National Technical University of Ukraine "Kyiv Polytechnic Institute" Faculty of Biomedical Engineering Department of Biomedical Engineering

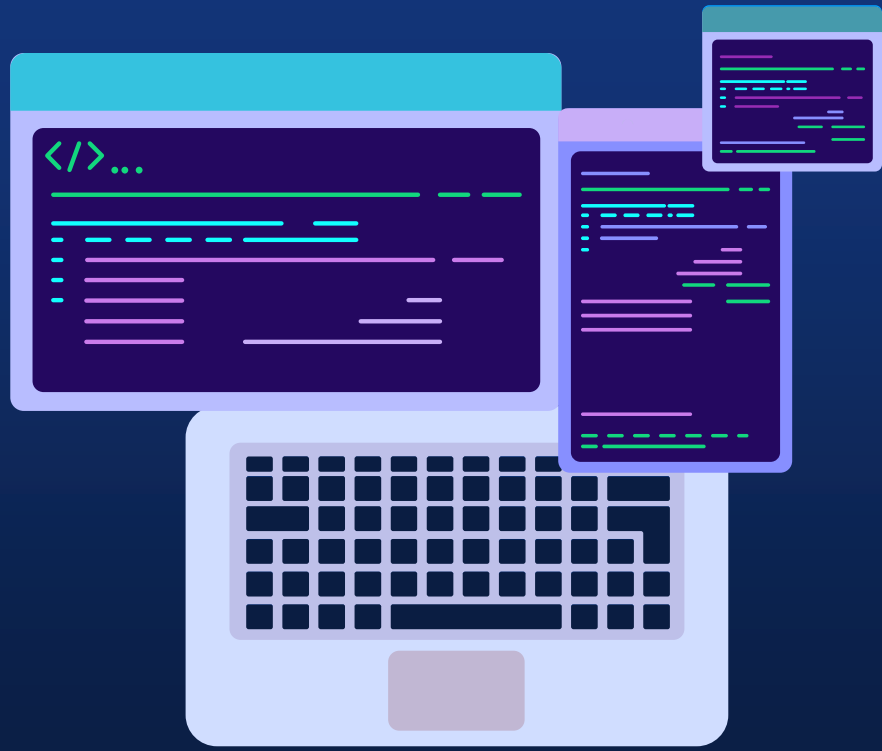
Diploma thesis for bachelor's degree
in the specialty 163 "Biomedical Engineering"
on the topic:

Statistical data analysis system for evidence-based medicine on the LabVIEW platform

Completed: 4th year student, BM-73i
AHMED SANI
Scientific adviser: Assoc. prof. BMI, Ph.D.
Solomin A.V

Kiev, 2021

LabVIEW statistical data analysis system for evidence-based medicine



Purpose

Purpose of practice: Ensuring the reliability of the statistical analysis conclusions by creating a software system of statistical processing modules with prior automatic verification of the relevant criteria of their usability.

Tasks

Tasks of practice:

- analytical review of existing statistical methods and algorithms for information processing for evidence-based medicine;
- development of tree-like architecture of software system with decision-making nodes and algorithms of statistical analysis modules;
- implementation in the NI LabVIEW environment of statistical analysis algorithms with preliminary automatic verification of the relevant criteria for their applicability

Projects Aim

The aim of the work was to create a software system that implements the most commonly used tools of statistical analysis, but at the same time provides verification of the necessary criteria and the limits of their use in automatic mode. Thus, the user is not required to know the appropriate mathematical apparatus, the system will automatically select the optimal statistical methods and draw conclusions according to specific conditions, significantly reducing the likelihood of errors.

MEDICINE-BASED RESEARCH

Identify and develop treatments for disease and estimate their effects.

Design, monitor, analyse, interpret, and report results of clinical studies

Identify risk factors for diseases.

Develop statistical methodologies to address questions arising from medical/public health data.

