ОŃŦÚS «Оңтүстік Қазақстан медицина ака	ПК-QAZAQSTAN MEDISINA AKADEMIASY демиясы» АҚ	медицинская академия»
Departments: "Medical	Biophysics and Information Technologies",	№ 35-11(Б)-2024
Guidelines for practical lessons	of the course "Introduction to Scientific Research"	
1.K. SKINA. SKINA. COLUNC. K. K. S. SKINA. SKINA. COLUNC. K. SKINA. COLUNC. K. SKINA. COLUNC. K. SKINA. COLUNC.	L SKINA. EULI. KL SKINA. EKINA.	na.edu.k. k. s. skn na.edu.k.k.
GUIDEI	INES FOR PRACTICAL LESSONS	na.edu.ku s skn

X skma.eo

NON AN	
Course: Introduction to Scient	tific Research
Course code de ISP 2212	
Educational program: 6B10115 «Medicine»	J.K. K. Skinona.eve
Number of academic hours/credits: 180/6	edu. 14 K2 Storman
Year/Term: 2/4	na. e. edu. u. Ki Ki Sh
Practical lessons 48 hours	skine a eduluki ki
N. H. K. SKI Mara edu, K. L. SKI Mara edu, K.	1 skning. edu
e en duit to stand a en duit to stand a en en	U.K. K. SKIII NO.
Mana. B. edu. K. K. Sk. Mana. Edu. U.K. J. Sk. Mana.	eachuite the skining
Stranger Store and the standard s	Mara.eocou.kr.K. S. S.
KL ST SKINN B. B. EULIN KL SK SKINB. EULINKL	SKIN 3. BOUNKIN
	1 241 23. 200.
all the skiller south that a skill as could be	1

ОŃTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ	SOUTH KAZAKHSTAN MEDICAL ACADEMY AO «Южно-Казахстанская л	медицинская академия»
Departments: "Medical Biophysics and	Information Technologies",	№ 35-11(Б)-2024
"Social health insurance an	d public health"	№ 582024
Guidelines for practical lessons of the course "	Introduction to Scientific Research"	p.2 out of 68

Guidelines for practical lessons were developed in accordance with the working curriculum of the discipline (syllabus) "Introduction to Scientific Research" and discussed at a meeting of the departments:

Medical biophysics and information technology

Protocol no. <u>11</u> from "<u>30</u>" <u>05</u> 202<u>4</u> g. Head of department, PhD, ass. prof, ______ M.B. Ivanova

Social health insurance and public health

Protocol no. <u>15</u> from "<u>10</u> "<u>D6</u> 202 g.

Head of department, PhD, ass. prof. Hcur G.Zh. Sarsenbayeva

MEDISINA АКАDEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ	SKMA -1979- ,,,,,, АО «Южно-Казахстанская ме	дицинская академия»
Departments: "Medical Biophysics and	Information Technologies",	№ 35-11(Б)-2024
"Social health insurance and	d public health"	№ 582024
Guidelines for practical lessons of the course "I	Introduction to Scientific Research"	p.3 out of 68

ONTÚSTIK-QAZAQSTAN

Lesson №1

1. The theme: Introduction to Biostatistics

2. The aim: formation in students of a basic understanding of the discipline "Biostatistics", its subject, tasks and stages of development; familiarizing students with the types of medical data, as well as the types of measurement scales.

3. The learning objectives:

- understand the role of biostatistics in medical education and the work of a practicing physician;
- get acquainted with the main stages of the development of biostatistics;
- learn to classify medical data;
- learn to distinguish between measurement scales used in medical research.

4. Key questions of the theme:

- 1. Subject and tasks of biostatistics.
- 2. Stages of development of biostatistics.
- 3. Classification of medical data.
- 4. The main types of measurement scales used in biomedical research.

5. Methods / technologies of teaching and learning: Practice

Tasks:

1. Learn the theory:

1.1. Subject and tasks of biostatistics

Statistics is a social science that studies the quantitative side of mass social phenomena.

It appeared as an independent science at the end of the 17th century.

Biostatistics- statistics that studies issues related to biology, medicine, pharmacy, hygiene and health care.

The science of bostatistics was formed in the middle of the 19th century.

The role of biostatistics in the practical and scientific work of a healthcare manager, doctor, epidemiologist, nurse, and pharmacist is great.

Biostatistics uses various methods: collecting data, summarizing it, analyzing and drawing conclusions based on the observations obtained.

Statistical analysis helps extract information from data and evaluate the quality of that information.

Objectives of biostatistics:

- quantitative representation of biological facts (measurement);
- generalized description of a set of facts (statistical assessment);

search for patterns (testing statistical hypotheses).

1.2. The role of scientists F. Galton, K. Pearson, R. Fisher in the development of biometrics

The founder of biometrics is considered to be the English scientist Francis Galton (1822-1911) (Figure 1.1, a).

A contemporary of F. Galton, Russian scientist K.A. Timiryazev said the following about him: "He was one of the most original scientists, researchers and thinkers of modern England."

Indeed, F. Galton was a very versatile person; he was interested in natural science, anthropology, heredity, psychology, the theory of evolution, meteorology and statistics.

In one of his books on heredity, the term biometry was first introduced. F. Galton's important achievements include the development of the foundations of correlation analysis.

A follower of F. Galton is Karl Pearson (1857-1936) (Figure 1.1,b) - English mathematician, biologist and philosopher. Author of more than 400 scientific papers on mathematical statistics.

He introduced the concept of standard deviation and coefficient of variation, developed the mathematical apparatus of the theory of conjugation of features, nonlinear correlation and regression,

ОŃTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ	SOUTH KAZAKHSTAN MEDICAL ACADEMY AO «Южно-Казахстанская м	едицинская академия»
Departments: "Medical Biophysics and	Information Technologies",	№ 35-11(Б)-2024
"Social health insurance an	d public health"	№ 582024
Guidelines for practical lessons of the course "	Introduction to Scientific Research"	p.4 out of 68

goodness-of-fit criteria, decision-making algorithms and parameter estimation.



Figure 1.1. a – F. Galton, b – K. Pearson, c – R. Fisher

Together with the English zoologist and biometrician W. Weldon, K. Pearson founded the journal Biometrics, dedicated to the application of mathematical and statistical methods in biology.

The successor and continuer of K. Pearson's work on biostatistics was Ronald Fisher (1890-1962) (Figure 1.1, c).

Danish statistician A. Hald described R. Fischer as "a genius who almost single-handedly laid the foundations of modern statistics," and English ethologist and biologist R. Dawkins called him "the greatest biologist like Darwin."

R. Fisher is the founder of the theory of sampling distributions, methods of variance and discriminant analysis, the theory of experimental design, the maximum likelihood method and much more, which forms the foundation of modern applied statistics and mathematical genetics.

1.3. Collection, classification and presentation of data

Regardless of what tasks are set when conducting scientific medical and biological research, it must be carried out step by step, in a certain sequence.

At the first stage, a research plan and program is drawn up, at the second, collection of statistical material, on the third - the analysis of the received data is carried out, on the fourth - the processing of the collected material takes place.

The stage of collecting statistical data is very important, because the correctness of all further research depends on the correctness of the collected data.

The process of obtaining information about the objects of the population under study and their properties is called collecting statistical data. This data is the subject of statistical processing and analysis.

At this stage, it is important to determine the type of data being considered.

Statistical data is divided into the following types: quantitative, qualitative and dates (Fig.1.2). *Quantitative data*are divided into two categories: discrete and continuous.

Discrete data- quantitative data that is presented as integers. For example, the number of children in the family, the number of ambulance calls per hour, heart rate, etc.

Continuous Data– quantitative data obtained by measurement on a continuous scale. For example, body weight, height, blood pressure, etc.

Continuous data can be interval or relative.

Interval data— continuous data that is measured in absolute values that have a physical meaning.



Figure 1.2. Types of statistics

Relative data- continuous data that reflects the proportion of change (increase or decrease) in the value of a characteristic in relation to the original (or some other) value of this characteristic. These data are dimensionless quantities or expressed as percentages.

Qualitative data are divided into nominal and ordinal.

Nominal data -qualitative data that reflects dummy codes for unmeasured categories. For example, diagnosis code, gender, marital status, nationality, etc.

Ordinal data- qualitative data that reflects the conditional degree of expression of any characteristic. For example, stages of cancer, degrees of heart failure.

Their main difference from discrete quantitative data is the absence of a proportional scale for measuring the severity of a trait.

Binary (or dichotomous) data -qualitative data that has only two possible meanings. For example, gender, the presence or absence of a disease, etc.

Dates- a special type of data, in some cases it may be necessary to perform some arithmetic operations with them, for example, to calculate the period of a patient's stay in the hospital.

1.4. Main types of measurement scales used in biomedical research

Measurement- this is a procedure for comparing objects according to certain indicators or characteristics (features).

Scale- a necessary, mandatory element of the measurement procedure.

The main types of measurement scales used in biomedical research include:

Nominal scale- used to classify the properties of an object, assign them numerical, alphabetic and other symbolic characteristics (gender, nationality, diagnosis, etc.) (Figure 1.3 shows a fragment of the electronic international classifier of diseases, where each type of disease is assigned a special code):

МКБ-10 Блоки.

Класс: Болезни органов дыхания

- Блок (J00-J06) Острые респираторные инфекции верхних дыхательных путей Блок (J10-J18) - Грипп и пневмония

- Блок (J10-J18) Грипп и пневмония Блок (J20-J22) Другие острые респираторные инфекции нижних дыхательных путей Блок (J30-J39) Другие болезни верхних дыхательных путей Блок (J40-J47) Хронические болезни нижних дыхательных путей Блок (J60-J70) Болезни легкого, вызванные внешними агентами Блок (J80-J84) Другие респираторные болезни, поражающие главным образом интерстициальную ткань Блок (J85-J86) Гнойные и некротические состояния нижних дыхательных путей Блок (J80-J94) Другие болезни плевры Блок (J90-J94) Другие болезни органов дыхания

Figure 1.3. International Classification of Diseases - example of a naming scale

ONTUSTIK-QAZAQSTAN MEDISINA АКАDEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ	SOUTH KAZAKHSTAN MEDICAL ACADEMY AO «Южно-Казахстанская м	едицинская академия»
Departments: "Medical Biophysics and	Information Technologies",	№ 35-11(Б)-2024
"Social health insurance an	id public health"	Nº 582024
Guidelines for practical lessons of the course "	Introduction to Scientific Research"	p.6 out of 68

• *ordinal*or ranking scale – orders the values of a characteristic (scale of stages of hypertension according to Myasnikov, scale of degrees of heart failure according to Strazhesko-Vasilenko-Lang, scale of severity of coronary insufficiency according to Fogelson (Figure 1.4), etc.);

Stage	Signs
	decreased myocardial contraction rate, decreased ejection fraction, shortness of breath, palpitations, fatigue during exercise
	circulatory failure is expressed moderately or significantly. The signs of circulatory failure indicated for the initial stage are detected not only during physical activity, but also at rest.
30	significant disturbances in cardiac activity and hemodynamics at rest, as well as the development of significant dystrophic and structural changes in organs and tissues

Figure 1.4. Fogelson coronary insufficiency severity scale - an example of an ordinal scale

• *interval scale*- shows the "range" of individual measurements of a characteristic (time, temperature scale, test scores, etc.) (Figure 1.5);



Figure 1.5. Stopwatch, thermometer - examples of interval scale

• *ratio scale*- reveals the relationship between the measured values of a trait (height, weight, reaction time, number of completed test tasks, etc.) (Figure 1.6).



Figure 1.6. Scales - an example of a ratio scale

1.5. Reliability and validity of measurements in biostatistics

During the measurement process, the question arises: *reliability* And *reliability*. *Reliability* measurements depends on:

- correctness (is the scale chosen correctly, are the readings recorded correctly, are systematic errors taken into account, etc.);
- stability (do the results coincide with repeated measurements);
- validity (It is the specified property of the object that is measured, and not another that is similar to it).

OŃTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY



«Оңтүстік Қазақстан медицина академиясы» АҚ

АО «Южно-Казахстанская медицинская академия» Departments: "Medical Biophysics and Information Technologies", № 35-11(Б)-2024 "Social health insurance and public health" № 58--2024Guidelines for practical lessons of the course "Introduction to Scientific Research" p.7 out of 68

Credibility measurement characterizes the accuracy of measurements of a quantity in relation to what exists in reality.

The main direction of checking the reliability of measurements is to obtain information from various sources.

- 2. Oral survey on the topic.
- 3. Take the test (completing test tasks using the "easyQuizzy" program)
- 6. Assessment methods / technologies: Oral questioning. Practical work. (assessment using a checklist).

7. Literature:

Main: •

1. Rosner Bernard Fundamentals of Biostatistics: Texbook/ B.Rosner - 8th ed.: GENGAGE learning, 2016.

- 2. Armitage P. Encyclopedia of Biostatistics. Wiley, 2016. 6100 p
- 3. Le C.T. Introductory biostatistics. Wiley, 2013. 536 p
- 4. Newman S.C. Biostatistical Methods in Epidemiology. Wiley, 2017. 388p
- 5. http://www.biometrica.tomsk.ru

8. Control:

- 1. What is "biostatistics"?
- 2. What is the role of scientists F. Galton, K. Pearson, R. Fisher in the development of biometrics?
- 3. What types of statistics do you know?
- 4. What types of measurement scales are used in biomedical research?

