

USE OF DIGITAL INSTRUMENTS DURING PROCESSING OF RESULTS OF PEDAGOGICAL RESEARCH

Tetiana V. Polishchuk¹, Dariia A. Voznosymenko¹

¹*Pavlo Tychyna Uman state pedagogical university, street Sadova 2, Uman, 20301, Ukraine*

Abstract

Pedagogical research is an important component of the qualification project for the attainment of masters by those who have chosen the profession of teaching. The article highlights the features of using digital tools to automate statistical analysis of data obtained during pedagogical research. Future teachers of physics, mathematics, computer science and natural sciences while studying the discipline «Mathematical apparatus pedagogical science» were invited to conduct a statistical analysis of data using ICT. The expediency of processing the results of a pedagogical experiment with the help of digital tools, in particular the package of dynamic mathematics GEOGEBRA, is investigated. An example showing the stages of implementation and the procedure of processing (calculations) of research results performed in the mathematical package GEOGEBRA is shown later in this paper. In particular, the expediency of using the capabilities of GeoGebra to calculate averages, statistical Student's criterion and Pearson, visualization of results on 2D-window in the form of a frequency histogram, table, etc. is substantiated. According to the results of the research, it is established that the use of GEOGEBRA dynamic mathematics package for calculation of arithmetic mean, standard deviation, empirical value of statistical criteria, etc. allows to automate statistical analysis of data obtained during pedagogical research.

Keywords

information and communication technologies, pedagogical experiment, statistical analysis, statistical criteria's, GEOGEBRA.

1. INTRODUCTION

1.1. The urgency of the problem.

The introduction of modern information and communication technologies (ICT) in the educational process has resulted in the ability to quickly and efficiently process a significant amount of data obtained as a result of the introduction of scientific research. It is especially important for the researcher to be able to perform such important tasks as obtaining truly scientific conclusions, provability of factual material, objectivity of experimental data. To solve them, it is advisable to use mathematical methods of processing experimental materials, which are aimed at eliminating some of the shortcomings of verbal methods of pedagogical research (descriptive nature, the operation of disparate facts, etc.). The specificity of statistical processing of the results of the pedagogical experiment is that the database being analysed is characterized by a large number of indicators of different types, their high variability under the influence of uncontrolled random phenomena, the complexity of the correlation processes between the variable samples.

Proceedings of the Second Workshop on Technology Enhanced Learning Environments for Blended Education, October 5–6, 2021, Foggia, Italy


EMAIL: polishchuk_t@ukr.net (T. V. Polishchuk); daryakholod@ukr.net (D. A. Voznosymenko)

ORCID: 0000-0001-7690-7723 (T. V. Polishchuk); 0000-0002-7557-643X (D. A. Voznosymenko)



© 2021 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

Successful combination of statistical methods with modern digital tools speeds up and simplifies the process of processing experimental data and is a mandatory stage of modern pedagogical experimentation.

1.2. Analysis of recent research and publications.

Statistical methods make it possible to systematize, scientifically process and submit research materials, check their scientific validity and reveal the links between the studied phenomena. Some aspects of the application of methods of mathematical statistics in pedagogical research were covered in the articles by domestic and foreign scientists [1 - 7].

M. Grabar believes that it is advisable to use non-parametric methods of mathematical statistics for pedagogical research because the theoretical probabilistic laws that allow us to test the hypothesis of the correspondence of experimental data to one or another distribution law are unknown [4]. The nonparametric methods are «free» from the distribution, which greatly expands their capabilities in contrast to parametric methods such as the Student's criterion and Pearson's linear correlation. This is because nonparametric methods can be applied to any data that is presented in numerical form [8].

There is a large number of scientific and methodical works on probability theory, mathematical statistics, which set out the basic methods of statistical research and consider examples of their application. However, the vast majority of these works do not focus on the capabilities of modern digital technologies, in particular dynamic mathematical packages.

The peculiarities of the use of information and communication technologies during the processing of statistical data of pedagogical research were studied by the following domestic and foreign scientists [9 - 18].

As noted in [18], over the years, several instructional methods have been employed to support students in learning Statistics. Among them, intelligent tutoring systems deserve particular interest, personalizing the learning activities according to individual characteristics. Generally, students enrolled in human and social sciences degrees exhibit a high level of Statistical anxiety and severe math shortcomings that negatively affect their success in Statistics.

