

ПОСТРОЕНИЕ ОДНОПАРАМЕТРИЧЕСКОЙ МОДЕЛИ РАША ДЛЯ ТЕСТА «ЛИНЕЙНАЯ АЛГЕБРА, АНАЛИТИЧЕСКАЯ ГЕОМЕТРИЯ»

*Л.А. Муратова*¹

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

Для теста «Линейная алгебра, аналитическая геометрия» в соответствии с современной теорией тестирования (IRT) строится однопараметрическая модель Раша. В качестве исследуемых величин выступают латентные параметры: уровень знаний испытуемых и трудность заданий. При этом оба параметра, а также их начальные значения, полученные непосредственно после тестирования, переводятся в единую линейную шкалу логитов. При вычислении вероятности правильного ответа на задание используется логистическое распределение. В результате выполненных действий построены характеристические кривые уровня знаний испытуемых, уровня трудности заданий, совмещенные гистограммы этих уровней, а также графики сравнения экспериментальных данных с моделью Раша. Проведен предварительный анализ полученных результатов.

Ключевые слова: *однопараметрическая модель Раша, уровень знаний, трудность заданий, логит, логистическое распределение.*

В работах [1–4] тест «Линейная алгебра, аналитическая геометрия», предназначенный для текущего контроля знаний студентов первого курса Самарского государственного технического университета, рассматривался с точки зрения классической теории тестов. В данной публикации с целью более тщательного исследования качества самого теста, а также для получения объективной оценки уровня знаний студентов будем использовать IRT (современная теория создания тестов) с моделями Раша [5–12].

В рамках этой теории результаты наблюдений (количество набранных баллов) трактуются как величины, лишь фиксирующие число положительных исходов случайных событий, и не могут давать объективную картину уровня знаний и трудности заданий. Модели Раша позволяют преобразовать исходный материал в такие оценки, которые можно рассматривать как результаты измерения уровня подготовленности испытуемых и уровня трудности зада-

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

ний. При этом используется единая линейная шкала логитов, а перечисленные параметры – уровень знаний испытуемого и трудность задания – полностью разделены, то есть не зависят друг от друга [12].

После построения модели проверяется, соответствуют ли ей результаты эксперимента. В теории Раша утверждается, что именно экспериментальные данные должны соответствовать модели, а не наоборот [7]. Следовательно, в случае выявленного несоответствия тест нужно переработать и добиться лучшего согласия с теорией.

Рассмотрим, как строится однопараметрическая модель Раша [5, 7, 9, 11]. Латентные (скрытые) параметры – уровень знаний (θ) и трудность задания (β) – считаются переменными величинами. Их начальные значения, полученные эмпирическим путем в результате тестирования, переводятся в единую интервальную шкалу логитов. В качестве математической модели для описания связи между латентными переменными и эмпирическими результатами теста используется логистическое распределение, хорошо согласующееся с педагогической практикой. При этом вероятность правильного ответа i -го студента на j -е задание выражается формулой

$$P_{ij} = (1 + \exp(-1,7(\theta_i - \beta_j)))^{-1},$$

где P_{ij} – вероятность успеха i -го студента в j -м задании.

Так как модель Раша описывает вероятность успеха как функцию одного параметра ($\theta_i - \beta_j$), она называется однопараметрической.

Построим такую модель для нашего теста [5, 7, 9, 11].

Обозначим n – число заданий теста, N – число студентов, участвовавших в тестировании. В рассматриваемом тесте $n = 19$, $N = 132$.

Начальные значения параметров θ и β в шкале логитов определяются формулами

$$\theta_i^0 = \ln \frac{p_i}{q_i}, \beta_j^0 = \ln \frac{p_j}{q_j}.$$

Здесь $p_i = \frac{X_i}{n}$, $p_j = \frac{R_j}{N}$, X_i – число правильных ответов i -го студента, R_j – число правильных ответов на j -е задание, $q_i = 1 - p_i$, $q_j = 1 - p_j$.