Lesson №2

1. The theme: Variation series.

1. The aim: developing skills in calculating the numerical characteristics of a variation series and their interpretation.familiarization with some of the capabilities of the STATISTICA program

2. The learning objectives:

- learn to calculate the numerical characteristics of a variation series;
- learn to interpret the numerical characteristics of a variation series
- -learn to create, edit, save source data tables;
- learn to build simple graphs in the STATISTICA program.

4. Key questions of the theme:

- 1. Measures of the central tendency of a variation series.
- 2. Measures of the spread of the variation series.
- 3. Interpretation of the numerical characteristics of the variation series.
- 4. Stages of statistical analysis in the "STATISTICA" program.
- 5. Types of documents created in the STATISTICA program. Their extensions.
- 6. Elements of the working window of the "STATISTICA" program.
- 7. Operations performed on columns and rows of a spreadsheet.
- 8. Construction of the simplest graphs using the "STATISTICA" program.

5. Methods / technologies of teaching and learning: Computer-based work / Solving situational tasks.

Example 1. The study of a new highly effective antibiotic involved patients in the purulent surgery department who were prone to obesity. The following distribution of patients by weight was obtained:

Body weight, kg	90	100	120	130	140	Total
Number of patients, people	4 5	4	8	65	2	21

ОŃTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ	SOUTH KAZAKHSTAN MEDICAL ACADEMY AO «Южно-Казахстанская ме	едицинская академия»
Departments: "Medical Biophysics and	d Information Technologies",	№ 35-11(Б)-2024
"Social health insurance a	nd public health"	№ 582024
Guidelines for practical lessons of the course '	'Introduction to Scientific Research"	p.8 out of 68

Calculate the indicators of the variation series: mean, mode, median, dispersion, standard deviation, coefficient of variation, range of variation, 25th and 75th percentiles. Does the presented variation series correspond to the law of normal distribution?

Solution.

Create a calculation table.

xi	vi	xi·vi	$xi-\overline{x}$	$(xi-\overline{x})2$	$(xi-\overline{x})2\cdot vi$
90	01 1	90	-29.5	870.25	870.25
100	0.460	400	-19.5	380.25	1521
120	8	960	0.5	0.25	2
130	6 0	780	10.5	110.25	661.5
140	2	280	20.5	420.25	840.5
Total	21	2510	92° . KI	5.0	3895.25

1) Calculate the average:

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i v_i}{\sum_{i=1}^{n} v_i} = \frac{2510}{21} \approx 119,5$$

those. average body weight of patients119.5 kg.

- 2) Determine the mode: Mo=120, i.e. the most common value found in observations is120 kg.
- 3) Determine the median: Me=120, i.e. meaning120 kgis in the middle of the variation series.
- 4) Calculate variance:

$$S^{2} = \frac{\sum_{i=1}^{n} (x_{i} - \bar{x})^{2} \cdot v_{i}}{n - 1} = \frac{3895, 2}{20} \approx 194, 76$$

5) Calculate standard deviation:

$$S = \sqrt{S^2} = \sqrt{194,76} \approx 13,96$$

those. standard deviation of patient weight13.6 kg.

6) Calculate the coefficient of variation:

$$V = \frac{S}{\bar{x}} \cdot 100\% = \frac{13,96}{119,5} \cdot 100\% \approx 11,7\%$$

because V \leq 33%, then the sample is homogeneous.

7) Calculate the range of variation: R = xmax - xmin = 140-90=50 those. difference between the largest and smallest mass values 50 kg.

Meh=120.

50th percentile number (median)

25th percentile (lowest quartile) number
$$N_{P_{-}} = \frac{n+1}{4} = \frac{21+1}{4} = 5.5$$
 P25=110.

75th percentile (top quartile) number
$$N_{P_{2}} = \frac{3(n+1)}{4} = \frac{3(21+1)}{4} = 16,5$$
 P75=130.

Example 2.In the "Statistica" program, create a sample of 100 random numbers that have a

МЕДІЗІЛА АКАДЕМІАЅҮ Оңтүстік Қазақстан медицина академиясы» АҚ Оңтүстік Қазақстан медицина академиясы» АҚ	цицинская академия»
Departments: "Medical Biophysics and Information Technologies",	№ 35-11(Б)-2024
"Social health insurance and public health"	№ 582024
Guidelines for practical lessons of the course "Introduction to Scientific Research"	p.9 out of 68

SOUTH KAZAKHSTAN

normal distribution with the parameters: μ =170, σ =7, which are the results of measuring the height of 1st year students at SKMA.

Calculate: mean, sum, median, geometric mean, harmonic mean, standard deviation, variance, standard error of the mean, confidence interval for the mean, skewness, kurtosis, largest and smallest sample value, lower and upper quartiles, range.

Construct a histogram and a normal probability plot, as well as a box-and-whisker plot. Explain the results.

1. Create a spreadsheet.

Open the "Statistica" program. A spreadsheet of size 10*10 will appear on the screen. If the table does not appear, then follow these steps: Home \rightarrow File \rightarrow New \rightarrow Spreadsheet.

2. Setting table sizes.

The data requires 1 column and 100 rows. Remove extra 9 columns and add 90 rows.

3. Table header design.

Enter the title of the table "Growth of 1st year students at SKMA" in the white field under the line "Data: Spreadsheet1 (1v by 100c)".

4. Setting the variable name.

Double-click the left mouse button on the variable name "Var1". Open the variable specification window. In the "Name" field, write "Student height, cm", click "OK". If the name is partially visible, then stretch the column.

5. Creation of a sample subject to the normal distribution law.

OŃTÚSTIK-QAZAQSTAN

While in the specification window for the variable "Var1", in the "Long name" field, write the formula "=RndNormal(7)+170" (Figure 1), and click the "OK" button. The program will automatically fill the cells with numbers.

	STATISTICA 64 - [Data: Sp	preadsheet1* (1v by 100c)]
Basic Multiple ANOVA Nonparametrics D Statistics Regression Base Department	Statistics Data Mining Graphs Tools Data Image: Statistics Data Mining Graphs Tools Data Istribution More Mult/Exploratory Image: Statistics Power Analysis Fitting Distributions Mult/Exploratory Image: Statistics Advanced/Multivariate	Scorecard Help Image: PLS, PCA, Image: PLS, PCA, Image: PLS, PCA, Image: PLS, PCA, </td
1 Рост студентов, см 4 5 6 7 7 8 9 9 10 10 11 12 13 14 15 16 17 18 19 19	A Arial A Arial A	OK Cancel
21 22 23 24 25 25 26 27 28	Long name (label or formula with <u>Eunctions</u>): =RndNormal(7)+170 Labels: use any text. Formulas: use variable names or v1, v2,, v0 is case Examples: (a) - mean(v1 v3, sqrt(v7), AGE) (b) - v1+v2, comment (after)	Function guide

Figure 1. Creating a sample subject to the normal distribution law

6. Changing the format of numeric data.

In the specification window for the variable "Var1", select "Number" in the "Display format" field, put "1" in the "Decimal places" field, and click the "OK" button.

7. Calculation of numerical characteristics of the sample.

Basic Statistics \rightarrow Descriptive statistics (Figure 2) \rightarrow OK.





Figure 2. Selecting the Descriptive statistics procedure

Set the variable "Variables" (Figure 3).

Descriptive Statistics: Spreadsheet1	
Variables: none	Summary
Quick Advanced Robust Normality Prob. & Scatterplots Categ. plots Options	Cancel
Su Select the variables for the analysis	D Options V
Image: Control of the second secon	Image: By Group Image: By Group Wight moments DF= Image: Wight moments
Select All Spread Zoom variables. Press F1	MD deletion
Select variables:	Casewise
Show appropriate variables only	NO. The shi was a which
ab, mark the desired numerical characterist e mean; mean; modard deviation;	
ard error of the mean; confidence interval for the mean;	edu, Kr. Kr. S. skrinens. er edu, Kr.
minimum and maximum; s - lower and upper quartiles;	na. equility of skinale et

Figure 3. Setting a variable

Select the "Advanced" tab, mark the desired numerical characteristics:

Valid N - sample size; Mean - average; Sum - amount; Median - median; Geom. mean - geometric mean; Harm. mean - harmonic mean; Standard Deviation - standard deviation; Variance - dispersion; Std. err. of mean - standard error of the mean; *Conf. limits for means - confidence interval for the mean;* Skewness - asymmetry; Kurtosis - excess: Minimum & maximum - minimum and maximum; Lower & upper quartiles - lower and upper quartiles; ma.edu.k2 Range - scope(Figure 4). Click the "Summary" button.

ОЙТÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ ЭЗЭ- ЭЗЭ- ЭЗЭ- ЭЗЭ- ЭЗЭ- ЭЗЭ- ЭЗЭ- ЭЗЭ	я медицинская академия»
Departments: "Medical Biophysics and Information Technologies",	№ 35-11(Б)-2024
"Social health insurance and public health"	Nº 582024
Guidelines for practical lessons of the course "Introduction to Scientific Research"	p.11 out of 68
	on the s. Mile



Figure 4. Selection of numerical characteristics

A summary t	table will appear on the screen (Figure 5).	Y. U.K
	STATISTICA - [Workbook1* - Descriptive Statistics (Spreadsheet1)]	_ 0 _ X
Home Edit View Insert	Format Statistics Data Mining Graphs Tools Data Workbook Scorecard Help O	ptions • 😭 🗕 🗗 🗙
Basic Multiple ANOVA Nonparametr Statistics Regression Base	Image: Non-Fitting Statusture Image: N	s• Block Data Stats•
Workbook1*	Descriptive Statistics (Spreadsheet1)	<u> </u>
Basic Statisti Descriptiv Variable	Valid N Mean Confidence Confidence Geometric Harmonic Median Sum Minimum Maximum Lower Upper Range Variance Std -95,000% 95,000% Mean Mean	.Dev. Stanc Err
Рост студентов, см	1 100 168,8230 167,3348 170,3113 168,6577 168,4918 169,1635 16882,30 151,8223 186,9792 163,9411 173,6095 35,15688 56,25610 7,5	00406 0,75

Figure 5. Final table with numerical characteristics of the sample

8. Constructing a histogram.

Return to the "Descriptive statistics" analysis window, select the "Normality" tab, click the "Histograms" button (Figure 6).

A Descriptive Statistics: Spreadsheet1	12 stinde e	0	? ×
🛃 Variables: Рост студентов, см			Summar⊻
Quick Advanced Robust Normality Prob. & S	catterplots Categ. plots Options		Cancel
Distribution Image: Erequency tables Image: Histograms Categorization Number of intervals:	Use Distribution Fitting, Process Analysis, or Graphs (P-P or Q-Q) to fit other distributions; use Survival Analysis to fit distributions to censored data.		Options
Integer intervals (categories) Normal expected frequencies Kolmogorov-Smirnov & Lilliefors test for norma Shapiro-Wilk's W test			s & w /ghtd momnts - V-1 N-1
3D histograms, bivariate distributions Categorized histograms	Stem and leaf Stem & leaf plot Compressed	-MD	deletion Casewise Pairwise

Figure 6. Plotting a histogram

A histogram will appear on the screen (Figure 7). The red line on the histogram is a graph of the density of a normal distribution.





9. Construction of a normal probability graph.

Return to the "Descriptive statistics" analysis window, select the "Prob. & Scatterplots" (Probability graphs and scatterplots), click the "Normal probability plot" button (Figure 8).

Contract Spreadsheet1	
Poct студентов, см	Summary
Quick Advanced Robust Normality Prob. & Scatterplots Ca	ateg. plots Options Cancel
2D scatterplot with names C Normal proba	ability plot
Up 3D scatterplot with names I Half-normal pro	obability plot 🔛 🔛 By Group
Categorized scatterplot	probability plot
Surface plot	1 ski no entr
E Scatterplot matrix	
is do. the st. the ser y	
3.0 du. Kt Stallo 2.00	@w-1
1 3. du. Hr St. Ma	MD deletion
the se do the se the	Casewise Pairwise

Figure 8. Construction of a normal probability graph

A normal probability graph will appear on the screen (Figure 9). The red line on the histogram is the density of the normal distribution, the blue dots are the sample observations. The closer the blue dots are to the red line, the more "normal" the distribution is.





Figure 9. Normal probability plot

10. Constructing a box-and-whisker plot.

Return to the "Descriptive statistics" analysis window, select the "Quick" tab, click the "Box & whisker plot for all variables" button (Figure 10).

🕼 Desci	riptive Statistics: Spread	sheet1		2, 1	e e	20		? ×
	⊻ariables: Poct cty,	дентов, см						Summary
Quick	Advanced Robust N	ormality Prob.	& Scatterplo	ts Categ. p	olots Options	r e		Cancel
	Summary: Statisti <u>c</u> s	🛛 🧱 Grapi	hs <u>1</u>	Graphs 2				Options -
	Erequency tables		Histogram	is in the second se		P. A		By Group
<u>Tãã</u>	Box & whis	er plot for all ve	ariables			24		
	Graphical com	parative summ	ary display	00		2.	2	<u>`</u> `
ð.	20111	1	N 20	· A			SELECT CRSES	s 🙃 🛛
						XL	W	ghtd momnts
							DF =	- /-1 ON-1
						1.1.1	MD	deletion
						20	OC	asewise
	10. 2.	<u>,0~ ,)</u>	F	2. 1) P	airwise

Figure 10. Plotting a box-and-whiskers plot

A box-and-whiskers plot will appear on the screen (Figure 11). The "legend" located at the bottom of the graph indicates: the mean and confidence intervals for the mean.