The scientists claim that the use of computer technology in scientific and pedagogical research, provides quality training for specialists with a new type of thinking, contributes to the effectiveness of their scientific and pedagogical activities. This is confirmed by the results of the introduction of the course «Information and Communication Technologies in Pedagogical Studies» to the curriculum for future doctors of philosophy in the field of knowledge «Education» [9].

The introduction of information technology in the educational process, according to researchers [10], allows the formation of a "new view" of students on pedagogical research. The use of digital tools contributes to a qualitative change in the approach to working with information obtained during the study at all stages: from the statement stage to the control, quantitative processing of measurement results by various statistical methods, without complex calculations and mathematical transformations. The use of various information technologies and software products during training switches the training of future teachers to a qualitatively new level.

Examples of using MS EXCEL spreadsheets to calculate the empirical value of the criterion (including χ^2), automate statistical analysis, to make decisions about the absence or presence of differences in certain characteristics of the experimental and control groups of pedagogical experiment [2, 11, 12, 18, 19]. The authors of these works point out the expediency and reveal the features of the use of ICT to automate the development of statistical analysis in the training of masters of humanities science.

The advantages of the STATISTICA package among the most well-known professional computer statistical packages for data processing of pedagogical experiment are noted [13, 17]. After all, it has a very wide set of commands for calculating statistical criteria (chi-square, sign criterion, T-Wilcoxon, Kolmogorov-Smirnov criterion, Spearman correlation criterion, Student's criterion, ranking method) and visualization of results (construction of graphs, charts, histograms and output of reports in the form of a table).

An overview of the features of selection, advantages and disadvantages of common statistical packages MS EXCEL, PSPP, PS, NMASTER, STATA, EPI-INFO, EZR and others for data analysis of these, research projects, data processing and report preparation is given [5, 21].

The problem of mathematical processing of results of pedagogical researches with use of separate dynamic mathematical packages remains insufficiently investigated needs further researches.

The purpose of the article is to highlight and analyse the possibilities of using statistical packages for statistical processing of experimental data.

2. RESULTS OF THE RESEARCH

The statistical analysis can be divided into two stages - primary and secondary. Primary processing involves ranking, scaling, descriptive statistics. Secondary - variance, regression, correlation, factor analysis, testing of statistical hypotheses and other methods. Today, there are a large number of computer statistical packages for processing experimental data. The most widespread and popular are professional computer statistical packages: MS EXCEL, NCSS-PASS-GESS, PSPP, STATISTICA PLUS MULTILANGUAGE, STATISTICA, MATHLAB, COSTAT, DELTAGRAPH, LEOSTATISTIC, SYSTAT, PROBABILITY AND PLATICS +.

However, the interactive dynamic mathematical environments GEOGEBRA and DUDAMATH are becoming increasingly popular among the educational community. Specialized programs usually contain tools to support both stages of statistical analysis. Dynamic mathematics packages benefit from the visualization of mathematical objects.

The above packages are affordable, easy to use and unpretentious to the power and specifications of the computer. They are quite popular, often used in teaching and research activities to process statistical and experimental data, including data obtained during a pedagogical experiment.

We propose to consider the main characteristics and features of processing the results of a pedagogical experiment on the examples of common statistical packages and resources. It is important that the list of mathematical packages is not chosen by chance. Students are familiar with the peculiarities of working in these environments, as they worked with them while studying computer science and mathematics courses within their educational programs.

MS EXCEL is an electronic table with powerful mathematical capabilities, in which some statistical functions are additional built-in formulas, that do not allow to fully perform mathematical and statistical processing of the results of the pedagogical experiment. In this package the descriptive statistics and parametric criterion of differences are quite correctly implemented. Investigation of the presence or absence of dependences of differences between two variables and quantitative description of these dependences (correlation, variance, regression analysis), determination of Pearson's linear correlation coefficient for two measured variables in the scale of relations can be implemented in MS EXCEL.

However, nonparametric criteria (Wilcoxon, Mann-Whitney, Fisher) for the two data samples measured on the ordinal scale in MS EXCEL are not implemented. It is necessary to use Spearman's rank correlation coefficients; this is possible when using the professional package. One-way analysis of variance is implemented in MS EXCEL, but in STATISTICA implemented all known methods of analysis of variance.

STATISTICA program is a statistical system for personal computer users, consisting of a number of modules (Basic statistics and tables, Nonparametric statistics, Analysis of variance, Multiple regression, Nonlinear estimation, Time series analysis and forecasting, Cluster analysis, Factor analysis, Discriminant functional analysis, Life expectancy analysis, Canonical correlation, Multidimensional scaling, modelling by structural equations, etc.), which work independently of each other. Each module includes a specific class of procedures. Graphs in this system are built both from the general menu and from the submenu of procedure which offers beginners a choice of adequate graphic representation of data.

