 \cdot 3^x + 8 > 0$.

3. Общий сравнительный анализ результатов теста

В таблице 1 представлены общие статистические данные по результатам тестирования в 2010-2012 гг.

Таблица 1

| год | Ср. значение / ср. кв. отклонение | Первичный балл (0-100) Бюджетники | Первичный балл (0-100) Контрактники |
|------|-----------------------------------|--------------------------------------|--|
| 2010 | Ср.значение | 58.5 | 47.2 |
| | Ср.кв.отклонение | 23.4 | 22.4 |
| 2011 | Ср.значение | 59.5 | 36.5 |
| | Ср.кв.отклонение | 22.7 | 16.5 |
| 2012 | Ср.значение | 61.5 | 37.5 |
| | Ср.кв.отклонение | 19.0 | 17 |

Из таблицы видно, что уровень подготовленности абитуриентов – контрактников традиционно гораздо хуже уровня подготовки бюджетников. И это понятно – на бюджетные места есть конкурс и не малый. Кроме того, имеет место в ряде случаев несоответствие результатов ЕГЭ результатам входного тестирования (расхождение в 20 баллов и более отмечается более чем у 12% от прошедших тестирование за 2011 год). Заметим, что у бюджетников таких несоответствий нет (2012год). Результаты тестирования использовались при формировании групп для дополнительных занятий по математике и для составления программы работы с этими группами, включавшей в себя, помимо текущей тематики, ряд разделов из элементарной математики.

4. Анализ результатов тестирования и качества заданий теста в рамках IRT (бюджетники 2012 г.)

Таблица 2. Результаты по испытуемым (всего участников 41)

| | Первичный балл (0-100) | Оценка уровня подготов (в логит.) | Ошибка измерения уровня подготов. | Статистики согласия | | | |
|-----------------|------------------------|-----------------------------------|-----------------------------------|---------------------|-----------|----------------------|----------------------|
| | | | | $U^{(1)}$ | $t^{(1)}$ | $U^{(2)}$ (взвеш) | $t^{(2)}$ (взвеш) |
| Ср.значение | 61.5 | 0.86 | 0.60 | 0.99 | 0.0 | 1.00 | 0.0 |
| Ср.квадр.отклон | 19.08 | 1.26 | 0.10 | 0.30 | 1.1. | 0.84 | 0.9 |
| Макс. значение | 85 | 3.74 | 1.06 | 1.69 | 2.5 | 4.15 | 2.4 |
| Мин.значение | 30 | 1.19 | 0.54 | 0.54 | -2.1 | 0.24 | -1.4 |

Таблица 3. Результаты по заданиям

| | Процент решенных задач, % | Оценка трудности заданий (в логит.) | Ошибка измерения трудности заданий | Статистики согласия | | | |
|-----------------|---------------------------|-------------------------------------|------------------------------------|---------------------|-----------|----------------------|----------------------|
| | | | | $U^{(1)}$ | $t^{(1)}$ | $U^{(2)}$ (взвеш) | $t^{(2)}$ (взвеш) |
| Ср.значение | 62 | 0.00 | 0.49 | 1.00 | 0.0 | 1.00 | 0.0 |
| Ср.квадр.отклон | 22.5 | 1.56 | 0.16 | 0.17 | 0.80 | 0.57 | 0.80 |
| Макс. значение | 92.5 | 2.39 | 1.036 | 1.51 | 2.1 | 2.85 | 1.9 |
| Мин.значение | 23.5 | -3.29 | 0.40 | 0.65 | -1.7 | 0.43 | -1.3 |

1. Тест включает в себя 20 заданий. Максимальный балл за выполнение теста равен 100. Среднее значение первичных баллов испытуемых равно 61,5. Следовательно, в среднем участник выполнил 61.5% заданий.

