

Эти значения используются для оценки уровня знаний и уровня усвоения знаний и построения карты знаний учащегося.

Модель решения тестового задания учащимся. Пусть запротоколированные действия пользователя выглядят так: « $x = \text{резСравнения}(23,31)$ ».

В этом случае мы можем сделать вывод, что ученик применил свойство сравнения, и исследовать соответствующую ветвь дерева решений.

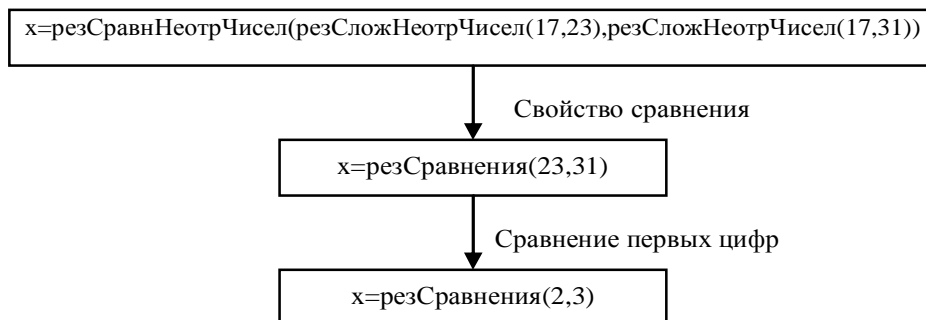


Рис. 3. Путь решения задания испытуемым

Если ученик отметил и дальнейшие действия – « $x = \text{резСравнения}(2,3)$ », то мы можем извлечь еще больше информации. В данном случае он правильно идентифицировал ситуацию сравнения первых цифр при равенстве порядков (рис. 3). На рисунке видно, какие знания применяет учащийся. Эту информацию можно использовать для уточнения карты знаний испытуемого. В свою очередь, карта знаний содержит оценки условных вероятностей правильного распознавания ситуации (знания ситуации), правильного разрешения ситуации и знания и использования УЭ в целом.

Выводы. Рассмотрен новый подход к тестированию, при котором модель тестового задания строится на основе сети знаний учебного модуля. Модель тестового задания строится на основе дерева путей решений. Дерево путей решений описывает структуру ситуаций, возникающих при применении операций или свойств УЭ. На основе дерева решений можно объективно оценить трудоемкость и сложность задания, а также вероятность его решения «средним испытуемым». На основе оформленной части решения можно сужать дерево возможных решений испытуемого. Для построения масштабируемой модели знаний учащегося можно использовать байесовскую сеть, условные вероятности для которой оцениваются по результатам тестирования.

Работа поддержана РФФИ, грант 08-07-00217а.



Литература

Нехаев И.Н. Математическая модель уровня усвоения знаний // Школьные технологии. 2009. № 6.

ФАКТОРНЫЙ ГРУППОВОЙ АНАЛИЗ РЕЗУЛЬТАТОВ ЕГЭ ПО МАТЕМАТИКЕ

Нехаев Игорь Николаевич (garry_nekhaev@mail.ru),

Красильников Михаил Игоревич (mi.krasilnikov@gmail.com)

ГОУ ВПО «Марийский государственный технический университет», г. Йошкар-Ола

АННОТАЦИЯ

В статье анализируются результаты тестирования ЕГЭ по математике с помощью факторного анализа и алгоритма экстремальной группировки параметров. Основная цель исследований – выявить, что можно измерить контрольно-измерительными материалами ЕГЭ по математике.

Введение. Многими авторами подчеркивается необходимость использования тестовых заданий (ТЗ) разного уровня сложности для создания таких тестов [1]. Однако иногда сложность предложенных в тесте заданий сильно превышает уровень подготовки учащихся. Так нужны ли в тесте такие задания? Чтобы дать аргументированный ответ, проанализируем, что же проверяют тестовые задания ЕГЭ по математике? О чем говорит итоговый балл ЕГЭ по математике? Какую еще информацию можно получить из результата тестирования? Зачем нужны задания части 3?

Для этого проведем факторный анализ результатов тестирования, выявим минимальное количество факторов, с достаточной степенью точности объясняющих эти результаты, и попытаемся проинтерпретировать их.

Для исследования были взяты результаты ЕГЭ по математике за 2006 год. Количество объектов (тестируемых выпускников) в матрице данных равнялось 577. Тестируемые – это абитуриенты второй волны, сдающие ЕГЭ в июле.

