

УДК 004.827

**Е. А. Городничая, Т. В. Онищенко,
В. А. Крисилов**, д-р техн. наук

ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ ОЦЕНКИ И ПОВЫШЕНИЯ КАЧЕСТВА АВТОМАТИЗИРОВАННЫХ СРЕДСТВ КОНТРОЛЯ ЗНАНИЙ В УЧЕБНОМ ПРОЦЕССЕ

Рассматривается методика анализа педагогических тестов без привлечения большого количества экспертов для повышения качества проверки знаний, умений и навыков студентов по изученному материалу за определенный период обучения.

Ключевые слова: дистанционное обучение, педагогические тесты, оценка педагогических тестов, повышение качества педагогических тестов.

**К. А. Gorodnichaya, T. V. Onishenko,
V. A. Krisilov**, ScD.

INFORMATION TECHNOLOGY OF ESTIMATION AND QUALITY IMPROVEMENT OF AUTOMATED KNOWLEDGE CONTROL METHODS IN THE EDUCATIONAL PROCESS

Considered the methods of analysis of pedagogical tests without the involvement of a large number of experts, to improve the quality of testing the knowledge and skills students learned material over a period of training.

Keywords: distance learning, educational tests, pedagogical tests, assessment pedagogical tests, improve the quality of pedagogical tests.

**К. О. Городничая, Т. В. Онищенко,
В. А. Крисілов**, д-р техн. наук

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ОЦІНКИ ТА ПІДВИЩЕННЯ ЯКОСТІ АВТОМАТИЗОВАНИХ ЗАСОБІВ КОНТРОЛЮ ЗНАНЬ У НАВЧАЛЬНОМУ ПРОЦЕСІ

Розглядається методика аналізу педагогічних тестів без залучення великої кількості експертів для підвищення якості перевірки знань, умінь і навичок студентів з вивченого матеріалу за певний період навчання.

Ключові слова: дистанційне навчання, педагогічні тести, оцінка педагогічних тестів, підвищення якості педагогічних тестів.

1. Введение

Одним из наиболее перспективных путей развития современных информационных технологий в сфере образования является внедрение дистанционного обучения. Поэтому на сегодняшний день большинство учебных заведений по всему миру ставят перед собой задачу использования в учебном процессе методов и средств, использующих мультимедиа, интернет и электронные ресурсы.

При этом важнейшей составляющей обучения остается контроль знаний.

Автоматизированный контроль знаний дает возможность решать многие проблемы: работать практически с любым количеством обучаемых, осуществлять различного рода контроль знаний, оперативно анализировать результаты контроля, осуществлять мониторинг качества обучения и повышать эффективность контроля знаний.

© Городничая Е.А., Онищенко Т.В.,
Крисилов В.А., 2012

Вопрос качества обучения неотделим от вопроса качества контроля знаний, поскольку именно контроль является одним из факторов, мотивирующих к непрерывному и глубокому изучению материала. В практике отечественной высшей школы основными его формами традиционно являлись опросы, домашние задания, курсовые и контрольные работы, письменные и устные экзамены. В последнее время в учебном процессе все более активно внедряется такая форма контроля, как тестирование [1].

Педагогический тест – это инструмент, предназначенный для измерения уровня знаний студента, состоящий из системы тестовых заданий (вопросов), стандартизированной процедуры проведения, обработки и анализа результатов [2].

Педагогические тесты позволяют провести объективную оценку достигнутого уровня знаний при массовой их проверке. Однако, как показывает практика, разработанные тесты могут содержать неэффектив-

ные задания и в этом случае не позволяют правильно определить степень и качество усвоения материала студентами. В большинстве тестов часто возникают проблемы, связанные со сложностью задаваемого вопроса.

