OPERATIONAL STATISTICAL ANALYSIS OF THE RESULTS OF COMPUTER-BASED TESTING OF STUDENTS

V.I. Nardyuzhev[1], I.V. Nardyuzhev[2], V.E. Marfina[1], I.N. Kurinin[1]

[1] Peoples' Friendship University of Russia (RUDN University)
Miklukho-Maklaya str., 6, Moscow, 117198, Russian Federation
[2] Software development department, JSC "The Seventh Continent"
Building, 21, MKAD 47, v. Govorovo, Moscow, Russia, 142784

The article is devoted to the issues of statistical analysis of results of computer-based testing for evaluation of educational achievements of students. The issues are relevant due to the fact that computer-based testing in Russian universities has become an important method for evaluation of educational achievements of students and quality of modern educational process. Usage of modern methods and programs for statistical analysis of results of computer-based testing and assessment of quality of developed tests is an actual problem for every university teacher. The article shows how the authors solve this problem using their own program "StatInfo". For several years the program has been successfully applied in a credit system of education at such technological stages as loading computer-based testing protocols into a database, formation of queries, generation of reports, lists, and matrices of answers for statistical analysis of quality of test items. Methodology, experience and some results of its usage by university teachers are described in the article. Related topics of a test development, models, algorithms, technologies, and software for large scale computer-based testing has been discussed by the authors in their previous publications which are presented in the reference list.

Key words: student, evaluation of educational achievements, computer-based testing, statistical analysis, quality of tests, program, technology

Issues of a development and usage of pedagogical tests are studied in a huge range of scientific researches and publications. Among these issues we can name theories of pedagogical tests [1; 21; 23] and their terminology [3], tests' design [2], methodological rules of pedagogical tests formation [4; 12], test administration [6; 21], statistical analysis and interpretation of obtained results [20; 22]. Researches of computer tests usage for assessing students' academic achievements were widely presented in the following publications covering such themes as a comparability of paper and computer testing [18], computer adaptive testing [5; 14], development of models, algorithms, complex technologies and information-computational systems of computer testing [10], administration of massive testing sessions of students in computer classes and by Internet [13; 17], statistical analysis of results of students' computer testing [19; 20], usage of computer testing in a practical work of teachers [8; 11], governmental registration of computer testing programs and test questions banks [16; 19].

Results of computer testing characterize students' academic achievements and support evaluation of educational process's quality. In this situation such questions as operational

statistical analysis and correct evaluation of test results as well as quality control of test materials have become of great importance for teachers.

Statistical Analysis in Computer-Based Testing. The collection and analysis of statistical information provide feedback, without which testing cannot be considered a scientific method for controlling knowledge and evaluating academic achievements. Conditionally, all statistics can be divided into two parts: 1) information considering participants and test results; 2) data on the quality of test materials.

When working with the first part the system of computer testing usually allows selecting the object of statistical processing. For the selected object you can get the number of participants by categories, the division of the participants by their scores, the percentage of correct answers for each test question. Based on the obtained data the system of computer testing can build diagrams, make a comparative analysis of the results of testing various chosen objects. It allows seeing the list of participants and, if needed, performing a quick search and desired sorting. It is possible to print out statistical forms and lists or parts of them. Finally, you can export the data of interest for analysis in a special statistics package.

To assess the quality of questions in the test you can use methods of statistical analysis. For this purpose the test must be performed by a sufficiently large number of participants. Based on the collected and processed protocols you can draw conclusions considering the quality of the tests. This should be done to improve the quality of the tests and to correct questions to use them in the following sessions of computer testing. Often with the help of the computer testing system the minimum, average, maximum value, standard deviation and coefficient of variation of the obtained test scores are determined. You can determine the number of tests, in which the question of interest appeared, and the number of participants who took part in the test. The following characteristics of the test tasks

of participants who took part in the test. The following characteristics of the test tasks can be calculated:

1) The difficulty of the question defined by the percentage of participants who have performed the task correctly. The range of the question difficulty should be from 20% to 80%. It is necessary to exclude from the bank (or correct the text of) the most difficult questions (which were performed correctly less than by 20% of the participants) and the simplest ones (which were performed correctly more than by 80% of the participants).

2) The differentiating ability of a question is a measurement of its validity. It is calculated as a difference between the two values: the difficulty of the question for a strong group of participants and its difficulty for a weak one. The higher the differentiating ability of a question, the better it divides the tested by the level of preparation.

3) Point-biserial correlation allows estimating, how the performance of this question is connected with the performance of the entire test. In other words it is an indicator of the suitability of the question for the test.