Figure 11. Box-and-whisker plot

ОЙТÚSTIК-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ SOUTH KAZAKHSTAN MEDICAL ACADEMY АО «Южно-Казахстанская медиц	цинская академия»
Departments: "Medical Biophysics and Information Technologies",	№ 35-11(Б)-2024
"Social health insurance and public health"	Nº 582024
Guidelines for practical lessons of the course "Introduction to Scientific Research"	p.14 out of 68

If you want the "legend" of the graph to contain information about the median, lower and upper quartiles, as well as the range, then you need to select the "Options" tab and check "Median/Quartiles/Range" (Figure 12), click the "Sammary" button, and then plot the box-and-whiskers plot again (Figure 13).



Figure 12. Selecting options for a box-and-whisker plot.



Figure 13. Box-and-whisker plot

11. Saving data.

Create a "Student" folder on your Desktop. In this folder, save the source data table under the name "Sample.sta": Home \rightarrow File \rightarrow Save \rightarrow Save As

In the same folder, save the workbook under the name "Analysis results.stw": Home \rightarrow File \rightarrow Save \rightarrow Save As.

ОŃTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ УЛУ- ОЙТÚSTIK-QAZAQSTAN SKMA -УЛУ- МЕДІСАL АСАДЕМҮ АО «Южно-Казахстанская медин	цинская академия»
Departments: "Medical Biophysics and Information Technologies",	№ 35-11(Б)-2024
"Social health insurance and public health"	№ 582024
Guidelines for practical lessons of the course "Introduction to Scientific Research"	p.15 out of 68

Tasks:

2) There are data on the duration (in seconds) of physical activity before the development of an angina attack in 12 people with coronary heart disease: 289, 203, 359, 243, 232, 210, 251, 246, 224, 239, 220, 211. Calculate the sample mean , variance, standard deviation, coefficient of variation, median, 25th and 75th percentiles. Can the sample be considered to be drawn from a population with a normal distribution? Check the results obtained in the Statistica program.

3) There are results for assessing the permeability of retinal vessels: 1.2; 1.4; 1.6; 1.7; 1.7; 1.8; 2.2; 2.3; 2.4; 6.4; 19.0; 23.6. Calculate sample mean, variance, standard deviation, coefficient of variation, median, 25th and 75th percentiles. Can the sample be considered to be drawn from a population with a normal distribution? Check the results obtained in the Statistica program.

4) For 25 days, the number of people who sought emergency medical help was recorded. The resulting sample is: 1, 0, 4, 2, 3, 5, 2, 4, 0, 1, 8, 5, 2, 4, 3, 3, 2, 5, 1, 3, 2, 5, 1, 3, 2. It is necessary to present the sample in the form of a variation series, calculate the sample mean, dispersion, standard deviation, coefficient of variation, mode and median. Check the results obtained in the Statistica program.

6. Assessment methods / technologies: Oral questioning. Practical work. (assessment using a checklist).

7. Literature:

1. Biostatistics: electronic textbook / K. Zh. Kudabaev [et al.]. - Shymkent: SKFFA, 2015.

2. Rosner Bernard Fundamentals of Biostatistics: Texbook/ B.Rosner - 8th ed.: GENGAGE learning, 2016.

- 3. Armitage P. Encyclopedia of Biostatistics. Wiley, 2016. 6100 p
- 4. Le C.T. Introductory biostatistics. Wiley, 2013. 536 p
- 5. Newman S.C. Biostatistical Methods in Epidemiology. Wiley, 2017. 388p
- 6. <u>http://www.biometrica.tomsk.ru</u>

8. Control:

- 1. What indicators of the variation series do you know?
- 2. What indicators are indicators of central tendency?
- 3. What are the indicators of dispersion?
- 4. What uhDo you know the steps for conducting statistical analysis in the STATISTICA program?
- 5. What types of documents can be created in the STATISTICA program? What extensions do they have?
- 6. List the main elements of the working window of the STATISTICA program.
- 7. What operations can be performed on the columns and rows of a spreadsheet?

Lesson №3

1. The theme: Fundamentals of the theory of testing statistical hypotheses. Consent criteria.

2. The aim: Develop the ability to use consent signs to check the assumption of a uniform distribution of the set.

3. The learning objectives:

- in what cases is the Pearson agreement sign and the Kolmogorov-Smirnov agreement sign used;

OŃTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY

SOUTH KAZAKHSTAN **SKM**A MEDICAL ACADEMY

«Оңтүстік Қазақстан медицина академиясы» АҚ

АО «Южно-Казахстанская медицинская академия» Departments: "Medical Biophysics and Information Technologies", № 35-11(Б)-2024 "Social health insurance and public health" Nº 58--2024Guidelines for practical lessons of the course "Introduction to Scientific Research' p.16 out of 68

- learn to form null and alternative forecasts;

- mastering the sign algorithm;
- learn to interpret the result;
- Implementation of the Pearson agreement and the Kolmogorov-Smirnov agreement in the program. STATISTICS.

4. Key questions of the theme:

- 1. Scheme for testing a statistical hypothesis.
- 2. Pearson test to test the hypothesis of normal distribution.
- 3. Kolmogorov-Smirnov test to test the hypothesis of normal distribution.
- 4. Testing the hypothesis of normal distribution in the STATISTICA program.

5. Methods / technologies of teaching and learning: Computer-based work / Solving situational tasks.

** Tasks:

Example 1.For statistical analysis of manufactured products, the fracture strength of the tablets was determined. The following results were obtained (in decinewtons):

Intervals	Frequencies, vi	Midpoints of intervals, xi
(373;421]	45	397
(421;445]	6 5	433
(445;469]	25	457
(469;493]	31	481
(493;517]	44	505
(517;541]	35	529
(541;565]	16	553
(565;589]	40.0	577

Test the hypothesis of normal distribution for tablet fracture using the χ^2 -Pearson goodness-offit test.

Solution.

1) Create a calculation table:

42436 26934
26934
1 1 1 1 1
46225
11191
1100
29435
44944
23716
225981

2) Calculate the average: $\overline{x} =$

MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ АО «Южно-Казахстанская медиц	инская академия»
Departments: "Medical Biophysics and Information Technologies",	№ 35-11(Б)-2024
"Social health insurance and public health"	№ 582024
Guidelines for practical lessons of the course "Introduction to Scientific Research"	p.17 out of 68

SOUTH KAZAKHSTAN

3) Calculate variance:

$$S^{2} = \frac{\sum_{i=1}^{n} (x_{i} - \bar{x})^{2} \cdot v_{i}}{n - 1} = \frac{225981}{164} \approx 1377,9$$

4) Calculate corrected sample variance:

$$s^{2} = \frac{n}{n-1}S^{2} = \frac{165}{165-1} \cdot 1377,9 \approx 1386$$

5) Calculate the corrected standard deviation:

OŃTÚSTIK-QAZAQSTAN

$$s = \sqrt{s^2} = \sqrt{1378} \approx 37$$

Using the goodness-of-fit criterion 2-Pearson test the hypothesis that the distribution is normal.

6) Determine the probability of a random variable falling into the intervals [xi, xi+1] using the formula:

$$p_i(x_i \leq X \leq x_{i+1}) = F\left(\frac{x_{i+1} - \overline{x}}{s}\right) - F\left(\frac{x_i - \overline{x}}{s}\right),$$

where F(x) is the distribution function of the normalized normal distribution(see table 1)

ex du	(1 0 (x)	2 . Edu	1. (a) (x) 36	ax of	• (x)
0,00	0,500000	1,00	0,841345	2,00	0,977250
0,05	0,519939	61,05	0,853141	2,05	0,979818
0,10	0,539828	1,10	0,864334	2,10	0,982136
0,15	0,559618	1,15	0,874928	2,15	0,984222
0,20	0,579260	1,20	0,884930	2,20	0,986097
0,25	0,589706	1,25	0,894350	2,25	0,987776
0,30	60,617911	1,30	0,903200	2,30	0,989276
0,35	0,636831	1,35	0,911492	2,35	0,990613
0,40	0,655422	1,40	0,919243	2,40	0,991802
0.45	0,673645	1,45	0,926471	2,45	0,992857
0.50	0,691463	1.50	0,933193	2,50	0,993790
0.55	0.708840	1.1.55 d	0,939429	2.55	0,994614
0.60	0.725747	1.60	0.945201	2.60	0,995339
0.65	0.742154	1.65	0.950528	2.65	0,995975
0.70	0.758036	1.70	0,955434	2.70	0,996533
0.75	0.773373	1.75	0.959941	2.75	0.997020
0.80	0.788145	1.80	0.964070	2,80	0,997445
0.85	0.802338	1.85	0.967843	2.85	0.997814
0.90	0.815940	1.90	S 0.971283	2.90	0.998134
0.95	0 828944	1.95	0.974412	2,95	0,998411
10.10	de do	10.100 H	1 2 2	3.00	0,998650

$$p_1(373 < X \le 421) = \Phi\left(\frac{421 - 500}{37}\right) - \Phi\left(\frac{373 - 500}{37}\right) = 0,02,$$

ОЙТÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ	цицинская академия»
Departments: "Medical Biophysics and Information Technologies",	№ 35-11(Б)-202- № 58 202
Guidelines for practical lessons of the course "Introduction to Scientific Research"	p.18 out of 68
$p_2(421 < X \le 445) = \Phi\left(\frac{445 - 500}{37}\right) - \Phi\left(\frac{421 - 500}{37}\right) = 0.05,$	du.r. K. Sky
$p_3(445 < X \le 469) = \Phi\left(\frac{469 - 500}{37}\right) - \Phi\left(\frac{445 - 500}{37}\right) = 0.15,$	3.03.000 du. K
$p_4(469 < X \le 493) = \Phi\left(\frac{493 - 500}{37}\right) - \Phi\left(\frac{469 - 500}{37}\right) = 0,21,$	KIN3. 3. edu
$p_5(493 < X \le 517) = \Phi\left(\frac{517 - 500}{37}\right) - \Phi\left(\frac{493 - 500}{37}\right) = 0,26,$	1. sknina.
$p_6(517 < X \le 541) = \Phi\left(\frac{541 - 500}{37}\right) - \Phi\left(\frac{517 - 500}{37}\right) = 0,19,$	MA SKU
$p_7(541 < X \le 565) = \Phi\left(\frac{565 - 500}{37}\right) - \Phi\left(\frac{541 - 500}{37}\right) = 0,09,$	edu. Kr. Kr. S.
$p_8(565 < X \le 589) = \Phi\left(\frac{589 - 500}{37}\right) - \Phi\left(\frac{565 - 500}{37}\right) = 0,03.$	13 eou du.k.
7) Create a calculation table:	K1. 3. 000

$$p_{2}(421 < X \le 445) = \Phi\left(\frac{445 - 500}{37}\right) - \Phi\left(\frac{421 - 500}{37}\right) = 0.05,$$

$$p_{3}(445 < X \le 469) = \Phi\left(\frac{469 - 500}{37}\right) - \Phi\left(\frac{445 - 500}{37}\right) = 0.15,$$

$$p_{4}(469 < X \le 493) = \Phi\left(\frac{493 - 500}{37}\right) - \Phi\left(\frac{469 - 500}{37}\right) = 0.21,$$

$$p_{5}(493 < X \le 517) = \Phi\left(\frac{517 - 500}{37}\right) - \Phi\left(\frac{493 - 500}{37}\right) = 0.26,$$

$$p_{6}(517 < X \le 541) = \Phi\left(\frac{541 - 500}{37}\right) - \Phi\left(\frac{517 - 500}{37}\right) = 0.19,$$

$$p_{7}(541 < X \le 565) = \Phi\left(\frac{565 - 500}{37}\right) - \Phi\left(\frac{541 - 500}{37}\right) = 0.03,$$

$$P_{8}(565 < X \le 589) = \Phi\left(\frac{589 - 500}{37}\right) - \Phi\left(\frac{565 - 500}{37}\right) = 0.03.$$

$$7) \quad \text{Create a calculation table:}$$

$$\boxed{\text{Interval} \quad \text{Frequencies vi} \quad \frac{\text{Probabilities}}{pi} \quad \frac{\text{Theoretic}}{\text{frequencie}} \\ p_{i}^{*} = np_{i}} \\ (373;421) = 4 \quad 0.02 \quad 3.3 \approx 3 \end{aligned}}$$