The PSPP program provides a full range of statistical analysis: frequency analysis and descriptive statistics, statistical analysis based on conjugation tables, comparison of averages by Student's criterion and one-way analysis of variance, regression analysis, nonparametric analysis methods, factor and cluster analysis, principal components analysis, etc. PSPP performs calculations fairly

quickly regardless of the size of the input data. The results of statistical calculations and graphs can be saved in ASCII, PDF, PostScript, SVG and HTML formats. The program also provides a wide range of histograms and different types of diagrams (columnar, point, circular). PSPP imports spreadsheet (Gnumeric and OpenDocument) data and PostgreSQL databases, as well as comma-separated values (CSV) and ASCII.

STATGRAPHICS + (PLUS) is a fairly powerful statistical program (contains more than 250 statistical functions). The procedures in this program are grouped by type of analysis (for example, «Compare», «Analyze relationships», «Describe»), which greatly facilitates the selection of the necessary procedures. Thus, techniques of parametric and nonparametric statistics are usually in one menu item and can be used at the review of options of this type of the analysis. Each analysis is followed by a brief commentary regarding what has been received and suggestions for the use of additional techniques. GEOGEBRA is the most popular free dynamic math program for all levels of education, combining in one easy-to-use package: geometry, algebra, tables, graphs, statistics and arithmetic, etc. This program differs from others by high-quality visualization of results and data. In the cloud storage free access contains a large number of applets that can be used as a template for calculating averages and parametric statistics.

DUDAMATH is a free digital environment that allows you to calculate and synchronously visualize averages, dispersion, study the normal distribution.

Here are examples of implementation of stages and procedures for processing (calculations) of research results performed in the package of dynamic mathematics GEOGEBRA. It was not by chance the dynamic mathematics package chosen. The future teachers of physics, mathematics, computer science and natural sciences are quite familiar with the tools of this interactive environment. Because in the previous bachelor's program and during the study of disciplines of professional and fundamental training, they actively used it to solve educational problems.

We will give examples of realization of stages and procedures of processing (calculations) of results of research executed in a package of dynamic mathematics GEOGEBRA.

The pedagogical measurements are the process of obtaining quantitative information about the knowledge of learners. Therefore, special evaluation criteria and less accurate evaluation scales (for example, nominal, ordinal and interval) are introduced in pedagogical measurements, which are sufficient for the characteristics of the educational process [1]. Literally, the school classroom in which most pedagogical experiments are conducted is not a random sample. There are always classes with a high level of success, medium and low. Different teaching methods work differently in different classes. There are many factors that affect the course of the experiment. Random choice in pedagogical experiments is difficult to ensure, so they are classified as experiments that involve some vagueness (quasi-experiments). The study of the experimental and control groups is an example of the most commonly used experiment. From the general population we form two groups at random (for example, an online number generator). Then we choose experimental and control from them. It is important that the groups enter the experiment under the same conditions: the same level of initial knowledge, self-esteem, interest, and so on. However, before starting the experiment, it is important to check how true this assumption is. Such a test can be performed by calculating some statistics of both groups, which should be approximately the same. Such indicators may be average.

Statistical methods do not interfere with the educational process, they serve to assess the reliability of the results obtained by the teacher-researcher.

We will follow the algorithm to compare the numerical characteristics of the two groups:

1. We enter the numerical characteristics of the control (G_C) and experimental (G_E) groups in the table.

2. Calculate the mean (arithmetic mean, dispersion, standard deviation).

3. Calculate the coefficient of variation.

4. We calculate the Student's criterion.

5. Calculate the Pearson correlation coefficient.

6. Conclusions that should indicate the presence or difference of advantages in group GE over group GC.

Once we have abstract results of test results that were carried out in control (G_C, column A) and experimental (G_E, column B) groups. Write the results in the form of an ungrouped variation series, in the table (Figure 1). In the command bar, enter the commands «list of point (<list>»», «Sun

(«List»), «Mean (<list>»), «Median (<list>»), «Moda (<list>»), «Variance (<list>»), «SampleVariance (<list>»), «TTest2», «ChiSquaredTest (<list>, <list>» and calculate the arithmetic mean, median, mode, dispersion, standard deviation, coefficient of variation and statistical criteria's of Student 's and Pearson in the G_C and G_E. Using the built-in command «Text» you can display the obtained values in a 2D window (Figure 2).