2. Во второй графе таблиц представлены значения оценок трудности заданий и уровней подготовленности испытуемых. Все оценки даны в логитах [1, 9, 11]. В соответствии с используемой моделью измерения каждое задание теста характеризуется одним числом – трудностью δ . Среднее значение оценок трудности всех заданий положено равным 0. Это сделано с целью фиксирования начала отсчета на метрической шкале логитов, на которой находятся все оценки – и заданий и испытуемых. Задания сильно варьируются по уровню трудности: самое легкое задание имеет уровень трудности $\delta = -3,29$ (задание №1), самое трудное - $\delta = 2,39$ (задание №19!). Участники тестирования также сильно отличаются по уровню подготовленности: лучший участник имеет оценку уровня подготовленности $\theta = 2,39$, худший - $\theta = -3,29$. В третьей графе таблиц представлены значения среднеквадратических ошибок вычисления оценок. Ошибка измерения для заданий невелика: среднее значение ошибки измерения по заданиям равно 0,49 по испытуемым 0.60. Заметим, что точность оценивания параметров заданий более высока, чем точность оценивания мер испытуемых. Это объясняется тем, что число участников тестирования превышает в 2 раза количество заданий в тесте, а точность результатов во многом зависит от объемов выборок, то есть от количества участников тестирования и количества заданий в тесте. Поэтому естественно ожидать, что ошибка измерения заданий теста будет ниже ошибки измерения испытуемых.

4. В таблицах 2 и 3 представлены данные о статистиках согласия, характеризующих согласие экспериментальных данных тестирования с используемой моделью измерения. Это – общие статистики согласия $U^{(1)}$ и $U^{(2)}$ и их стандартизованные версии $t^{(1)}$ и $t^{(2)}$ [1]. В соответствии с моделью математические ожидания значений общих статистик согласия равны 1, стандартизованных – 0. Из таблиц 2 и 3 мы видим, что средние значения всех статистик согласия близки к их ожидаемым значениям, что означает, что большинство заданий и испытуемых находятся в хорошем согласии с моделью. Однако есть задания и испытуемые, находящиеся в плохом согласии с моделью, о чем говорят максимальные значения статистик согласия, значительно превышающие критические. Эти задания и испытуемые будут проанализированы ниже.

Коэффициент надежности (KR-20) измерения характеристик испытуемых данным тестом равен 0,78. Коэффициент надежности измерения характеристик заданий равен 0,89 .

На рисунках 1 и 2 приведены графики характеристической и информационной [1,9,11] функций теста.

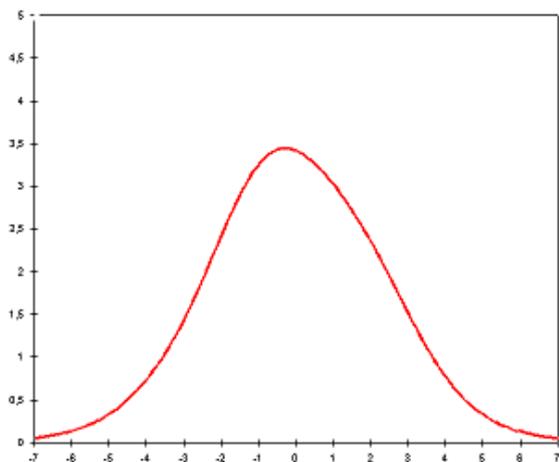


Рис. 1. Информационная функция теста

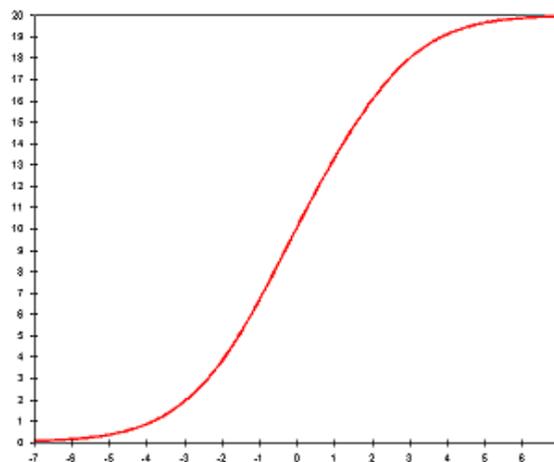


Рис.2. Характеристическая функция теста

5. Анализ заданий теста.

В таблице 4 приведены статистические данные по заданиям теста. Задания расположены в порядке их следования в тесте. В первой колонке представлен номер задания, В следующих трех колонках представлены оценка трудности задания на шкале логитов и ошибка измерения трудности задания. Так, например, задания 16,19 – самые трудные в заданном тесте: их трудность $\delta = 2,01$ и $2,39$, ошибка измерения $\sigma = 0,43$ и $0,45$ соответственно. Задание №1 - самое легкое: его трудность $\delta = -3,29$ ошибка измерения $\sigma = 1,03$. Далее следуют значения статистик согласия $U(1)_i$, $t(1)_i$, $U(2)_i$, $t(2)_i$. Средние значения статистик согласия представлены в табл. 2 и 3. В предпоследней колонке представлен коэффициент корреляции баллов по заданию с общим баллом по тесту. В последней колонке представлен коэффициент дискриминативности задания D_i (одна из версий D_i). В идеальном случае $D_i = 1$.