Факторный анализ тестовых заданий. Будем предполагать линейную связь между измеряемыми параметрами и факторами:

$$x^j = \sum \alpha_{kj} \cdot f^k + \xi_j.$$

Здесь x^j – это столбец матрицы результатов тестирования, соответствующий j -му заданию теста (параметр), f^k – k -й общий фактор, а α_{kj} – факторная нагрузка, показывающая степень влияния конкретного фактора на параметр [3]. Если для объяснения результатов измерений было задействовано меньше общих факторов, чем измерено параметров (что, собственно, и является желаемым результатом факторного анализа), то модуль ξ_j (вектор ошибки аппроксимации j -го параметра или характерный фактор) будет отражать погрешность такой аппроксимации. Для вычисления факторных нагрузок используем метод главных компонент [2].

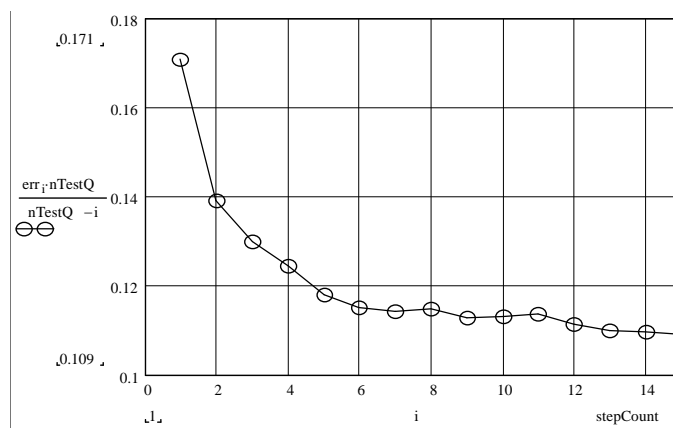


Рис. 1. Зависимость среднего квадрата характерных факторов от количества используемых для аппроксимации общих факторов по формуле (1)

На рисунке 1 показано убывание погрешности аппроксимации с ростом количества общих факторов. Из рисунка видно, что погрешность снижается значительно при использовании 5–6 общих факторов для аппроксимации параметров. Анализ распределения абитуриентов по значениям факторов показывает, что явно выделяются первые три фактора.

По остальным факторам получается распределение, близкое к нормальному (рис. 2а, б). На рисунках 3–4 представлены диаграммы факторных нагрузок для первых шести по значимости главных факторов. По оси абсцисс указаны номера заданий теста ЕГЭ: с 1-го по 10-е задания – это блок А, с 11-го по 21-е – блок В, а с 22-го по 26-е – блок С.

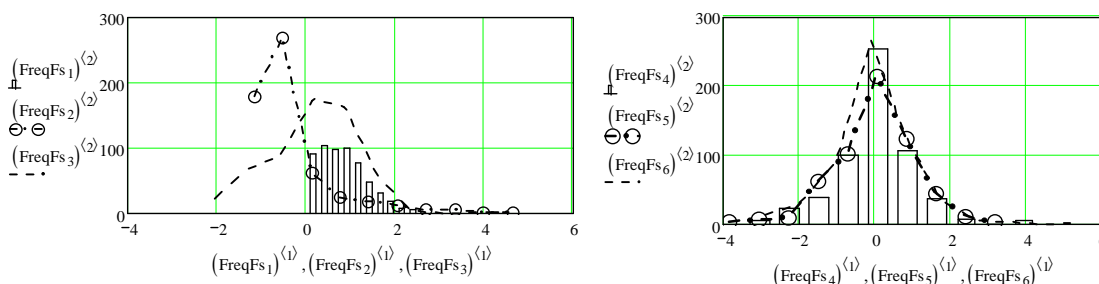


Рис. 2. Полигоны распределения частот абитуриентов по значениям факторов а) 1-го, 2-го, 3-го; б) 4-го, 5-го, 6-го