	A	B	C	D
1	5	8		
2	6	8		
3	9	7		
4	8	9		
5	11	9		
6	9	11		
7	7	8		
8	8	7		
9	7	10		
10	10	9		
11	8	9		
12	9	8		
13	11	7		
14	7	6		
15	9	7		
16	8	5		
17	10	10		
18	7	9		
19	6	7		
20	9	6		

Figure 1. Variation series of test results G_E and G_C.

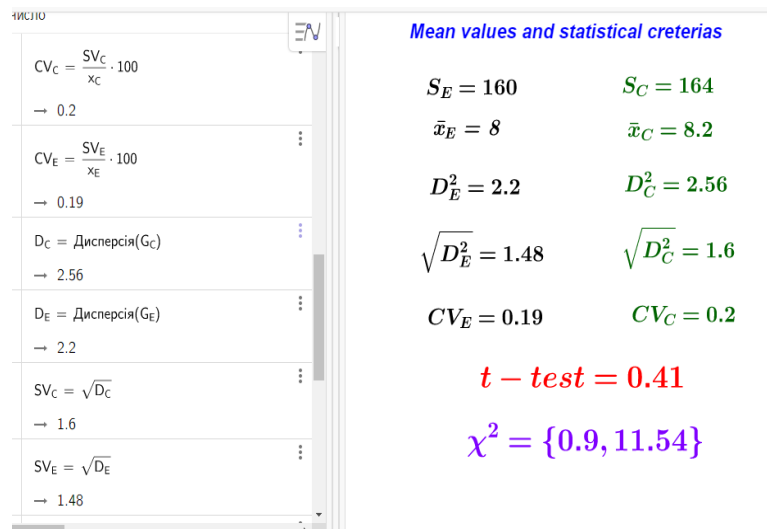


Figure 2. Calculation of means in G_E and G_C and statistical criteria's.

The program allows using the built-in commands «One Variable Analysis» (Figure 3, 4) or «Multiple Variable Analysis» (Figure 5) to build a histogram in various formats, visualize statistics as a table (frequency of the corresponding value, arithmetic mean, mode, median, variance, standard deviation, minimum and maximum value), calculate Student's and Pearson's criteria and present the results on a 2D window in the form of a frequency histogram, table, etc. Graphic data of pedagogical measurements in nominal and ordinal scales are conveniently presented in the form of a frequency polygon or a relative frequency polygon. With the GEOGEBRA applet you can see by following the link <https://www.geogebra.org/classic/swrtmnmk>.

