DOI: 10.17117/na.2016.07.01.278

<http://ucom.ru/doc/na.2016.07.01.278.pdf>

Поступила (Received): 02.07.2016

Муратова Л.А.
Норма и критерий на примере теста «Линейная алгебра, аналитическая геометрия»

Muratova L.A.
Norm and criterion on the example of the "Linear Algebra, Analytical Geometry" test

В статье исследуется вопрос о возможности использования критериально-ориентированного теста в качестве нормативно-ориентированного на примере теста «Линейная алгебра, аналитическая геометрия». Рассматриваются характеристики теста: надежность, дискриминативность, трудность заданий, распределение индивидуальных тестовых баллов. Делается вывод о том, что между указанными подходами нет неразрешимых противоречий, и в результате некоторой переработки может быть получен тест, удовлетворяющий и требованиям нормативности и критериальности

Ключевые слова: тест критериально-ориентированный, нормативно-ориентированный, надежность, дискриминативность

In article the question of possibility of use of the criteria focused dough as standard focused on the example of the "Linear Algebra, Analytical Geometry" test is investigated. Characteristics of dough are considered: reliability, diskriminativnost, difficulty of tasks, distribution of individual test points. The conclusion that between the specified approaches there are no unsoluble contradictions is drawn, and as a result of some processing the test meeting and to requirements of a normativnost and kriterialnost can be received

Key words: test criteria focused, standard focused reliability, a diskriminativnost

Муратова Лидия Александровна

Кандидат технических наук, доцент
Самарский государственный технический университет
г. Самара, ул. Молодогвардейская, 244

Muratova Lidiya Aleksandrovna

Candidate of Technical Sciences, Associate Professor
Samara state technical university
Samara, Molodogvardeyskaya st., 244

Тестируемый контроль имеет неоспоримые преимущества перед прежними способами проверки знаний. Использование тестируемого контроля позволяет получить объективную целостную картину усвоенных знаний каждого студента по каждому разделу, выяснить, какие темы усвоены лучше, какие хуже, и впоследствии скорректировать дальнейшее обучение. В связи с этим на кафедре высшей математики и прикладной информатики СамГТУ были созданы тесты для промежуточного и итогового контроля знаний по всем разделам курса в соответствии с рабочими программами. Ставилась следующая задача – определить степень усвоения пройденного материала и соответствующим образом оценить знания студентов. Составители, руководствуясь здравым

смыслом и накопленным опытом преподавания высшей математики в вузе, старались охватить весь материал курса высшей математики. Накопленный статистический материал результатов тестирования дал возможность рассмотреть вопрос качества используемых тестов с целью их усовершенствования. При этом появилась проблема, связанная со способами проверки качества в зависимости от принадлежности теста к той или иной категории. Известна классификация тестов на две группы: нормативно-ориентированные и критериально-ориентированные. Их главное отличие состоит в том, что нормативно-ориентированные тесты предполагают сравнение уровней знаний тестируемых между собой, в то время как критериально-ориентированные позволяют выяснить, в какой степени испытуемые владеют учебным материалом [1, 2, 6, 7]. Известна точка зрения, согласно которой нельзя использовать критериально-ориентированный тест в качестве нормативно-ориентированного и наоборот; нельзя использовать один и тот же тест и в качестве нормативно-ориентированного и в качестве критериально-ориентированного. Считается, что нарушение этих правил приводит к получению тестовых результатов, обладающих низкой надежностью и большой ошибкой измерения [2]. В то же время существует и другая точка зрения: утверждается, что нормативный и критериальный подходы при составлении теста могут взаимно дополнять друг друга, так что в результате будет получен более информативный тест [6, 7]. Придерживаясь этой позиции, в качестве примера, рассмотрим тест «Линейная алгебра, аналитическая геометрия». Этот тест используется для текущего контроля знаний студентов первого курса, он является тестом закрытого типа и содержит 20 заданий с выбором единственного правильного ответа из пяти предложенных. Для анализа качества данного теста были взяты результаты тестирования 106 студентов факультета автоматизации и информационных технологий.

При составлении теста применялся критериально-ориентированный подход, то есть на первый план выдвигалась задача – выяснить, знает ли студент стандартный учебный материал. Поэтому содержание теста имеет высокий уровень детализации и достаточно полно отражает все темы рассматриваемых разделов. Возникает вопрос: можно ли этот тест использовать не только как критериально-ориентированный, но и как нормативно-ориентированный? В работе [3] проведен соответствующий статистический анализ материалов тестирования. Воспользуемся его итогами.

Одним из самых важных показателей качества теста является его надежность. Для нормативно-ориентированного теста, как правило, вычисляют коэффициенты надежности следующим образом: по формуле KR-20, с помощью среднего коэффициента корреляции и путем расщепления теста по четным и нечетным заданиям [1, 2, 4, 5, 7]. В исследуемом тесте коэффициенты надежности соответственно равны: 0,753, 0,824, 0,774 [3]. Поскольку считается [4, 5], что нормативно-ориентированный тест можно использовать, если значение коэффициента надежности не меньше 0,7 [1, 2, 5, 7], то это вполне приемлемые результаты.