Находим средние значения логитов уровня знаний $\bar{\theta}$ и трудности заданий $\bar{\beta}$:

$$\bar{\theta} = \frac{\sum_{i=1}^N \theta_i^0}{N}, \bar{\beta} = \frac{\sum_{j=1}^n \beta_j^0}{n},$$

а также дисперсии V и U по множествам θ_i^0 и β_j^0 соответственно:

$$V = \frac{1}{N-1} (\sum_{i=1}^N (\theta_i^0)^2 - N(\bar{\theta})^2), U = \frac{1}{n-1} (\sum_{j=1}^n (\beta_j^0)^2 - n(\bar{\beta})^2).$$

Получаем $\bar{\theta} = 0,644177$, $\bar{\beta} = -0,55848$, $V = 1,349436$, $U = -0,166361839$.

Далее вычисляем поправочные коэффициенты X и Y :

$$X = \sqrt{\frac{1+U/2,89}{1-UV/8,35}} = 0,957998, Y = \sqrt{\frac{1+V/2,89}{1-UV/8,35}} = 1,19521.$$

В результате для оценок параметров θ и β в единой интервальной шкале логитов согласно формулам

$$\theta_i = \bar{\beta} + X\theta_i^0, \beta_j = \bar{\theta} + Y\beta_j^0$$

имеем соотношения

$$\theta_i = -0,55848 + 0,957998\theta_i^0, \beta_j = 0,644177 + 1,19521\beta_j^0.$$

Стандартные ошибки измерения уровня знаний $S_e(\theta_i)$ и трудности задания $S_e(\beta_j)$ находим по формулам

$$S_e(\theta_i) = \frac{X}{\sqrt{np_iq_i}}, S_e(\beta_j) = \frac{Y}{\sqrt{Np_jq_j}}.$$

Для анализа и корректировки тестовых заданий строим две группы характеристических кривых, используя формулу вероятности P_{ij} логистического распределения. В одном случае фиксируем трудность β_j , полагая θ независимой переменной. Характеристическая кривая задания j строится согласно формуле

$$P_j(\theta) = (1 + \exp(-1,7(\theta - \beta_j)))^{-1}.$$

Во втором случае при фиксированном θ_i и независимой переменной β , согласно соотношению

$$P_i(\beta) = (1 + \exp(-1,7(\beta - \theta_i)))^{-1}$$

получаем характеристическую кривую i -го студента.

Выполним предварительный анализ полученных результатов.

На рис. 1 представлены характеристические кривые 1-й группы для 19 заданий теста. По горизонтали отложены логиты уровня знаний, по вертикали – вероятности выполнения заданий в зависимости от уровня знаний студента.

Задания теста по шкале логитов перекрывают диапазон от -5 до +5 логитов. По рисунку видно, что некоторые задания дублируют друг друга (кривые практически совпадают), поэтому их можно исключить из теста. С другой стороны, есть некоторая неравномерность (провалы) в расположении кривых, значит, некоторые задания теста следует переработать.

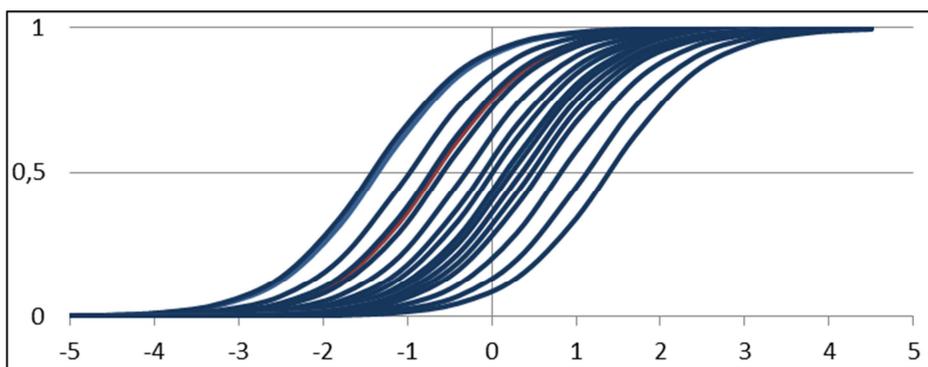


Рис. 1. Характеристические кривые заданий

На рис. 2 показаны характеристические кривые 2-й группы для 132 студентов. По горизонтали отложены логиты уровня трудности заданий, по вертикали – вероятности выполнения заданий в зависимости от трудности.

