Application of a method for student academic performance control

R. Kudarov¹*, I. Mikulik¹, and E. Blagoveschenskaya¹

¹Petersburg State Transport University, 9 Moskovsky pr, 190031, Saint Petersburg, Russian Federation

Abstract. A method for monitoring the results of academic performance of students in a mathematical discipline is considered. As part of the monitoring method, a mathematics testing system was developed, implemented on the basis of the Moodle distance learning system platform. A bank of questions was compiled, including more than 600 tasks in elementary Mathematics and more than 900 tasks in Further Mathematics, on the basis of which tests were created to check the progress of students. To analyze the results of intermediate testing in mathematics, the rank correlation coefficients of Spearman and Kendall were calculated. A comparison of the results of tests conducted in the classroom and remotely.

Keywords: Monitoring methods, statistical analysis, distance learning system.

1 Introduction

The article is devoted to the analysis of the grades in mathematics received by the students of PSUPS at various stages of education. The assessments for the Unified State Exam, initial testing, current control and intermediate attestation are considered.

The data analyzed in this article is the result of a huge collective work of the teaching staff of the Department of Higher Mathematics. In this regard, research mainly affects the results of education in mathematics at the university. However, given the role of Mathematics in education, the conclusions drawn in the work may apply not only to the study of one discipline. This is due to the fact that the course of the discipline «Higher Mathematics» is a basic general educational course.

As initial data, it is advisable to consider:
- the results of student learning at the university (control cuts of the current attestation, intermediate attestation, testing for knowledge);
- school achievements of students (exam, entrance examinations);
- the basis of the student's education agreement (budgetary, targeted, commercial).

Assessment of students upon entrance can be confirmed or refuted only by dynamic monitoring of students' progress during the course of study [1].

An essential factor is the inevitability of the introduction of distance learning. Obviously, distance learning at the university can be considered only as an additional form of training

* Corresponding author: r.s.kudarov@gmail.com

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).
and knowledge control. Monitoring students in distance learning systems in the education process is an effective technique for detecting students at risk [2]. But we must understand that distance learning is a new form of education, along with classical and correspondence. The development of appropriate techniques, technical support, teaching materials is required. However, the problem requires a systematic approach [3].

The compilation of a bank of questions is the object of study. The aim of the work is to analyze the developed monitoring (testing) system. It is necessary to understand how adequately the proposed testing model reflects the actual knowledge of students. Whether the results of intermediate testing are able to predict the result of the final attestation. It is also necessary to find out whether the initial data given above affect the result of the final exam, as well as intermediate testing.

2 Development of an electronic testing system

To register progress at the university, a special testing system in Mathematics has been developed. The testing system is implemented on the basis of the Moodle distance learning system. The Moodle distance learning system is used in many countries [4][5][6]. The testing system covers most of the sections of elementary mathematics, as well as several sections of further mathematics.

In the course of the work, an extensive bank of questions was programmed (more than 600 in elementary mathematics and more than 900 in further mathematics), on the basis of which tests were created.

The special testing system has a single platform, which is the educational course. Along with access to the developed bank of questions, ready-made type tests are provided for individual sections of mathematics.

The same unified educational course contains materials necessary for working in Moodle and errors found in the bank of questions.

The collection of data for subsequent analysis in the work involved the obligatory accounting of test results within the framework of current control (summer semester, 2020-2021 academic year) and grades for other controls at the discretion of teachers conducting research in their group.

The current control test is planned in such a way that the developed special testing of the distance learning system will gain experience in implementation in various scenarios:

- in display classes;
- on smartphones;
- on paper;
- at home.

The issue of receiving and storing students' video identification files has been resolved separately. The server has e-mail addresses for each study group, with a single password.

3 Analysis of intermediate test results in Mathematics

The section provides an analysis of the results obtained using a special testing system. Correlation coefficients was calculated.

3.1 Data obtained from the test

Testing was carried out in three groups of the first course.

There are 23 students in the first group. Students of this group were tested on smartphones (tablets and other gadgets) at home during the hours of independent work. Twenty students
took part in the test. The test results were distributed as follows: 9 students solved 9 to 12 tasks; 8 students solved 12 to 16 tasks, 2 students solved 16 to 20 tasks, 1 student solved 6 tasks. Figure 1 shows the percentage of students in the first group who coped with different numbers of tasks.

![Graph showing percentage of students in the first group who coped with different numbers of tasks.]

**Fig. 1.** Percentage of students of the first research group who completed a certain number of tasks.

In the second group there are 22 students. Testing in this group took place in the classroom during practical classes.

The test results were distributed as follows: 8 students solved from 9 to 12 tasks; 5 students solved from 12 to 16 tasks; 8 students solved from 16 to 20 tasks; one student solved 3 tasks. Figure 2 shows the percentage of students in the second group who coped with different numbers of tasks.