Несмотря на то, что различные приемы тестирования начали изучаться давно, в настоящий момент не существует однозначного мнения относительно отбора и формирования тестовых заданий. Хотя эта работа является подготовительной, значение ее достаточно высоко, так как объективность, качество и трудоемкость теста, безусловно, зависят от правильного подбора вопросов. На практике выбор тестовых заданий очень часто выполняется произвольным образом и во многом зависит от лица, разрабатывающего тест. Это существенно влияет на качество тестирования.

Экспертиза тестовых заданий является обязательным условием для создания тестов вне зависимости от сферы их применения и проводится группой экспертов. Оценка одним экспертом не дает гарантии отсутствия целого ряда недостатков. Под экспертизой тестовых заданий понимается проверка объективности и корректности формулировки вопросов [2].

Разработка общих принципов оценки качества тестов – достаточно актуальная проблема, потому что трудно осуществить последовательное сопоставление целей образования с достигаемыми результатами обучения [1] и для этого необходимо формализовать методику оценки теста [7].

Проблема объективного подхода к определению уровня сложности задаваемых вопросов и соответственно определение интегральной оценки знаний является недостаточно изученной.

В существующих системах дистанционного обучения, таких как Moodle, ILIAS, WebTutor, CourseLab, aTutor, Blackboard Learning System, оценка обучаемых в ходе тестирования зависит от объективности выбора вопросов, которые не контролируются системой с точки зрения правильности.

Целью данной работы является разработка методики анализа педагогических

тестов, без привлечения большого количества экспертов, для повышения качества проверки знаний, умений и навыков студентов по изученному материалу за определенный период обучения.

Существуют два подхода оценки качества теста: априорный и апостериорный. По данной методике расчет критериев качества теста будет проводиться по результатам пробного тестирования, т.е. по апостериорным оценкам.

Согласно данной методике к критериям качества теста будем относить следующие показатели: [4]

1. **Сложность** – определяется сложностью каждого вопроса и сложностью теста в целом.

2. **Надежность** – позволяет получить устойчивую оценку способностей испытуемых и показывает, что результаты проводимого исследования близки к истине, а также получить устойчивость результатов теста к посторонним случайным факторам. Для определения надежности используется коэффициент корреляции тестовых заданий и результата за тест. Результат за тест – это оценка, которая равна сумме баллов, полученных при ответе на каждый вопрос.

3. **Валидность** – позволяет определить насколько качественно проводится обучение. Валидность проверяется на основе сравнения результатов оценок по разработанному тесту с результатами, полученными другими методами (устный опрос, традиционная контрольная работа, экзамен, РГР, лабораторные работы и др.), т.е. опорной оценки. В данном случае предлагается в качестве опорной оценки использовать средний бал по всем предметам за весь период обучения.

Расчет критериев качества теста проводится с использованием методов теории вероятностей и математической статистики и предлагается проводить по нижеприведенной схеме.

1. Анализ сложности теста и тестовых вопросов

2. Анализ надежности теста

2.1. Вычисление коэффициента корреляции тестовых заданий и результата за тест

2.2. Построение диаграммы коэффициента корреляции тестовых заданий и результата за тест

3. Анализ валидности теста

3.1. Формирование презентативной группы студентов

3.1.1. Построение функции распределения баллов за весь период обучения

3.1.2. Выбор студентов, успешность которых подчиняется унимодальному нормальному распределению

3.2. Расчет отклонения тестовых оценок по данной дисциплине от среднего

2. Описание исследования

2.1. Анализ сложности

2.1.1. Анализ сложности теста (плотность распределения оценок студентов за тест)

Для анализа сложность теста строим диаграмму плотности распределения оценок за тест.

При построении диаграммы разбиваем шкалу оценок на диапазоны: «неудовлетворительно», «удовлетворительно», «хорошо», «отлично» в равных пропорциях. Вычисляем количество оценок, которые студенты получили за тест в каждом диапазоне.