PROBLEM VS SOLUTION

PROBLEM

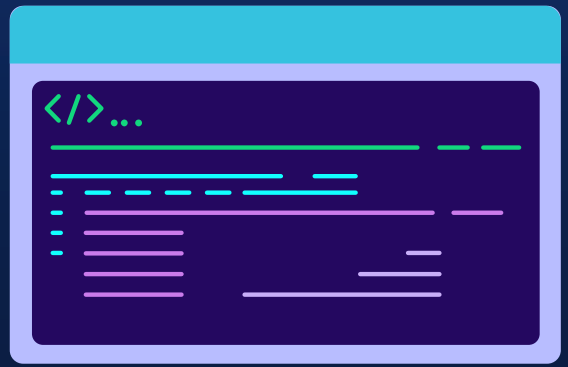
Typical statistical studies involve the detection of the influence of a factor on the experimental group. Moreover, depending on the situation, the effect can be investigated using control and experimental groups, or on the basis of one group before and after the action of this factor.

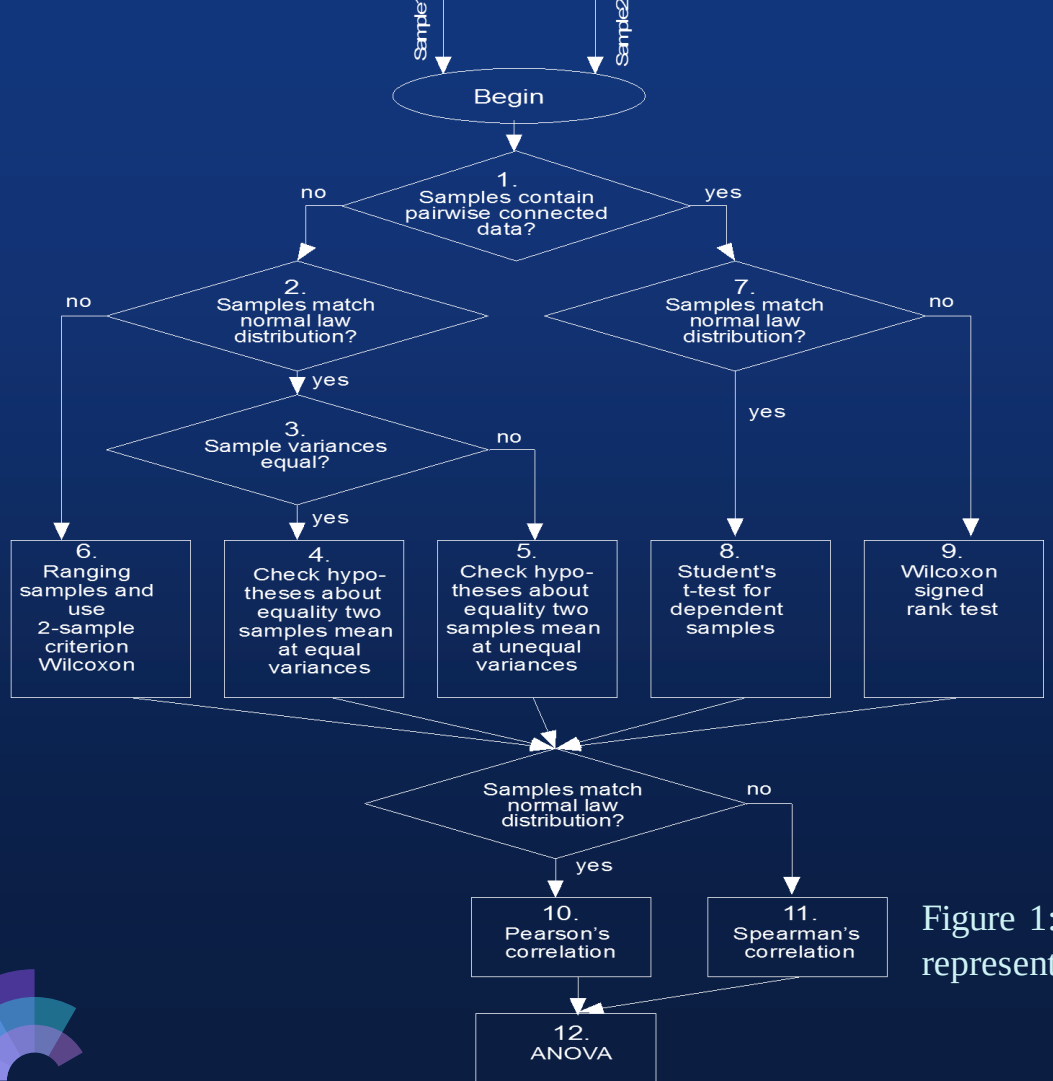
SOLUTION

To solve this problem, it is proposed to construct a tree-like structure of subroutines representing specific statistical methods, and in the nodes of this tree there are decision-making tools - branches based on the analysis of the conditions of their applicability

Formulation of the problem in the applied setting	Formulation of the problem in a statistically statement	Additional terms		Applied method
Comparison of control and experimental samples	Testing the hypothesis of equality of means (distribution centers) two independent samples	Normal law distribution	Dispersions of the samples are equal	Student's t-test for equal variances
			Dispersions of the samples unequal	Student's t-test for unequal variances
		Law distribution different from normal, or data is measured in discrete scale	Dispersions of the samples are equal	Mann - Whitney (U- test of Wilcoxon - Mann - Whitney)
			Without the assumption about dispersions	Wilcoxon two-samples median test
Comparison of samples before and after experiment	Testing the hypothesis of equality of means in two dependent samples	Normal distribution law		Student's t-test for dependent samples
		The distribution law is different from the normal or data are measured in a discrete scale		Wilcoxon signed rank test
Comparison of indicator dispersion in two samples	Testing the hypothesis of equality of variances (belonging of variances to the same general population)	Normal distribution law		F-test (Fisher)
		The distribution law is different from the normal or data are measured in a discrete scale		Siegel - Tukey, Moses
Testing the hypothesis about the presence of correlation between the variables	Correlation analysis	Normal distribution law		Pearson’s correlation
		The distribution law is different from the normal or data are measured in a discrete scale		Spearman’s correlation
Testing the hypothesis about the presence of the influence of a factor on a	Analysis of variance (ANOVA)			ANOVA, F-test (Fisher)