Таблица 4. Результаты по заданиям теста

| Номер задачи | Трудность задания (лог.) | Сделали задание (в %) (Из 41) | Ошибка σ | Статистики согласия | | | | Коеф. Корреляции (бис.) | Коеф-т дискр-сти D_i |
|--------------|--------------------------|-------------------------------|-----------------|---------------------|-----------|-----------|-----------|-------------------------|------------------------|
| | | | | $U^{(1)}$ | $t^{(1)}$ | $U^{(2)}$ | $t^{(2)}$ | | |
| 1 | -3,29 | 97 | 1,03 | 1,00 | 0,30 | 0,04 | 0,19 | 0,19 | 1,02 |
| 2 | -0,49 | 73 | 0,44 | 1,50 | 2,07 | 1,61 | 0,09 | 0,09 | 0,22 |
| 3 | -0,30 | 71 | 0,43 | 1,22 | 1,03 | 0,25 | 0,35 | 0,35 | 0,73 |
| 4 | -0,30 | 71 | 0,43 | 1,03 | 0,23 | -0,16 | 0,46 | 0,46 | 0,99 |
| 5 | 1,33 | 41 | 0,40 | 1,07 | 0,45 | 0,32 | 0,44 | 0,44 | 0,85 |
| 6 | 1,33 | 41 | 0,40 | 0,90 | -0,59 | -0,65 | 0,57 | 0,57 | 1,25 |
| 7 | 0,54 | 56 | 0,40 | 0,92 | -0,37 | -0,41 | 0,56 | 0,56 | 1,16 |
| 8 | -2,05 | 91 | 0,63 | 1,01 | 0,17 | 1,12 | 0,46 | 0,24 | 0,96 |
| 9 | -2,05 | 91 | 0,63 | 1,01 | 0,17 | 1,12 | 0,46 | 0,24 | 0,96 |
| 10 | 0,22 | 62 | 0,41 | 1,03 | 0,23 | 0,91 | -0,19 | 0,49 | 0,99 |
| 11 | -1,39 | 85 | 0,52 | 1,13 | 0,51 | 2,86 | 1,95 | 0,17 | 0,72 |
| 12 | 1,49 | 38 | 0,40 | 0,88 | -0,68 | 0,85 | -0,45 | 0,56 | 1,24 |
| 13 | -2,52 | 94 | 0,75 | 0,96 | 0,12 | 0,44 | -0,17 | 0,30 | 1,07 |
| 14 | -0,69 | 76 | 0,45 | 0,65 | -1,68 | 0,43 | -1,29 | 0,68 | 1,50 |
| 15 | 1,33 | 41 | 0,40 | 0,99 | 0,01 | 0,90 | -0,30 | 0,51 | 1,06 |
| 16 | 2,01 | 29 | 0,43 | 0,90 | -0,45 | 0,68 | -0,82 | 0,55 | 1,24 |
| 17 | 1,02 | 47 | 0,40 | 0,86 | -0,84 | 0,86 | -0,51 | 0,59 | 1,29 |
| 18 | 0,22 | 62 | 0,41 | 0,80 | -1,01 | 0,70 | -1,00 | 0,63 | 1,38 |
| 19 | 2,39 | 24 | 0,45 | 1,23 | 1,00 | 1,18 | 0,52 | 0,28 | 0,69 |
| 20 | 1,21 | 41 | 0,41 | 0,95 | -0,22 | 0,92 | -0,22 | 0,52 | 1,10 |

Примечание. Трудности задний приведены в логитах., $U^{(1)}$, $U^{(2)}$ - простая и взвешенная статистики согласия, статистики $t^{(1)}$, $t^{(2)}$ – это статистики $U^{(1)}$, $U^{(2)}$, приведенные к стандартизованному виду $N(0,1)$ [1].