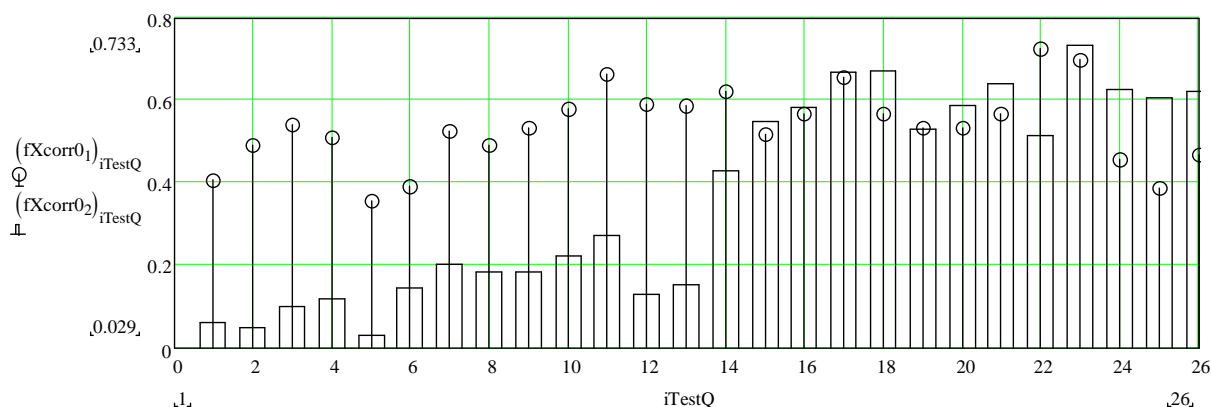


Рис. 3. Диаграмма факторных нагрузок для первых двух по значимости факторов. По оси абсцисс указаны номера заданий теста ЕГЭ: с 1-го по 13-е задания – это часть 1, с 14-го по 23-е – часть 2, а с 24-го по 26-е – часть 3

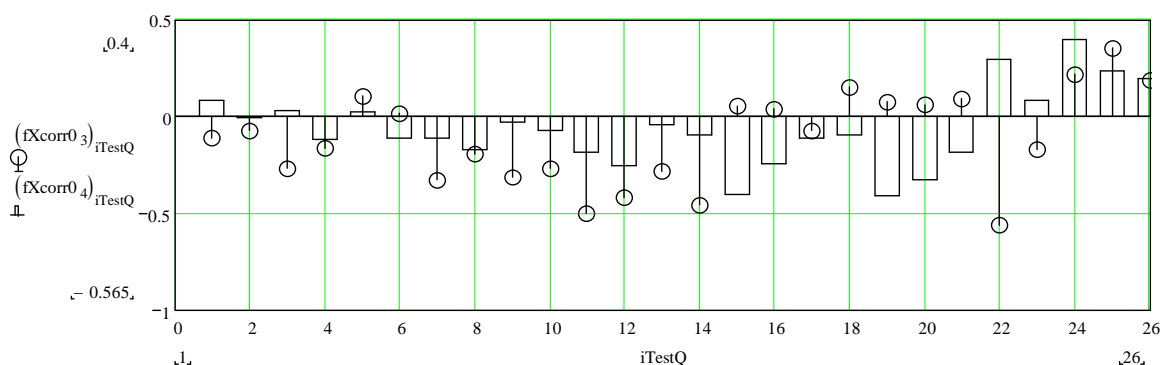


Рис. 4. Диаграмма факторных нагрузок для 3,4-го по значимости факторов. По оси абсцисс указаны номера заданий теста ЕГЭ: с 1-го по 10-е задания – это блок А, с 11-го по 21-е – блок В, а с 22-го по 26-е – блок С

Чтобы проинтерпретировать факторы, необходимо проанализировать содержание заданий теста. Мы используем значения факторов для выделения групп заданий теста и проанализируем каждую группу вместо анализа каждого из 26 заданий.

Экстремальная группировка параметров (заданий теста). Другой подход к обработке результатов тестирования заключается в рассмотрении корреляций заданий теста. В методе «экстремальная группировка параметров» параметры делятся на фиксированное количество групп. Для оценки качества разбиения выбирается функционал, значение которого возрастает по мере увеличения взаимной корреляции параметров внутри групп и уменьшения связей между параметрами в разных группах. На каждом последующем шаге алгоритма предпринимается попытка переместить очередной параметр из текущей группы в другую таким образом, чтобы это привело к возрастанию функционала. Если при проверке всех параметров не было произведено ни одного перемещения, то алгоритм завершает работу. На получаемые таким методом результаты влияет как выбор функционала, так и начальное разбиение.

Функционал разбиения представляет собой сумму частных функционалов. Для каждой группы параметров (ГЗ) вычисляется частный функционал, равный сумме квадратов (или модулей) коэффициентов корреляции каждого параметра с главным фактором для группы. Рассматривалось количество групп, равное 3, 5 и 7. При начальном разбиении на три группы задания можно разделить так же, как они группируются в билете ЕГЭ: группа А (1–10), группа В (11–19), группа С (20–26). Или задания части 1 (1–13), части 2 (14–23) и части 3 (24–26). Для получения начального разбиения на большее количество групп использовалось равномерное разбиение или учитывались результаты предварительного анализа содержания заданий (табл. 1). В результате работы алгоритма были получены следующие результаты для 5 и 7 групп (табл. 2).