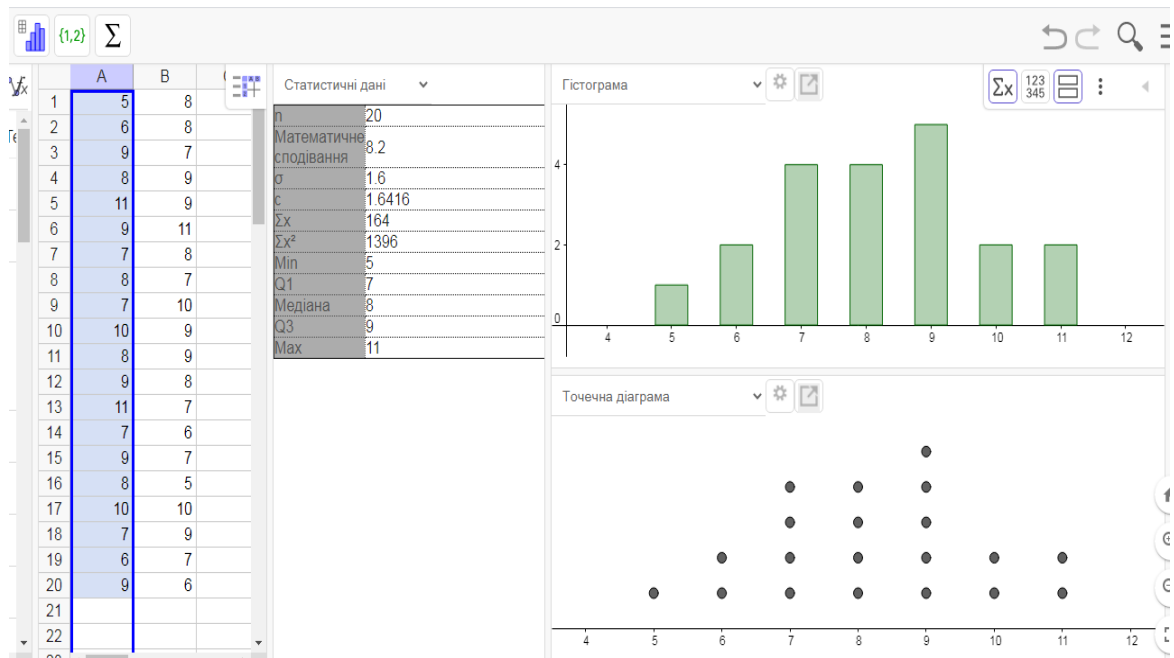


Figure 3. Visualization of statistical data G_C.

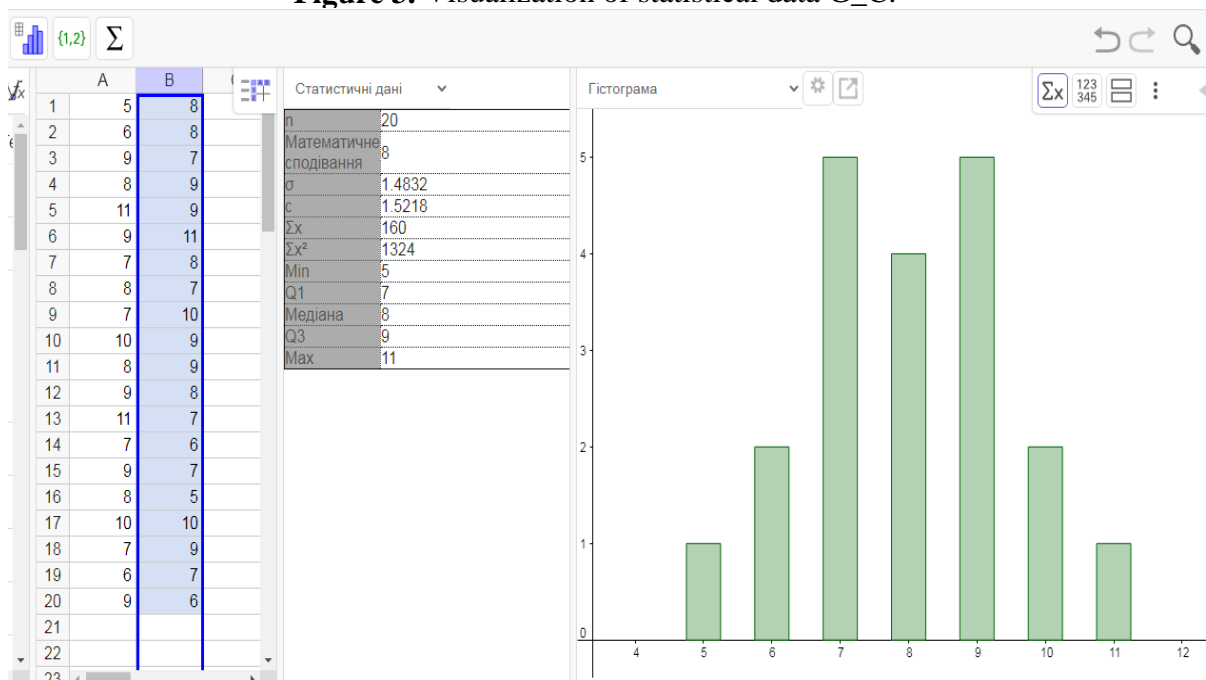


Figure 4. Visualization of statistical data G_E.