1.Kr 1 3		$7 \qquad (37 \qquad 37 \qquad 37 \qquad (160 \qquad 37 \qquad 3$			na. edu
$p_4 (469 < X$	$\leq 493) = \Phi\left(\frac{493}{3}\right)$	$\left(\frac{469-3}{7}\right) - \Phi\left(\frac{469-3}{37}\right)$	$\left(\frac{500}{2}\right) = 0,21,$	1. 1. St	KU3.3.80
$v_5 (493 < X)$	$\leq 517 = \Phi \left(\frac{517 - 37}{37} \right)$	$\left(\frac{-500}{7}\right) - \Phi\left(\frac{493 - 5}{37}\right)$	$\left(\frac{500}{2}\right) = 0,26,$	edu. Kr. Kr.	
$p_6(517 < X)$	$\leq 541 \Big) = \Phi \left(\frac{541 - 37}{37} \right)$	$\left(\frac{500}{7}\right) - \Phi\left(\frac{517 - 5}{37}\right)$	$\left(\frac{500}{2}\right) = 0,19,$	13. equi	KL SKI
$p_7 (541 < X)$	$\leq 565) = \Phi\left(\frac{565 - 37}{37}\right)$	$\frac{500}{7} - \Phi\left(\frac{541-5}{37}\right)$	$\left(\frac{00}{0}\right) = 0,09,$	KI Marae	edu, Kr. Kr.
$p_8(565 < X)$	$\leq 589 = \Phi \left(\frac{589}{25} \right)$	$\left(\frac{500}{7}\right) - \Phi\left(\frac{565 - 5}{27}\right)$	$\left(\frac{600}{2}\right) = 0.03.$		3. e0t 11.
7) 0	Create a calculation	n table:	Kna. 2. edu	J.KI VI SK	manaed
CON K	KI SKI M	edu tr	Theoretical	100° V1	
Interval	Frequencies vi	Probabilities <i>pi</i>	frequencies $v_i^* = np_i$	$\left(\boldsymbol{\nu}_i - \boldsymbol{\nu}_i^*\right)^2$	$\frac{\left(\nu_i - \nu_i^*\right)^2}{\nu_i^*}$
Interval (373;421]	Frequencies vi	Probabilities pi 0.02	frequencies $v_i^* = np_i$ $3.3\approx 3$	$\frac{\left(\nu_i - \nu_i^*\right)^2}{1}$	$\frac{(v_i - v_i^*)^2}{v_i^*}$
Interval (373;421] (421:445]	Frequencies vi 4 6	Probabilities pi 0.02 0.05	frequencies $v_i^* = np_i$ $3.3\approx 3$ $8.25\approx 8$	$\frac{\left(\nu_i - \nu_i^*\right)^2}{1}$	$\frac{(v_i - v_i^*)^2}{v_i^*}$ 0.3 0.5
Interval (373;421] (421;445] (445;469]	Frequencies vi 4 6 25	Probabilities <i>pi</i> 0.02 0.05 0.15	frequencies $v_i^* = np_i$ $3.3 \approx 3$ $8.25 \approx 8$ $24.75 \approx 25$	$\frac{\left(\nu_i - \nu_i^*\right)^2}{\frac{1}{4}}$	$ \begin{array}{c} \frac{\left(v_i - v_i^*\right)^2}{v_i^*} \\ \hline 0.3 \\ 0.5 \\ 0 \end{array} $
Interval (373;421] (421;445] (445;469] (469;493]	Frequencies vi 4 6 25 31	Probabilities pi 0.02 0.05 0.15 0.21	frequencies $v_i^* = np_i$ $3.3 \approx 3$ $8.25 \approx 8$ $24.75 \approx 25$ $34.65 \approx 35$		$ \begin{array}{c} \frac{(v_i - v_i^*)^2}{v_i^*} \\ \hline 0.3 \\ 0.5 \\ 0 \\ 0.46 \end{array} $
Interval (373;421] (421;445] (445;469] (469;493] (493;517]	Frequencies vi 4 6 25 31 44	Probabilities pi 0.02 0.05 0.15 0.21 0.26	frequencies $v_i^* = np_i$ 3.3 \approx 3 8.25 \approx 8 24.75 \approx 25 34.65 \approx 35 42.9 \approx 43		$ \begin{array}{c} \frac{(v_i - v_i^*)^2}{v_i^*} \\ \hline 0.3 \\ 0.5 \\ \hline 0 \\ 0.46 \\ 0.02 \end{array} $
Interval (373;421] (421;445] (445;469] (469;493] (493;517] (517;541]	Frequencies vi 4 6 25 31 44 35	Probabilities pi 0.02 0.05 0.15 0.21 0.26 0.19	frequencies $v_i^* = np_i$ $3.3 \approx 3$ $8.25 \approx 8$ $24.75 \approx 25$ $34.65 \approx 35$ $42.9 \approx 43$ $31.35 \approx 31$		$ \begin{array}{c} \frac{(v_i - v_i^*)^2}{v_i^*} \\ \hline 0.3 \\ 0.5 \\ 0 \\ 0.46 \\ 0.02 \\ 0.52 \\ \end{array} $
Interval (373;421] (421;445] (445;469] (469;493] (493;517] (517;541] (541;565]	Frequencies vi 4 6 25 31 44 35 16	Probabilities pi 0.02 0.05 0.15 0.21 0.26 0.19 0.09	frequencies $v_i^* = np_i$ $3.3 \approx 3$ $8.25 \approx 8$ $24.75 \approx 25$ $34.65 \approx 35$ $42.9 \approx 43$ $31.35 \approx 31$ $14.85 \approx 15$		$ \begin{array}{c} \frac{(v_i - v_i^*)^2}{v_i^*} \\ \hline 0.3 \\ 0.5 \\ \hline 0 \\ 0.46 \\ \hline 0.02 \\ \hline 0.52 \\ \hline 0.07 \\ \end{array} $
Interval (373;421] (421;445] (445;469] (469;493] (493;517] (517;541] (541;565] (565;589]	Frequencies vi 4 6 25 31 44 35 16 4	Probabilities pi 0.02 0.05 0.15 0.21 0.26 0.19 0.09 0.03	frequencies $v_i^* = np_i$ $3.3 \approx 3$ $8.25 \approx 8$ $24.75 \approx 25$ $34.65 \approx 35$ $42.9 \approx 43$ $31.35 \approx 31$ $14.85 \approx 15$ $4.95 \approx 5$		$ \begin{array}{c} \frac{(v_i - v_i^*)^2}{v_i^*} \\ \hline 0.3 \\ 0.5 \\ 0 \\ 0.46 \\ 0.02 \\ 0.52 \\ 0.07 \\ 0.2 \\ \end{array} $
Interval (373;421] (421;445] (445;469] (469;493] (493;517] (517;541] (541;565] (565;589] Sum	Frequencies vi 4 6 25 31 44 35 16 4	Probabilities pi 0.02 0.05 0.15 0.21 0.26 0.19 0.09 0.03 1 1	frequencies $v_i^* = np_i$ $3.3 \approx 3$ $8.25 \approx 8$ $24.75 \approx 25$ $34.65 \approx 35$ $42.9 \approx 43$ $31.35 \approx 31$ $14.85 \approx 15$ $4.95 \approx 5$ 165		$ \begin{array}{c} \frac{(v_i - v_i^*)^2}{v_i^*} \\ \hline 0.3 \\ 0.5 \\ 0 \\ 0.46 \\ 0.02 \\ 0.52 \\ 0.07 \\ 0.2 \\ 2.07 \\ \end{array} $
Interval (373;421] (421;445] (445;469] (469;493] (493;517] (517;541] (541;565] (565;589] Sum 8) Ca	Frequencies vi 4 6 25 31 44 35 16 4 Lculate:	Probabilities pi 0.02 0.05 0.15 0.21 0.26 0.19 0.09 0.03 1 1	frequencies $v_i^* = np_i$ $3.3 \approx 3$ $8.25 \approx 8$ $24.75 \approx 25$ $34.65 \approx 35$ $42.9 \approx 43$ $31.35 \approx 31$ $14.85 \approx 15$ $4.95 \approx 5$ 165		$ \begin{array}{c} \frac{(v_i - v_i^*)^2}{v_i^*} \\ \hline 0.3 \\ 0.5 \\ 0 \\ 0.46 \\ 0.02 \\ 0.52 \\ 0.07 \\ 0.2 \\ 2.07 \\ \end{array} $
Interval (373;421] (421;445] (445;469] (469;493] (493;517] (517;541] (541;565] (565;589] Sum 8) Ca	Frequencies vi 4 6 25 31 44 35 16 4 Lculate:	Probabilities pi 0.02 0.05 0.15 0.21 0.26 0.19 0.09 0.03 1	frequencies $v_i^* = np_i$ 3.3 ≈ 3 8.25 ≈ 8 24.75 ≈ 25 34.65 ≈ 35 42.9 ≈ 43 31.35 ≈ 31 14.85 ≈ 15 4.95 ≈ 5 165 $k_i (v_i - v_i^*)^2$	$ \begin{pmatrix} \nu_i - \nu_i^*)^2 \\ 1 \\ 4 \\ $	$ \begin{array}{c} \frac{(v_i - v_i^*)^2}{v_i^*} \\ \hline 0.3 \\ 0.5 \\ 0 \\ 0.46 \\ 0.02 \\ 0.52 \\ 0.07 \\ 0.2 \\ 2.07 \\ \end{array} $
Interval (373;421] (421;445] (445;469] (469;493] (493;517] (517;541] (541;565] (565;589] Sum 8) Ca	Frequencies vi 4 6 25 31 44 35 16 4	Probabilities pi 0.02 0.05 0.15 0.21 0.26 0.19 0.09 0.03 1 $\chi^2_{pacy} =$	frequencies $v_i^* = np_i$ 3.3 ≈ 3 8.25 ≈ 8 24.75 ≈ 25 34.65 ≈ 35 42.9 ≈ 43 31.35 ≈ 31 14.85 ≈ 15 4.95 ≈ 5 165 $k_i (v_i - v_i^*)^2$ $v_i^* = 2,0^{\circ}$		$ \begin{array}{c} \frac{(v_i - v_i^*)^2}{v_i^*} \\ \hline 0.3 \\ 0.5 \\ 0 \\ 0.46 \\ 0.02 \\ 0.52 \\ 0.07 \\ 0.2 \\ 2.07 \\ \end{array} $
Interval (373;421] (421;445] (445;469] (469;493] (493;517] (517;541] (541;565] (565;589] Sum 8) Ca	Frequencies vi 4 6 25 31 44 35 16 4 Lculate:	Probabilities pi 0.02 0.05 0.15 0.21 0.26 0.19 0.09 0.03 1 $\chi^2_{pacy} =$	frequencies $v_i^* = np_i$ 3.3~3 8.25~8 24.75~25 34.65~35 42.9~43 31.35~31 14.85~15 4.95~5 165 $k_i \frac{(v_i - v_i^*)^2}{v_i^*} = 2,0^{10}$	$ \begin{array}{r} (\nu_i - \nu_i^*)^2 \\ 1 \\ 4 \\ 0 \\ 16 \\ 1 \\ 16 \\ 1 \\ 1 7. $	$ \begin{array}{c} \frac{(v_i - v_i^*)^2}{v_i^*} \\ 0.3 \\ 0.5 \\ 0 \\ 0.46 \\ 0.02 \\ 0.52 \\ 0.07 \\ 0.2 \\ 2.07 \\ \end{array} $

$$\chi^{2}_{pacy} = \sum_{i=1}^{k} \frac{(\nu_{i} - \nu_{i}^{*})^{2}}{\nu_{i}^{*}} = 2,07.$$

$$\chi 2_{table}(p; f), \chi 2_{table}(0.05;5)=11.1.$$
 (see table 2)

8) Calculate				5105		2.07
o) Calculate		10. 2.	an Ar	*)2	a co do	· VI St
41. 3. 000		2	${}^{2}_{nacu} = \sum_{k=1}^{k} \frac{(\nu_{i})}{(\nu_{i})}$	$(-v_i) = 2.0$	\overline{n}	XV. VI
. Kn 3. d		Sign		v_i^*	and e	1, 1.1.1.
9) Define:				11. 11		60, 11. L.
	$\chi 2_t$	able $(p; f), \chi$	$2_{table}(0.05;5)$)=11.1. (see t	table 2)	3. 600 ()
). VI		601 11.K		2. du
Чисто	N. 3.	<u>-200'' 12</u>	Уровень	значимости 🖉	et de G	A COLO
степеней свободы	0,01	0,05	0,1	0,90	0,95	6 0,99
	6,6	3,8	2,71 9	0,02	0,004	0,0002
3. 0 2	9,2	6,0	4,61	5 0,21	0,1	0,02
	11,3	9.5	6,25	0,58	0,35	0,12
5	15,1	11.1	9,24	1,61	1,15	0,55
35 110 00	Nº VI	St a	0. 60 M	in the d	4 2.	200 141
	1. 1.0 × 1		A. 60.			
	e0. '''		A			2.0 00.

e «C	Оңтүстік Қазақста	О́́л ан медицина	TÚSTIK-QAZAQS MEDISI AKADEMI академиясы»	TAN INA ASY AK	SOUTH KAZAKHS' MEDICAL ACADEMY AO «Южно-Ка:	ГАN захстанская ме;	цицинская академи	у w
	Depart	tments: "Medi	cal Biophysics	s and Information	on Technologies	s", X	№ 35-11(Б)-20	024
2	Ch. Kr	"Social	health insurance	ce and public he	ealth"	· (,)), '	Nº 582	024
0	Guidelines for	p.19 out of 6	58					
			126	C INC	1 220		0.07	
		10,0	12,0	12.0	2,20	2,17	1,24	
		18,5	15.5	12,0	2,05	2,17	1.65	
	0.0	G1 7	16.9	147	4 17	5 3 33	0 2 09	
	. 10	23.2	183	160 8	4.87	3.94	236	
		24.7	10,7	17.3	5 58	4 57	3.05	
	12	26.2	21.0	18.5	6.30	5.23	3.57	
	13	27.7	224	19.8	7.04	5 89	A 11 00	
	14 6	29.1	23.7	21.1	7.79	6.57	4.66	
	15	30.6	25.0	22.3	8.5	7.26	5.23	
\mathcal{D}	16	32.0	26.3	23.5	9.31	7.98	5.81	
	17	33.4	27.6	24.8	10.1	8.67	6.41	γ_{\cdot}
	18	34.8	28.9	26.0	10.9	9.39	7.01	
	19	36.2	30.1	27.2	11.7	10.1	7.63	
	20	37.6	31.4	28.4	12.4	10.9	8.26	
	210	38,9	32,7	29,6	13,2	11.6	8,90	
	22	40,3	33,9	30,6	14,0	12,63	9,54	
	23	41,6	35,2	32,0	14,8	13,1	10,2	
	24	(43,0)	36,4	33,2	15,7	13,8	10,9	
$\sum_{i=1}^{n}$	25	6 44,3	37,7	34,4	16,5	14,6	11,5	
	26	45,6	38,9	9 35,6	17,3	15,4	12,2	1.
	275	47,0 0	40,1	36,7	0 18,1	16,2	12,9	
	28	48,3	041,3	37,9	18,9	16,9	13,6	
	29	49,6	42,6	39,19	19,8	17,7	614,3	
	30	50.9	43.8	40.3	20.6	18.5	15.0	

1) Compare $\chi^{2}_{calculation}$ And χ^{2}_{table} : $\chi^{2}_{pacy} < \chi^{2}_{ma\delta n}$, which means the hypothesis that the distribution is normal is accepted.

