

# АВТОМАТИЗИРОВАННЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ В ОБРАЗОВАНИИ И НАУКЕ

УДК 004.43 : 378.146  
DOI: 10.17223/16095944/63/10

Ю.В. Осипов, Г.Л. Сафина, Ф.Я. Ветухновский  
Национальный исследовательский Московский государственный строительный университет,  
Москва, Россия

## МОДЕЛИРОВАНИЕ ТЕСТОВ ПО МАТЕМАТИКЕ

Разработанный по оригинальной методике тест по математике используется для оценки уровня и структуры знаний студентов. Авторская компьютерная программа генерирует индивидуальные варианты теста для каждого студента. По выбору преподавателя тестовые задания могут предлагаться в открытой или закрытой форме. При проведении тестирования одна группа испытуемых выполняла задания в открытой форме, а две группы – в закрытой форме со списком из пяти вариантов ответов. Исследованы причины различия результатов тестирования с открытыми и закрытыми заданиями. На основе анализа результатов тестирования трех групп испытуемых изучаются статистические характеристики теста. Тестовые баллы учащихся подчиняются нормальному распределению. По результатам проведенного тестирования отдельно для каждой группы испытуемых построены шкалы сложности заданий. Проверена гипотеза о сохранении порядка в шкале сложности заданий при выполнении тестов в открытой и закрытой формах испытуемыми с различным уровнем знаний. Исследуются профили знаний студентов, построенные по результатам тестирования. Каждому профилю сопоставляется регулярность – число, характеризующее степень правильности профиля. Изучаются распределение регулярности, его статистические характеристики.

**Ключевые слова:** тест, задания открытого и закрытого типа, тестовый балл, сложность задания, профиль испытуемого.

**Введение.** Задания в тестовой форме используются в педагогике, психологии, социологии, медицине и многих других областях [1–5]. Тесты могут разрабатываться как для отдельных организаций, ведомств, регионов, так и для целого ряда стран [6–9]. В системе образования контроль знаний является неотъемлемой частью процесса обучения и оценки качества знаний учащихся [10, 11]. В настоящее время преподаватели вузов широко используют тесты в различных формах для контроля уровня подготовки студентов [12, 13]. Педагогический тест является набором тщательно подобранных заданий различной сложности по определенной тематике [14]. Тесты позволяют оперативно и объективно оценить знания учащихся и сократить время проверки работ. При помощи тестов можно определить уровень знаний студентов как перед началом изучения дисциплины, так и после прохождения разделов университетского курса [15, 16]. Моделирование тестов и статистический анализ результатов тестирования позволяют определить не только уровень знаний обучающихся, но и эффективность используемого теста [17–21]. В НИУ МГСУ

тесты используются не только для текущего контроля знаний студентов, но и на вступительных экзаменах для поступающих в магистратуру. В отличие от ЕГЭ для выпускников школ тесты для отбора магистрантов состоят только из заданий единственного выбора с несколькими вариантами ответов. В связи с этим представляет интерес сравнение результатов тестов с заданиями открытого и закрытого типов. Кроме того, в настоящее время тестирование студентов проводится в рамках аттестации вузов [22]. В работе анализируются свойства разработанного теста по высшей математике, проводится сравнение результатов тестирования в нескольких группах испытуемых, оцениваются уровень и структура знаний студентов, моделируются статистические характеристики теста.

**Подготовка и проведение теста. Обработка результатов.** Авторы разработали тест по разделу «Введение в анализ» дисциплины «Математический анализ». Тест состоит из 20 заданий по темам: свойства и пределы последовательностей, пределы функций, бесконечно малые и бесконечно большие величины, непрерывность и разрывы функций. Для каждого из 20 видов