5.1. Распределение трудностей заданий теста

Выпишем из таблицы 3 эмпирическое распределение заданий теста по трудностям: 1-13-8-9-11-14-2-3-4-18-7-10-17-5-6-15-20-12-16-19. Сразу видно, что теоретическое распределение трудностей заданий и распределение трудностей, полученное в результате апробации теста, не совпадают. В основном это связано с фактической многомерностью теста, вызванного многообразием проверяемых разделов курса (групп навыков и умений), зачастую слабо связанных между собой. Возможна и другая причина - неправильная схема оценивания заданий как дихотомических, хотя на самом деле эти задания было бы правильно трактовать как политомические (например задание №5). В любом случае необходим анализ получившегося рассогласования и корректировка целой группы заданий и новой ранжировки заданий по уровням трудности. В первую очередь это задания 5, 6, 13, 18.

Заметим, что близость теоретического и эмпирического распределения трудностей свидетельствует о том, что в тесте правильно оценены когнитивные процессы, необходимые для выполнения заданий теста и то, что тест измеряет то, что задумано авторами (валидность конструкта). Впрочем, по нашему мнению, распределение задач в тесте не должно оказывать большого влияния на результативность тестирования.

4.2. Валидность заданий теста

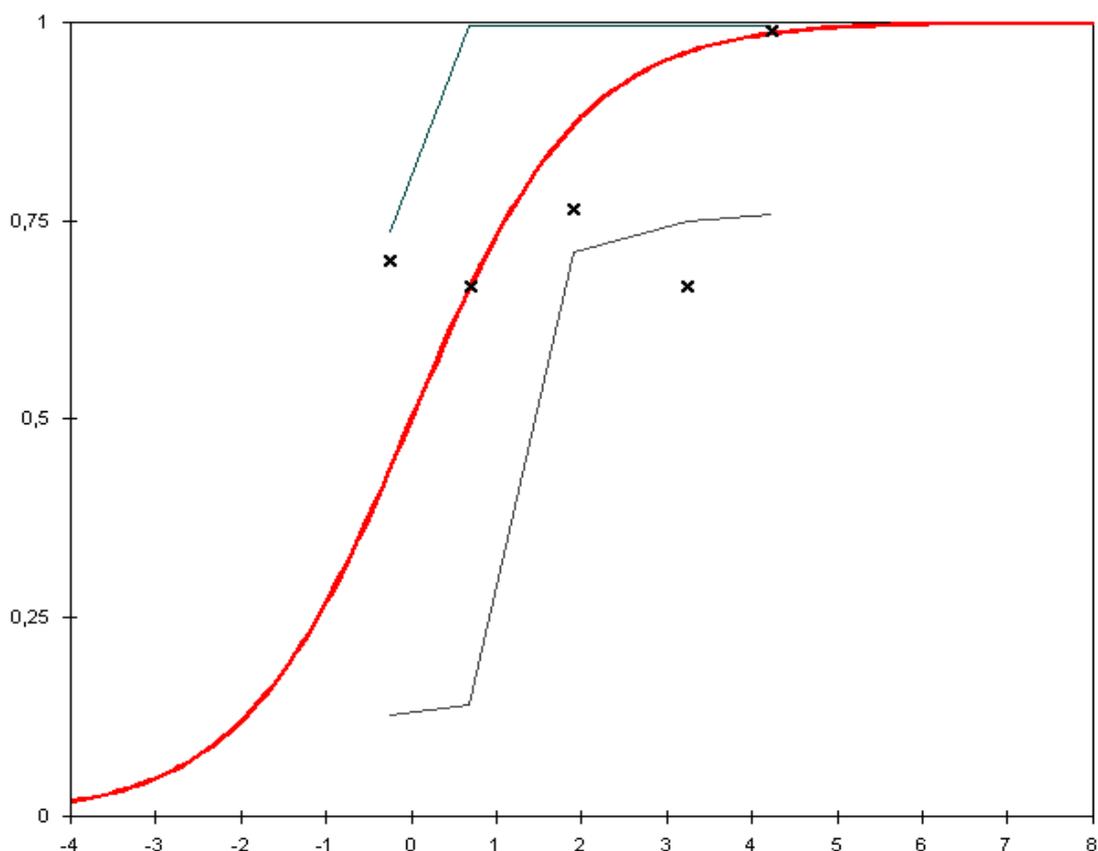
Применение моделей Раша позволяет определить вероятность правильного ответа любого испытуемого на любое задание теста. Тем самым возможно определить вероятность профиля ответов на любое задание. Если профиль ответов на любое задание имеет неправильный характер – слабые испытуемые выполнили его правильно, а сильные неправильно, то вероятность такого профиля очень мала. В таком случае разумно предположить, что данное задание не функционирует правильно и вносит дезинформацию в оценивание испытуемых. Такое задание не отвечает цели тестирования, то есть не обладает валидностью конструкта. Выявить такие задания можно с помощью статистик согласия (см. табл 4). Задания, значения статистик согласия которых имеют значения большие правых критических, обладают более низкой дифференцирующей способностью, чем другие задания теста. Задания, статистики согласия которых имеют значения меньшие левых критических, обладают более высокой дифференцирующей способностью, чем другие задания теста. И в том и другом случае происходит нарушение одного из основополагающих требований моделей Раша: характеристические кривые заданий не должны пересекаться, т.е. дифференцирующая сила заданий заранее предполагается примерно одинаковой. Это свойство обеспечивает важные преимущества моделей Раша.