![Graph showing percentage of students in the second group who coped with different numbers of tasks.]

**Fig. 2.** Percentage of students of the second research group who completed a certain number of tasks.

17 students participated in the third group. The test results were distributed as follows: 9 students solved from 9 to 12 tasks; 5 students solved from 12 to 16 tasks; 3 students solved from 16 to 20 tasks. Figure 3 shows the percentage of students in the third group who coped with different numbers of tasks.

![Graph showing percentage of students in the third group who coped with different numbers of tasks.]

**Fig. 3.** Percentage of students of the third research group who completed a certain number of tasks.
Combining the results, you can get the following histogram: there are 59 participants in total. It is shown in Figure 4.

Thus, it can be seen that about half of the students: 26 out of 59 (44%) showed satisfactory results, 18 people (30.6%) good results, 13 people (22%) very good results and 2 people (3.4%) poor results.

The arithmetic mean number of solved tasks is $\bar{X} = 12.78$.

Given the complexity of the test, 13 tasks can serve as a guideline for assessing the student's knowledge. 8-13 tasks are satisfactory result. From 13 to 16 tasks are a good score, 17-20 tasks are an excellent score.

In each of the groups, conversations were held, during which all students noted that classroom education significantly helps the assimilation of the material, creates a positive energy that motivates to better study. Most of the students noted that a clear understanding of the prospects for further work with decent pay after graduation would be a serious motivation to study. This circumstance undoubtedly affects the interest in learning.

3.2. Calculating the rank correlation coefficients

Note that the mark for the exam is a qualitative indicator. It’s a rank parameter, so it, in particular, serves to organize the examinees by gradation, but is not a numerical expression of the amount of knowledge, it cannot be said that a student who received 4 scores knows exactly twice as much than the one who received 2 ones.
The ranking system makes sense. One of the known methods for calculating the correlation of such quantities is the Spearman correlation and Kendall correlation [7].

Method for calculating Spearman's rank correlation coefficient $\rho$ includes:
1. Each student is assigned a rank according to one and another criterion, that is, the result of one or another method of attestation. The ranks are assigned in ascending order. In this case, in the group of identical test results, all ranks are assigned equal. The arithmetic mean of the ranks is taken, which would have been assigned if the initial data were slightly different.
2. The vector of the difference of rank vectors is calculated.
3. To obtain the coefficient $\rho$, the sum of the squares of the coordinates of the difference vector divided by the number of combinations from $n + 1$ to 3, where $n$ is the dimension of the array:

$$\rho = 1 - \frac{\sum_{i=1}^{n}(X_i - Y_i)^2}{C_{n+1}^3},$$

Where $X$ is a vector of ranks according to one criterion, $Y$ is a vector of ranks according to another criterion, $n$ is the dimension of vectors $X$ and $Y$.

Kendall coefficient $\tau_A$ is inapplicable in this case due to repeated values for both $X$ and $Y$. So, we need to use $\tau_B$. The calculation of it includes the next.

The first of all, introduce definitions.
1) pairs are called concordant (by indices $i$ and $j$), which are strictly ordered in both arrays in the same direction ($X_i > X_j \land Y_i > Y_j \lor X_i < X_j \land Y_i < Y_j$).

Pairs (by indices $i$ and $j$) are called discordant if they are strictly ordered in both arrays in different directions ($X_i > X_j \land Y_i < Y_j \lor X_i < X_j \land Y_i > Y_j$).

Pairs whose values in at least one of the arrays are equal ($X_i = X_j \lor Y_i = Y_j$) are neither concordant nor discordant.

The coefficient $\tau_B$ is calculated as:

$$\tau_B = \frac{XY_c - XY_d}{\sqrt{X_d Y_d}},$$

Where $XY_c$ is the number of concordant pairs between $X$ and $Y$, $XY_d$ is the number of discordant pairs between $X$ and $Y$, $X_d$ is the number of pairs that differ in $X$, that is, $X_i \neq X_j$, $Y_d$ is the number of pairs that differ in $Y$, that is, $Y_i \neq Y_j$.

Correlation coefficients were calculated programmatically.

Rank correlation coefficients of test passing and attestation results were calculated programmatically. For this, a Python data analysis library Pandas was used. The following lines of code had to be written:

```python
DataFrame.corr(method='spearman')
DataFrame.corr(method='kendall')
Spearman's coefficient $\rho = 0.82$.
Kendall's coefficient $\tau_B = 0.66$.
```

According to the results, we can say that the results of passing the compiled electronic tests are in good agreement with the result of the intermediate attestation.
4 Analysis of intermediate attestation in mathematics

The section provides an analysis of the results of passing the winter session exam. The section describes the dependence of the success of passing the exam, on the parameters: the results of entrance examinations, the form of education, scores for individual achievements.

