

АНАЛИЗ НАДЕЖНОСТИ ПЕДАГОГИЧЕСКОГО ТЕСТА КУРСА МАТЕМАТИКИ ПО ТЕМЕ «РЯДЫ»

*Л.В. Лиманова*¹

Самарский государственный технический университет
443100, г. Самара, ул. Молодогвардейская, 244
E-mail: llv-1@mail.ru

Данная работа посвящена анализу качества педагогического теста по теме «Ряды» курса математики, который применяется для проверки уровня знаний студентов на первом курсе СамГТУ. Для этого проводились исследования трех характеристик теста: степень трудности каждого задания, вариация тестовых заданий и надежность. Коэффициент надежности вычислялся по трем формулам: формуле KR-20, среднему коэффициенту корреляции всех заданий между собой и формуле Спирмена – Брауна. При анализе полученных результатов оказалось, что значения коэффициентов надежности недостаточно высоки. Поэтому для повышения качества теста рекомендуется либо изменить задания, негативно влияющие на коэффициент надежности, либо увеличить количество заданий в тесте (как показали исследования – на два задания).

Ключевые слова: педагогический тест, коэффициент корреляции, коэффициент надежности, вариация тестовых заданий, степень трудности каждого задания.

Для текущего и рубежного контроля знаний студентов на кафедре «Высшая математика и прикладная информатика» Самарского государственного технического университета были составлены тесты по всем разделам высшей математики в соответствии с рабочей программой для бакалавров всех специальностей университета. Некоторые из них уже проверялись на качество и были исправлены согласно рекомендациям [1–6]. Данная статья посвящена проверке качества педагогического теста по теме «Ряды» (табл. 1). Этот тест содержит 9 заданий, решив которые, студент должен выбрать правильный ответ из пяти предложенных.

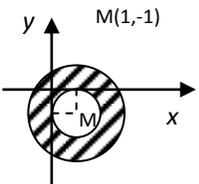
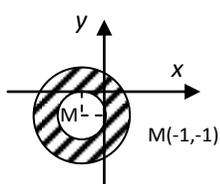
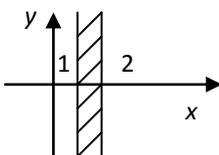
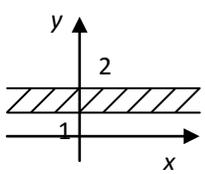
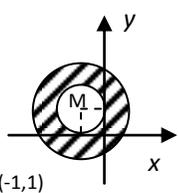
К основным характеристикам, оценивающим качество педагогического теста, относятся объективность, надежность и валидность [7–12]. Поскольку проверка качества теста является объемным процессом, то в данной статье

¹ Лариса Владимировна Лиманова, кандидат технических наук, доцент кафедры высшей математики и прикладной информатики.

мы рассмотрим только его часть и проведем анализ надежности педагогического теста.

Таблица 1

Тест по теме «Ряды»

№	Задания
1	Представить в показательной форме число $1 - \sqrt{3}i$. Ответы: 1) $2e^{-\frac{\pi}{6}}$; 2) $2e^{\frac{\pi}{3}}$; 3) $\sqrt{3}e^{-\frac{\pi}{6}}$; 4) $2e^{\frac{\pi}{3}}$; 5) $4e^{-\frac{\pi}{3}}$
2	Вычислить $\frac{1-5i}{5+i}$. Ответы: 1) $-2+i$; 2) $1+i$; 3) $4i$; 4) $5-i$; 5) $-$
3	Указать множество точек на комплексной плоскости, удовлетворяющее условию $1 \leq z-i \leq 2$. Ответы: 1)  2)  3)  4)  5)  R1=1 R2=2
4	$\operatorname{Ln}(-) = \dots$ Ответы: 1) $\ln 2 + k\pi i$; 2) $i(\pi + k\pi)$; 3) $i \ln 2$; 4) $\ln 2 + i(\pi + k\pi)$; 5) $\ln 2 + i\left(-\frac{\pi}{2} + 2k\pi\right)$.
5	Решить уравнение $z^4 + z + i = 1$ и найти возможные значения $\arg z$. Ответы: 1) $-\frac{3\pi}{16} + \frac{k\pi}{2}, k = \overline{1,3}$; 2) $\frac{3\pi}{16} + \frac{k\pi}{2}, k = \overline{1,3}$; 3) $\frac{3\pi}{8} + \frac{k\pi}{2}, k = \overline{1,3}$; 4) $-\frac{3\pi}{8} + \frac{k\pi}{2}, k = \overline{1,3}$; 5) $-\frac{3\pi}{4} + \frac{k\pi}{2}, k = \overline{1,3}$.
6	Для аналитической функции $f(z) = u(x,y) + v(x,y)$ найти $v(x,y)$, если известно, что $u(x,y) = x^2 - y^2, f(0)=0$. Ответы: 1) $x^2 + y^2$; 2) $-x^2 - y^2$; 3) $x - y$; 4) $y - x$; 5) x .
7	Вычислить интеграл $\int_{AB} (xi - z) dz$, где AB – отрезок прямой между точками $z_1 = 1+i, z_2 = -1-i$. Ответы: 1) $1+i$; 2) $-1-i$; 3) $-2+i$; 4) $2-i$; 5) $1-i$.