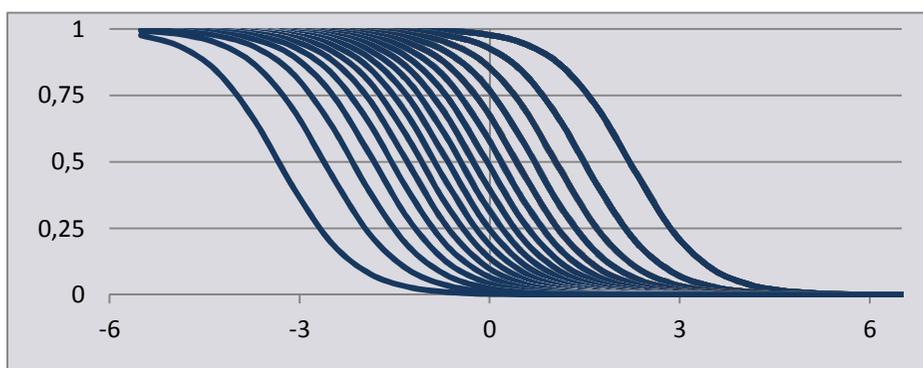


Рис. 2. Характеристические кривые уровня знаний

Лучшим критерием качества измерений в IRT считаются совмещенные гистограммы уровней знаний и трудности заданий [5]. Построенные в единой шкале логитов, такие гистограммы выявляют все недостатки теста. Для исследуемого теста эти гистограммы изображены на рис. 3: сверху – гистограмма частот правильных ответов в зависимости от уровня знаний, снизу – гистограмма частот правильных ответов в зависимости от трудности заданий. Что касается нормального распределения, которое принимается в качестве критерия [7], то визуально здесь наблюдается неплохое соответствие. В то же время диапазон значений логитов трудности меньше диапазона уровня знаний. Это означает, что в тесте нет заданий, соответствующих как уровню наиболее, так и уровню наименее подготовленных студентов, то есть не хватает трудных и легких заданий.

Проверим соответствие эмпирических данных модели Раша. С этой целью выполняются следующие действия [7]. Все тестируемые делятся на J групп в зависимости от уровня подготовленности θ . Выбирается задание. Для каждой группы вычисляется вероятность правильных ответов на выбранное задание теста:

$$p_j = \frac{r_j}{n_j},$$

где j – номер группы ($j = 1, 2, \dots, J$), n_j – число студентов в группе, r_j – число правильных ответов на задание. Эта величина – экспериментальное значение вероятности. Затем строятся два графика: экспериментальная кривая и характеристическая кривая, полученная по модели Раша.

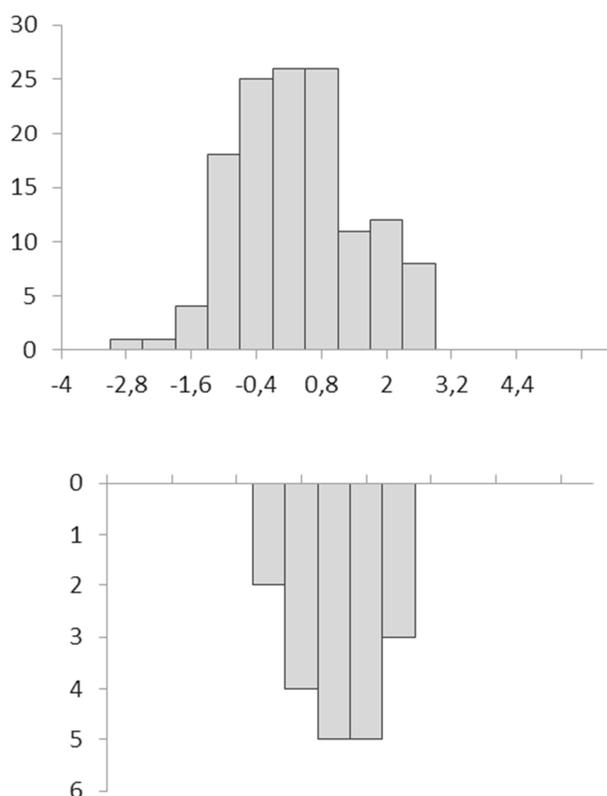


Рис. 3. Совмещенные гистограммы уровней знаний и трудности заданий

Для рассматриваемого теста число групп с одинаковым уровнем подготовки оказалось равным $J = 18$. На рис. 4 представлены результаты вычислений в единой интервальной шкале логитов для заданий различной трудности. Это задания: 14 – слева и 15 – справа. Сплошными линиями изображены гра-

фики характеристических кривых, построенных по модели Раша, точками – экспериментальные значения.