There are three basic forms of higher education in Russian Federation: public budget sponsored education or budget places, employer-sponsored education or target places, self-sponsored education or commercial. In Figure 5 is a histogram showing the average total score for the results of entrance examinations of students of various forms of education.

![Histogram of entrance examination scores](image)

**Fig. 5.** Average total score for the results of entrance test.

The dotted line indicates the average total score for the entire sample. As you can see, the highest total score is among students of the budgetary form of education. The average score of targeted students is lower than that of budget students, but higher than that of commercial students.

On Figure 6 is a bar graph showing the percentage of students in each of the three groups who successfully passed a Mathematics test or exam at the university. As can be seen from the histogram, in percentage terms, the students of the targeted education did the best.

![Bar graph of mathematics test pass rates](image)

**Fig. 6.** Percentage of students who successfully passed a Mathematics test.

Figure 7 shows a histogram showing the percentage of students in each of the three groups who failed the Mathematics exam.
The modules of the correlation coefficients of the mark were obtained from the parameters: the form of education, the total score of the entrance tests, the score for individual achievements. The correlation coefficients were obtained using the Python script:

- 0.213039 is for the form of education,
- 0.1893 is for the total score,
- 0.056968 is for the score for individual achievements.

The modulus of the coefficient indicates the degree of dependence: the closer its value is to 0, the weaker the linear dependence. The closer the correlation coefficient is from zero to one, the stronger the direct linear dependence.

It can be seen that, in general, there is no strong direct dependence on the presented coefficients. Nevertheless, it can be noted that the mark depends more on the form of study, less on the total score of the entrance tests.

5 Comparison of test results in classrooms and remotely

To compare the results of classroom and distance testing for students, testing on the studied material was carried out twice: in the classroom on forms and in the Moodle, system using a smartphone.

The average score obtained on the forms is 12.43, which is higher than the Moodle score (10.12). The proposed tasks were selected from one bank of tasks. If students' knowledge has not changed, then we have two pairwise dependent samples X and Y of the same size, since we measured the same value in different ways. Now it is necessary to establish whether the obtained pairs of points differ statistically significantly.

Let’s formulate this problem with statistically terms. It is required at the level of significance $\alpha$ to test the null hypothesis $H_0$ about the equality of general means $M (X) = M (Y)$ with a competing hypothesis $H_1$ $M (X) \neq M (Y)$ for two dependent samples of the same size. Consider the difference of random variables $Z = X - Y$. If the null hypothesis is true, i.e., $M (X) = M (Y)$, then $M (X) - M (Y) = 0$ and $M (Z) = 0$.

In order to test the null hypothesis $H_0$ ($M (X) = M (Y)$) about the equality of two mean normal populations in the case of dependent samples of the same size, in order to test the null hypothesis $H_0$ ($M (X) = M (Y)$) at a given significance level $\alpha$, the observed value of the criterion must be calculated:

$$T_{obs} = \bar{Z} \sqrt{n} / s_Z,$$

Where $\bar{Z}$ is the mean value of the differences;
$n$ is the sample size. Standard deviation of the differences:
\[ s_z = \sqrt{\frac{\sum_{i=1}^{n} z_i^2 - (\sum_{i=1}^{n} z_i)^2/n}{n-1}} \]

The value of T has a Student's t distribution [8] with \( k = n - 1 \) degrees of freedom. Calculate this value programatically:

\[ T_{\text{obs}} = 2.42 \]
\[ T_{\text{cr}} = 2.07 \]

The null hypothesis is rejected because \( |T_{\text{obs}}| > T_{\text{cr}} \).

Other useful values have also been calculated.

The correlation between the scores of the first test and the USE is 0.404.

The correlation between the scores of the second test and the USE is 0.492.

The correlation between the scores for the initial testing on the forms and in the remote platform is -0.371.

### 6 Conclusions

Initial testing was organized and its analysis was carried out for knowledge of school mathematics. Initial testing is used to assess the level of training, both an individual student and the group as a whole. The initial testing made it possible to implement a number of measures for the successful mastering by students of sections of mathematics and related special disciplines. Such activities include, for example, conducting additional classes in mathematics. This is especially important when reducing the number of hours devoted to the study of mathematics, as well as when introducing distance learning [9].

A special testing system for mathematics in the distance learning system is proposed. With its help, the current attestation in Mathematics was carried out in all groups of the first year in the summer semester 2020-2021.

An analysis of the results of the winter examination session (2020-2021) in mathematics showed that students of targeted education form are better with intermediate attestation than students of budgetary and commercial education form. There was no strong linear relationship between the result of passing the exam and the form of education, entrance scores and individual achievements. However, the form of education is the more influential of the three.

### References

1. E.N. Ogor, Electronics, robotics and automotive mechanics conference (2007) DOI: 10.1109/CERMA.2007.4367712
5. Z.P. Aminova, Modern Science 5-1, 276-278 (2020)