Таблица 1

Начальные разбиения заданий на группы

Начальное разбиение	Кол-во групп	Группы заданий теста	Значение функционала
№ 2a	5	1 гр.: = (1, 2, 3, 5, 13, 22), 2 гр.: = (4, 7, 9, 10, 11, 14), 3 гр.: = (6, 8, 12, 15), 4 гр.: = (16, 17, 18, 19), 5 гр.: = (20, 21, 23, 24, 25, 26)	Std = 12,1
№ 2b	5	1 гр.: = (1, 2, 3, 4, 5), 2 гр.: = (6, 7, 8, 9, 10), 3 гр.: = (11, 12, 13, 14, 15), 4 гр.: = (16, 17, 18, 19, 20), 5 гр.: = (21, 22, 23, 24, 25, 26)	Std = 12,03
№ 3a, porog = 1; porog = 3;	7	1 гр.: = (1, 2, 3), 2 гр.: = (4, 5, 6, 7), 3 гр.: = (8, 9, 10, 11), 4 гр.: = (12, 13, 14), 5 гр.: = (15, 16, 17, 18), 6 гр.: = (19, 20, 21, 22), 7 гр.: = (23, 24, 25, 26)	Std = 13,5 Std = 13,07
№ 3b, porog = 1; porog = 3;	7	1 гр.: = (1, 2, 3, 5, 13, 22), 2 гр.: = (4, 7, 9, 10, 11, 14), 3 гр.: = (6, 8, 12), 4 гр.: = (15, 16, 19, 20), 5 гр.: = (24, 25, 26), 6 гр.: = (17, 18, 21), 7 гр.: = (23)	Std = 13,7 Std = 13,27

Параметр porog служит для задания порога удаления из матрицы результатов тестирования строк с результатами слабых учащихся, которые в сумме баллов набрали меньше, чем пороговое значение. По умолчанию порог равен 1.

Таблица 2

Результаты группировки заданий теста

Нач-ое разб.	Колич. измен.	Группы заданий теста	Значение функционала
Любое	–	Гр. 1 = (1–14, 22), Гр. 2 = (15–21, 23), Гр. 3 = (24–26)	Std = 10,99
№ 2a	13	1 гр.: = (2–5 , <u>12, 13</u>), 2 гр.: = (1, <u>7</u> , 8, 9, <u>10, 11, 14, 22</u>), 3 гр.: = (6), 4 гр.: = (<u>15, 16, 17, 18, 19, 20, 21, 23</u>), 5 гр.: = (<u>24, 25, 26</u>)	Std = 12,78
№ 2b	11	1 гр.: = (2–5, 12, 13), 2 гр.: = (1, 6, 7, 10), 3 гр.: = (8, 9, <u>11, 14, 22</u>), 4 гр.: = (<u>15–21, 23</u>), 5 гр.: = (<u>24, 25, 26</u>)	Std = 12,76
№ 3a, porog = 0	16	1 гр.: = (<u>5</u>), 2 гр.: = (2, <u>3</u> , 4, <u>12, 13</u>), 3 гр.: = (1, 6, <u>7, 10</u>), 4 гр.: = (8, 9, <u>11, 14, 22</u>), 5 гр.: = (15, <u>16, 17, 18, 23</u>), 6 гр.: = (<u>19, 20, 21</u>), 7 гр.: = (<u>24, 25, 26</u>)	Std = 14,25
№ 3a, porog = 3	14	1 гр.: = (<u>5</u>), 2 гр.: = (2, <u>3</u> , 4, <u>12, 13</u>), 3 гр.: = (1, 6, <u>10</u>), 4 гр.: = (7, 8, 9, <u>11, 14, 22</u>), 5 гр.: = (15, <u>16, 17, 18, 23</u>), 6 гр.: = (<u>19, 20, 21</u>), 7 гр.: = (<u>24, 25, 26</u>)	Std = 13,84
№ 3b, porog = 0	16	1 гр.: = (2, <u>3</u> , 4, 5, <u>12, 13</u>), 2 гр.: = (1, <u>7, 10</u>), 3 гр.: = (<u>6</u>), 4 гр.: = (<u>15, 19, 20</u>), 5 гр.: = (<u>24, 25, 26</u>), 6 гр.: = (8, 9, <u>11, 14, 22</u>), 7 гр.: = (<u>16, 17, 18, 21, 23</u>)	Std = 14,275
№ 3b, porog = 3	16	1 гр.: = (2, <u>3</u> , 4, 5, <u>12, 13</u>), 2 гр.: = (1, 6, <u>10</u>), 3 гр.: = (<u>8, 4</u>), 4 гр.: = (<u>15, 19, 20</u>), 5 гр.: = (<u>24, 25, 26</u>), 6 гр.: = (7, 9, <u>11, 14, 22</u>), 7 гр.: = (<u>16, 17, 18, 21, 23</u>)	Std = 13,78

В таблице выделены и подчеркнуты самые характерные задания группы. Видим, что, несмотря на различие в количествах групп, в начальных разбиениях и в пороговых значениях, существуют закономерности в экстремальной группировке заданий. Отчетливо выделяются 5 групп заданий. Проанализируем их содержание.