Example 2. Using the goodness-of-fit criterion Kolmogorov-Smirnov to test the hypothesis that the distribution is normal.

1) Determine the values of the theoretical distribution function using the formula:

$$F(x) = \frac{1}{2} + \Phi\left(\frac{x_{i+1} - \bar{x}}{s}\right),$$

where $\Phi(x)$ is the Laplace function (see table 3)

x	0	21	2	3	4	59	6	7	8	9
o,	K.	.1	12M	C	отые ,	доли	x	<u></u>	<u> </u>	5. 1
0,0	0,0000	0040	0080	0112	0160	0199	0239	0279	0319	0359
0,1	0398	0438	0478	0517	0557	0596	0636	0675	0714	0754
0,2	0793	0832	0871	0910	0948	0987	1026	1064	1103	1141
0,3	1179	1217	1255	1293	1331	1368	1406	1443	1480	1517
0,4	1554	1591	1628	1664	1700	1736	1772	1808	1844	1879
0,5	1915	1950	1985	2019	2054	2088	2123	2157	2190	2224
0,6	2258	2291	2324	2357	2389	2422	2454	2486	2518	2549
0,7	2580	2612	2642	2673	2704	2734	2764	2794	2823	2852
0,8	2881	2910	2939	2967	2996	3023	3051	3079	3106	3133
0,9	3159	3186	3212	3238	3264	3289	3315	3340	3365	3389

«Он	атустік Қазақс	стан меди		ТІК-QAZA MED AKADEJ адемияс	QSTAN ISINA MIASY ы» АҚ	-1979- 	SOUTH H MEDICA ACADE AO «Ю	(AZAKHST AL MY жно-Каз	AN ахстансі	кая меди	цинская акад	цемия»
0	Depa	artments: "	'Medical	Biophys	sics and	Informat	ion Tech	nologies	", 00	K	№ 35-11(1	5)-2024
$\sqrt{0}$	Guidelines f	or practica	al lesson	s of the c	ourse "In	ntroducti	on to Sci	entific R	esearch'		58- p.20 out	-2024 of 68
		2412	2420	2461	2405	2500	9591	2552	9577	2500	2621	
		3413	3430	2606	3400	2790	2740	2770	3011	2010	2021	
	1,1	3043	3000	3080	3100	3129	3749	3/10	3790	2007	3030	
	5 1,2	3849	3869	3888	3907	3925	3944	3962	3980	3990	4015	Yr.1
	1,3	4032	4049	4066	4082	4099	4115	4131	4147	4162	4177	
\mathcal{V}	1,4	4192	4207	4222	4236	4251	4265	4279	4292	4306	4319	
	1,5	4332	4345	4357	4370	4382	4394	4406	4418	4430	4441	
	1.6	4452	4463	4474	4485	4495	4505	4515	4525	4535	4545	2.00
	1.7	4554	4564	4573	4582	4591	4599	4608	4616	4625	4633	0.0
	1.8	4641	4649	4656	4664	4671	4678	4686	4693	4700	4706	
	1,9	4713	4719	4726	4732	4738	4744	4750	4756	4762	4767	
	Десятые доли х								Y.V			
	2,	4773	4821	4861	4893	4918	4938	4953	4965	4974	4981	
	5 3.	4987	4990	4993	4995	4997	4998	4998	4999	4999	5000 ⁸	

$$F_{1}(x) = \frac{1}{2} + \Phi\left(\frac{421 - 500}{37}\right) = 0,02,$$

$$F_{2}(x) = \frac{1}{2} + \Phi\left(\frac{445 - 500}{37}\right) = 0,07,$$

$$F_{3}(x) = \frac{1}{2} + \Phi\left(\frac{469 - 500}{37}\right) = 0,20,$$

$$F_{4}(x) = \frac{1}{2} + \Phi\left(\frac{493 - 500}{37}\right) = 0,43,$$

$$F_{5}(x) = \frac{1}{2} + \Phi\left(\frac{517 - 500}{37}\right) = 0,68,$$

$$F_{6}(x) = \frac{1}{2} + \Phi\left(\frac{541 - 500}{37}\right) = 0,87,$$

$$F_{7}(x) = \frac{1}{2} + \Phi\left(\frac{565 - 500}{37}\right) = 0,96,$$

$$F_{8}(x) = \frac{1}{2} + \Phi\left(\frac{589 - 500}{37}\right) = 0,99.$$
1) Create a calculation table:

Interval [xi, xi+1]	Frequencies vi	Cumulative frequencies vi, accumulated	$F_n(x) = \frac{V_{i, \text{ накопл}}}{n}$	F(x)	$\left F_n(x)-F(x)\right $
(373;421]	4	40	4/165=0.02	0.02	
(421;445]	6	4+6=10	10/165=0.06	0.07	0.01
(445;469]	25	10+25=35	-0.21	0.20	0.01
(469;493]	31	35+31=66	0.4	0.43	0.03
(493;517]	44	66+44=110	0.67	0.68	0.01
(517;541]	35	110+35=145	0.88	0.87	0.01
(541;565]	16	145+16=161	0.98	0.96	0.02
(565;589]	4	161+4=165	100 100	0.99	0.01

$$d_{\max} = \max |F_n(x) - F(x)| = 0.03.$$

2) Calculate:
$$\lambda_{pacy} = d_{\max} \sqrt{n} = 0.03 \cdot \sqrt{165} \approx 0.39$$
.

3) $\lambda_{maon} = 1,36$.

4) Compare λ_{pacu} And $\lambda_{ma\delta n}$: $\lambda_{pacu} < \lambda_{ma\delta n}$, which means the hypothesis that the distribution is normal is accepted.

OŃTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY



«Оңтүстік Қазақстан медицина академиясы» АҚ

АО «Южно-Казахстанская медицинская академия» Departments: "Medical Biophysics and Information Technologies", № 35-11(Б)-2024 "Social health insurance and public health" № 58--2024

Guidelines for practical lessons of the course "Introduction to Scientific Research" p.21 out of 68

Tasks.

1) The sample is presented in the form of a statistical series (n=200):

xi	0.3	0.5	0.7	0.9	1.1	1.3	1.5	1.7	1.9	2.1	2.3
Vi	6	9	26	25	30	26	21	24	20	8	5

Test the hypothesis about the normal distribution of the sample using the γ 2-Pearson goodnessof-fit test.

6. Assessment methods / technologies: Oral questioning. Practical work. (assessment using a checklist).

7. Literature:

Main:

1. Biostatistics: electronic textbook / K. Zh. Kudabaev [et al.]. - Shymkent: SKFFA, 2015.

2. Rosner Bernard Fundamentals of Biostatistics: Texbook/ B.Rosner - 8th ed.: GENGAGE learning, 2016.

- 3. Armitage P. Encyclopedia of Biostatistics. Wiley, 2016. 6100 p
- 4. Le C.T. Introductory biostatistics. Wiley, 2013. 536 p
- 5. Newman S.C. Biostatistical Methods in Epidemiology. Wiley, 2017. 388p
- 6. http://www.biometrica.tomsk.ru

8. Control:

- 1. What is a statistical hypothesis? What types of statistical hypothesis do you know?
- 2. What is the general scheme for testing statistical hypotheses?
- 3. What are the consent criteria used for?
- 4. What is the scheme for applying the Pearson χ^2 goodness-of-fit test?
- 5. What is the scheme for using the Kolmogorov-Smirnov goodness-of-fit test?
- 6. How are Pearson's χ^2 goodness-of-fit tests implemented in the Statistica program?

Lesson №4

1. The theme: Parametric methods of comparative statistics.

2. The aim: studying the methodological foundations and conditions for using the two-sample and paired Student's t-test and introducing them into the "STATISTICA" program.

3. The learning objectives:

- understand in what cases two-sample and paired sampling is usedt-Student's test;
- learn to formulate null and alternative hypotheses;
- master the algorithm of the criterion;
- learn to interpret the result;

-cdevelop skills in applying two-sample and paired Student's t-tests in the STATISTICA program.

4. Key questions of the theme:

- 1. Student's t-test for analyzing biomedical data.
- 2. Conditions for using the Student t-test.
- 3. Scheme for applying the two-sample Student t-test.
- 4. Scheme for applying the paired Student's t-test.
- 5. Interpretation of the result.
- 6. Checking the conditions for applying the Student's test (normal distribution of samples, equality of variances).
- 7. The procedure for implementing the Student's t-test in the "STATISTICA" program.

ОŃTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ	цинская академия»
Departments: "Medical Biophysics and Information Technologies",	№ 35-11(Б)-2024
"Social health insurance and public health"	№ 582024
Guidelines for practical lessons of the course "Introduction to Scientific Research"	p.22 out of 68

8. Interpretation of results.

Example 1. If the cervix does not dilate for a long time during childbirth, the duration of labor increases and a caesarean section may be necessary. Scientists decided to find out whether a gel with prostaglandin E2 accelerates the dilatation of the cervix. The study included 2 groups of women in labor. Women in labor in the first group were injected into the cervix with prostaglandin E2 gel, while women in the second group were injected with a placebo gel. In both groups there were 21 women in labor, the age, height and duration of pregnancy were approximately the same. Labor in the group receiving prostaglandin E2 gel lasted on average 8.5 hours (standard deviation 4.7 hours), in the control group - 13.9 hours (standard deviation 4.1 hours). Is it possible to say that the gel with prostaglandin E2 shortened the duration of labor?

Solution.

1) H₀:
$$\bar{x}_1 = \bar{x}_2$$
.

$$\mathbf{H}_1: x_1 \neq x_2.$$

3)
$$t_{pacy} = \frac{13.9 - 8.5}{\sqrt{(21-1) \cdot 4.1^2 + (21-1) \cdot 4.7^2}} \cdot \sqrt{\frac{21 \cdot 21}{21+21}} (21+21-2) \approx 4.$$

4)
$$t_{ma\delta n}(0,05;40) = 2,02$$
 .(See table 5)

Erkindik darijeler sana	KI SKIIINO	Mynyzdyl	yk money, r	na.edu.u.kl
f	0.10	0.05	0.02	0.01
12.	6.31	12.7	31.82	63.7
2	2.92	4.30	6.97	9.92
53	2.35	3.18	4.54	5.84
4	2.13	2.78	3.75	4.60
5 5	2.01	2.57	3.37	4.03
6 9	1.94	2.45	3.14	3.71
207 V	5 1.89	2.36	3.00	3.50
8	1.86	2.31	2.90	3.36
9	1.83	2.26	2.82	3.25
10	1.81	2.23	2.76	3.17
M CM	1.80	2.22	2.72	3.11
5 12	1.78	2.18	2.68	3.05
413	(1.77)	2.16	2.65	3.01
14	1.76	2.14	2.62	2.98
15	1.75	2.13	2.60	2.95
16	1.75	2.12	2.58	2.92
17	1.74	2.11	2.57	2.90
0 18	1.73	2.10	2.55	2.88
19	1.73	2.09	2.54	2.86
20	1.73	2.09	2.53	2.85

OŃTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY



«Оңтүстік Қазақстан медицина академиясы» АҚ

 Departments: "Medical Biophysics and Information Technologies",
 № 35-11(Б)-2024

 "Social health insurance and public health"
 № 58- -2024

 Guidelines for practical lessons of the course "Introduction to Scientific Research"
 p.23 out of 68

30	1.70	2.04	2.46	2.75
40	1.68	2.02	2.42	2.70
60	1.67	2.00	2.39	2.66
120	1.66	1.98	2.36	2.62
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	1.64	1.96	2.33	2.58

5) Because  $t_{pacy} > t_{ma\delta a}$ , then "H0" is rejected, i.e. gel with prostaglandin E2 shortened the duration of labor.

**Example 2**. To evaluate the effectiveness of the new hypoglycemic drug, blood glucose levels were measured in patients with diabetes mellitus before and after taking the drug:

Blood glucose level, mol/l				
before taking the drug	after taking the drug			
9.6	5.7			
8.1	4.2			
8.8	6.4			
7.9	5.5			
9.2	5.3			
8.0	5.2			
8.4	5.1			
5 00.1	5.9			
7.8	7.5			
8.1	5.0			
8.6	5.6			
	Blood glucos           before taking the drug           9.6           8.1           8.8           7.9           9.2           8.0           8.4           10.1           7.8           8.1           8.6			

Can we assume that after taking the drug, the level of glucose in the blood of patients decreases? *Solution.* 

1) H₀:  $\bar{x}_1 = \bar{x}_2$ .