The table shows how many factors a researcher must consider to make a statistically significant conclusion. This requires a fairly deep knowledge of mathematical principles, which is not always present in medical professionals.





Computer software system built in form of a tree, where the nodes automatically check conditions and criteria for specific statistical methods. The result of the software system is formation of a conclusion about existence or absence of a statistically significant difference between average values of two data sets.In practice, this may relate, for example, to the detection of statistically significant effects of a pharmacological agent, drug procedure, regimen and other factors

Figure 1: Tree-like structure of subroutines representing specific statistical methods

METHODS

The convenience of combining with hardware for entering information provides:

- Opportunities to create automated workstations for statistical processing
- Ease of programming
- The ability to improve, configure and adapt the software system to specific situations

a computer software system was built in the form of a tree, where:

- The nodes automatically check the conditions and criteria for specific statistical methods
- Branching to the next nodes of the algorithm
- The final node of chain draws statistical conclusions

The software development framework was selected to be ***NI LabVIEW***, which is currently a common standard in the field of biomedical instrumentation and biomedical analysis and is readily implemented with most contemporary software and hardware systems.

- ***NI LabVIEW*** (National Instrument Laboratory Virtual Instrument Engineering Workbench) is ideal for software developed and used by engineering experts, including biomedical engineers, rather than programmers. Instead of text programming languages with complex syntactic laws, the graphic language G is used here, which has a kind of block diagram that engineers are more familiar with.

the software system is formation of a conclusion about existence or absence of a statistically significant difference between average values of two data sets and the statistical relationship between them. In practice, this may relate, for example, to the detection of statistically significant effects of a pharmacological agent, drug procedure, regimen and other factors