Полученная диаграмма наглядно демонстрирует количество студентов, не прошедших тестирование, т.е. не набравших нужное количество баллов, а также количество студентов, которые набрали удовлетворительные, хорошие и отличные баллы, т.е. получили оценки в заданных диапазонах.

Принимаем во внимание существующие в педагогической практике рекомендации:

если после прохождения тестирования более 10–15 % студентов набрали отличные оценки, то тест требует усложнения;

если более 5–10 % студентов набрали за тест неудовлетворительные оценки, то тестовые вопросы следует упростить [3].

Пример теста, в котором требуется усложнить тестовые задания, показан на рис. 1.

В данном примере неудовлетворительных оценок 0 %, т.е. тест не следует упрощать. Отличные баллы набрало 53,3 % сту-

дентов. Из этого делаем вывод, что тест следует усложнить.



Рис. 1. Гистограмма, отображающая тест, требующий усложнения

Пример теста, в котором требуется упростить тестовые вопросы, показан на рис.2. В данном примере неудовлетворительных оценок 43,3 %, т.е. тест следует упрощать. Отличные баллы набрало 6,7 % студентов. Из этого делаем вывод, что тест не следует усложнять.



Рис. 2. Гистограмма, отображающая тест, требующий упрощения

Пример «оптимального» теста показан на рис. 3. В данном примере неудовлетворительных оценок 6,7 %, т.е. тест не следует упрощать. Отличные баллы набрало 13,3 % студентов. Из этого делаем вывод, что тест не следует дорабатывать.



Рис. 3. Гистограмма, отображающая, «оптимальный» тест

2.1.2. Анализ сложности тестовых вопросов

Под сложностью тестового вопроса будем понимать показатель, характеризующий долю студентов, неправильно ответивших на тестовый вопрос. Рассчитаем его по формуле

$$P_j = \frac{n_j}{n}, \quad (1)$$

где n_j – число неправильных ответов на j -й тестовый вопрос; n – количество ответов на тест;

P_j – сложность j -го тестового вопроса.

В педагогической практике существуют рекомендации по оптимальному соотношению в процентном соотношении успевающих и неуспевающих студентов [3]. Поэтому считаем, что сложность вопроса является удовлетворительной, если она находится в пределах от 0,15 до 0,9. Вопросы, в которых сложность находится в пределах от 0,9 до 1 считаются сложными, их рекомендуется упростить. Вопросы, в которых сложность находится в пределах от 0 до 0,15, считаются легкими, их рекомендуется переработать.

После вычисления сложности вопросов строим диаграмму, отображающую сложность вопросов теста. Диаграмма наглядно демонстрирует вопросы, которые следует исправить в зависимости от набранных результатов, рассчитанных по формуле (1). Например, по рис.4. рекомендуется усложнить вопросы номер 1, 3, 4, 5, 8, 10, 15, 16, 17, 22, 24, 30 и упростить вопрос номер 28.



Рис. 4. Диаграмма, отображающая сложность каждого вопроса

2.2. Анализ надежности теста

Надёжностью называется один из критериев качества теста, характеризующий его устойчивость по отношению к погрешностям измерения (случайные ответы на тест, сбои в системе). Анализируем надежность по значе-

нию коэффициента корреляции тестовых заданий и общего результата за тест.

2.2.1. Вычисление коэффициента корреляции тестовых заданий и общего результата за тест

Коэффициент корреляции, как степень линейной зависимости двух случайных величин, в данном случае зависимость между результатом каждого тестового задания и общим результатом за тест, позволяет определить целесообразность включения задания в тест [5].

Для анализа по данной методике будем использовать коэффициент корреляции Пирсона, который изменяется в диапазоне $[-1;1]$.

Коэффициенте корреляции равен:

- а) 1, при строгой линейной зависимости между коррелируемыми величинами;
- б) 0, если нет связи между коррелируемыми величинами;
- в) -1 , при строгой обратной связи между коррелируемыми величинами.