 $\mathbf{H}_1: \overline{x}_1 \neq \overline{x}_2 \,.$ 

2)  $\alpha$ =0.05- significance level.

Kr. S.	Blood glucos	e level, mol/l	Difference of	J. 11 . 64
Patient no.	before taking the drug	after taking the drug	values $d = x_i - y_i$	$(d_i - \overline{d})^2$
H 1 5	9.6	5.7	3.9	0.77
2 9	8.1	4.2 0	3.9	0.77
3	8.8	6.4	2.4	0.38
4	7.9	5.5	2.4	0.38
5	9.2	5.3	3.9	0.77
6	8.0	5.2	2.8	0.048
7 07	8.4	5.1	3.3	0.078
80	10.1	5.9	4.2	1.39
9 0	7.8	7.5	0.3	7.4
10	8.1	5.0	5 3.1 0	0.006

омти́ятік-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ АО «Южно	.KHSTAN о-Казахстанская медицинская академия»
Departments: "Medical Biophysics and Information Technolo "Social health insurance and public health"	ogies", № 35-11(Б)-2024 № 582024
Guidelines for practical lessons of the course "Introduction to Scienti	ific Research" p.24 out of 68
Sum A Sum	30.2 12

1) 
$$\overline{d} = \frac{30,2}{10} = 3,02 \cdot \overline{d} = \frac{30,2}{10} = 3,02$$

2) 
$$S_d = \sqrt{\frac{\sum_{i=1}^n (d_i - \overline{d})^2}{n-1}}$$
  $S_d = \sqrt{\frac{12}{9}} = 1.15$ 

3) 
$$t_{pacy} = \frac{\overline{d}}{S_d / \sqrt{n}} t_{pacy} = \frac{3,02}{1,15/\sqrt{10}} = 8,39$$

4)  $t_{ma\delta n}(0,05;9) = 2,26$ . (See Table 5)

5) Because  $t_{pacy} > t_{ma\delta n}$ , then "H0" is rejected, i.e. The blood glucose level has decreased after taking the drug, which means the new drug is effective.

**Example 3.** Medical students had their pulse examined before and after passing the exam. Pulse rate before the exam was  $98.8\pm4.0$ , and after the exam  $84.0\pm5.0$ . Can we assume that after the exam the heart rate decreases and approaches normal?

$$\mathbf{H} \cdot \mathbf{\bar{x}} \cdot \mathbf{\bar{x}}$$

2) 
$$\alpha = 0.05$$
.

3)  $t_{pacy} = \frac{98,8-84}{\sqrt{4^2+5^2}} \approx 2,3.$ 

4) Because tcalc>2, then "H0" is rejected, i.e. After the exam, the heart rate decreases and approaches normal.

**Example 4**. Compare test results in two groups.

No.	Group No. 1 results (sec.)	Group No. 2 results (sec.)
1	thirty	46 0 1
2	45	49
3	41 5 00 0	52
4	38 24 00	550 00
55	34 1	56
6	36	40
5.7.1	31	46
8	1. 24 3.30,00 4. 5	51 4 5
9	49	58
10	50	46
NI C	S 51 S 51	46
12	46	56
13	15 AV 12 415 00 00	53 0 00 1

интүстік Қазақстан ме	ОŃTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY едицина академиясы» АҚ	SOUTH KAZAKHSTAN MEDICAL ACADEMY AO «Южно-Казахстанская м	иедицинская академия»
Departmen	ts: "Medical Biophysics and Info "Social health insurance and put	rmation Technologies",	№ 35-11(Б)-2024 № 582024
Guidelines for prac	tical lessons of the course "Introd	duction to Scientific Research"	p.25 out of 68
14	1 37 3.00	57	LUT IT SKUT
15	36	44	C' NOT A
16	34	42	200 JUN 1
17	-33	40	(0, 0, 1), 1, 1
18	49	58	100. 600 1. Kr.
19 9	32 5	54	24. 3. 60
20 5	46	531	24° ~ 2. 000
21 5	41	51	1 24 2
22	St 0 44 80 10.	57	H 3. 1811 - 3.0
23	38	56	Kr S. Mar
24	50	44	N. Kr St. U.

«C

Create a data table "Test Results" of size 2*58 in the Statistica program, enter the initial data.
 Schot Statistica Paris Statistica enter independent has severe (4 tot for independent)

2. Select Statistics  $\rightarrow$  Basic Statistics  $\rightarrow$  t-test independent by groups (t-test for independent samples) (Figure 1).



Figure 1. Selecting the "t-test independent by groups" procedure

3. In the dialog box, click the "Variables" button, indicate the grouping characteristic (column containing group codes) in the right part of the window, and the column containing the analyzed characteristic in the left part of the window (Figure 2), click the "OK" button.





Figure 2. Setting variables

#### 4. Check samples for normal distribution.

Select the "Advanced" tab, build "Categorized normal plots" and "Categorized histograms" Draw a conclusion about the normality of the distribution (Figure 3).



Figure 3. Testing samples for normal distribution

5. Check the equality of sample variances.

To check this condition, the program automatically uses Fisher's F-test, but you can also use Levene's test, to do this you need to select the "Options/Levene's test" option (Figure 4).

🖓 T-Test for Independent Samples by Gr	oups: Spreadsheet1	? 📈	
Mariables: Dependent: результать Grouping: группа	Cancel		
Code for Group 1: 1 Coc	le for Group 2: 2	Doptions -	
Quick Advanced Options	By Group		
Display long variable names     Test w/ separate variance estimates     Multivariate test (Hotelling's T*)	Stitter s & w CASES S & w Weighted moments		
p-value for highlighting: _05 € Cl for estimates 95,00 € %	source in the second	DF = W-1 N-1 MD deletion Casewise Pairwise	

Figure 4. Selecting the "Levene's test" option

ойти́зтік-qazaqstan MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ Оңтүстік Қазақстан медицина академиясы» АҚ	инская академия»
Departments: "Medical Biophysics and Information Technologies",	№ 35-11(Б)-2024
"Social health insurance and public health"	№ 582024
Guidelines for practical lessons of the course "Introduction to Scientific Research"	p.27 out of 68

6. Click the "Summary" button, a final table will appear on the screen with the results of comparing two independent samples using the Student's t-test (Figure 5).

	T-tests; Gr	ouping: rpy	ипа (Spre	adsh	eet1)	K		$\mathcal{O}$	11	6	~~~	67		
CF AC	Group 1: 1													Kr
1 1 2	Group 2: 2													$\sum ($
VLV ST A	Mean	Mean	t-value	df	р	Valid N	Valid N	Std.Dev.	Std.Dev.	F-ratio	р	Levene	df	р
Variable	1	2	. 4		1 7	1	2	. 1	2	Variances	Variances	F(1,df)	Levene	Levene
результаты тестирования	40,20690	49,96552	-6,06567	56	0,000000	29	29	6,444152	5,790897	1,238340	0,575457	0,222792	56	0,638755

Figure 5. Final table with the results of comparison of two independent samples using Student's t-test

Names of the columns of the final table:

Mean1, Mean 2 – average values of variables;

*t-value*- t-test value;

*df*– number of degrees of freedom;

*R*-level of significance of the t-test;

Valid1, Valid 2 - number of observations in groups;

Std. Dev.- standard deviations of variable values;

F-ratio Variances-F-criterion value;

*p*Variances - level of significance of the F-test;

*Levene F*(*1*,*df*) –Lewin's criterion value;

df Levene- number of degrees of freedom of Lewin's criterion;

r Levene -level of significance of Lewin's criterion.

If Levine's test p<0.05, a conclusion should be drawn about the difference in variances in the compared groups.

If for Lewin's test p>0.05, one should conclude that the variances in the compared groups are equal.

Similarly for the F-test.

If the t-test value is p>0.05, then the null hypothesis of equality of means is accepted.

If the t-test has a p-value of <0.05 (such results are highlighted in red), then the null hypothesis of equality of means is rejected.

For this example:

• *p*Variances>0.05, which means the variances are equal (Fisher's F test);

• *r Levene* >0.05, which means the variances are equal (Levine's criterion);

• p < 0.05, which means the null hypothesis of equality of means is rejected (Student's t-test).

**Example 5**.Compare the test results of students before and after training.

No.	Results before training (sec.)	Results after training (sec.)
1 2	thirty	46
2	5 45 1 5	49
34	St and Call	52 20 3. 0
4	38	55
5	34 00 1	56
6	36	40
~~7 e	31 3. 0	46
8	30	Kr S. ( 51 W. Q St
9	49	58 6 10 1
10	50 50	46

«Оңт	MEDISINA AKADEMIASY устік Қазақстан медицина академиясы» АҚ	MEDICAL ACADEMY АО «Южно-Казахстанская меди	цинская академия»
3.00	Departments: "Medical Biophysics and Info "Social health insurance and pu	ormation Technologies",	№ 35-11(Б)-2024 № 582024
	Buidelines for practical lessons of the course "Intro	oduction to Scientific Research"	p.28 out of 68
ele	ven 51	46	3. VI SA
A	2 46	56	N. V. St
1	3 41	53	e 11. 11
1	4 37	57	- C
1	5 36	44	(0, 60, 11).
1	6 34 9	42	· · · · · · · · · · · · · · · · · · ·
1	7 33	se 40	X41
K .	8 49	58	Kin di
01	9 32	54	1 3. 15. 3.
2	0 46	53	Fr S. Mis
2	1 2 41 0 0	51	VI ST N
2 2	2 44	57	N. KI St.
2	3 38	56	Nº VI G
2	4 50	44	8 x x x x x
2	5 37	\$42	· · · · · · · · · · · · · · · · · · ·
2	6 39	49	20. 60 11 H
2	7 40 40	50	N. 2. 00
2	8 46	55	Kr. 3. 01
2	9 42	43	3. 12. 3.
>			

-cdb2-

SOUTH KAZAKHSTAN

ΟΝΤΙΊSTIK-ΟΑΖΑΟSTAN

*1*. Create a data table "Test results before and after training" of size 2*29 in the Statistica program, enter the initial data.

2. Check samples for normal distribution.

Statistics  $\rightarrow$  Basic statistics  $\rightarrow$  Descriptive statistics  $\rightarrow$  Normality. Normality testing can be done using the Kolmogorov-Smirnov & Lilliefors test of normality or the Shapiro-Wilk's W test. To do this, select the appropriate options and click the "Histograms" button (Figure 1).



Figure 1. Testing samples for normal distributions

The histograms that appear (Figure 1) at the top show the results of tests for normality, which can be explained as follows:

- if in these tests p>0.05, then the hypothesis of normal distribution is accepted;
- if p<0.05, then the hypothesis of normal distribution is rejected.

О́́́ИТÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ О́́Ю́Х САДЕМУ АО «Южно-Ка	STAN азахстанская медицинская академия»
Departments: "Medical Biophysics and Information Technologie	es", № 35-11(Б)-2024
"Social health insurance and public health"	Nº 582024
Guidelines for practical lessons of the course "Introduction to Scientific	Research" p.29 out of 68

3. To test the null hypothesis about the equality of means in dependent groups using the Student t-test, select: Statistics  $\rightarrow$  Basic statistics  $\rightarrow$  t-test dependent samples (t-test for dependent samples) (Figure 2).



Figure 2. Selecting the t-test dependent samples procedure

- 4. Set variables "Variables".
- 5. Click the "Summary" button, a summary table will appear on the screen (Figure 3).

T-test for I Marked dit	Dependent fferences a	Sam ire s	ples (Spre ignificant a	adsheet1) at p < ,0500	0 0		1 3	KIND	eduin
Mean	Std.Dv.	N	Diff.	Std.Dv.	et y	df	p	Confidence	Confidence
10 20600	6 444152		YV.	S Dill.				-33,00070	.33,00070
50,00000	5,769377	29	-9,79310	7,537427	-6,99675	28	0,000000	-12,6602	-6,92602
	T-test for I Marked di Mean 40,20690 50,00000	T-test for Dependent           Marked differences a           Mean         Std.Dv.           40,20690         6,444152           50,00000         5,769377	T-test for Dependent Sam Marked differences are s Mean Std.Dv. N 40,20690 6,444152 50,00000 5,769377 29	T-test for Dependent Samples (Spre Marked differences are significant a MeanMeanStd.Dv.NDiff.40,206906,44415250,000005,76937729-9,79310	T-test for Dependent Samples (Spreadsheet1)           Marked differences are significant at p < ,0500           Mean         Std.Dv.         N         Diff.         Std.Dv.           40,20690         6,444152         0.00000         0.7537427	T-test for Dependent Samples (Spreadsheet1)           Marked differences are significant at p < ,05000	Marked differences are significant at p < ,05000	Marked differences are significant at p < ,05000	Marked differences are significant at p < ,05000

Figure 3.Final table with the results of comparison of two dependent samples using Student's t-test

The results obtained can be interpreted as follows:

- if p<0.05, then the null hypothesis is rejected (such results are highlighted in red font);
- if p>0.05, then the null hypothesis is accepted.

For this example R < 0.05, which means the null hypothesis of equality of means is rejected.