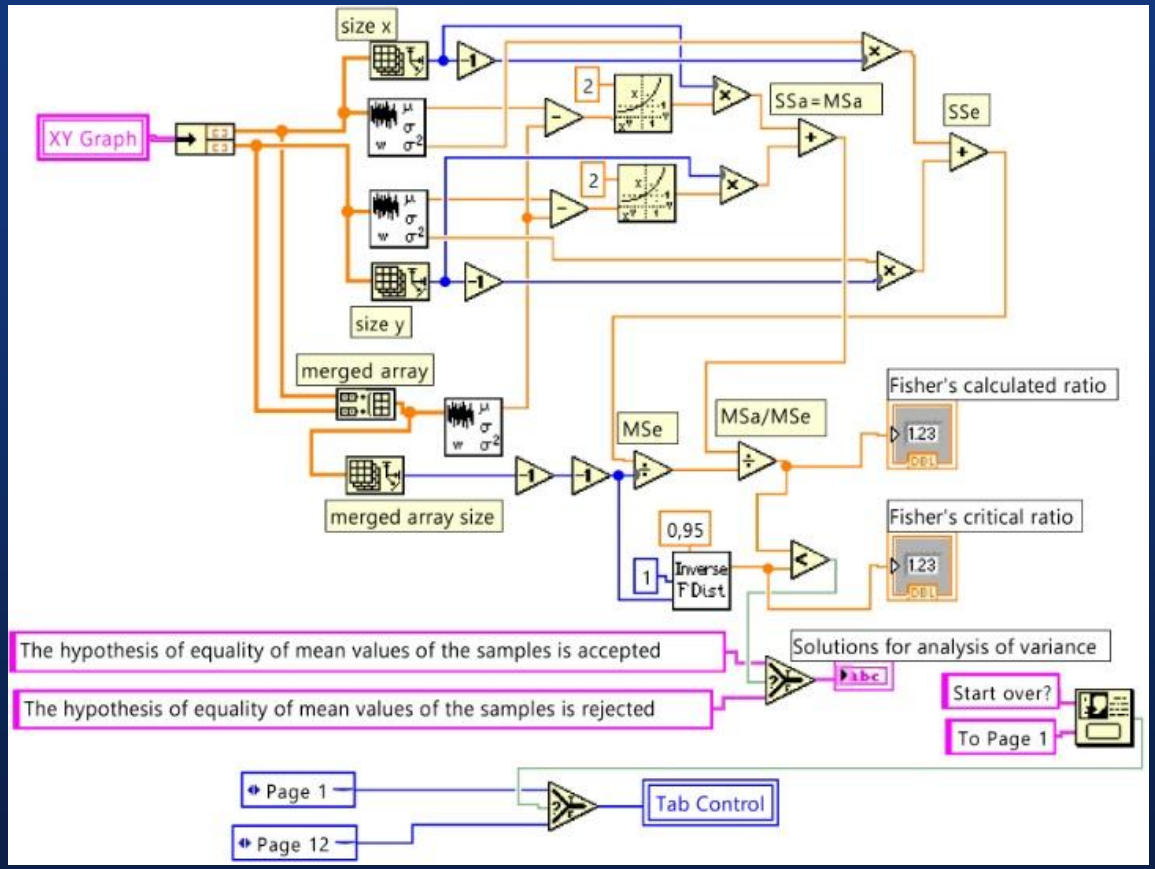


Figure 2: Block diagram of the module for analysis of variance

The module calculates actual Fisher ratio MSa / MSe – the ratio of variance, which is explained by the influence of factor (intergroup), and unexplained variance (intragroup), and compares with the critical value of Fisher for the appropriate number of degrees of freedom and a given level of significance