№	Задания
8	<p>Вычислить $\int_{\Gamma} \frac{z-2}{z(z-3)} dz$, $\Gamma: z+1 =r$.</p> <p>Ответы: 1) $\frac{4\pi}{3}$; 2) $-\frac{4\pi}{9}$; 3) $\frac{4\pi}{9}$; 4) 0; 5) $-\frac{4\pi}{3}$.</p>
9	<p>Вычислить $\int_{\Gamma} \frac{z-2}{z(z-3)} dz$, $\Gamma: z-1 =r$.</p> <p>Ответы: 1) $\frac{4\pi}{9}$; 2) 0; 3) $-\frac{4\pi}{9}$; 4) $\frac{4\pi}{3}$; 5) $-\frac{4\pi}{3}$.</p>

При исследовании качества теста используются методы математической статистики. Для обработки были выбраны работы студентов нефтетехнологического факультета. Объем выборки составил 235 работ, из которых были исключены 33 работы. Авторы этих работ либо решили все задания теста, либо не справились с заданиями и не решили ни одной задачи. Такие работы исключаются из выборки, так как никакой информации о качестве теста не несут.

На первом этапе анализа оценим степень трудности каждого задания p_j . Эта величина находится по формуле

$$p_j = \frac{m_j}{n},$$

где m_j – количество правильных ответов на j -задание; n – общее количество студентов. Графически эта характеристика изображена на рис. 1. Наиболее трудным для решения оказалось девятое задание теста, с ним справилось меньше всего студентов (26 %). А наиболее легким – второе задание теста, но однако не все студенты справились и с ним (лишь 70 %).

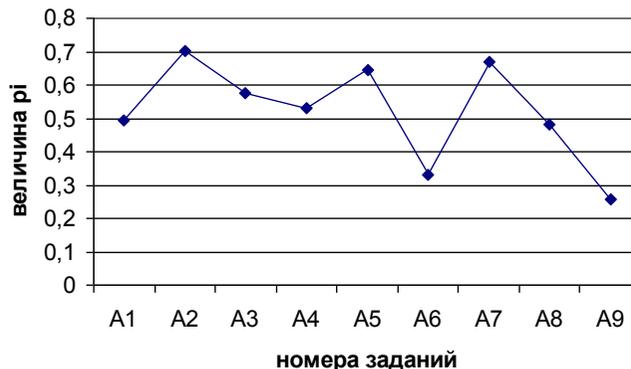


Рис. 1. Степень трудности задания

Данный тест относится к нормативно-ориентированным тестам. Для этого вида тестов при исследовании качества важной является следующая характеристика – вариация тестовых заданий $p_j \cdot q_j$, где ($q_j = 1 - r_j$). Отсортируем номера заданий в порядке убывания количества правильных ответов и построим ее график (рис. 2). Эксперты считают, что ее величина должна находиться около 0,25. Менее всего удовлетворяют этим требованиям опять второе и девятое задания. Поэтому рекомендуется их изменить.