При высоком коэффициенте корреляции студенты, имеющие высокий бал за тест, ответили на данный вопрос правильно, а большинство студентов с низким балом за тест ответили не правильно. При низком коэффициенте корреляции студенты с высоким балом за тест ошибались чаще, чем студенты с низким балом за тест.

Таким образом, чем выше значение коэффициента корреляции, тем больше вероятность включения данного задания в тест. Задания со значением коэффициента корреляции меньше или равному 0,4 предлагается исключать из теста как не выдержавшие эмпирической проверки.

Вычисление коэффициента корреляции сопровождается вычислением дисперсии, которая отражает меру неоднородности результатов теста [5], т.е измеряет разброс баллов, полученных студентами при ответе на конкретное задание теста.

Задания с нулевой дисперсией и дисперсией близкой к нулю показывают слабое разделение студентов по уровню подготовки, т.е. все студенты дали на вопрос правильный ответ, или наоборот, все ответы были неправильными.

Задания с дисперсией равной 0,25 (среднеквадратическое отклонение равно 0,5) показывает, что половина студентов отвечает правильно, а половина неправильно.

Чем выше дифференцирующая способность тестового задания (больше величина дисперсии), тем выше качество теста.

2.2.2. Построение диаграммы коэффициента корреляции тестовых заданий

После вычисления коэффициента корреляции строится диаграмма. Эта диаграмма наглядно демонстрирует перечень вопросов, которые следует заменить или исправить.

На рис. 5. представлена диаграмма, отображающая коэффициент корреляции тестовых вопросов. По ней можно определить, что нужно переделать задания номер 3, 4, 5, 10, 12, 13, 14, 15, 16, 19, 21, 24, 25, 28, 30.



Рис. 5. Диаграмма, отображающая коэффициент корреляции

2.3. Анализ валидности теста

Для анализа валидности теста проводится расчет отклонения оценок за тест от некоторой опорной оценки. В качестве опорной оценки можно использовать средний бал, либо оценки по данному предмету, но полученные в рамках других форм учебного процесса (РГР, лабораторные работы и др.). В данной методике предлагается использовать средний бал по всем предметам за весь период обучения.

2.3.1. Формирование презентативной группы студентов

Использование среднего бала в качестве опорной оценки накладывает некоторые ограничения на состав презентативной группы студентов, для которой будет проверяться валидность.

В состав этой группы предлагается включать только студентов со стабильным уровнем знаний, закон распределения оценок

которых соответствует нормальному (уни-модальному). Такая функция имеет только одну моду (один экстремум). Примеры таких функций показаны на рис. 6 и 7.

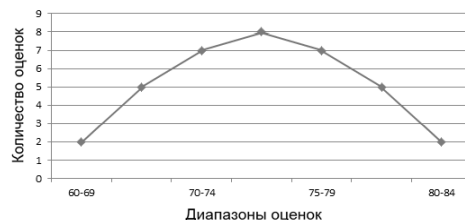


Рис. 6. Функция, характеризующая студента 1, набравшего за рассматриваемый период обучения большее количество «хороших» оценок

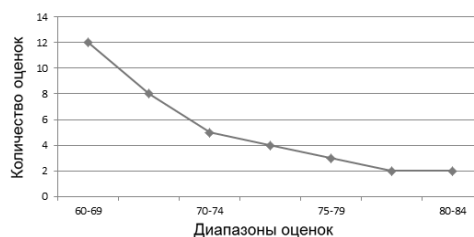


Рис. 7. Функция, характеризующая студента 2, набравшего за рассматриваемый период обучения большее количество «удовлетворительных» оценок

Студенты, у которых наблюдается полимодальный закон распределения оценок, в презентативную группу не включаются, т.е. уровень оценок студентов имеет значительный разброс и расчет арифметического среднего для них не является корректным. Такая функция имеет две или более мод (т.е. два или более экстремумов). Пример такой функции показан на рис. 8.