заданий составлено 30 однотипных задач с числовыми ответами. Варианты теста образуются путем случайного выбора задач на основе оригинальной методики [23]. Создано 20 различных файлов в формате Microsoft Word, соответствующих двадцати различным тестовым заданиям, в каждом файле содержится 30 однотипных задач. Авторская программа в виде макроса на языке Visual Basic генерирует 30 параллельных индивидуальных вариантов тестов, распределяя задания между испытуемыми случайным образом, что обеспечивает уникальность набора заданий для каждого студента во всех учебных группах. Программа позволяет преподавателю выбирать тип тестовых заданий. Они могут быть либо закрытого типа, когда испытуемый должен выбрать готовый ответ из пяти предложенных вариантов ответов, либо открытого типа, когда испытуемый самостоятельно решает задачу и записывает результат. Студенты выполняют тесты в письменной форме. Это позволяет проводить тестирование во время учебных занятий при изучении дисциплин, не предусматривающих использование компьютерных классов. Преподаватель может быстро проверить задания, используя файл ответов, который генерируется программой одновременно с индивидуальными вариантами тестов.

Апробация теста была проведена в девяти учебных группах студентов первого курса Московского государственного строительного университета, обучающихся по направлениям бакалавриата «Экономика» и «Информатика и вычислительная техника». Первые три группы экономистов образовали группу А, другие три – группу В, а студенты-системотехники – группу С. Студенты групп А и В, составляющие один лекционный поток и обучающиеся по единой методике, близки по уровню знаний; студенты группы С изучали математический анализ в том же объеме, что и экономисты, но имели в среднем более высокий балл ЕГЭ по математике. Испытуемые группы А выполняли тестовые задания в открытой форме, а учащиеся групп В и С – задания в закрытой форме.

Результаты тестирования были статистически обработаны с помощью табличного процессора Microsoft Excel [24]. Каждая задача оценивалась по двухбалльной шкале: 0 – задача решена неверно, 1 – задача решена верно. Согласно классической теории тестов результат тестирования

испытуемого определяется тестовым баллом – суммой набранных баллов при решении тестовых заданий.

Для каждой группы были вычислены основные характеристики распределения тестовых баллов: математическое ожидание  $\bar{X}$ , дисперсия  $s_x^2$ , стандартное отклонение  $s_x$  и коэффициент асимметрии  $A$ . Результаты представлены в табл. 1.

Таблица 1

Распределение тестовых баллов

Группа	Кол-во студентов	$\bar{X}$	$s_x^2$	$s_x$	$A$
А	51	7,2	8,4	2,9	0,75
В	66	10,8	5,7	2,4	0,04
С	60	11,4	6,8	2,6	0,3

Результаты тестирования оказались предсказуемыми: лучший результат показали наиболее сильные студенты группы С, среди близких по уровню подготовки студентов групп А и В лучшими оказались учащиеся группы В, выполнявшие тестовые задания в закрытой форме.

Для оценки дифференцирующей способности теста, т.е. возможности разделять при помощи набранного тестового балла учащихся с различным уровнем подготовки, используется отношение стандартного отклонения к математическому ожиданию. В группах В и С значение стандартного отклонения меньше среднего балла приблизительно в 4,5 раза, в группе испытуемых А – в 2,5 раза, что указывает на достаточно высокую дифференцирующую способность теста как в закрытой, так и открытой форме.

**Проверка гипотезы о нормальном распределении тестовых баллов.** Известно, что для нормативно-ориентированного теста, предназначенного для ранжирования испытуемых по уровню знаний, кривая распределения тестовых баллов испытуемых симметрична и близка к гауссовой кривой.

На рис. 1 представлены гистограммы распределения индивидуальных баллов для трех групп испытуемых в зависимости от количества решенных задач и теоретические кривые распределения.