Из таблицы видно, что большинство заданий теста демонстрируют хорошее согласие с моделью: но у одного из заданий – №2 – значения более двух статистик согласия больше правых критических, у задания 11, значения статистики согласия $U^{(2)}$ больше правого критического значения, а у заданий 13,14 значения статистик согласия $U^{(2)}$ меньше левых критических значений (в качестве правого критического значения для статистик $U^{(1)}$ и $U^{(2)}$ выбрано значение 1,5, левого-0,5; для статистик $t^{(1)}$ и $t^{(2)}$ - значения (-2) и 2 соответственно [1]),

Самым проблемным является задание №2: Задание имеет уровень трудности ниже среднего: $\delta = -0,49$. Три статистики согласия $U^{(1)}$, $U^{(2)}$ и $t^{(1)}$ имеют значения, большие правых критических. Это означает, что задание вызывает большие затруднения у хорошо подготовленных испытуемых, чем у подготовленных более слабо. Следовательно, это задание вводит в заблуждение сильных учащихся, которые в результате выполняют его неправильно. Причина, возможно в том, что хорошо подготовленные испытуемые прибегают к вычислениям и при этом ошибаются, а подготовленные более слабо опираются на интуицию и на известное отношение площадей подобных фигур (см. задание в п.2.).

Заметим, что корреляция баллов по этому заданию с общим баллом по тесту самая низкая(0.09) и самый низкий коэффициент дискриминативности.

Приведем характеристическую кривую анализируемого задания (рисунок 3).



**Рис. 3. Характеристическая кривая задания №2.
(тонкой линией указан доверительный коридор с $p=0,95$)**

Из рисунка 3 видно, что эмпирическая характеристическая кривая задания сильно отличается от теоретической кривой: четыре эмпирические точки находятся вне модельной кривой и часть из них вне доверительного коридора. Более того, испытуемые с высоким уровнем подготовки выполнили это задание хуже испытуемых из слабой группы. Задание абсолютно не дифференцирует испытуемых, внося дезинформацию в их оценивание. Это задание находится в плохом согласии с ответами испытуемых на другие задания теста и, кроме того, демонстрирует смещение оценок трудности по отношению к различным группам испытуемых, о чем свидетельствуют данные проведенного анализа. Такое задание следует удалить из теста.

На рисунке 4 для сравнения приведена характеристическая кривая задания №4, значения, все критерии согласия которого находятся в допустимых пределах.