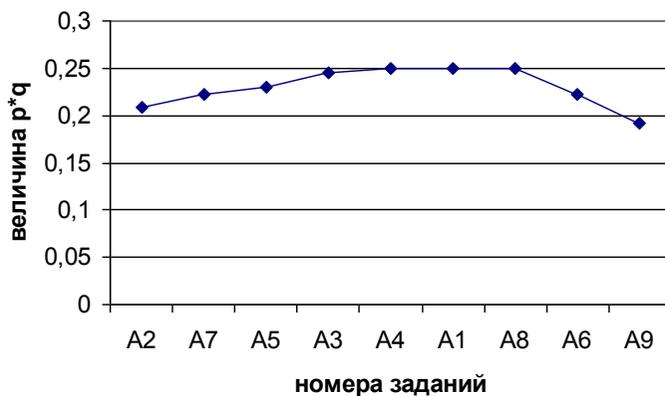


Рис. 2. Вариация тестовых заданий

Для исследования надежности теста применяется коэффициент надежности r_t [7 – 10], который вычисляется с помощью формулы KR-20, среднего коэффициента корреляции \bar{R} и формулы Спирмена – Брауна.

Применим формулу KR-20 [8, 9]:

$$r_t = \frac{M}{M-1} \left(1 - \frac{\sum_{j=1}^M p_j q_j}{s_x^2} \right),$$

где M – количество заданий; s_x^2 – исправленная дисперсия индивидуальных баллов студентов: $s_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$; x_i – индивидуальный балл испытуемого; $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ – средний балл всех студентов; n – общее количество студентов.

Коэффициент надежности, вычисленный с помощью этой формулы, равен $r_t = 0,65$.

Формула для нахождения коэффициента надежности с помощью среднего коэффициента корреляции \bar{R} всех заданий между собой имеет вид

$$r_i = \frac{M\bar{R}}{1 + (M-1)\bar{R}},$$

где $\bar{R} = \frac{1}{M} \sum_{j=1}^M r_{xy}^j$ – средний коэффициент корреляции всех заданий между собой; r_{xy}^j – среднее значение коэффициентов корреляции для j -го задания.

Поэтому сначала надо найти средние значения коэффициентов корреляции для каждого задания r_{xy}^j . Их значения приведены в табл. 2. Считается [8], что они должны быть меньше 0,3. Как видно из таблицы, вычисленные коэффициенты удовлетворяют требованиям экспертов.

Таблица 2

Средние значения коэффициентов корреляции

	A2	A7	A5	A3	A4	A1	A8	A6	A9	A2
r_{xy}^j	0,24	0,29	0,29	0,3	0,28	0,16	0,3	0,22	0,26	0,24

Далее вычисляем средний коэффициент корреляции $\bar{R} = 0,26$, а затем и коэффициент надежности $r_i = 0,76$.

Формула Спирмена – Брауна [8, 11] имеет вид

$$r_i = \frac{2r_{1/2}}{1 + \frac{1}{2}},$$

где $r_{1/2}$ – коэффициент корреляции между группами, на которые необходимо разделить тест по четным и нечетным заданиям. Он вычисляется следующим образом:

$$r_{1/2} = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{\sqrt{\left(n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right) \left(n \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i \right)^2 \right)}}$$

где x_i и y_i – индивидуальные баллы i -го испытуемого в четных и нечетных заданиях соответственно.

Прделав необходимые вычисления, получаем $r_{1/2} = 0,5$. Тогда коэффициент надежности $r_i = 0,67$.

Итак, мы нашли значения коэффициента надежности с помощью трех формул и получили результат: 0,65, 0,76 и 0,67. Эксперты считают, что его величина должна быть больше 0,7. Только один из коэффициентов удовлетворяет этим требованиям. Для повышения надежности теста эксперты [7, 8, 12] предлагают либо изменить задания, которые снижают коэффициент надежности (а в нашем случае это задания № 2 и № 9), либо увеличить количество заданий в тесте. Посмотрим, на сколько заданий надо увеличить тест. Для этого применяется формула [1]

$$r'_i = \frac{kr_i}{1 + k - r_i},$$

где r_i – коэффициент надежности до изменения длины теста; r'_i – коэффициент надежности после изменения; k – кратность изменения.

В качестве r'_i возьмем требуемую величину коэффициента надежности – 0,7, а величине r_i зададим наименьшее из полученных значений коэффициента надежности $r_i = 0,65$. Проведя необходимые вычисления, получим $k = 1,26$. Значит, необходимо увеличить тест на две задачи.