Согласованность теоретического и статистического распределений была проверена с помощью критерия Пирсона  $\chi^2$ . В табл. 2 приведены расчетные и критические значения статистики Пирсона для всех трех групп при уровне значимости

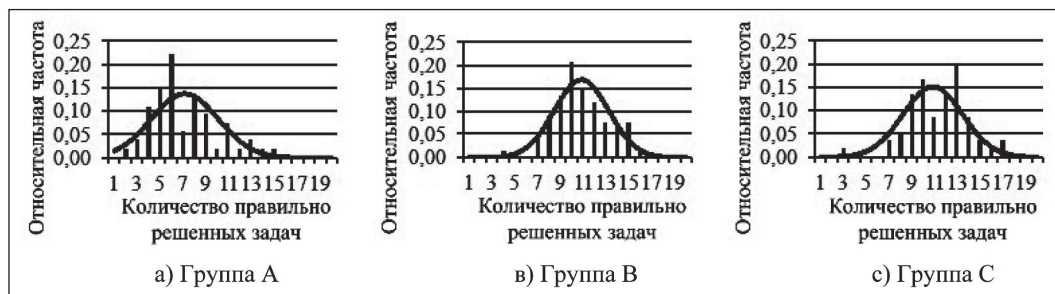


Рис. 1. Распределение тестовых баллов

$\alpha=0,05$ . Для каждой группы расчетное значение статистики  $\chi^2$  меньше критического, значит, гипотеза о нормальном распределении тестовых баллов испытуемых во всех трех группах принимается. Следовательно, разработанные тесты с заданиями как открытого, так и закрытого типа могут быть использованы при изучении раздела «Введение в математический анализ» в качестве нормативно-ориентированных тестов. Данные тесты позволяют сравнить уровень знаний студентов и выяснить, кто из них лучше, а кто хуже владеет пройденным материалом.

Таблица 2

Расчетные и критические значения

Группа	$\chi^2_{\text{расч}}$	$\chi^2_{\text{кр}}$
А	8	9,49
В	3,8	12,6
С	10,6	11,1

**Сравнение результатов открытых и закрытых тестов.** Сравним результаты тестирования студентов групп А и В. Как отмечалось выше, студенты этих двух групп имеют примерно одинаковый уровень знаний, однако средние тестовые баллы в группах существенно различаются (см. табл. 1). Испытуемые двух групп выполняли одни и те же тестовые задания, но студенты группы А должны были самостоятельно получить ответ (задачи открытого типа), а студенты группы В должны были выбрать правильный ответ из предлагаемого списка (задания закрытого типа). Если испытуемый группы В не смог выполнить задание, он мог попытаться угадать ответ, выбрав произвольно один из пяти предлагаемых вариантов. Если студент группы А не решил задачу, он не мог угадать верный ответ.

Различие средних тестовых баллов может объясняться как возможностью угадывания от-

ветов при выполнении закрытых тестов, так и различиями в поведении учащихся при решении открытых и закрытых заданий. Если студент не пытается угадать ответ, то сам по себе список числовых ответов в закрытом задании не может способствовать нахождению правильного ответа, но помогает заметить ошибку. Поскольку значительная часть ошибок в математических задачах связана с вычислениями, возможные ошибки носят случайный характер и их количество существенно превышает число вариантов ответа в тесте. В связи с этим найденный студентом неправильный ответ часто не совпадает ни с одним ответом из списка закрытого задания. Если учащийся не находит полученный ответ в списке, он понимает, что допустил ошибку, и снова решает задачу. Испытуемый, выполняющий задание в открытой форме, лишен такой возможности. Относительная простота закрытых заданий по сравнению с открытыми заданиями связана с возможностями как угадывания ответа, так и самостоятельного обнаружения ошибки в решении задания. Варианты коррекции тестовых баллов для заданий в закрытой форме рассматриваются в [25–27].

**Исследование сложности заданий.** Сложностью  $j$ -й задачи  $q_j$  называется отношение числа испытуемых, не решивших данную задачу, к общему числу испытуемых. Простотой  $j$ -й задачи назовем число  $p_j = 1 - q_j$  – отношение числа решивших данную задачу к общему числу испытуемых. Если задачу решили все студенты, то  $p_j = 1$ ,  $q_j = 0$ ; если никто не выполнил задание, то  $p_j = 0$ ,  $q_j = 1$ . Такие излишне простые или слишком сложные задания следует исключить при рассмотрении результатов тестирования, поскольку они не позволяют дифференцировать студентов по уровню знаний. В общем случае все числа  $p_j$  и  $q_j$  лежат в диапазоне от 0 до 1.