The analysis of the work of distractors in the question determines how each of the proposed answers "operates" in each task. The data of the distraction analysis makes it possible to draw conclusions about the existence of "good" questions, "badly" formulated questions, which the participants do not understand, question that need material that students have not studied so far, and questions that do not contain the correct answer. The distraction analysis includes obtaining the following list of characteristics: 1) the frequency of the choice of the correct response to the task (in percentage); 2) response

ИННОВАЦИОННЫЕ ПЕДАГОГИЧЕСКИЕ ТЕХНОЛОГИИ В ОБРАЗОВАНИИ

rate (in percentage); 3) the differentiating ability of the response; 4) deviation of frequency of the wrong answer selection from the average value, etc.

To determine complex characteristics, such as reliability, validity, test effectiveness, special statistical packages are used. For example, reliability analysis in SPSS [12] can be used to select the most suitable test questions. In order to do so a preliminary version of the test is developed which has an excessive number of questions. It is tested on a fairly representative number of students. Then a reliability analysis is carried out, which allows eliminating unsuitable questions.

The bank of test questions for computer testing should clearly reflect the structure of the test for a specific subject. Each question of the test should have a set of variants of different level. All statistical parameters of these test question variants must be real. They can be determined on the basis of statistical analysis of the quality of trial tests used in different groups of students. Normally, the volume of the analyzed sample should go up to several hundred results for each question. The algorithm of inserting questions to the bank is based on the results of the specified statistical parameters calculation and on excluding inappropriate questions.

"StatInfo" Program. To solve the problems that a teacher faces the authors of the article developed the technology of the operative statistical analysis of the computer testing results [9]. It is based on the functionality of the "StatInfo" program, which has the state registration certificate]. The effectiveness and reliability of the technology has been proven by many years of applying it to assess the academic achievements of students in the credit system of education. In the article [19], the authors showed as an example the results of an operative statistical analysis of the real results of the final computer testing of students at the "Informatics" course for 15 semesters. The "StatInfo" program runs under the operating system MS Windows. The informational basis of this program is the table-based (relational) database. The material accumulated in the database can be used for scientific and practical work in the field of computer testing and monitoring the quality of education. The program can be installed locally or in a network.

The "StatInfo" program allows a teacher: to select the object of statistical processing (group, faculty, etc.) and get the number of participants for this object, get their distribution by obtained scores or by points taking into account the transfer of test scores to the credit points, get the percentage of correct answers to the test question. Based on the received data, it is possible to build diagrams, perform a comparative analysis of the results of testing various objects of statistical processing, look at the lists of participants, and print out statistical forms and lists (or parts of them). If necessary, any statistical form or list can be saved in a separate text file or transported to electronic worksheets to build necessary diagrams and tables.

The list of table fields with the results of computer testing is available for the formation of queries. An essential feature of the program is the "flexibility" of a possible query. The "Query" dialog contains a filter table in which, for each column of the information database, you can specify the values range allowed in the query. To form the received results the "Statistics" button opens a dialog where you can select a necessary statistical form. The following options are available: 1) the number of test participants; 2) received scores; 3) received points; 4) correct answers.

An important role in the system is assigned to the lists of participants. To switch to working with lists of participants, use the "List" button. In this mode, you can perform the following operations: 1) select the necessary attributes of the test participants and determine the order of their succession for viewing on the screen (printing on the printer) by pressing the "Columns" button; 2) by pressing the "Sort" button, order the list of participants by any combination of attributes (by default, entries in the list are sorted by last name); 3) by pressing the "Search" button, search participants in the database by asking any combination of attributes (by name, subject code, number of points scored, etc.). Before starting the search, the list must be sorted by the according attributes.

The storage of all necessary information (results of computer testing) in a single relational (table-based) database makes it possible to significantly speed up the process of calculating basic values and allows the author of the test to quickly move from the figure received in the report to the test question it came from.

To perform calculations in MS Excel, a response matrix is often required. It is represented by a rectangular table, in each position of which the responses of the testing participant are indicated. Usually the line number corresponds to the number of the participant, and the column number corresponds to the number of the test question. To make a matrix of answers in the "StatInfo" program you must select the "Responses" column (students answers to the test questions) or the "Verification" column (the results of checking students' answers to the tests questions) and in the "Export" mode make a text file (with the separator " ; "). The matrix of responses can additionally determine the median, mode, excess, asymmetry and other statistical characteristics of test tasks.

Conclusion. The technology of the operative statistical analysis of the results of computer testing developed by the authors for assessing students' academic achievements is based on the functional capabilities of the "StatInfo" program. Using the outcome of the statistical processing of the results of students computer testing errors in test questions (long formulation or lack of a correct answer) are identified and the sections of the course that are well or poorly apprehended by students are determined.

LITERATURE

[1] *Анастази А.* Психологическое тестирование. Кн. 1—2. М.: Педагогика, 1982. 320 с.