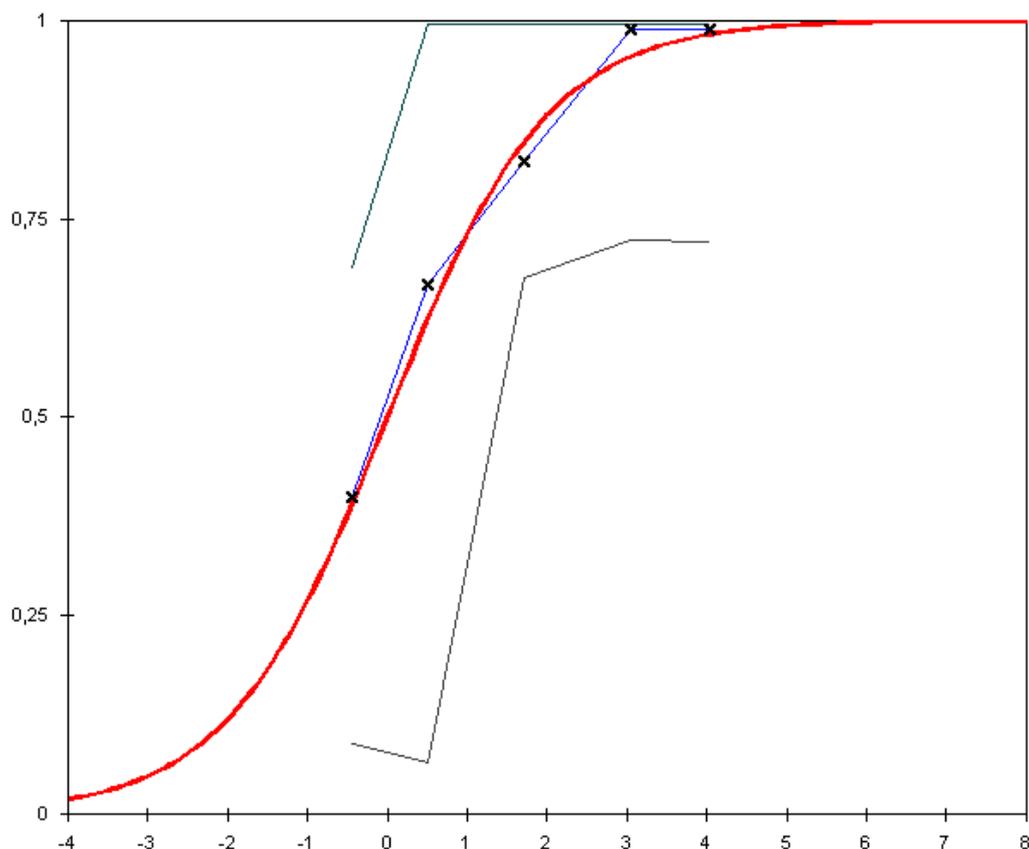


Рис. 4. Характеристическая кривая задания № 4
(тонкой линией указан доверительный коридор с $\rho=0,95$)

Функционирование задания №2, как впрочем, и других заданий теста, было изучено на двух различных группах испытуемых. С этой целью исходная выборка была разбита на 2 выборки: в первую группу вошли первые 21 участников, показавших лучшие результаты, во вторую группу – 20 испытуемых, показавших худшие результаты

В первой группе первичные баллы варьировались от 9 до 19; среднее значение первичных баллов равно 15,2; Во второй группе первичные баллы варьировались от 2 до 10; среднее значение первичных баллов равно 8,8.

Задание №2 плохо согласуется с моделью и в первой и второй группах. С этим заданием необходимо работать. Обращают на себя внимание задания № 3, 7, 10, 14, 18, чья дифференцирующая способность для разных групп испытуемых различна. Видимо, необходимо проанализировать и, в дальнейшем, изменить эти задания (см.п. 2.).

Распределения задач по трудности в сильной и слабой подгруппах сильно отличаются (см. рисунок 5). Коэффициент ранговой корреляции Спирмана между распределением задач по трудности в первой подгруппе равен 0,56, во второй группе 0,17, что говорит о разном функционировании заданий для разных групп испытуемых. Но имеет место совпадение параметров испытуемых (уровней подготовки) полученных при тестировании исходной (полной группы) и сильных и слабых студентов.

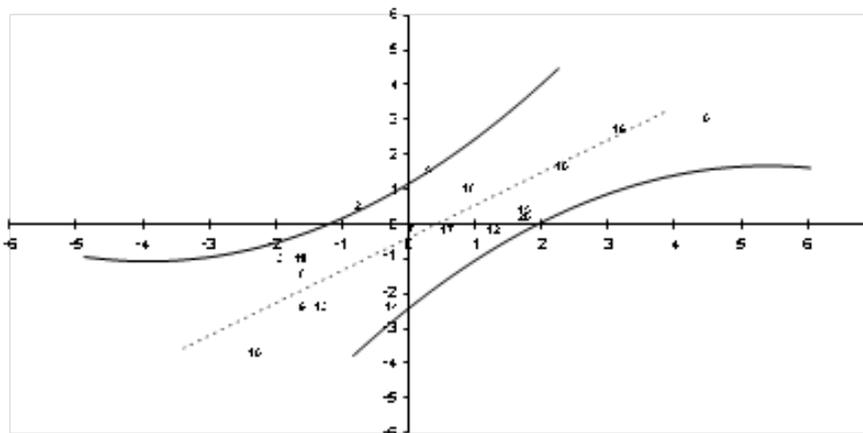


Рис. 5. Сравнение трудностей заданий по результатам обсчета данных тестирования слабой подгруппы (ось ОУ) и сильной (ось ОХ)