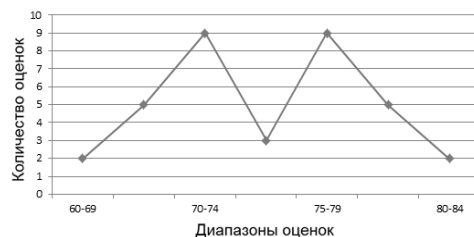


Рис. 8. Функция, характеризующая студента 3, набравшего за рассматриваемый период обучения одинаковое количество «хороших» и «удовлетворительных» баллов

2.3.2. Отклонение оценок по тесту от среднего бала

По результатам тестирования для студентов, выбранных в презентативную группу, строится диаграмма, показывающая от-

клонение оценок за тест от среднего бала каждого студента (рис. 9).

Величина этого отклонения определяет качество обучения и позволяет контролировать уровень знаний студентов по тестируемому предмету.

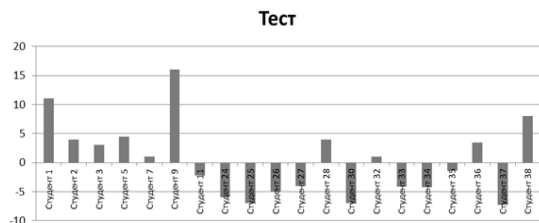


Рис. 9. Отклонение оценок по тесту от среднего бала

3. Рекомендации по повышению качества теста

Проанализировав значения критериев качества теста, рассматриваемых данной методикой, можно предложить некоторые рекомендации по повышению качества теста.

1. Анализ сложности

1.1. Переработать тест в целом, если:

более 10–15 % студентов набрали за тест отличные оценки;

- более 5 – 10 % набрали неудовлетворительные балы.

1.2. Заменить либо доработать тестовые вопросы, у которых сложность выходит за пределы (0,15–0,9).

2. Анализ надежности теста.

Исключить либо доработать тестовые вопросы, у которых коэффициент корреляции меньше или равен 0,4.

3. Анализ валидности теста.

Для студентов, входящих в презентативную группу по тестируемому предмету на основании диаграммы, показывающей отклонение оценок за тест от среднего бала каждого студента, сделать выводы о сложности теста и возможном изменении ряда вопросов.

4. Выводы

Приведенная методика позволяет каждому преподавателю на основе представленных диаграмм и таблиц наглядно проанализировать качество разработанных им тестов по своему предмету, заменить или исключить проблемные вопросы, правильно расставить акценты на главных темах в изучаемом мате-

риале, обратить внимание на пробелы в знаниях конкретных студентов.

Данная методика была использована для оценивания качества теста по предмету «Моделирование и анализ ПО» кафедры «Системное программное обеспечение» для студентов, обучающихся по направлению 6.050103 «Программная инженерия». Количество студентов, проходящих пробное тестирование – 46, количество тестовых вопросов – 30. Анализ сложности каждого вопроса показал, что рекомендуется усложнить двенадцать вопросов и упростить один вопрос (рис. 4). По результатам анализа надежности тестовых вопросов рекомендуется заменить или доработать пятнадцать вопросов (рис. 5). Анализ валидности показал, что из 30 студентов, попавших в презентативную группу, 1/3 студентов получили примерно на 10 баллов выше их среднего значения.

По результатам пробного тестирования тестовые вопросы были изменены. Анализ сложности показал, что изменений не требуется (рис. 10). Исходя из анализа надежности следует пересмотреть вопрос № 5 (рис. 11). Анализ валидности показал, что все студенты, которые попали в презентативную группу, набрали балы, не отличающиеся от среднего больше, чем на 5 баллов.



Рис. 10. Диаграмма, отображающая сложность каждого вопроса после использования методики и перефразирования нечетких вопросов



Рис. 11. Диаграммы, отображающая коэффициент корреляции каждого вопроса после использования методики и изменения «плохих» вопросов

Представленная выше методика расчетов позволяет повысить качество тестов, а также оценить степень и качество усвоения материала студентами.