The appearance of the system interface

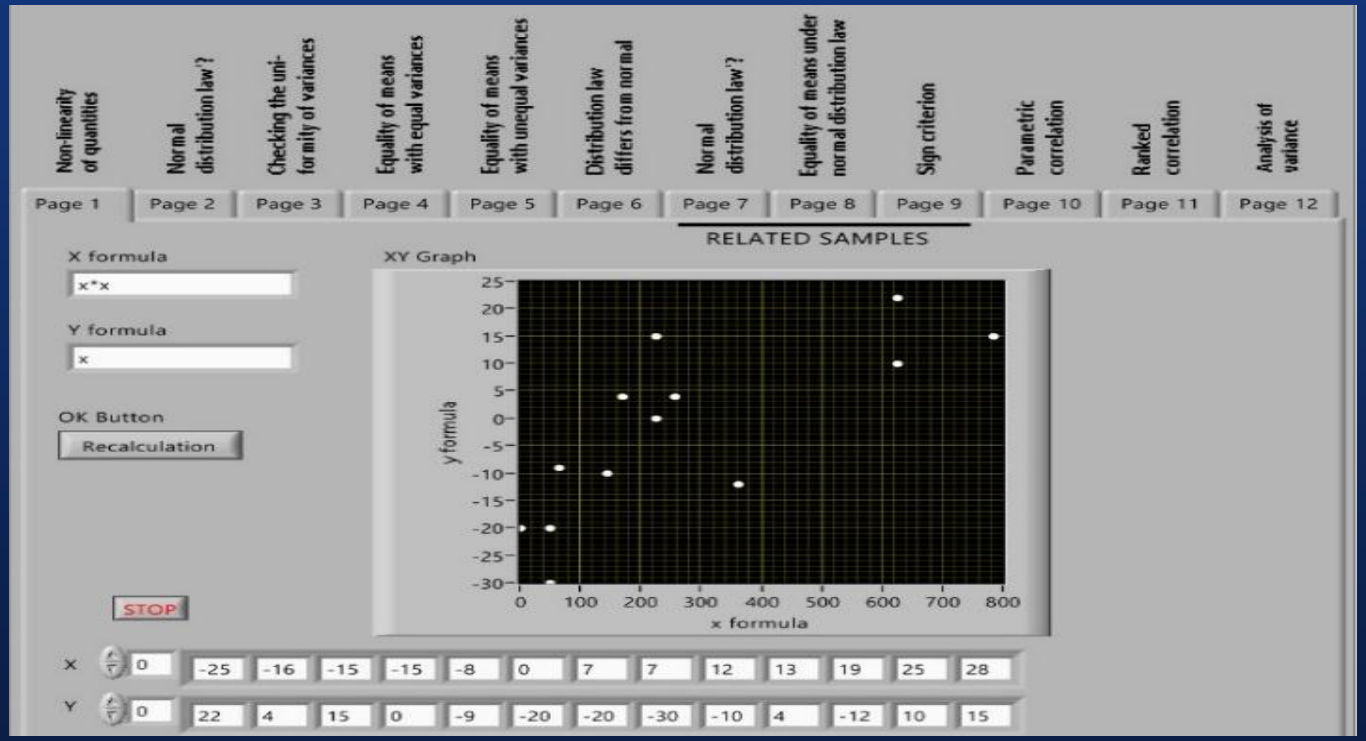


Figure 3: Display in the interface of the statistical conclusion of analysis of variance