6. Заключение

Проведенный анализ позволил создать новый улучшенный вариант теста для студентов – первокурсников 2013 года.

Заметим в заключение, что технология IRT может быть с успехом использована при анализе больших массивов экспериментальных данных.

Литература

1. Нейман Ю. М., Хлебников В. А.. Введение в теорию моделирования и параметризации педагогических тестов. Москва, 2000. 165 с.
2. Карданова Е. Ю. Согласие экспериментальных данных тестирования с моделью измерения // Вопросы тестирования в образовании. 2007. №18. С.5-18
3. Карпинский В. Б. Исследование эффективности общих статистик согласия для обнаружения искажений при массовых тестирования // Вопросы тестирования в образовании. 2006. №1(17). 7 с.
4. Аванесов В. С. Основы теории педагогических заданий // Педагогические Измерения. 2006. №2. С.26- 62.
5. Аванесов В. С. Композиция тестовых заданий. М.: Центр тестирования, 2002. 215 с.
6. Аванесов С. В. Теоретические основы разработки заданий в тестовой форме. М.: Исследовательский центр, 1995. 95 с.
7. Васильев В. И., Тягунова Т. Н Основы культуры адаптивного тестирования М; Изд – во ИКАР, 2003. 370 с.
8. Майоров А. Н. Теория и практика создания тестов для системы образования. М.: Народное образование, 2000. 352 с.
9. Чельшкова М. Б. Теория и практика конструирования педагогических тестов. Учебное пособие. М.: Исследовательский центр проблем качества подготовки специалистов, 2001. 74 с.
10. Звонников В. И., Чельшкова М. Б. Контроль качества обучения при аттестации. Компетентностный подход. М.: Из-во Университетская книга, Логос, 2009. 272 с.
11. Звонников В. И, Чельшкова М. Б. Современные методы оценивания результатов обучения. Уч.пособие. М.: Из-во Академия, 2009. 224 с.
12. Новые педагогические и информационные технологии в системе образования: Учеб. Пособие / Полат Е. С, Бухаркина М. Ю., Моисеева М. В, Петров А. Г; Под ред. Полат Е. С. М. Издательский центр «Академия», 2002. 272 с.
13. Иванов Б. С. Основы педагогической диагностики и мониторинг образовательной деятельности в техническом вузе : Монография. СПбГПУ, 2003. 120 с.
14. Baker F. V. The Basics of Item Response Theory. Heinemann, 1985. 137 p.
15. Hambleton R. K., Swaminathan H. Item Response Theory: Principles and Applications. Boston, 1985. 327 p.
16. Bimbaam A. Some Latent Trait Models and Their Use is Statistical Theories of Mental Test Sc Reading//Journal of Educational and Behavioral Statistics, 1994. Vol. 19. No. 3. Pp. 293-295

17. Lord F. M. Application of Item Response Theory to Practical Testing Problems. Hillsdale, N.,J., Erlbaum Ass., 1980. 296 p.
18. John M. Linacre. A User's Guide to "WINSTEPS", Rasch -Model computer program 1.680. P.O. BOX 811322, Chicago, 60681-1322. 2012. [электронный ресурс]. URL: <http://www.winsteps.com/winman/index.htm?copyright.htm> (дата обращения: 06.04.2013).
19. Baxter J-A. The Impact of Professional Learning on the Teaching Identities of Higher Education Lecturers//European journal of open, distance and e-learning. 2012. Vol. 10. Pp. 1-11
20. Хватов Ю. А., Речинский А. В., Кетов Д. В. Технологии дистанционного обучения в очном образовании при изучении курса «Математика» // Строительство уникальных зданий и сооружений. 2013. №3. С. 60-67.