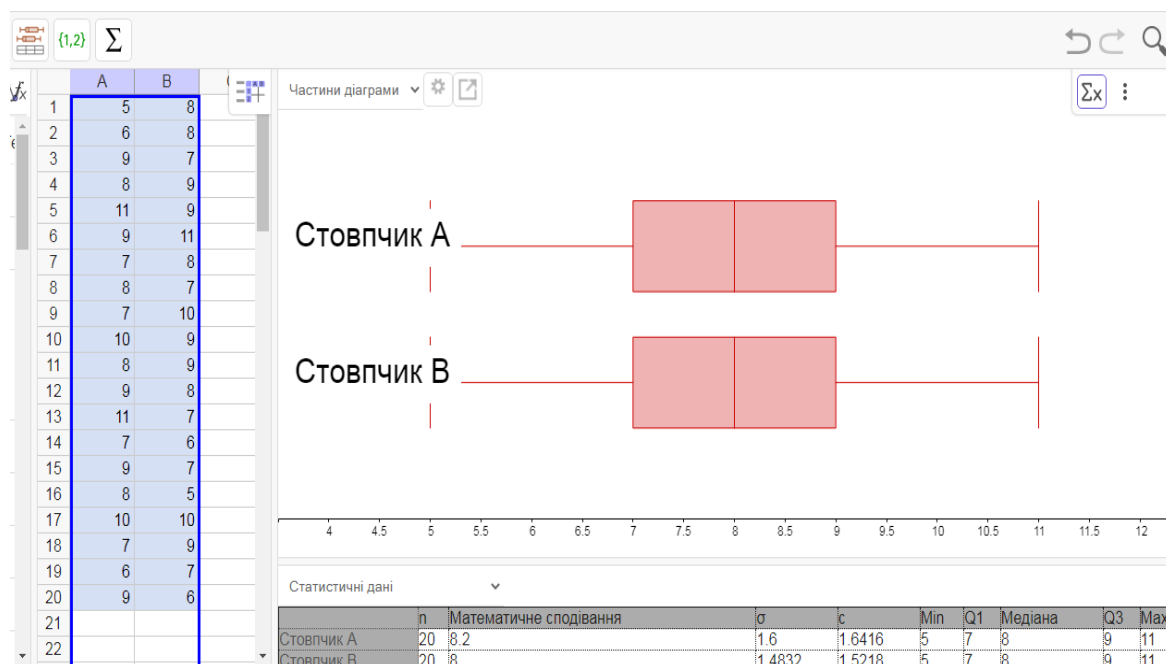


Figure 5. Visualization of statistical data G_E and G_C.

As seen in the figure 2-5:

1. The arithmetic mean in the experimental group more than in the control group $x_E > x_C$. This means that in the experimental group the overall level of indicators is higher than in the control. The experimental group promises a higher probability of achieving better results.

2. The dispersion in the experimental group is more than in the control group $(D_E)^2 > (D_C)^2$. Therefore, the experimental group has less trait variation and is heterogeneous. The control group has a greater variation of the trait, so it is more homogeneous.

3. The standard deviation in the experimental group is more than in the control group $D_E > D_C$. That is, in the control group, the fluctuations of the deviation from the mean value are less.

4. The coefficient of variation in the experimental group is more than in the control group $CV_E > CV_C$. That is, in the experimental group the variability of the results is less. Therefore, the results of the experimental group are more reliable.

5. The tabular (theoretical) value of the Student's criterion $t_t = 2,04$ at degrees of freedom ($N=38$). The value we obtained as a result of calculations (experimental) $t_e = 0,41$, that is the calculated value is greater than the table, so accordingly the experiment gives a positive result.

6. The value of the Pearson criterion that we obtained as a result of calculations (experimental) $t_e = 0,9$. This value is less than 1, accordingly the experiment gives a positive result.

In 2020-2021 a.y. 84 masters of educational and professional programs in the field of knowledge 0114 Teacher training with subject specialization during the study of the discipline "Mathematical apparatus of pedagogical sciences" were asked to process and present the results of the experiment with the help of the above packages and traditionally. All 100% of higher education students used digital tools during the task, of which 25 students (30%) preferred GEOGEBRA, 34 students (40%) used MS EXCEL, 23 students (28%) STATISTICA and 2 students (2%) DUDAMATH (Figure 6).