Сложность и простота всех 20 заданий определялись по результатам тестирования отдельно в каждой группе испытуемых. Среднее значение  $\bar{p}$  простоты заданий в группе равно отношению среднего числа выполненных заданий  $\bar{X}$  к объему теста  $n = 20$ ; среднее значение сложности  $\bar{q} = 1 - \bar{p}$ . В исследуемом тесте

$$\bar{p}^A = 0,36; \bar{q}^A = 0,64; \bar{p}^B = 0,54;$$

$$\bar{q}^B = 0,46; \bar{p}^C = 0,57; \bar{q}^C = 0,54.$$

По результатам тестирования в каждой группе тестовые задания были упорядочены по убыванию сложности; соответствующие шкалы сложности представлены на рис. 2.

Все три шкалы оказались близкими, попарные корреляции трех массивов простоты заданий приблизительно равны 0,9. Таким образом, хотя сложность задач зависит от формы заданий и уровня подготовки учащихся, относительная сложность/простота заданий, т.е. их взаимное расположение на шкале сложности примерно одинаково для всех трех групп испытуемых.

**Распределение чисел регулярности профилей испытуемых.** Профилем знаний испытуемого называется индивидуальный результат выполнения теста [28, 29]. В случае дихотомической оценки выполнения заданий профиль испытуемого представляет собой двоичный код – набор единиц и нулей, получаемых студентом за выполненные тестовые задания. Для оценки качества индивидуальных профилей производится сравнение профилей со шкалой сложности заданий. Профиль называется правильным, если испытуемый дает верные ответы на легкие задания и неверные ответы на сложные задания, т.е. вектор-строка его оценок в матрице тестовых результатов имеет вид 11...100...0. Правильному профилю присваивается значение 0. Если же испытуемый правильно отвечает на трудные задания и неправильно на легкие, т.е. строка баллов испытуемого имеет

вид 00 ... 011 ... 1, то профиль называется инвертированным, и ему присваивается значение, равное 1.

Чаще всего студент делает и сложные, и простые задачи, и его профиль занимает промежуточное положение между 0 и 1. Такому профилю сопоставляется число по формуле

$$C_i = \frac{\sum_{j=1}^{x_i} (1 - S_{ij}) R_j - \sum_{j=x_i+1}^n S_{ij} R_j}{\sum_{j=1}^{x_i} R_j - \sum_{j=n+1-x_i}^n R_j} = \frac{\sum_{j=1}^{x_i} R_j - \sum_{j=1}^n S_{ij} R_j}{\sum_{j=1}^{x_i} R_j - \sum_{j=n+1-x_i}^n R_j}, \quad (1)$$

где  $x_i$  – индивидуальный балл испытуемого;  $R_j$  – количество правильных ответов на  $j$ -е задание;  $n$  – общее число заданий;  $S_{ij} = \{0, 1\}$  – балл, получаемый испытуемым за  $j$ -е задание (индикатор выполнения задания). Число  $C_i$ , характеризующее степень отклонения профиля от правильного, называется регулярностью профиля.

Исследование профилей позволяет оценить степень упорядоченности знаний студентов и помогает выявить обучающихся, фрагментарно изучивших материал либо нарушивших правила проведения тестирования (подсказки, списывание и т.п.).

Используя простоту заданий  $p_j$ , перепишем формулу (1) в виде

$$C_i = \frac{\sum_{j=1}^{x_i} p_j - \sum_{j=1}^n S_{ij} p_j}{\sum_{j=1}^{x_i} p_j - \sum_{j=n+1-x_i}^n p_j}. \quad (2)$$

Во всех группах регулярности профилей студентов принимают значения в диапазоне от 0 до 0,4, что указывает на правильную структуру