**Example 6.** Compare test results in two groups. The samples are normally distributed and have equal variances. It is known that  $n_1 = 29$ ,  $n_2 = 29$ ,  $\overline{x_1} = 40$ ,  $\overline{x_2} = 50$ ,  $s_1 = 6,4$ ,  $s_2 = 5,8$ .

1. Select Statistics  $\rightarrow$  Basic Statistics  $\rightarrow$  Difference tests: r, %, means (Other tests).

2. Substitute the appropriate parameters for each of the two groups into the dialog box and perform the calculation to obtain the "p" value (Figure 1).

ОЙТÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ	SOUTH KAZAKHSTAN MEDICAL MA ACADEMY AO «Южно-Казахстанская медицинская академия»
Departments: "Medical Biophysics and Inf	ormation Technologies", № 35-11(Б)-2024
"Social health insurance and pu	ublic health" № 582024
Guidelines for practical lessons of the course "Intr	oduction to Scientific Research" p.30 out of 68

🖣 Dif	ference to	ests:	r, %, means: :	Spread	Isheet7	AL SPI	? X
Diffo	Send/print	resu	ults for each Co	mpute t	o Report w	indow P g	Cancel
r1:	0,00	een T	N1: 10		n: 1,0000	One-sided	Compute
r2:	0,00	8	N2: 10		p. 1,0000	Two-sided	
Diffe	rence betv	veen	two means (no	rmal di	stribution)	$\overline{U}_{\alpha}$ $\overline{C}_{\alpha}$ $\overline{V}_{\alpha}$	
M 1:	40	Ð	StD∨1: 6,4		N1: 29	p: 0000	Compute
M 2:	50	Ì	StD∨ 2: 5,8		N2: 29	One-sided	
S	ingle mea	n 1 v	s .population m	iean 2		• I wo-sided	
Diffe	rence betv	veen	two proportion	s	. I.F.	1 KI	<u>.</u>
Pr.1:	,500000		N1: 10		p: 1.0000	One-sided	Compute
Pr.2:	,500000		N2: 10		p. 1,0000	Two-sided	

Figure 1. Dialog box of the procedure "Difference tests: r, %, means"

The results obtained can be interpreted as follows:

- if p>0.05, then the null hypothesis of no differences in means is accepted;
- if p<0.05, then the null hypothesis is rejected.

For this example, p < 0.05, which means the null hypothesis of equality of means is rejected.

#### Tasks

1. Two batches of tablets were produced at different compression pressures (80 and 100 MPa). The fracture strength of tablets of the first group turned out to be 50.4; 53.6; 54.4; 46.4; 44.0; 48.2; 49.4 (in Newtons), in the second group -47.2; 62.4; 64.8; 62.4; 58.9; 55.4; 66.2; 49.5; 67.8; 68.9 (in Newtons). Using the Student's test at p=0.05, check the hypothesis about the equality of means (the alternative hypothesis is about their inequality).

2. The psychological test measured choice reaction time in two groups. The first group included athletes, the second group included people not involved in sports. In the first group the following results were obtained: 0.42, 0.52, 0.48, 0.46, 0.55, 0.62, 0.58, 0.64, 0.56 (seconds). In the second: 0.51, 0.67, 0.54, 0.52, 0.56, 0.66, 0.68 (seconds). Using the Student's test at p=0.05, check the hypothesis about the equality of means (the alternative hypothesis is about their inequality).

3. Using paired Student's t-test, compare the results of performing logical tasks before and after the training course. The initial data is presented in the table.

No.	Results of performing logical tasks before the course (sec.)	Results of performing logical tasks after the course (sec.)					
42	25	22					
2	23 23 23	25 5					
3	28 4 5 6	23 5 10 10					
4	29	22					
5	35 20 1	30 1 2					
6	VI 5 (131 0 1) 1	27					
7	24 0 10	20					
8	x 24 0 0 x 1	219 00 14					
9	38	32					
10	26 20 20 20	25					
11	20 20 20	20					

МЕDISINA АКАDEMIASY Оңтүстік Қазақстан медицина академиясы» АҚ	MEDICAL ACADEMY АО «Южно-Казахстанская медицинская академия»
Departments: "Medical Biophysics and Informa	ation Technologies", № 35-11(Б)-2024
"Social health insurance and public	health" № 582024
Guidelines for practical lessons of the course "Introduc	tion to Scientific Research" p.31 out of 68

SOUTH KAZAKHSTAN

4. Using a paired Student's t test, determine whether smoking affects platelet function. The initial data are shown in the table.

OŃTÚSTIK-QAZAQSTAN

No.	Platelet aggregation									
. Mur	Before smoking	After smoking								
21	25 5									
25	25 25	29 29								
3	5 27 1 2	37								
4	5 44	56								
5	30 00 1	46								
6	1 2 670 0° 1 K	82								
7	53	57 80 4								
8	53	80 80								
9	52	No ver St 61 Co white								
10	60	59 60 00								
2 11	28	43								

5. Using a paired Student's t test, check whether the drug under study functions as a diuretic. The initial data are shown in the table.

No.	Daily diuresis, ml									
3. 1	Before taking the drug	After taking the drug								
1.	1490	1600								
02	1300	1850								
3	1400	1300								
24 0	1410	1500								
5	1350	× 51400 × 1								
6	1000	1010								

6. Using paired Student's t-test, check the effectiveness of a special diet to get rid of excess weight. The initial data is presented in the table.

No.	Weight (kg) before experiment	Weight (kg) after experiment
1	93.2	88.9
2	98.2	94.5
3	105.6	106.1
4	86.8	84.3
5	95.5	92.5
2		i vil at a construction

6. Assessment methods / technologies: Oral questioning. Practical work. (assessment using a checklist).

#### 7. Literature:

Main:

1. Biostatistics: electronic textbook / K. Zh. Kudabaev [et al.]. - Shymkent: SKFFA, 2015.

2. Rosner Bernard Fundamentals of Biostatistics: Texbook/ B.Rosner - 8th ed.: GENGAGE learning, 2016.

3. Armitage P. Encyclopedia of Biostatistics. - Wiley, 2016. - 6100 p

4. Le C.T. Introductory biostatistics. - Wiley, 2013. - 536 p

- 5. Newman S.C. Biostatistical Methods in Epidemiology. Wiley, 2017. 388p
- 6. <u>http://www.biometrica.tomsk.ru</u>

#### 8. Control:

1. Why is the Student's t-test very popular in statistical analysis of biomedical data?

OŃTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY



«Оңтүстік Қазақстан медицина академиясы» АҚ

АО «Южно-Казахстанская медицинская академия» Departments: "Medical Biophysics and Information Technologies", № 35-11(Б)-2024

"Social health insurance and public health' № 58--2024Guidelines for practical lessons of the course "Introduction to Scientific Research" p.32 out of 68

- 2. What conditions must be met when using the Student's t-test?
- 3. How are the null and alternative hypotheses formulated for the Student's t-test?
- 4. In what ways can it be implemented t-Student's test in the program "STATISTICA"?
- 5. How to check the conditions for applying the Student's test in the STATISTICA program?
- 6. How is the information contained in the summary table interpreted?

#### Lesson №5

1. The theme: Nonparametric methods of comparative statistics.

2. The aim: studying the methodological foundations and conditions for using the Mann-Whitney Utest and the Wilcoxon W(T)-test, becoming familiar with some of the capabilities of the "STATISTICA" program

#### 3. The learning objectives:

- understand in what cases it is used UMann-Whitney test and Wilcoxon W(T) test;
- learn to formulate null and alternative hypotheses;
- master the algorithm of the criterion;
- learn to interpret the result.

#### . Key questions of the theme:

1. Mann-Whitney U test is a nonparametric analogue of the two-sample Student t test.

- 2. Wilcoxon W (T) test is a nonparametric analogue of the paired Student t test.
- 3. Conditions for using the Mann-Whitney U test and the Wilcoxon W (T) test.
- 4. Scheme of application of the Mann-Whitney U-test and Wilcoxon W (T)-test.
- 5. The procedure for implementing the Mann-Whitney criterion in the "STATISTICA" program.

6. The procedure for implementing the Wilcoxon test in the "STATISTICA" program.

7. Interpretation of results.

5. Methods / technologies of teaching and learning: Computer-based work / Solving situational tasks.

#### * Tasks:

Example 1. Researchers decided to find out the effectiveness of the drug, which allows obese patients to lose excess weight. In this case, a group of patients was prescribed a certain diet.

A month later, the amount of weight lost was recorded. 8 people were selected to conduct the experiment. 3 of them received the study drug (experimental group), and 5 received placebo (control group). The selection of 3 out of 8 subjects into the experimental group was carried out randomly. All participants in the experiment believed that they were taking the drug.

all the start	10° 0	Lost	t weight,	kg	<u>y.</u> 60	
Experimental group	6,2	3,0	3,9	241	<i>9</i> .	
Control group	4,0	-0,5	3,3	1,5	3,0	· di
and see going the						-0 ² . 6
The second				11. 1.1		· · 67.
Structure du				W JU.F		
cance level.	10. 1				1.1.1	
igle row.	~ ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~			9. 60.	111.	
). I st do	<u> </u>	1.1	24	2.	<u>,0~ (</u> )	H A

#### Solution.

1) H₀:  $\bar{x}_1 = \bar{x}_2$ .

 $H_1: \overline{x}_1 \neq \overline{x}_2$ .

2)  $\alpha$ =0.05 - significance level.

3) Let's make a single row.

Solution. $\overline{x}_1 = \overline{x}_2$ . $\overline{x}_1 \neq \overline{x}_2$ . 0.05 - significance leases of the single row	evel.	3.edu.K	1. SKI	Maria e Skina e KL Skin	a.edu.k	J.K2 K2 edu.k2	SKIN RL SKIN NL SKIN	a.e. edi	Ng. Sqn
Lost weight, kg	6.2	3.0	3.9	4.0	-0.5	3.3	1.5	3.0	7
Rank	8	3.5	06	7	17	5	2	3.5	5

о́мти́sтік-QazaQstan MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ О́мти́sтік-QazaQstan MEDICAL АСАDEMY АО «Южно-Казахстанская меди	цинская академия»
Departments: "Medical Biophysics and Information Technologies",	№ 35-11(Б)-2024
"Social health insurance and public health"	№ 582024
Guidelines for practical lessons of the course "Introduction to Scientific Research"	p.33 out of 68

Let us divide the single ranked series into two, consisting of units of the first and second samples.

Experimental	l group	Control group				
Lost weight, kg	Rank	Lost weight, kg	Rank			
6.2	8 6	4.0	1. 7.0			
3.0	3.5	-0.5	1			
3.9	6	3.3	5			
1 ct 20.	80. "I'F.	1.5	2			
12 24 00	- 00 K	3.0	3.5			
U.M. J. Hur	<i>T1</i> =17.5	Kr 2. Min	<i>T</i> 2=18.5			

T1 and T2 – sums of ranks; T1<T2, which means T2=Tx, nx=n2=5.

$$U_{pacy} = 3 \cdot 5 + \frac{5 \cdot (5+1)}{2} - 18,5 = 11,5$$

4) Utable (0,05;3;5)=1 (see Appendix 5).

l v	$\mathcal{L} \xrightarrow{\mathcal{L}} \mathcal{L} \xrightarrow{\mathcal{L}} \xrightarrow{\mathcal{L}} \mathcal{L} \xrightarrow{\mathcal{L}} \xrightarrow{\mathcal{L}} \mathcal{L} \xrightarrow{\mathcal{L}} \mathcal$																		
N1	2	3	D4	5.	6	78	8	9	10	11	12	13	14	15	16	17	18	19	20
N2																			
3	·	0	-	2	10	<u>`</u>			$\mathcal{O}$			6	1			S.	11		11
A		- <del>-0</del>	1/	(	2		0	$\odot$	ľ.,	$\langle \cdot \rangle$	. /					•	<u>60°</u>		Y
5	<b>)</b> (	1	2	.4		$\Delta$		2.		<u> </u>	. /-		2	<u> </u>	$\mathcal{O}$	0		$\Delta O$	<u>```\</u>
<u> </u>	0	2-	<b>3</b>	5	70		$\left  A \right\rangle$	<u> </u>	0		7.	XÉ	/	6		0	0	<u> </u>	
$\sim T$	0	<u>⁄2</u>	4	6	8	11	6	L d	<u>) (</u>	l ov				- ×			2.		0×.
8	1	3	5	8	-10	13	15	$\langle \cdot \rangle$		<u> </u>	$\Box \Delta$	D-		/	5		10		<u> </u>
9.(	1	4	6	9	12	15	18	21	2		<u> 27</u>			. 4		26		0.	
10	1	4	7	211	14	17	20	24	27	2	6	<u>_0'</u>	Ľ			2.	0		
11	1	5	8	12	16	19	23	27)	31	34		K) _	121		11		1		01
12	2	5	9	13	17	21	26	30	34	38	42		<u>50</u>		N.	4	× 1	$\langle \rangle$	- 0
13	2	6	10	15	19	24	28	33	37	42	47	51	Ľ	$\Delta V$			6		
14	3	D7.	11	16	21	26	31	36	41	46	51	56	61	27	1.1	<u> </u>	И		1
15	3	70	12	18	23	28	33	39	44	50	55	61	66	72	07.	. N	JV .	2	
16	3	8	14	19	25	30	36	42	48	54	60	65	71	77	83	$\nabla$	N		
17	3	9	15	20	26/	33	39	45	51	57	64	70	77	83	89	96	1.	·	~
18	4	9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109		/
19	4	10	/17	23	30	37	44	/51	58	65	72	80	87	94	101	109	116	123	1
20	4	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138

5)  $U_{pacy} > U_{ma\delta \pi}$ , then "H0" is accepted, i.e. the drug is ineffective.