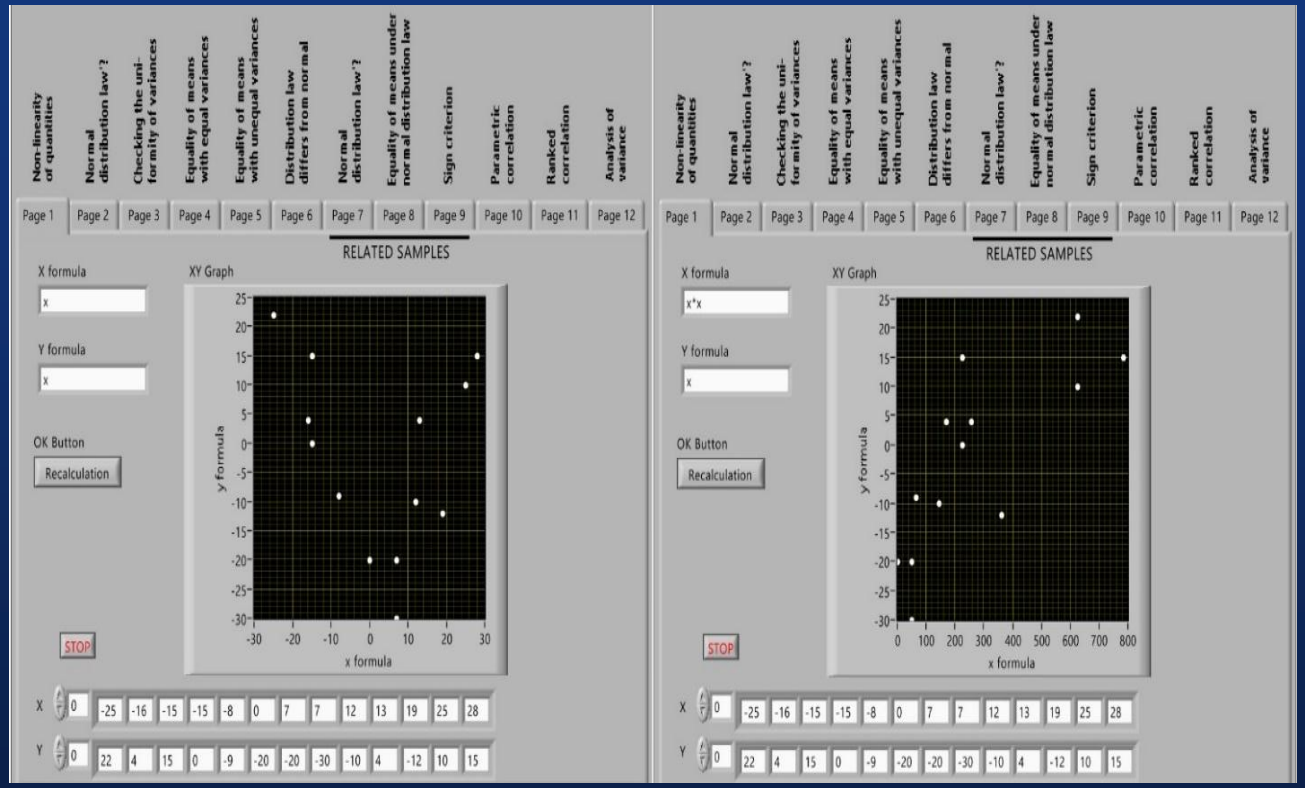


Figure 3: View of the software system interface with direct display of input data sets and with the results of the previous functional transformation

Two options for displaying input data on the left and right screens demonstrate a specially designed ability of the system to analyze either the input data arrays directly, or the results of their preliminary conversion by almost any function that can be simply entered in the interface window

In the example shown in Figure 3 on the first screen you can see signs that between the input data arrays there is a relationship similar to quadratic law, so on the second screen instead of the data of one of the arrays enter their squares, then see the trend of linearity of such data connection law.

