

КАЧЕСТВЕННЫЕ ИЗМЕНЕНИЯ ПЕДАГОГИЧЕСКОГО ТЕСТА «ЛИНЕЙНАЯ АЛГЕБРА, АНАЛИТИЧЕСКАЯ ГЕОМЕТРИЯ»

Л.А. Муратова¹

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

Согласно классической теории оценены характеристики валидности и дискриминативности педагогического теста «Линейная алгебра, аналитическая геометрия» курса высшей математики Самарского государственного технического университета. Результаты исследований показали, что все коэффициенты корреляции заданий друг с другом положительны и не превосходят величины 0,3, что считается хорошим показателем. Изменились значения точечного бисериального коэффициента корреляции заданий с тестовым баллом студента. Это повлияло на такие характеристики теста, как валидность и дискриминативность. И, наконец, с помощью критерия Пирсона получен положительный результат о нормальном распределении тестовых баллов.

Ключевые слова: педагогический тест, коэффициент корреляции, коэффициент валидности, коэффициент дискриминативности.

Проверка знаний студентов – это важная и ответственная часть учебного процесса. В последнее время проверка знаний учащихся осуществляется с помощью тестовых методов. Результаты оценивания должны быть объективными, поэтому педагогический тест должен отвечать определенным требованиям. В номере 2 (26) – 2015 научного журнала «Вестник Самарского государственного технического университета» (СамГТУ) серии «Психолого-педагогические науки» была представлена статья, посвященная анализу качества педагогического теста по теме «Линейная алгебра, аналитическая геометрия» из курса высшей математики [1]. В ходе исследований согласно классической теории тестов [2–14] для повышения качества теста было предложено исключить из него или заменить задания 1, 4 и 12, поскольку они либо имели отрицательную корреляцию со многими другими заданиями, либо слишком низкую корреляцию с тестовым баллом студента. Авторы изменили первое задание (табл. 1), исключили задание № 12 (теоретический вопрос), а задание № 4 (умножение матриц) было решено оставить без изменения. Новый тест теперь состоит из 19 задач (нумерация осталась прежней, чтобы

¹ Лидия Анатольевна Муратова, кандидат технических наук, доцент кафедры высшей математики и прикладной информатики.

не было путаницы). Для чистоты эксперимента тестирование проводилось на факультете автоматики и информационных технологий СамГТУ (как и в прошлый раз). В тестировании участвовало 136 студентов. Четыре студента справились со всеми заданиями и никакой информации для анализа основных характеристик качества теста не давали, поэтому их тесты были исключены из дальнейшего рассмотрения.

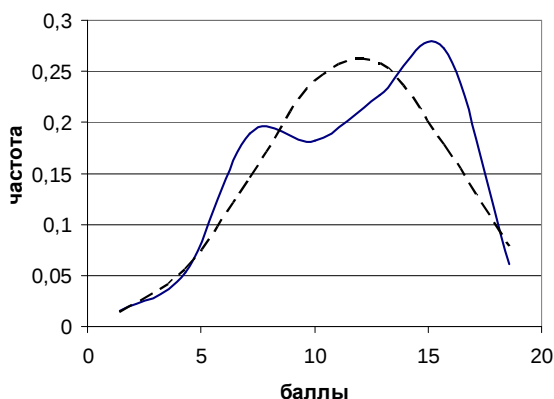
Таблица 1

Задание № 1, представленное в тестах 2014 и 2015 гг.

2014 г. (старый тест)		2015 г. (новый тест)
Вычислить:	$\begin{vmatrix} 2 & -1 & 0 \\ 3 & 0 & -2 \\ 0 & 2 & 1 \end{vmatrix}$	Для системы уравнений $\begin{cases} x - 2z = 1, \\ y - z = 4, \\ 2x - y = 5 \end{cases}$ найти Δ, Δ_x, x .