Таким образом, в ходе исследований вычислялись такие характеристики, как степень трудности каждого задания, вариация тестовых заданий и коэффициент надежности, который определялся с помощью трех формул. Оценка первых двух характеристик показала необходимость изменения второго и девятого заданий теста. Они также влияют на величину коэффициента надежности (два результата не удовлетворяют требованиям экспертов). Поэтому для повышения качества теста рекомендуется либо изменить эти задания (второе – усложнить, а девятое – упростить), либо увеличить количество заданий в тесте.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Лиманова Л.В., Муратова Л.А. Статистический анализ качества теста из курса высшей математики по теме «Пределы. Производные» // Вестник Самарского государственного технического университета. Сер. Психолого-педагогические науки. – 2015. – № 1(25). – С. 143–151.
2. Лиманова Л.В., Муратова Л.А. Анализ качества теста из курса высшей математики по теме «Линейная алгебра, аналитическая геометрия» // Вестник Самарского государственного технического университета. Сер. Психолого-педагогические науки. – 2015. – № 2(26). – С. 113–122.
3. Муратова Л.А. Валидность и дискриминативность при исследовании и оценке качества теста «Интегральное исчисление» // Научный альманах. – 2016. – № 6-1(19). – С. 323–326.
4. Лиманова Л.В. Повышение надежности педагогического теста «Линейная алгебра, аналитическая геометрия» // Вестник Самарского государственного тех-

- нического университета. Сер. Психолого-педагогические науки. – 2016. – № 2(30). – С. 75–81.
5. *Лиманова Л.В.* Улучшение качества педагогического теста из курса высшей математики по теме «Интегральное исчисление» // Вестник Самарского государственного технического университета. Сер. Психолого-педагогические науки. – 2016. – № 3(31). – С. 54–60.
 6. *Лиманова Л.В.* Проверка качества педагогического теста по высшей математике по теме «Теория функции комплексной переменной» // Вестник Самарского государственного технического университета. Сер. Психолого-педагогические науки. – 2017. – № 2(34). – С. 56–65.
 7. *Звонников В.И., Челышкова М.Б.* Современные средства оценивания результатов обучения. – М.: Академия, 2007. – 224 с.
 8. *Ким В.С.* Тестирование учебных достижений. – Уссурийск: Изд-во УГПИ, 2007. – 214 с.
 9. *Карпенко А.П., Домников А.С., Белоус В.В.* Тестовый метод контроля качества обучения и критерии качества образовательных тестов // Наука и образование: электронное научно-техническое издание. Вып. 04/2011. – 28 с.
 10. *Олейник Н.М.* Тест как инструмент измерения уровня знаний и трудности заданий в современной технологии обучения: Учеб. пособие. – Донецк: ДонГУ, 1991. – 168 с.
 11. *Ащепкова Л.Я.* Материалы к семинару по обработке результатов тестирования // Региональный центр проблем качества при ДВГУ. – Владивосток, 2001.
 12. Психологическая диагностика: Учеб. пособие / Под ред. К.М. Гуревича, Е.М. Борисовой. – М.: Изд-во УРАО, 1997. – 304 с.

Поступила в редакцию 11.10.17
В окончательном варианте 16.11.17

UDC 378.14

THE ANALYSIS OF RELIABILITY OF PEDAGOGICAL TEST OF A COURSE OF MATHEMATICS ON THE SUBJECT "RANKS"

*L.V. Limanova*¹

Samara State Technical University
244, Molodogvardeyskaya St., Samara, 443100
E-mail: llv-1@mail.ru

This work is devoted to the analysis of quality of pedagogical dough on the subject "Ranks" of a course of mathematics which is applied to check of level of knowledge of students at the first year of SAMGTU. Researches of three characteristics of dough were for this purpose conducted: degree of difficulty of each task, variation of test

¹ *Ludmila V. Limanova* – Cand. Tech. Sci., Associate Professor of Higher Mathematics and Applied Informatics Department.

tasks and reliability. The coefficient of reliability was calculated on three formulas: by means of a formula KR-20, average coefficient of correlation of all tasks among themselves and on Spirmena-Brown's formula. In the analysis of the turned-out results it appeared that values of coefficients of reliability aren't rather high. Therefore for improvement of quality of dough it is recommended or to change the tasks influencing decrease of coefficient of reliability or to increase quantity of tasks in the test (as showed researches – on two tasks).

Keywords: pedagogical test, correlation coefficient, reliability coefficient, variation of test tasks, degree of difficulty of each task.