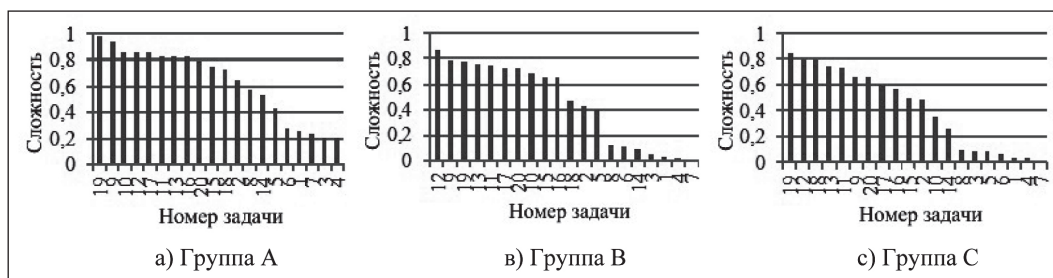


Рис. 2. Шкала сложности заданий





Рис. 3. Распределение регулярности профилей

знаний испытуемых. Средние значения регулярности профилей равны

$$\bar{C}^A = 0,18, \bar{C}^B = 0,11, \bar{C}^C = 0,13.$$

Гистограммы распределений значений регулярности профилей в трех группах приведены на рис. 3.

Значение регулярности профиля испытуемого зависит от его тестового балла. Как было показано выше, тестовые баллы испытуемых распределены по нормальному закону. Известно, что линейное преобразование нормальной случайной величины приводит к нормальной случайной величине. Однако коэффициенты линейного преобразования (2) зависят от величины тестового балла студента, что может деформировать нормальное распределение. При проверке статистической гипотезы о нормальном распределении регулярности профилей на основании критерия Пирсона выяснилось, что гипотеза принимается для двух групп экономистов и отвергается для студентов-системотехников.

**Заключение.** В работе изучаются возможности создания и использования параллельных тестов для оценки знаний студентов по математике на примере теста по теме «Введение в анализ». Тест предлагался студентам в открытой и закрытой формах. Сравнение результатов тестирования показало, что различие набранных баллов в закрытых и открытых тестах нельзя объяснить исключительно случайным угадыванием ответов. Возможно, неудачно подобранные дистракторы создают дополнительный перекосяк в пользу верного ответа. Также не исключено, что испытуемые, решающие трудные задания самостоятельно, получают дополнительную подсказку в виде списка возможных вариантов ответа. Изучение причин различий тестовых баллов в закрытых и открытых заданиях может стать предметом отдельного исследования.

Результаты тестирования показали, что обе формы теста успешно дифференцируют испытуемых по уровню знаний и могут использоваться для контроля знаний учащихся на занятиях по математическому анализу.

Сравнение сложности заданий для трех групп испытуемых позволило сделать вывод о сохранении порядка в шкале сложности заданий при выполнении тестов в открытой и закрытой формах испытуемыми с различным уровнем знаний.

#### ЛИТЕРАТУРА

1. Fisher W.P. The Central Theoretical Problem of the Social Sciences // *Rasch Measurement Transactions*. – 2014. – Vol. 28:2. – P. 1464–1466.
2. Kenneth D.R. Measuring Liberal / Conservative Voting Tendencies among U.S. Senators // *Rasch Measurement Transactions*. – 2012. – Vol. 26:2. – P. 1366–1367.
3. Reeve B.B. Item response theory modeling in heart outcomes measurement // *Expert Review of Pharmacoeconomics and Outcomes Research*. – 2003. – Vol. 3(2). – P. 131–145.
4. Bartram D. The development of standards for the use of psychological tests in occupational settings: The competence approach // *The Psychologist*. – 1995. – Vol. 5. – P. 219–223.
5. Baayen R.H., Davidson D.J., Bates D.M. Mixed-effects modeling with crossed random effects for subjects and items // *Journal of Memory and Language*. – 2008. – Vol. 59, is. 4. – P. 390–412.
6. Беспалько В. Быть или не быть тестам в образовании? // *Педагогические измерения*. – 2012. – № 1. – С. 17–41.
7. Muniz J., Bartram D. Improving International Tests and Testing // *European Psychologist*. – 2007. – Vol. 12(3). – P. 206–219.
8. Bartram D. The development of international guidelines on test use: the International Test Commission Project // *International Journal of Testing*. – 2001. – Vol. 1. – P. 33–53.
9. Андриенко А.В. Современная практика использования тестирования в России и за рубежом // *Открытое и дистанционное образование*. – 2013. – № 2 (50). – С. 78–83.
10. Артюшова Е.К. Об инструментари педагогической диагностики в учебном процессе вуза // *Вестник Российского государственного университета им. И. Канта*. – Калининград: Изд-во РГУ им. И. Канта, 2007. – Вып. 4. – С. 27–34.
11. Кокшарова Е.А. Педагогическая экспертная система как средство оценки качества обучающих тестов // *Вестник Челябинского государственного педагогического университета*. – Челябинск, 2008. – № 12. – С. 79–86.