Анализ содержания групп тестовых заданий

Таблица 3

Содержательная характеристика групп заданий теста

Группы заданий	Содержание
1 гр. = (2, <u>3</u> , 4, 5, <u>12</u> , <u>13</u>)	Простые задания на знание и самое простое воспроизведение
2 гр. = (8, 9, <u>11</u> , <u>14</u> , <u>22</u>)	Типовые задания, требующие небольшого количества формально-логических выкладок-преобразований
3 гр. = (15, <u>16</u> , <u>17</u> , <u>18</u> , <u>23</u>),	Вообще говоря, нетиповые задания средней сложности, требующие понимания, анализа и умения применять знания
5 гр. = (19, 20, 21)	Нетиповые текстовые и геометрические задания средней сложности, требующие понимания, анализа и умения применять знания, строить модель задания.
6 гр. = (24, 25, 26)	Нетиповые задания высокого уровня сложности, кроме всего вышперечисленного требуют умения синтезировать решение, умение оценивать знания и выбирать стратегию решения задания.

Результатом анализа содержания тестовых заданий явилась следующая интерпретация общих факторов (табл. 4).

Таблица 4

**Характеристика групп заданий теста
(средние значения коэффициентов корреляции тестовых заданий группы с факторами)**

Группы заданий теста	Факт. 1	Факт. 2	Факт. 3	Факт. 4	Факт. 5	Факт. 6
1 гр.: = (2–5, 12, 13)	+0,5	+0,1	–0,18	–0,1	–0	–0
2 гр.: = (8, 9, 11, 14, 22)	+0,6	+0,3	–0,4	–	–	+0,1
3 гр.: = (1, 6, 7, 10)	+0,5	+0,15	–0,2	–0	–0	0
4 гр.: = (15–18, 23)	+0,6	+0,7	0	–	+0	–0
5 гр.: = (19, 20, 21)	+0,55	+0,6	+0,1	–0,3	+0,1	+0
6 гр.: = (24, 25, 26)	+0,45	+0,62	+0,27	+0,27	–0,2	+0

Фактор № 1 является интегральным и отвечает за общий уровень знаний. Данный фактор проверяется всеми заданиями теста и пропорционален набранному баллам. Именно он и отображается как результат теста. Коэффициент корреляции с суммарным баллом составляет 0,998. Однако следует отметить, что явно выделяются остаточные «тени» еще, по меньшей мере, двух факторов. Фактор № 2 можно интерпретировать как отражение уровня понимания и умения применять имеющиеся знания. Поэтому он хорошо коррелирует именно с заданиями частей 2 и 3. Он положителен только для тех тестируемых, которые характеризуются небольшим объемом знания или для очень хорошо подготовленных абитуриентов, для которых уровень знания превышает 60–70 % от общего [3]. Анализируя корреляцию фактора № 3 с заданиями теста, можно сказать, что он характеризует дихотомию – склонность к формально-логическому решению (отрицательные значения) или к «мягкому» способу решения, к образному мышлению (положительные значения).

Выводы. Анализ содержания тестовых заданий показал, что задания первой части (A1–A10, B1–B3) и некоторые задания части 2 проверяют общий уровень знания и способности к формальному мышлению. Задания частей 2, 3 проверяют еще анализ и умение применять знания. Задания 3-й части проверяют еще и способность к синтезу решения, к образному математическому мышлению. Показано, что «многознание не есть знание» и что большинство поступающих 2-й волны (более 80 %) либо мало знают, либо не усвоили свои знания и склонны к формальному мышлению.

Работа поддержана РФФИ, грант 08-07-00217а.

Литература

1. Аванесов В.С. Знания как предмет педагогического измерения // Школьные технологии. 2006. № 3.
2. Браверманн Э.М., Мучник И.Б. Структурные методы обработки эмпирических данных. М.: Наука, 1983. 464 с.
3. Беспалько В.П. Параметры и критерии диагностической цели // Школьные технологии. 2006. № 1.