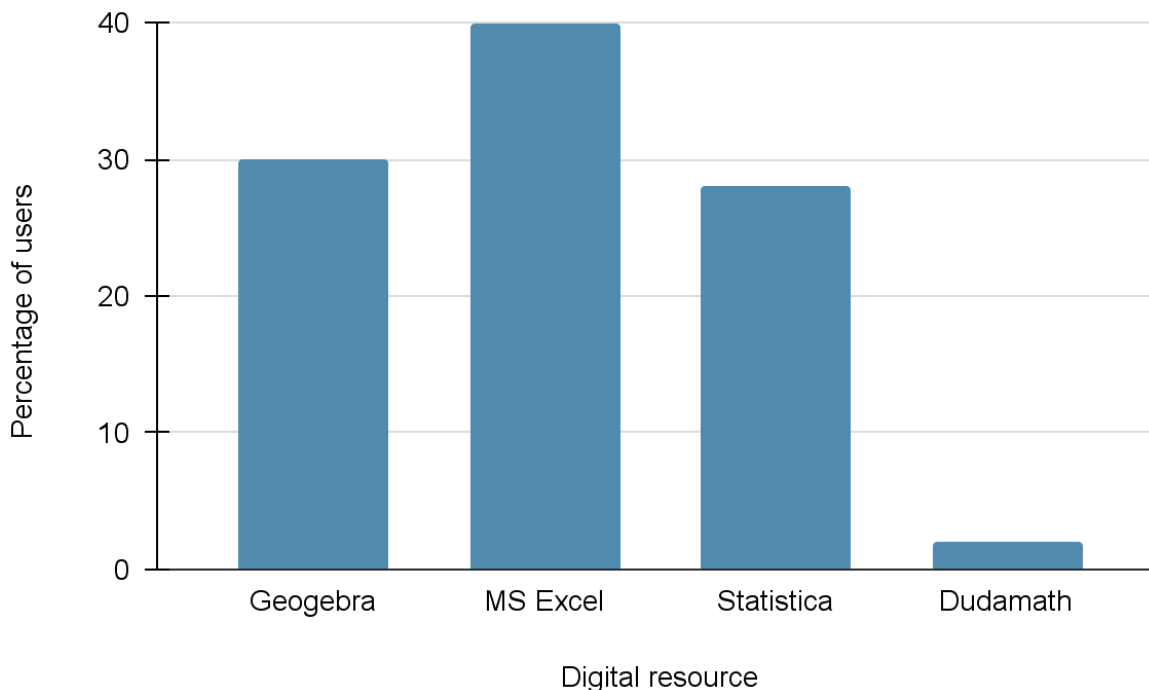


Figure 6. The distribution of users by digital resource.

3. CONCLUSIONS AND PROSPECTS OF FURTHER RESEARCH

The use of ICT in the study of the discipline «Mathematical apparatus of pedagogical sciences», gives students the opportunity to deepen the theoretical knowledge on which mathematical methods of statistical data are based, to receive the necessary training for the free use of methods of applied statistics of pedagogical analysis, psychological data, to form skills of solving professional applied problems using software tools. Of the many non-specialized mathematical environments, the GEOGEBRA dynamic mathematics package has the widest possibilities for processing experimental data based on the results of a pedagogical experiment. This program has a very wide set of commands for calculating statistical values and visualizing the results: mean values (arithmetic mean, mode, dispersion, median, standard deviation), minimum and maximum value, frequency of occurrence of values, Student's and Pearson's criteria's and construction of histograms in various formats, tables with statistical data, simultaneous presentation of results in a 2D window in the form of an inscription, histograms of frequencies, tables, etc.

References

- [1]. R. M. Khusainova, Z. V. Shilova & O. V. Curteva, "Selection of Appropriate Statistical Methods for Research Results Processing", *International Electronic Journal of Mathematics Education*, no. 1 (11), pp. 303-315, 2016. [Online]. Available: <https://www.iejme.com/article/selection-of-appropriate-statistical-methods-for-research-results-processing>. https://doi.org/10.1007/978-981-13-0890-1_26. [Accessed on: May 20, 2021].
- [2]. M. O. Rubtsov, E. H. Murtaziev & N. M. Rubtsova, "Methods of selection of statistical criterion and its application to the results of pedagogical experiments", *Scientific Bulletin Melitopol State Pedagogical University. Series: Pedagogy*, no. 2 (21), pp. 164-172, 2018. Available: <http://eprints.mdpu.org.ua/id/eprint/3669/1/%E2%84%96%2022.pdf>.

[3]. M. Cichoń, "Reporting statistical methods and outcome of statistical analyses in research articles", *Pharmacological Reports*, no. 72, pp. 481–485, 2020. [Online]. Available: <https://doi.org/10.1007/s43440-020-00110-5> [Accessed on: June 20, 2021].

[4]. M. I. Grabar, *The use of mathematical statistics in pedagogical research. Nonparametric Methods*, Moscow: Pedagogy, 1977.

[5]. N. S. Nair, K.T. Harichandrakumar & N. Ravishankar *Statistical Packages for Data Analysis*. In: Parija S., Kate V. (eds) *Thesis Writing for Master's and Ph.D. Program*. Available: https://link.springer.com/chapter/10.1007/978-981-13-0890-1_26#citeas https://doi.org/10.1007/978-981-13-0890-1_26/. [Accessed on: May 20, 2021].