Еще одной характеристикой, с помощью которой проверяется качество нормативно-ориентированного теста, является дискриминативность – способность задания дифференцировать студентов на лучших и худших. Индекс дискриминативности можно найти следующим образом. По результатам теста выделяют две группы студентов: 27% «лучших» и 27% «худших», подсчитывают относительные частоты правильных ответов в каждой группе, а затем находят их разность ($r_{диск}^j$). Это и есть индекс дискриминативности [1, 4, 5, 7]. Индекс не должен быть менее 0,2 [7]. В рассматриваемом тесте 1, 4, 12 задания имеют индекс дискриминативности 0,07, 0,1 и 0,14 соответственно и значит, не удовлетворяют этому условию.

О дискриминативной способности задания можно также судить и по точечному бисериальному коэффициенту - коэффициенту корреляции r_{pb}^j каждого задания с тестовым баллом студента [1, 4, 5, 7]. В данном случае «плохими» оказались лишь два задания (1, 12) со значениями 0,15 и 0,13.

В качественном нормативно-ориентированном тесте должно быть подавляющее большинство заданий средней трудности (от 0,4 до 0,7). В этом случае вариация (дисперсия) достигает максимального значения (0,25) либо близка к нему [1,5,7]. Это, в свою очередь, означает, что индивидуальные тестовые баллы в достаточной степени отличаются друг от друга и тест хорошо дифференцирует испытуемых. В рассматриваемом тесте таких заданий хоть и не подавляющее большинство, но тоже много (50%). Как и полагается в нормативно-ориентированном тесте, есть очень простые (трудность 0,93) и достаточно сложные задания (трудность 0,2).

Рассмотрим распределение индивидуальных тестовых баллов. У нормативно-ориентированного теста оно должно быть близко к нормальному, в то время, как у критериально-ориентированного может быть произвольным.

На рис.1 слева изображена кривая распределения индивидуальных баллов испытуемых (сплошная линия) и кривая, соответствующая нормальному распределению (пунктирная линия).

Если выбирать среди самых распространенных распределений (нормального, показательного и равномерного), то самым подходящим в данной ситуации будет нормальное распределение. Однако проверка показала: согласно критерию Пирсона при уровне значимости 0,01 и числе степеней свободы, равном 4, гипотезу о нормальном распределении следует отвергнуть.

Суммируем получившиеся результаты относительно использования теста как нормативно-ориентированного. С точки зрения надежности нет противоречий. Анализ дискриминативности выявил 3 задания, которые требуют переработки. При этом необходимо отслеживать изменения кривой распределения индивидуальных баллов на предмет соответствия нормальному распределению. Например, можно просто исключить «плохие» 1, 4, и 12 задания, в результате получится кривая, изображенная на рис.1 справа сплошной линией. В этом случае согласно критерию Пирсона при уровне значимости 0,01 и числе степеней свободы, равном 4, уже нет оснований отвергать гипотезу о нормальном распределении. Но тогда в тест не попадут важные темы изучаемого раздела. Значит,

нужно искать компромиссные решения: изменения теста должны быть выполнены не в ущерб содержанию.

Рис. 1. Распределение индивидуальных баллов

В целом, можно сделать вывод: проведенное исследование не выявило неразрешимых противоречий между двумя подходами, а соответствующая доработка даст возможность получить тест, отвечающий требованиям и нормативности и критериальности.

Список используемых источников:

1. 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.
2. Ким В.С. Тестирование учебных достижений. Уссурийск: Изд-во УГПИ, 2007. 214 с.
3. Лиманова Л.В., Муратова Л.А. Анализ качества теста из курса высшей математики по теме «Линейная алгебра, аналитическая геометрия» // Вестник СамГТУ. № 2(26)-2015. С. 113-122.
4. Олейник Н.М. Тест как инструмент измерения уровня знаний и трудности заданий в современной технологии обучения. Донецк: ДонГУ, 1991. 168 с.

5. Переверзев В.Ю. Критериально-ориентированное педагогическое тестирование. М.: Логос, 2003. 120 с.
6. Сафонцев С.А. Технология критериально ориентированного тестирования потенциальных возможностей учащихся: Дис. ... канд. пед. наук: 13.00.01. Ростов н/Д, 2001 156 с.
7. Чельшкова М.Б. Теория и практика конструирования педагогических тестов. М.: Логос, 2002. 432 с.

© 2016, Муратова Л.А.

Норма и критерий на примере теста «Линейная алгебра, аналитическая геометрия»

© 2016, Muratova L.A.

Norm and criterion on the example of the "Linear Algebra, Analytical Geometry" test