12. Аванесов В.С. Форма тестовых заданий. – М.: Центр тестирования, 2005. – 156 с.
13. Чельщикова М.Б. Теория и практика конструирования педагогических тестов. – М.: Логос, 2002.
14. Майоров А.Н. Теория и практика создания тестов для системы образования. – М.: Интеллект-центр, 2001. – 296 с.
15. Кузьмина Л.И., Осипов Ю.В. О тестировании студентов по «школьной» математике // Качество. Инновации. Образование. – 2014. – № 2. – С. 9–13.
16. Кузьмина Л.И., Осипов Ю.В. Новые технологии преподавания и «старые» дисциплины // Качество. Инновации. Образование. – 2013. – № 12. – С. 3–7.
17. Duncan K.A., MacEachern S.N. Nonparametric Bayesian modelling for item response // Statistical Modelling. – 2008. – Vol. 8, № 1. – P. 41–66.
18. Liang L., Browne M.W. A Quasi-Parametric Method for Fitting Flexible Item Response Functions // Journal of Educational and Behavioral Statistics. – 2015. – Vol. 40. – P. 5–34.
19. Fujimoto K.A., Karabatsos G. Dependent Dirichlet Process Rating Model Applied // Psychological Measurement. – 2014. – Vol. 38. – P. 217–228.
20. Захаров А.А. Количественные и структурные характеристики комплексных тестовых композиций, сформированных на основе бинома Ньютона // Вестник Саратовского государственного технического университета. – 2004. – № 1. – С. 19–24.
21. Кирьяков Б.С. Статистическая модель многократного тестирования учащихся // Вестник Рязанского государственного университета им. С.А. Есенина. – 2008. – № 1/18. – С. 3–23.
22. Карнаухов В.М. Статистическое моделирование интернет-экзамена, проводимого в рамках аттестации вуза // Открытое и дистанционное образование. – 2014. – № 1 (53). – С. 60–68.
23. Сафина Г.Л., Осипов Ю.В., Керимова Д.Х., Красовская И.А. Полуавтоматическая система тестирования по математике // Открытое и дистанционное образование. – 2015. – № 2 (58). – С. 56–62.
24. Ким В.С. Тестирование учебных достижений. – Уссурийск: УГПИ, 2007. – 214 с.
25. Ким В.С. Коррекция тестовых баллов на угадывание // Педагогические измерения. – 2006. – № 4. – С. 47–55.
26. Кроммер В.В. Еще раз о коррекции тестовых баллов // Педагогические измерения. – 2007. – № 1. – С. 89–94.
27. Кузьмина Л.И., Осипов Ю.В. Коррекция тестовых баллов с учетом отказов от угадывания // Alma-mater (Вестник высшей школы). – 2014. – № 12. – С. 85–91.
28. Gutman L. A basis for analyzing test-retest reliability // Psychometrika. – 1945. – Vol. 10. – P. 255–282.
29. Linacre J.M. Disconnected Subsets, Guttman Patterns and Data Connectivity // Rasch Measurement Transactions. – 2013. – Vol. 27:2. – P. 1415–1457.