[2] *Аванесов В.С.* Композиция тестовых заданий. М.: Арена, 2002. 240 с.

[3] *Балыхина Т.М.* Словарь терминов и понятий тестологии. М.: Изд-во РУДН, 2000. 164 с.

[4] *Васильев В.И., Демидов А.Н., Малышев Н.Г., Тягунова Т.Н.* Методологические правила конструирования компьютерных педагогических тестов. М.: Изд-во ВТУ, 2000. 64 с.

[5] *Васильев В.И., Тягунова Т.Н.* Основы культуры адаптивного тестирования. М.: Икар, 2003. 584 с.

[6] *Ефремова Н.Ф.* Современные тестовые технологии в образовании: учеб. пособие. М.: Логос, 2003. 176 с.

[7] *Куринин И.Н., Нардюжев В.И., Нардюжев И.В.* Компьютерное тестирование в оценке учебных достижений студентов: учеб.-метод. пособие. М.: Изд-во РУДН, 2008. 308 с.

[8] *Куринин И.Н., Нардюжев В.И., Нардюжев И.В.* Сборник тестовых заданий по курсам «Информатика» и «Компьютерные технологии в науке и образовании». М.: Изд-во РУДН, 2010. 306 с.

[9] *Куринин И.Н., Нардюжев В.И., Нардюжев И.В.* Оперативный статистический анализ результатов компьютерного тестирования в кредитной системе обучения // Вестник Рос-

сийского университета дружбы народов. Серия: Информатизация образования. 2013. № 1. С. 115—125.

[10] *Куринин И.Н., Нардюжев В.И., Нардюжев И.В.* Комплексная технология компьютерного тестирования // Вестник Российского университета дружбы народов. Серия: Информатизация образования. 2013. № 2. С. 112—121.

[11] *Куринин И.Н., Марфина В.Е., Нардюжев В.И., Нардюжев И.В.* Информатизация практической работы преподавателя // Вестник Российского университета дружбы народов. Серия: Информатизация образования. 2015. № 1. С. 42—52.

[12] *Майоров А.Н.* Тесты школьных достижений: конструирование, проведение, использование. СПб.: Образование и культура, 1996. 304 с.

[13] *Нардюжев В.И., Нардюжев И.В.* Тестирование на компьютерах через Интернет // Тр. Центра тестирования. Выпуск 2. М.: Прометей, 1999. С. 139—157.

[14] *Нардюжев В.И., Нардюжев И.В.* Достоинства и недостатки компьютерного адаптивного тестирования / Развитие системы тестирования в России: тезисы док. 2 Всеросс. науч.-практ. конф. М.: Прометей, 2000. Часть 3. С. 47—48.

[15] *Нардюжев В.И., Нардюжев И.В.* Модели и алгоритмы информационно-вычислительной системы компьютерного тестирования: монография. М.: Прометей, 2000. 148 с.

[16] *Нардюжев В.И., Нардюжев И.В.* Свидетельство РОСПАТЕНТ № 2000610068 от 27 января 2000 г. об официальной регистрации программы для ЭВМ «Оперативный статистический анализ результатов тестирования на компьютерах через Интернет (СТАТИНФО)».

[17] *Нардюжев В.И., Нардюжев И.В.* Анализ уровня компьютерной грамотности участников централизованного компьютерного тестирования // Вопросы тестирования в образовании. М.: Центр тестирования Минобразования России, 2001. Вып. 1. С. 86—93.

[18] *Нардюжев В.И., Нардюжев И.В.* Сравнение результатов централизованного тестирования на бланках и на компьютерах / Педагогическая диагностика. М.: Народное образование, 2002. 10 с.

[19] *Нардюжев В.И., Нардюжев И.В., Марфина В.Е., Куринин И.Н.* Свидетельство № 2016620662 от 24 мая 2016 г. о государственной регистрации базы данных «Тестовые задания. Компьютерные технологии в науке и образовании». Выдано Федеральной службой по интеллектуальной собственности (РосПатент).

[20] *Наследов А.* SPSS 19: профессиональный статистический анализ данных. СПб.: Питер, 2011. 400 с.

[21] *Нейман Ю.М., Хлебников В.А.* Введение в теорию моделирования и параметризации пе-

дагогических тестов. М.: Прометей, 2000. 168 с.

[22] *Тюрин Ю.Н., Макаров А.А.* Статистический анализ данных на компьютере. М.: ИНФРА-М, 1998. 528 с.

[23] *Челышкова М.Б.* Разработка педагогических тестов на основе современных математических моделей: учеб. пособие. М.: Исследовательский центр проблем качества подготовки специалистов, 1995. 32 с.

INNOVATION PEDAGOGICAL TECHNOLOGIES IN EDUCATION