Построенные графики выявляют значительное расхождение экспериментальных данных и результатов, полученных с помощью однопараметрической модели Раша. Это может быть обусловлено [7] несовершенством задания, нарушением в процедуре тестирования, но, возможно, означает, что для более адекватного отображения действительности следует переходить к двух- и трехпараметрическим моделям [9].

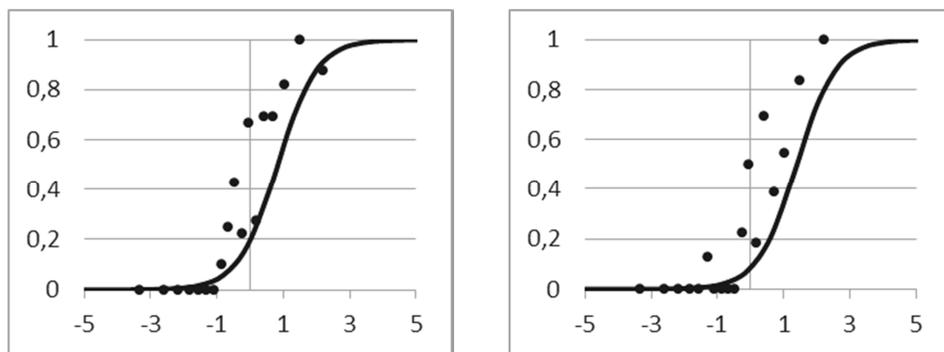


Рис. 4. Сравнение экспериментальных данных с моделью Раша

Итак, для теста «Линейная алгебра, аналитическая геометрия» построена однопараметрическая модель Раша, проиллюстрированная графиками характеристических кривых уровней знаний и трудности заданий, совмещенными гистограммами этих уровней, графиками сравнения модели Раша с экспериментальными данными по результатам тестирования. Предварительный анализ позволил сделать следующие выводы: требуется исправить неоднородность расположения кривых уровня знаний, переработав задания теста; устранить совпадение кривых, исключив задания, дающие повторения. Кроме того, необходимо добавить в тест задания, соответствующие как уровню наиболее, так и уровню наименее подготовленных студентов. Расхождение экспериментальных данных с результатами, полученными по модели Раша, следует проверить по известным критериям теории вероятностей.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Лиманова Л.В., Муратова Л.А. Анализ качества теста из курса высшей математики по теме «Линейная алгебра, аналитическая геометрия» // Вестник СамГТУ. Сер. Психолого-педагогические науки. – 2015. – № 2(26). – С. 113–122.

2. *Лиманова Л.В.* Повышение надежности педагогического теста «Линейная алгебра, аналитическая геометрия» // Вестник СамГТУ. Сер. Психолого-педагогические науки. – 2016. – № 2(30). – С. 75–81.
3. *Муратова Л.А.* Качественные изменения педагогического теста «Линейная алгебра, аналитическая геометрия» и их оценка // Вестник СамГТУ. Сер. Психолого-педагогические науки. – 2016. – № 2 (30). – С. 82–86.
4. *Муратова Л.А.* Норма и критерий на примере теста «Линейная алгебра, аналитическая геометрия» // Научный альманах: По материалам международной научно-практической конференции «Теоретические и прикладные вопросы науки и образования»: Тамбов, 30 июля 2016. – № 7-1 (21). – С. 278–282.
5. *Аванесов В.С.* Критерии качества педагогических измерений. // Педагогические измерения. – 2012. – № 1. – С. 55–68.
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. *Ким В.С.* Тестирование учебных достижений. – Уссурийск: Изд-во УГПИ, 2007. – 214 с.
8. *Нейман Ю.М., Хлебников В.А.* Введение в теорию моделирования и параметризации педагогических тестов. – М.: Прометей, 2000. – 168 с.
9. *Олейник Н.М.* Тест как инструмент измерения уровня знаний и трудности заданий в современной технологии обучения: учеб. пособие. – Донецк: Изд-во ДонГУ, 1991. – 168 с.
10. *Переверзев В.Ю.* Критериально-ориентированные педагогические тесты для итоговой аттестации студентов. – М.: НМЦ СПО Минобразования РФ, 1999. – 152 с.
11. *Чельщикова М.Б.* Теория и практика конструирования педагогических тестов: учеб. пособие. – М.: Логос, 2002. – 432 с.
12. [portal.tpu.ru/SHARED/m/MURATOVA/...](http://portal.tpu.ru/SHARED/m/MURATOVA/) копия