Osipov Yu.V., Safina G.L., Vetuhnovskii F.Ya.  
National Research Moscow State University  
of Civil Engineering, Moscow, Russia  
**MODELING MATHEMATICS TESTS**

**Keywords:** test, open and closed forms of tasks, test score, task complexity, examinee's pattern.

The test on mathematics developed by an original technique is used for an assessment of level and

structure of knowledge of university students. The author's computer program generates individual versions of a test for each student. The teacher can offer test tasks with open or closed forms. In the course of testing one group was performing a test with tasks of the open form, and two other groups - with tasks of the closed form with a list of five response selection. Based on the analysis of test results of three groups of examinees the statistical characteristics of the test are studied.

Test scores of students have normal distribution, the test has a high differentiating ability, it can be used as a standard-based test, that makes it possible to separate students with different levels of knowledge. The reasons of differences in test results with open or closed form of test tasks are investigated. The relative simplicity of closed tasks in comparison with open tasks is connected with both possibilities of guessing the answer, and self-detection of error in the task when the response received does not match any answer in the list of options. The testing results enabled to construct the scales of task complexity for each group of examinees.

The hypothesis of maintaining order in the scale of task complexity for the tests in the open and closed forms and for students with different levels of knowledge was checked and confirmed. The patterns of student knowledge constructed due to the test results are studied. In the case of a dichotomous assessment of assignments (0 - the task is solved incorrectly, 1 - the task is solved correctly) the pattern is a set of ones and zeros, that student receives for performed test tasks. Investigation of patterns makes it possible to estimate the degree of ordering of students' knowledge and helps to identify the students, who has studied the material fragmentary or violated the test rules (tips, cheating, etc.). Each pattern is associated with regularity - a number which characterizes the degree of correctness of the pattern. It was found that in all groups the pattern of students' regularities takes values ranging from 0 to 0,4, that indicates the correct structure of the knowledge of the examinees. The regularity of distribution and its statistical characteristics are studied.

#### REFERENCES

1. Fisher W.P. The Central Theoretical Problem of the Social Sciences // Rasch Measurement Transactions. – 2014. – Vol. 28:2. – P. 1464–1466.
2. Kenneth D.R. Measuring Liberal / Conservative Voting Tendencies among U.S. Senators // Rasch Measurement Transactions. – 2012. – Vol. 26:2. – P. 1366–1367.