References

1. Nejman Yu. M., Xlebnikov V. A.. Vvedenie v teoriyu modelirovaniya i parametrizacii pedagogicheskix testov. Moskva, 2000. 165 p. (rus)
2. Kardanova E. Yu. Soglasie e'ksperimental'nyx dannyx testirovaniya s model'yu izmereniya // Voprosy testirovaniya v obrazovanii. 2007. No. 18. Pp.5-18. (rus)
3. Karpinskij V. B. Issledovanie e'ffektivnosti obshhix statistik soglasiya dlya obnaruzheniya iskazhenij pri massovyx testirovaniya // Voprosy testirovaniya v obrazovanii. 2006. No. 1(17). 7 p. (rus)
4. Avanesov V. S. Osnovy teorii pedagogicheskix zadaniy // Pedagogicheskie Izmereniya. 2006. No. 2. Pp.26- 62. (rus)
5. Avanesov B. C. Kompoziciya testovyx zadaniy. M.: Centr testirovaniya, 2002. 215 p. (rus)
6. Avanesov S. V. Teoreticheskie osnovy razrabotki zadaniy v testovoj forme. M.: Issledovatel'skij centr, 1995. 95 p. (rus)
7. Vasil'ev V. I., Tyagunova T. N Osnovy kul'tury adaptivnogo testirovaniya. M; Izd – vo IKAR, 2003. 370 p. (rus)
8. Majorov A. N. Teoriya i praktika sozdaniya testov dlya sistemy obrazovaniya. M.: Narodnoe obrazovanie, 2000. 352 p. (rus)
9. Chelyshkova M. B. Teoriya i praktika konstruirovaniya pedagogicheskix testov. Uchebnoe posobie. M.: Issledovatel'skij centr problem kachestva podgotovki specialistov, 2001. 74 p. (rus)
10. Zvonnikov V. I., Chelyshkova M. B. Kontrol' kachestva obucheniya pri attestacii. Kompetentnostnyj podxod.M.: Iz-vo Universitetskaya kniga, Logos, 2009. 272 p. (rus)
11. Zvonnikov V. I, Chelyshkova M. B. Sovremennye metody ocenivaniya rezul'tatov obucheniya. Uch.posobie. M.: Iz-vo Akademiya, 2009. 224 p. (rus)
12. Noveye pedagogicheskie i informacionnye texnologii v sisteme obrazovaniya: Ucheb. Posobie / Polat E. S, Buxarkina M. Yu., Moiseeva M. V, Petrov A. G; Pod red. Polat E. S. M. Izdatel'skij centr «Akademiya», 2002. 272 p. (rus)
13. Ivanov B. S. Osnovy pedagogicheskoy diagnostiki i monitoring obrazovatel'noj deyatel'nosti v texnicheskom vuze: Monografiya. SPbGPU, 2003. 120 p.
14. Baker F. B. The Basics of Item Response Theory. Heinemann, 1985. 137 p. (rus)
15. Hambleton R. K., Swaminathan H. Item Response Theory: Principles and Applications. Boston, 1985. 327 p.
16. Bimbaam A. Some Latent Trait Models and Their Use is Statistical Theories of Mental Test Sc Reading//Journal of Educational and Behavioral Statistics, 1994. Vol. 19. No. 3. Pp. 293-295
17. Lord F. M. Application of Item Response Theory to Practical Testing Problems. Hillsdale, N.,J., Erlbaum Ass., 1980. 296 p.
18. John M. Linacre. A User's Guide to "WINSTEPS", Rasch -Model computer program 1.680. P.O. BOX 811322, Chicago, 60681-1322. 2012. [web source]. URL: <http://www.winsteps.com/winman/index.htm?copyright.htm> (date of reference: 06.04.2013).
19. Baxter J-A. The Impact of Professional Learning on the Teaching Identities of Higher Education Lecturers // European journal of open, distance and e-learning. 2012. Vol. 10. Pp. 1-11
20. Khvatov Yu. A., Rechinskiy A. V., Ketov D. V. Distance learning technologies in the full-time education course "Mathematics" // Construction of unique buildings and structures. 2013. No. 3. Pp. 60-67. (rus)

Данная статья публикуется в рамках работы по проекту
530603-TEMPUS-1-2012-1-LT-TEMPUS-JPCR

This article is published in the framework of project
530603-TEMPUS-1-2012-1-LT-TEMPUS-JPCR