REFERENCES

1. *Limanova L.V., Muratova L.A.* Statisticheskii analiz kachestva testa iz kursa vysshei matematiki po teme «Predely. Proizvodnyye» [Statisticheskyy the analysis of quality of dough from a course of the higher mathematics on a subject "Limits. Derivatives"] Vestnik SamGTU, Seriya «Psihologo-pedagogicheskie nauki» [The Messenger SAMGTU, the Psychology and Pedagogical Sciences Series]. No. 1(25)–2015, pp. 143–151.
2. *Limanova L.V., Muratova L.A.* Analiz kachestva testa iz kursa vysshei matematiki po teme «Lineinaya algebra, analiticheskaya geometriya» [The analysis of quality of dough from a course of the higher mathematics on the subject "Linear Algebra, Analytical Geometry"] Vestnik SamGTU, Seriya «Psihologo-pedagogicheskie nauki» [The Messenger SAMGTU, the Psychology and Pedagogical Sciences Series]. No. 2(26)–2015, pp. 113–122.
3. *Muratova L.A.* Validnost' i diskriminativnost' pri issledovanii i ocnke kachestva testa «Integral'noe ischislenie» [Validnost and a diskriminativnost at research and an assessment of quality of the Integral calculus test] Nauchnyi al'manah [The Scientific almanac]. 2016. No. 6–1(19). P. 323–326.
4. *Limanova L.V.* Povyshenie nadezhnosti pedagogicheskogo testa «Lineinaya algebra, analiticheskaya geometriya» [Increase of reliability of the pedagogical test "Linear Algebra, Analytical Geometry"] Vestnik SamGTU, Seriya «Psihologo-pedagogicheskie nauki» [The Messenger SAMGTU, the Psychology and Pedagogical Sciences Series]. No. 2(30)–2016, pp. 75–81.
5. *Limanova L.V.* Uluchshenie kachestva pedagogicheskogo testa iz kursa vysshei matematiki po teme «Integral'noe ischislenie» [Improvement of quality of pedagogical dough from a course of the higher mathematics on the subject "Integral calculus"] Vestnik SamGTU, Seriya «Psihologo-pedagogicheskie nauki» [The Messenger SAMGTU, the Psychology and Pedagogical Sciences Series]. No. 3(31)–2016, pp. 54–60.
6. *Limanova L.V.* Proverka kachestva pedagogicheskogo testa po vysshei matematike po teme «Teoriya funktsii kompleksnoi peremennoi» [Quality check of pedagogical dough on the higher mathematics on a subject "Theory of function of a complex variable"] Vestnik SamGTU, Seriya «Psihologo-pedagogicheskie nauki» [The Messenger SAMGTU, the Psychology and Pedagogical Sciences Series]. No. 2(34)–2017, pp. 56–65.

7. *Zvonnikov V. I., Chelyshkova M. B.* Sovremennye sredstva ocenivaniya rezul'tatov obucheniya [Modern means of estimation of results of training]. Moscow: Academy, 2007. 224 p.
8. *Kim V. S.* Testirovanie uchebnyh dostizhenii [Testing of educational achievements]. Ussuriisk: Izd-vo UGPI, 2007. 214 p.
9. *Karpenko A.P., Domnikov A.S., Belous V. V.* Testovyi metod kontrolya kachestva obucheniya i kriterii kachestva obrazovatel'nyh testov [Test method of quality control of training and criteria of quality of educational tests] Nauka i obrazovanie: elektronnoe nauchno-tehnicheskoe izdanie. No. 04/2011, 28 p.
10. *Oleynik N. M.* Test kak instrument izmereniya urovnya znaniy i trudnosti zadaniy v sovremennoi tehnologii obucheniya [The test as the instrument of measurement of level of knowledge and difficulty of tasks in modern technology of training]. Donetsk, DonGU, 1991. 168 p.
11. *Ashchepkova L.Ya.* Materialy k seminaru po obrabotke rezul'tatov testirovaniya [Materials to a seminar on processing of results of testing] Regional'nyi centr problem kachestva pri DVGU. Vladivostok, 2001.
12. Psihologicheskaya diagnostika [Psychological diagnostics] [Under the editorship of K.M. Gurevich and E.M. Borisova]. M.: Izd-vo URAO, 1997. 304 p.

Original article submitted 11.10.17

Revision submitted 16.11.17