3. *Reeve B.B.* Item response theory modeling in heart outcomes measurement // *Expert Review of Pharmacoeconomics and Outcomes Research*. – 2003. – Vol. 3(2). – P. 131–145.
4. *Bartram D.* The development of standards for the use of psychological tests in occupational settings: The competence approach // *The Psychologist*. – 1995. – Vol. 5. – P. 219–223.
5. *Baayen R.H., Davidson D.J., Bates D.M.* Mixed-effects modeling with crossed random effects for subjects and items // *Journal of Memory and Language*. – 2008. – Vol. 59, is. 4. – P. 390–412.
6. *Bespal'ko V.* Byt' ili ne byt' testam v obrazovanii? // *Pedagogicheskie izmereniya*. – 2012. – № 1. – S. 17–41.
7. *Muniz J., Bartram D.* Improving International Tests and Testing // *European Psychologist*. – 2007. – Vol. 12(3). – P. 206–219.
8. *Bartram D.* The development of international guidelines on test use: the International Test Commission Project // *International Journal of Testing*. – 2001. – Vol. 1. – P. 33–53.
9. *Andrienko A.V.* Sovremennaja praktika ispol'zovaniya testirovaniya v Rossii i za rubezhom // *Otkrytoe i distancionnoe obrazovanie*. – 2013. – № 2 (50). – S. 78–83.
10. *Artishheva E.K.* Ob instrumentarii pedagogicheskoy diagnostiki v uchebnom processe vuza // *Vestnik Rossijskogo gosudarstvennogo universiteta im. I. Kanta*. – Kaliningrad: Izd-vo RGU im. I. Kanta, 2007. – Vyp. 4. – S. 27–34.
11. *Koksharova E.A.* Pedagogicheskaja jekspertnaja sistema kak sredstvo ocenki kachestva obuchajushhih testov // *Vestnik Cheljabinskogo gosudarstvennogo pedagogicheskogo universiteta*. – Cheljabinsk, 2008. – № 12. – S. 79–86.
12. *Avanesov V.S.* Forma testovyh zadaniy. – M.: Centr testirovaniya, 2005. – 156 s.
13. *Chelyshkova M.B.* Teorija i praktika konstruirovaniya pedagogicheskikh testov. – M.: Logos, 2002.
14. *Majorov A.N.* Teorija i praktika sozdaniya testov dlja sistemy obrazovaniya. – M.: Intellect-centr, 2001. – 296 s.
15. *Kuz'mina L.I., Osipov Ju.V.* O testirovanii studentov po «shkol'noj» matematike // *Kachestvo. Innovacii. Obrazovanie*. – 2014. – № 2. – S. 9–13.
16. *Kuz'mina L.I., Osipov Ju.V.* Novye tehnologii prepodavaniya i «starye» discipliny // *Kachestvo. Innovacii. Obrazovanie*. – 2013. – № 12. – S. 3–7.
17. *Duncan K.A., MacEachern S.N.* Nonparametric Bayesian modelling for item response // *Statistical Modelling*. – 2008. – Vol. 8, № 1. – P. 41–66.
18. *Liang L., Browne M.W.* A Quasi-Parametric Method for Fitting Flexible Item Response Functions // *Journal of Educational and Behavioral Statistics*. – 2015. – Vol. 40. – P. 5–34.
19. *Fujimoto K.A., Karabatsos G.* Dependent Dirichlet Process Rating Model Applied // *Psychological Measurement*. – 2014. – Vol. 38. – P. 217–228.
20. *Zaharov A.A.* Kolichestvennye i strukturnye harakteristiki kompleksnyh testovyh kompozicij, sformirovannyh na osnove binoma N'jutona // *Vestnik Saratovskogo gosudarstvennogo tehničeskogo universiteta*. – 2004. – № 1. – S. 19–24.
21. *Kir'jakov B.S.* Statisticheskaja model' mnogokratnogo testirovaniya uchashhihsja // *Vestnik Rjazanskogo gosudarstvennogo universiteta im. S.A. Esenina*. – 2008. – № 1/18. – S. 3–23.
22. *Karnauhov V.M.* Statisticheskoe modelirovanie internet-jekzameni, provodimogo v ramkah attestacii vuza // *Otkrytoe i distancionnoe obrazovanie*. – 2014. – № 1 (53). – S. 60–68.
23. *Safina G.L., Osipov Ju.V., Kerimova D.H., Krasovskaja I.A.* Poluavtomaticheskaja sistema testirovaniya po matematike // *Otkrytoe i distancionnoe obrazovanie*. – 2015. – № 2 (58). – S. 56–62.
24. *Kim V.S.* Testirovanie uchebnyh dostizhenij. – Ussurijsk: UGPI, 2007. – 214 s.
25. *Kim V.S.* Korrekciya testovyh ballov na ugadyvanie // *Pedagogicheskie izmereniya*. – 2006. – № 4. – S. 47–55.
26. *Kromer V.V.* Eshhe raz o korrekcii testovyh ballov // *Pedagogicheskie izmereniya*. – 2007. – № 1. – S. 89–94.
27. *Kuz'mina L.I., Osipov Ju.V.* Korrekciya testovyh ballov s uchetom otkazov ot ugadyvaniya // *Alma-mater (Vestnik vysshej shkoly)*. – 2014. – № 12. – S. 85–91.
28. *Gutman L.* A basis for analyzing test-retest reliability // *Psychometrika*. – 1945. – Vol. 10. – P. 255–282.
29. *Linacre J.M.* Disconnected Subsets, Guttman Patterns and Data Connectivity // *Rasch Measurement Transactions*. – 2013. – Vol. 27:2. – P. 1415–1457.