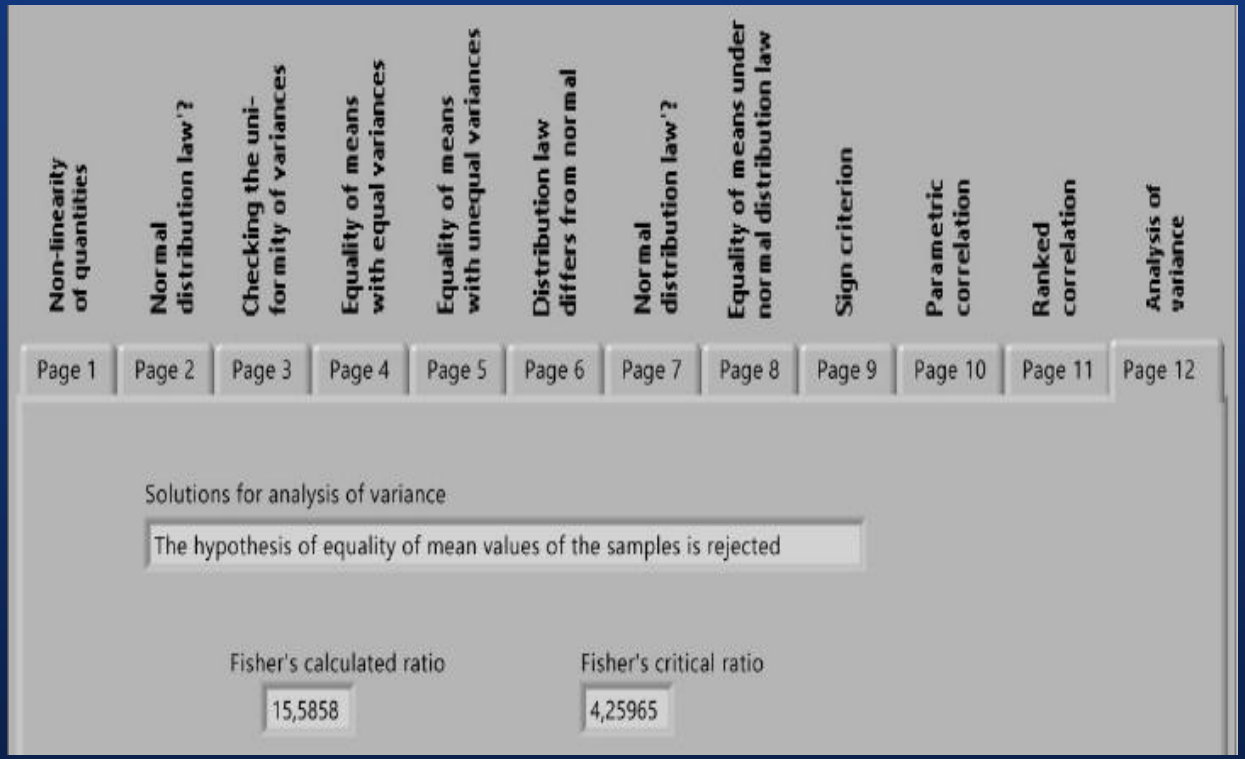
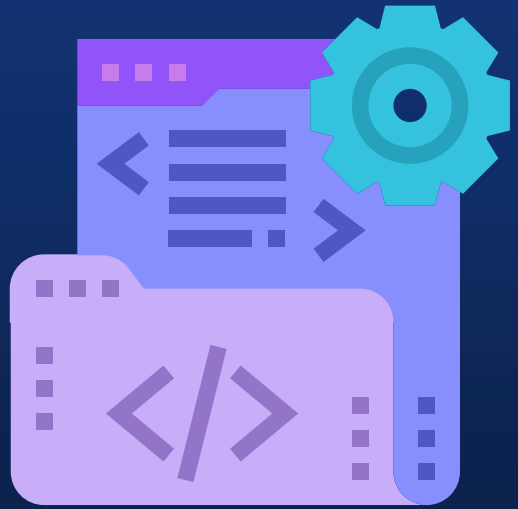


Figure 4: Display in the interface of the statistical conclusion of analysis of variance

the normality of the distribution law of both samples is checked, then the equality of variances; each combination of features uses its own statistical tool. Checks in the modules are performed automatically using the appropriate statistical criteria.

The given examples of the implementation of the modules of the software system demonstrate the principles of achieving the formulated goal of work. The main distinguishing feature of the system is the automatic check of criteria's for the applicability of the corresponding statistical methods, as well as the tree structure of the system, where the decision on branching is made at the nodes in accordance with the results of the check.



BENEFITS

The use of an automated system to detect statistical associations and statistically relevant factor effects will reduce the number of incorrect findings in evidence-based medicine, especially if such a system is readily accessible, which can be easily supported on the NI LabVIEW platform by connecting to the Internet.

The use of software environment **NI LabVIEW** is a promising area of building information processing and analysis systems. One of the important advantages is the ability to modify algorithms even by non-programmers.

CONCLUSIONS

Prospects for improving the system imply further addition of modules with the implementation of additional methods of statistical analysis, as well as providing easy access to the system by users.

One of the significant advantages of the NI LabVIEW platform is the simplicity of interaction with research equipment and other software systems. This provides the basis for creating flexible automated workplaces.

the article has been accepted for publication in the journal "Biomedical Engineering and Technology": Solomin A.V., Ahmed Sani. LabVIEW statistical data analysis system for evidence-based medicine / Biomedical Engineering and Technology. – 2021. v.5.p.10-16.

THANKS