Многие эксперты в области исследования качества тестов говорят о том, что нормативно-ориентированный тест должен хорошо дифференцировать испытуемых [8, 9, 14], то есть индивидуальные тестовые баллы должны в достаточной степени отличаться друг от друга, а их распределение должно быть близко к нормальному. О дифференцирующей способности теста позволяет судить такая величина, как дисперсия тестовых баллов. М.Б. Чельшкова [14] считает, что если среднее арифметическое индивидуальных баллов $\bar{x}$ примерно равно утроенному стандартному отклонению $3s_x$, то можно считать дисперсию оптимальной, а распределение тестовых баллов близким к нормальному. В рассматриваемом тесте получены именно такие значения: $\bar{x} = 11,9$, $3s_x = 12,6$.

На рисунке изображен график относительных частот тестовых баллов (сплошная линия), полученных по выборке, и кривая нормального распределения (пунктирная линия).



Относительные частоты

Гипотеза о нормальном распределении генеральной совокупности проверялась с помощью критерия Пирсона для уровня значимости 0,01. Полученное значение оказалось меньше критического. Значит, данные наблюдений согласуются с нормальным распределением генеральной совокупности.

Для анализа валидности теста были вычислены значения коэффициентов корреляции φ_{mk} заданий друг с другом. По материалам экспертов [7, 9, 12], для качественного теста их величины должны лежать в пределах [0; 0,3]. В новом тесте нет отрицательных значений φ_{mk} и все $\varphi_{mk} \leq 0,3$.

По величине средних значений коэффициента корреляции r_{xy}^j (табл. 2) для каждого задания также можно судить о качестве теста. Как видим, все значения $r_{xy}^j \leq 0,3$, следовательно, удовлетворяют требованиям экспертов.

Таблица 2

**Коэффициенты корреляции и индексы дискриминативности
по результатам тестирования, проведенного в 2015 г.**

	A4	A1	A3	A7	A2	A14	A10	A13	A8	A5	A9	A6	A17	A19	A18	A11	A15	A20	A16
r_{xy}^j	0,2 3	0,1 5	0,1 8	0,2 1	0,2 4	0,2 7	0,2 5	0,1 9	0,2 6	0,2 4	0,1 9	0,1 3	0,2 8	0,2 7	0,3 0	0,1 6	0,3 0	0,1 7	0,2 8
r_{pb}^j	0,4 5	0,2 9	0,3 6	0,4 4	0,5 5	0,5 5	0,5 2	0,3 9	0,5 4	0,5 1	0,4 2	0,2 9	0,6 0	0,5 8	0,5 9	0,3 5	0,6 4	0,3 7	0,5 9
$r_{диск}^j$	0,4 2	0,2 2	0,3 1	0,5 0	0,5 8	0,5 0	0,4 4	0,6 4	0,5 6	0,5 3	0,6 4	0,3 6	0,7 5	0,7 2	0,7 5	0,3 9	0,8 1	0,5 3	0,6 9

Как и в предыдущей статье [1], в качестве коэффициента валидности использовался точечный бисериальный коэффициент корреляции r_{pb}^j – коэффициент корреляции каждого задания с тестовым баллом студента (индивидуальным баллом). Для вычислений этих коэффициентов применялась формула [3, 9]

$$r_{pb}^j = \frac{\bar{X}_1 - \bar{X}_0}{s_x} \sqrt{\frac{n_1 \cdot n_0}{n(n-1)}},$$

где n_1 – число студентов, выполнивших данное задание;

n_0 – число студентов, не выполнивших его;

$n = n_1 + n_0$ – общее количество студентов;

$\bar{X}_1$ – средний индивидуальный балл студентов, справившихся с данным заданием (отношение суммы индивидуальных баллов студентов, справившихся с данным заданием, к n_1);

$\bar{X}_0$ – средний индивидуальный балл студентов, не справившихся с данным заданием (отношение суммы индивидуальных баллов студентов, не справившихся с данным заданием, к n_0);

s_x – стандартное отклонение для индивидуальных баллов всех студентов.

Для высококачественного теста значения точечных бисериальных коэффициентов корреляции должны быть не менее 0,5 [9]. В новом тесте их величины (табл. 2) лежат в средних пределах от 0,3 и выше.

Общий коэффициент валидности теста вычисляется по формуле

$$r_{pb} = \frac{\sum_{j=1}^M r_{pb}^j}{M}.$$

Оценить его значение помогают интервалы [13]: 0,2–0,3 – низкое; 0,3–0,5 – среднее; свыше 0,5 – высокое.