Поступила в редакцию 23.11.16;
в окончательном варианте 5.12.16

UDC 378.14

CREATION OF ONE-PARAMETRICAL MODEL RUSSIA FOR TESTS "LINEAR ALGEBRA, ANALYTICAL GEOMETRY"

L.A. Muratova¹

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

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

In this article for the "Linear Algebra, Analytical Geometry" test, according to the modern theory of testing (IRT) the one-parametrical Russia model is under construction. Latent parameters act as the studied sizes: level of knowledge of examinees and difficulty of tasks. Thus both parameters, and also their initial values received directly after testing are transferred to a uniform linear scale of logits. At calculation of probability of the correct answer on a task logistic distribution is used. As a result of the executed actions characteristic curves of level of knowledge of examinees, level of difficulty of tasks, the combined histograms of these levels, and also graphics of comparison of experimental data with model Russia are constructed. The preliminary analysis of the received results is carried out.

Keywords: *one-parametrical Russia model, level of knowledge, difficulty of tasks, logit, logistic distribution.*

REFERENCES

1. *Muratova L.V., Limanova 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». 2015. № 2(26). P. 113–122.*
2. *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» 2016. № 2(30). P. 75–81.*
3. *Muratova L.A. Kachestvennye izmeneniya pedagogicheskogo testa «Lineinaya algebra, analiticheskaya geometriya» i ih ocenka [High-quality changes of the pedagogical test "Linear Algebra, Analytical Geometry" and their assessment] // Vestnik SamGTU, Seriya «Psihologo-pedagogicheskie nauki» 2016. № 2(30). P. 82–86.*
4. *Muratova L.A. Norma i kriterii na primere testa «Lineinaya algebra, analiticheskaya geometriya» [Norma and criterion on the example of the "Linear Algebra, Analytical Geometry" test]. Nauchnyi al'manah 2016· N 7-1(21). Po materialam mezhdunarodnoi nauchno-prakticheskoi konferencii «Teoreticheskie i prikladnye voprosy nauki i obrazovaniya» [Scientific almanac 2016 – N 7-1(21). On materials of the international scientific and practical conference "Theoretical and Applied Questions of Science and Education"] Russia, Tambov, on July 30, 2016 g.s. 278–282.*
5. *Avanesov V.S. Kriterii kachestva pedagogicheskikh izmerenii [Kriteriya of pedagogical measurements quality] // Pedagogical measurements, 2012. No. 1. P. 55–68.*
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 – 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. P. 53–62.*

7. *Kim V.S.* Testirovanie uchebnyh dostizhenii [Testing of educational achievements]. Ussuriisk: Publishing house of UGPI, 2007. 214 p.
8. *Neumann U.E.M., Khlebnikov V.A.* Vvedenie v teoriyu modelirovaniya i parametrizatsii pedagogicheskikh testov [Introduction to the theory of modeling and parametrization of pedagogical tests]. Moscow: Prometheus, 2000. 168 p.
9. *Oleynik N.M.* Test kak instrument izmereniya urovnya znaniy i trudnosti zadaniy v sovremennoi tehnologii obucheniya. Uchebnoe posobie [The test as the instrument of measurement of level of knowledge and difficulty of tasks in modern technology of training]. – Manual: Donetsk, to DONG, 1991. 168 p.
10. *Pereverzev V.Yu.* Kriterial'no-orientirovannye pedagogicheskie testy dlya itogovoi attestatsii studentov [The criteria focused pedagogical tests for total certification of students]. Moscow: NMTs SPO of the Ministry of Education of the Russian Federation, 1999. 152 p.
11. *Chelyshkova M.B.* Teoriya i praktika konstruirovaniya pedagogicheskikh testov: Uchebnoe posobie [Theory and practice of designing of pedagogical tests: Manual]. – Moscow: Logos, 2002. 432 p.
12. portal.tpu.ru/SHARED/m/MURATOVA/... copy

Original article submitted 23.11.16;
revision submitted 1.12.16