Список использованной литературы

1. Зорин, С. Ф. Разработка автоматизированной системы контроля знаний / С. Ф. Зорин. – М. :2007. – 36 с.

2. Майоров, А. Н. Теория и практика создания тестов для системы образования: Как выбирать, создавать и использовать тесты для целей образования./ А. Н. Майоров. – М. : Интеллект-Центр, 2002. – 296 с.

3. Мелецинек, А. Международное общество по инженерной педагогике / А. Мелецинек, В. Приходько, В. Жураковский// Высшее образование в России: Науч. пед. журн. – М. : 2004. – № 3. – С. 53–60

4. Русинова, Н.В., Методика проведения анализа тестов по результатам тестирования. / Н. В. Русинова, Т. В. Онищенко, В. А. Крисилов – Одесса: Тр. Одесского политехн. ун-та. – 2007. – Вып. 1(13). – С. 125 – 128.

5. Теория вероятностей и математическая статистика в задачах: Учеб. пособие для вузов / В. А. Ватутин, Г. И. Ивченко, Ю. И. Медведев, В. П. Чистяков— 2-е изд., испр. – М. : Дрофа, 2003.– 328 с.

6. Тertyshnaya, Т. И. Автоматизированная система контроля знаний / Т. И. Тertyshnaya, Е. В. Колесникова, В. Д. Гогунский – Одесса : Тр. Одесского политехн. ун-та. – 2001. – Вып. 1(13). – С. 125 – 128.

7. Чельшкова, М. Б. Теория и практика конструирования педагогических тестов: Уч. пособие / М. Б. Чельшкова – М. : Логос, 2002. – 432 с.

Получено 10.09.2012

References

1. Zorin, S. F. Development of the automated control system of knowledge / S. F. Zorin. – Moscow : 2007. – 36 p. [in Russian].

2. Major, A. Theory and practice of tests for the education system: How to choose, create, and use tests for educational purposes / A. Major. – Moscow : Intelligence Center, 2002. – 296 p. [in Russian].

3. Melezinek, A. International Society for Engineering Education / A. Melezinek, V. Prikhodko, V. Zhurakovsky // Higher Education in Russia. – 2004. – № 3. – P. 53–59 [in Russian].

4. Rusinova, N.V. The methodology of the analysis of the tests on the test results.: Proc. Odessk. Polytechnic. University / N. V. Rusinova, T. V. Onishchenko, V. A. Krisilov. – Odessa: – 2007. – Issue. 1 (13.) – P. 125–128 [in Russian].

5. The theory of probability and mathematical statistics in tasks: Manual for high schools – 2nd ed., rev. / V. A. Vatutin, G. I. Ivchenko, U. I. Medvedev, V. P. Chestyakov. – Moscow: Drofa, 2003. – 328 p. [in Russian].

6. Tertyshnaya, T. I. Automated monitoring system of knowledge / T. I. Tertyshnaya, E. V. Kolesnikova, V. D. Gogunsky. – Odessa: Proc. Odessk. Polytechnic. University. – 2001. – Issue. 1 (13). – P.125–128 [in Russian].

7. Chelyshkova, M. Theory and practice of designing pedagogical tests: Tutorial / M. Chelyshkova. – Moscow : Logos, 2002. – 432 p. [in Russian].



Городничая
Екатерина Александровна,
магистр каф. Системного
программного обеспечения
Одесского нац. политехн. ун-та
пр. Шевченко, 1 0630204303
e-mail:
katherine.gorodnichaya@ukr.net



Онищенко
Татьяна Викторовна,
ст. преп. каф. Системного
программного обеспечения
Одесского нац. политехн.
ун-та



Крисилов
Виктор Анатолиевич,
д.т.н., зав.каф. Системного
программного обеспечения
Одесского нац. политехн.
ун-та