Для нового теста величина этого коэффициента составляет 0,475 (в прошлый раз $r_{pb} = 0,434$), то есть в обоих случаях общий коэффициент валидности теста имеет среднее значение, но его величина возросла.

Также в прошлой статье анализировалось значение еще одного коэффициента – дискриминативности, т. е. способности задания дифференцировать студентов на лучших и худших [7, 8, 12, 14]. Для определения $r_{диск}^j$ использовалась формула

$$r_{диск}^j = p_1^j - p_0^j,$$

где p_1^j – отношение количества правильных ответов на j -е задание к 27 % «лучших» студентов по результатам выполнения теста;

p_0^j – отношение количества правильных ответов на j -е задание к 27 % «худших» студентов по результатам выполнения теста.

Значения индекса дискриминативности $r_{диск}^j$ [7] лежат в интервале [-1; 1]. Из теста рекомендуется исключать те задания, у которых величина $r_{диск}^j \leq 0$. В данном тесте индекс дискриминативности для всех заданий положителен (четвертая строка табл. 2), значит, удовлетворяет критериям экспертов.

В качестве индекса дискриминативности можно рассматривать точечный бисериальный коэффициент корреляции r_{pb}^j [14]. В экспертной литературе говорится, что необходимо удалить из теста те задания, для которых $r_{pb}^j < 0,2$. Теперь в нашем тесте таких заданий нет (в прошлом тесте это были задания с номерами 1 и 12), все $r_{pb}^j > 0,2$.

Таким образом, после рекомендованной доработки качество теста по теме «Линейная алгебра, аналитическая геометрия» повысилось и его можно использовать в качестве инструмента при оценке знаний студентов при текущем контроле.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *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.
2. *Аванесов В.С.* Десятилетний юбилей журнала «Педагогические измерения» // Педагогические измерения. – 2014. – № 3. – С. 3–43.
3. *Ащепкова Л.Я.* Материалы к семинару по обработке результатов тестирования // Региональный центр проблем качества при ДВГУ, Владивосток. – 2001.
4. *Гласс Дж., Стенли Дж.* Статистические методы в педагогике и психологии. – М.: Прогресс, 1976. – 496 с.
5. *Ефремова Н.Ф., Звонников В.И., Чельшкова М.Б.* Педагогические измерения в системе образования // Педагогика. – 2006. – № 2. – С. 14–22.
6. *Жилина Е.В.* Анализ применяемых моделей и методов тестирования для оценки знаний специалиста // Zprávy vědecké ideje – 2011: materiály VII mezinárodní vědecko-praktická konference. 27 října – 05 listopadu 2011 roku. Díl 4. Ekonomické vědy. Praha: Publishing House «Education and Science» s.r.o. 2011. – С. 53–62.
7. *Звонников В.И., Чельшкова М.Б.* Современные средства оценивания результатов обучения. – М.: Академия, 2007. – 224 с.
8. *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. – 2011. – № 04. – 2011. – 28 с.
9. *Ким В.С.* Тестирование учебных достижений. – Уссурийск: Изд-во УГПИ, 2007. – 214 с.
10. *Крокер Л., Алгина Дж.* Введение в классическую и современную теорию тестов / Под общ. ред. В.И. Звонникова, М.Б. Чельшковой. – М.: Логос, 2010. – 668 с.
11. *Морев И.А.* Образовательные информационные технологии. Ч. 2 // Педагогические измерения. – Владивосток: Изд-во Дальневост. ун-та, 2004. – 174 с.
12. *Олейник Н.М.* Тест как инструмент измерения уровня знаний и трудности заданий в современной технологии обучения. – Донецк: ДонГУ, 1991. – 168 с.
13. *Психологическая диагностика* / Под ред. К.М. Гуревича, Е.М. Борисовой. – М.: Изд-во УРАО, 1997. – 304 с.
14. *Чельшкова М.Б.* Теория и практика конструирования педагогических тестов. – М.: Логос, 2002. – 432 с.
15. *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.