[6]. C. A. Peters, "Statistics for Analysis of Experimental Data". [Online]. Available: https://www.princeton.edu/~cap/AEESP_Statchap_Peters.pdf. [Accessed on: May 19, 2021].

[7]. J. Glass, J. Stanley, *Statistical Methods in Pedagogy and Psychology*, Moscow: Progress, 1976.

[8]. E.V. Sidorenko, *Methods of the mathematical processing in Psychology*, SPb.: OOO "Rech'", 2002.

[9]. L. H. Havrilova, Y. V. Topolnik, L. O. Kuhar & L. A. Matviychuk, "Content and structure of the course "Information and communication technologies in pedagogical research", *Information technologies and teaching aids*, no. 4(66), pp. 245-259, 2018. Available: <https://doi.org/10.33407/itlt.v66i4.2220>

[10]. P. V. Nikitin, R. I. Bazhenov, I. V. Petukhov, L. A. Stesina & M. S. Gorokhov "Modern approaches in teaching informatics" *Proceedings of SOCIOINT» 2018- 5th International Conference on Education, Social Sciences and Humanities*, 2-4 July 2018- Dubai, UAE pp. 358-364 .

[11]. O. Sikora, U. Kogut, T. Vdovychyn, "Aspects Of The Use Of Ict For Automation Of Dragonization Of Databases Of Statistical Analysis In Training Of Humanitarian Profile Masters". *Physical and Mathematical Education*, no. 4(18), pp. 143-148, 2018.

[12]. Ya. O. Ostapenko, "Use of ICT for statistical analysis of psychological processes of future psychologists", *Engineering and Educational Technologies*, no. 1(9), pp. 102–113, 2021. [Online]. Available: [http://eetecs.kdu.edu.ua/2021_01/EETECs2021_009\(1\)_09](http://eetecs.kdu.edu.ua/2021_01/EETECs2021_009(1)_09). [Accessed on: May 21, 2021].

[13]. A.V. Chorna "Treatment of experimental data on the results of pedagogical experiment by means of computer statistical packages", *Pedahohichnyy dyskurs*, no. 14, pp. 453-457, 2013.

[14]. T. A. Kramarenko, O. M. Iie, "Automation of Treatment of Data of Statistical Analysis in Pedagogical Researches by Facilities of Tabular Processor of Microsoft Excel", *Bulletin of Luhansk Taras Shevchenko National University*, no. 18(1), pp. 85-93, 2013. [Online]. Available: [http://nbuv.gov.ua/UJRN/vlup_2013_18\(1\)_11](http://nbuv.gov.ua/UJRN/vlup_2013_18(1)_11). [Accessed on: May 22, 2021].

[15]. M. P. Shishkina, "Requirements for informatization of educational research". [Online]. Available: <https://journal.iitta.gov.ua/index.php/itlt/article/view/19/112001>. [Accessed on: June 24, 2021].

[16]. G. Kyryletska, "Computer technologies in processing the results of a pedagogical experiment", *New pedagogical thought*, no. 1. pp. 23-25, 2015. [Online]. Available: http://nbuv.gov.ua/UJRN/Npd_2015_1_7. [Accessed on: May 23, 2021].

[17]. V. G. Khomenko, L. V Pavlenko, "Problems of preparation of future engineers-teachers for the use of computer technologies of statistical processing of experimental data", *Problems of engineering-pedagogical education*, no. 26-27. pp. 149-155, 2010. [Online]. Available: http://nbuv.gov.ua/UJRN/Pipo_2010_26-27_24. [Accessed on: May 21, 2021].

[18]. R. Fabbriatore, A. Parola, G. Pepicelli, F. Palumbo "A latent class approach for advising in learning statistics: implementation in the ALEAS system", *Proceedings of the First Workshop on Technology Enhanced Learning Environments for Blended Education (teleXbe2021)*, January 21–22, 2021, Foggia, Italy

[19]. A.S. Korchakova, O. M. Nikitenko "Features of statistical data processing using a computer", *Metrology and instruments*, no. 1, pp. 138 - 142, 2014.

[20]. N. A. Siaska, "Using the methods of mathematical statistics in physical education and sports by means of the latest information technologies", *Innovation in education*, no. 4, pp. 257-264, 2016.

[21].M. V. Roik, O. I. Prysiazhniuk & V. O. Denysiuk, "Review of software for statistical data analysis", *Efektyvna ekonomika*, no. 7, 2017 [Online]. Available: <http://www.economy.nayka.com.ua/?op=1&z=5676>. [Accessed on: December 12, 2019].