Поступила в редакцию 13.05.16;
в окончательном варианте 30.05.16

HIGH-QUALITY CHANGES OF THE PEDAGOGICAL TEST "LINEAR ALGEBRA, ANALYTICAL GEOMETRY"

L.A. Muratova¹

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

In this article according to the classical theory characteristics of validity and discriminative of the pedagogical test "Linear Algebra, Analytical Geometry" of a course of the higher mathematics of Samara State Technical University were estimated. Results of researches showed that all coefficients of correlation of tasks are with each other positive and don't surpass the size 0,3 that is considered a good indicator. Values of dot biserial coefficient of correlation of tasks with test point of the student changed. It influenced such characteristics of dough as a validity and a diskriminativnost. And, at last, by means of Pearson's criterion the positive result about normal distribution of test points is received.

Keywords: pedagogical test, correlation coefficient, validity coefficient, diskriminativnost coefficient.

REFERENCES

1. 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.
2. Avanesov V.S. Desyatiletanii yubilei zhurnala «Pedagogicheskie izmereniya» [Ten-year anniversary of the Pedagogical Measurements magazine]. Pedagogicheskie izmereniya. No. 3. – pp. 3–43.
3. 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.
4. Glass J., Stenli J. Statisticheskie metody v pedagogike i psihologii [Statistical methods in pedagogics and psychology]. M.: Izd-vo Progress [Moscow: Publishing house Progress]. – 496 p.

¹ Lidiya A. Muratova, Cand. Tech. Sci., Associate professor of Advanced Mathematics and Applied Information Science Department.

5. *Efremova N.F., Zvonnikov V.I., Chelyshkova M.B.* Pedagogicheskie izmereniya v sisteme obrazovaniya [Pedagogical measurements in an education system] *Pedagogika* [Pedagogics]. No. 2. – pp. 14–22.
6. *Zilina E.V.* Analiz primenyaemykh modelei i metodov testirovaniya dlya ocenki znanii specialista [The analysis of the applied models and methods of testing for an assessment of knowledge of the expert]. Zprávy vědecké ideje: materiály VII mezinárodní vědecko-praktická konference. 27 října – 05 listopadu 2011 roku. Díl 4. Ekonomické vědy. Praha: Publishing House "Education and Science" s.r.o. – pp. 53–62.
7. *Zvonnikov V.I., Chelyshkova M.B.* Sovremennye sredstva ocenivaniya rezul'tatov obucheniya [Modern means of estimation of results of training]. Moscow: Academy. – 224 p.
8. *Karpenko A.P., Domnikov A.S., Belous V.V.* Testovyi metod kontrolya kachestva obucheniya i kriterii kachestva obrazovatel'nykh 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.
9. *Kim V.S.* Testirovanie uchebnykh dostizhenii [Testing of educational achievements]. Ussuriisk: Izd-vo UGPI. – 214 p.
10. *Crocker L., Algina J.* Vvedenie v klassicheskuyu i sovremennuyu teoriyu testov [Introduction to the classical and modern theory of tests] [Under the general editorship of V. I. Zvonnikov, M. B. Chelyshkova]. Moscow: Logos. – 668 p.
11. *Morev I.A.* Obrazovatel'nye informacionnye tehnologii [Educational information technologies. Part 2]. Pedagogicheskie izmereniya [Pedagogical measurements]. Vladivostok: Publishing house of Dalnevost. un-ta. – 174 p.
12. *Oleynik N.M.* Test kak instrument izmereniya urovnya znanii 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. – 168 p.
13. Psihologicheskaya diagnostika [Psychological diagnostics] [Under the editorship of K.M. Gurevich and E.M. Borisova]. M.: Izd-vo URAO. – 304 p.
14. *Chelyshkova M. B.* Teoriya i praktika konstruirovaniya pedagogicheskikh testov [Theory and practice of designing of pedagogical tests]. Moscow: Logos. – 432 p.
15. *Limanova L.V., Muratova L.A.* Statisticheskii analiz kachestva testa iz kursa vysshei matematiki po teme «Predely. Proizvodnye» [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.

Original article submitted 13.05.16;
revision submitted 30.05.16