**Example 2.**To check whether there is a difference in the blood sugar level on an empty stomach before work and three hours after work in 12 workers working on ultrasonic units. The initial data are shown in the table.

No.	Sugar content before work	Sugar content after work
1	112	54 54
2	82	67
3	101	96
4	72 20 1	59
5	79	79
6	82 82	76 0
07 X	64	66
8	70	66

TÚSTIK-QAZAQSTAN	
MEDISINA	
AKADEMIASY	
A1/	

2960 SOUTH KAZAKHSTAN SKMA MEDICAL ACADEMY

«Он

ΟŃ

Departments: "Medical Biophysics and Information Technologies",	№ 35-11(Б)-2024
"Social health insurance and public health"	№ 582024
Guidelines for practical lessons of the course "Introduction to Scientific Research"	p.34 out of 68

è			
2	9	88	48
	10	81	50
	110	66	61 6
	12	88	61 61

лh,

Solution.

1) H₀:  $\bar{x}_1 = \bar{x}$ 

$$H_1: x_1 \neq x_2$$

2)  $\alpha \approx 0.05$  - significance level.

No.	Y.	5	Ma.	<u>, 67</u>	٢٧, ١	Blood	sugar le	evel	60	111.	1	24
Sugar content BEFORE work	112	82	101	72	79	82	64	70	88	81	66	88
Sugar content AFTER work	54	67	96	59	79	76	66	66	48	50	61	61
Sugar difference	58 0	15	5	13	0	6	-2	4	40	31	5	27
Ranking. row	0	-2	4	5	5	6	13	15	27	31	40	58
Ranks	161 0	0.1	2	3.5	3.5	5	6	7	8	9	10	01
T+	65	2.		Y		3 1	10		Y. K	1 5		0.6
T-3. 0	1	10	2.	Yr.	Y	55	. do	0	10.	VL.	St.	<u>~</u> 0.

Trasch=1 3)

4) *Ttable* (0,052; 12)=50 (see Appendix 6).

3.00	T S	a	V. KIN SK	1 00 TO	α
5 0.	15	0,062	du 13	65 .	0,022
6 3	21	0,032	a gr. Kr	5 57	0,048
1 3. 191	19	0,062	~ ° 14	573	0,020
7	28	0,016	15 10	63	0,050
KL ST	24	0,046		80	0,022
8	32 0	0,024	16	70	0,048
11.1.1	28	0,054	der de	88	0,022
9.	39	0,020	17 17	76	0,050
· e0.	33	0,054		97	0,020
010 00	45	0,020	185	83 20	0,050
	39 5	0,048	J. VI S	1050	0,020
11 2.	52	0,018	XV 19	91 0	0,048
3 ville	44 1	0,054	ev 10. 1	114 0'	0,020
(12)	58	0,020	20	98	0,050
A St	(10° (50) XV)	0,052	2° 60°	124	0,020
11 6	1. <u>2. 60</u>		N. 2. 00	106	0,048

Trasch< Ttable, then "H0" is rejected, which means there is a difference in the blood 5) sugar levels of workers before and after work.

Example 3. The effectiveness of a drug that allows obese patients to lose excess weight is being studied. In this case, a group of volunteers is prescribed a certain diet.

After a month, in order to check compliance with the diet and regular use of the drug, the

Ś	омти́зтік-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ АОңтүстік Қазақстан медицина академиясы АҚ	ицинская академия»
(Ū	Departments: "Medical Biophysics and Information Technologies",	№ 35-11(Б)-2024
	"Social health insurance and public health"	№ 582024
	Guidelines for practical lessons of the course "Introduction to Scientific Research"	p.35 out of 68

amount of weight lost (kg) is recorded. A group of 8 people was selected to conduct the experiment. 3 of them received the study drug (experimental group), and 5 received placebo (control group). The selection of 3 out of 8 subjects into the experimental group was carried out randomly. All participants in the experiment believed that they were taking the drug.

Mar Co XUITI	SK	Lost	t weight,	kg	, Ku,
Experimental group	6,2	3,0	3,9		3.1
Control group	4,0	-0,5	3,3	1,5	3,0

*1.* Create a data table "Drug Efficacy" in the "Statistica" program with a size of 2*8 and enter the initial data.

2. Select Statistics $\rightarrow$ Nonparametrics (Figure 1) $\rightarrow$ Comparing two independent samples (groups) (Figure 2), click the "OK" button.



#### Figure 1. Selecting the "Nonparametrics" module

							STATIS	TICA - [[	Data: Sp	preadsh	eet1*
<u>F</u> ile <u>E</u> c	lit <u>V</u> iew <u>I</u> nsert F <u>o</u> rn	nat <u>S</u> tatistics Da	ta <u>M</u> ini	ing <u>G</u> rap	hs <u>T</u> ools	Data	Window	Scorecard	d <u>H</u> elp	<u> </u>	LV
ו 🛋 🗖	13 A A X 🗈		84	Add to V	/orkbook -	Add to	Report -	Add to N	AS Word	- 22	≥ ▶?
				2			le luu	[ +.0 .m [			A 1
Ariai		у в у п		e = X	V 🕁 . 📈			100 +10	1¥ 🖽	<b>•••</b>	Z + *=
	Эффективность пре	парата	70		r		1	N.	$-\Delta$	,	<u>}</u>
<i>U</i> g.	1 группа	2 потерянная мас кг	ca,	3.	edu	Y		5	<u>Ur</u>	3.	~
1	контрольная		74	M No	paramet	ric Sta	tistics: S	spreadsh	leet1		
2	контрольная		-0,5		6	11	7. 1	11	C E		
3	контрольная		3,3	Quick	2			Hr.		OK	
2 4	контрольная		1,6	I mm		6				C	
5	контрольная			2x:	2 Tables (X?/	W?/Phi?	, McNema	ir, Fisher exa	ac l	Lancei	
6	экспериментальная		6,2	X Obs	erved versus	expecte	dX?	( <b>( )</b> • * * *		Ontione	÷
7	экспериментальная		1	Corr	elations (Spe	arman, K	endall tau	, gamma)		optiona	
8	экспериментальная		3,9	Con	paring two in	depende	ent sample	s (aroups)			
				Con	paring multin	le inden	samples (	ากมารไ	- I V		51
	2				paring two d	opondor	t espondes	(upriphles)	$\langle   \rangle \langle \rangle$		
	1. 1				ipaning two u	ependen	t sampies	(vanabies)	MY.		
	Fr 2.			Lon	.paring multip	ile dep. s	ampies (va	snables]	M		
			DU	iiii Coc	hran Q test	2	VV.		~	Open Da	ata
	Lr S		N I	🔐 Ordi	nal descriptiv	e statisti	cs (mediar	, mode,)			
			A				1	<u> </u>	SELECT	s 6	w

Figure 2. Selecting the "Comparing two independent samples (groups)" procedure

ОЙТÚSTIК-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ SOUTH KAZAKHSTAN MEDICAL ACADEMY АО «Южно-Казахстанская медиц	цинская академия»
Departments: "Medical Biophysics and Information Technologies",	№ 35-11(Б)-2024
"Social health insurance and public health"	Nº 582024
Guidelines for practical lessons of the course "Introduction to Scientific Research"	p.36 out of 68

*3*. In the dialog box, click the "Variables" button, indicate the grouping characteristic (column containing group codes) on the right side of the window, and the column containing the analyzed characteristic on the left side of the window (Figure 3) and click the "OK" button.



Figure 3. Setting variables

4. In the dialog box, click the "Mann-Whitney U test" button. A summary table will appear on the screen (Figure 4).

eardu.k.k	Mann-Whitney U By variable rpyr Marked tests ar	J Test (Spreadsh ina e significant at p	eet1) <,05000	Y VI	Sid	Kno	0.00	201 K	SK.	4man
ariable	Rank Sum контрольная	Rank Sum экспериментал ьная	0.0	Z	p-value	Z adjusted	p-value	Valid N контрольная	Valid N экспериментал ьная	2*1sided exact p
отерянная масса, кг	18,50000	17,50000	3,500000	-1,04350	0,296718	-1,04977	0,293827	5	3	0,250000

Figure 4. Final table with the results of comparing two independent samples using the Mann-Whitney test

The first and second columns indicate the Rank Sum.

The results obtained can be interpreted as follows:

- if p<0.05, then the null hypothesis is rejected (such results are highlighted in red font);
- if p>0.05, then the null hypothesis is accepted.

In this example, p = 0.296718 and p = 0.293827, which means the hypothesis of equality of means is accepted, i.e. the drug is ineffective.

5.For clarity, you can build a "box with whiskers" plot by clicking the "Box & whisker plots for all variables" button in the analysis window (Figure 5).



Figure 5. Box-and-whisker plot

	ОŃTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY Оңтүстік Қазақстан медицина академиясы» АҚ	-1979-	SOUTH KAZAKHSTAN MEDICAL ACADEMY AO «Южно-Казахстанская ме;	дицинская академия»
0	Departments: "Medical Biophysics and In	format	ion Technologies",	№ 35-11(Б)-2024
2	"Social health insurance and p	bublic l	nealth"	Nº 582024
0	Guidelines for practical lessons of the course "Intr	roduct	on to Scientific Research"	p.37 out of 68

**Example 4.** To check whether there is a difference in the blood sugar level on an empty stomach before work and three hours after work in 12 people working on ultrasound machines.

No.	Sugar content before work	Sugar content after work
1	112 5	54
2	82 82	67 1 20
53	101	96
4	72	59
5	79	79
6	82	76
7	64	66
8	70	66 , 66
9	88	548
10	81	XX 50 XO
- 11	66	61 200
12	88 5	61 61

*I*. Create a data table "Blood sugar content" in the Statistica program, size 2*12, and enter the initial data.

2. Select Statistics  $\rightarrow$  Nonparametrics  $\rightarrow$  Comparing two dependent samples (variables) (Figure 1) and click the "OK" button.

נ 😂 🖪	🔁 🖨 🖪 🐰	s 🗈 🛍 ダ 🗠 S	Add to Workbook - Add to Report - Add to MS	Word - 🛟 🛛 🤣 🎀
Arial S		10 🗸 <b>В</b> І <u>Ц</u>	│ <b>三</b> ☰ ☰ ☎ <mark>▲ · △ · ⊡ · 🃎 Ⅲ</mark> अ 🕫 ≠	≠ 📾 🚰 🔂 👌 ∗=
	содержание сах	кара в крови	K. S. O. K. S.M.	
1.1	1 до работы	2 после работы	1 3 char 3. South the st	Luca en d
1	112	54	Nonparametric Statistics: Spreadshoe	4c ? ×
2	5 82	67		
3	101		We with the second seco	
04	72	59		
5	79	<u> </u>	IIII 2.4.2 Tables (V2A/2/PH2 MeMorrer Eisber eus)	Cancel
<u> </u>	82	76	2 X 2 Tables (A //Y //Philit, Micheliai, Pisher exat	
	64	66	X Ubserved versus expected X?	🔊 Options 🔻
8	/0	66	Correlations (Spearman, Kendall tau, gamma)	
9		48	Comparing two independent samples (groups)	N. VI
		50	Comparing multiple indep. samples (groups)	$e^{\gamma}$
- <u>S 11</u>	00	01	Comparing two dependent samples (variables)	
12			Comparing multiple dep_samples (variables)	
			11111 Coobran D test	
	Ch' a'a'			🗁 Open <u>D</u> ata
	Stype		Crdinal descriptive statistics (median, mode,)	

Figure 1. Selecting the procedure Comparing two dependent samples (variables)

3. In the dialog box, click the "Variables" button, indicate the first variable on the left side of the window, and the second variable on the right side of the window (Figure 2) and click the "OK" button.





Figure 2. Setting variables

4. In the dialog box, click the "Wilcoxon matched pairs test" button. A summary table will appear on the screen (Figure 3).

			( )		Kr.
K. S. M. S. S	Wilcox	on Matche	d Pairs Te	st (Spreadsh	ieet6)
	Market		Significant	at p <,0500	
Pair of Variables	N	V.	52	p-value	20.
до работы & после работы	V 11	1.000000	2.845147	0.004439	

Figure 3. Final table with the results of comparison of two dependent samples using the Wilcoxon test

The results obtained can be interpreted as follows:

• if p<0.05, then the null hypothesis is rejected (such results are highlighted in red font);

• if p>0.05, then the null hypothesis is accepted.

In this example, p = 0.004439, which means the null hypothesis of equality of means is rejected, i.e. There is a difference in the blood sugar levels of workers before and after work.

5.For clarity, you can build a "box with whiskers" plot by clicking the "Box & whisker plots for all variables" button in the analysis window (Figure 4).



Figure 4. Box-and-whisker plot

мерізіна АКАДЕМІАЅҮ Оңтүстік Қазақстан медицина академиясы» АҚ АО «Южно-Ка	ззахстанская медицинская академия»
Departments: "Medical Biophysics and Information Technologie	ès", № 35-11(Б)-2024
"Social health insurance and public health"	Nº 582024
Guidelines for practical lessons of the course "Introduction to Scientific	Research" p.39 out of 68

SOUTH KAZAKHSTAN

OŃTÚSTIK-QAZAQSTAN

Tasks:

1. Using the Mann-Whitney U test, compare the level of intelligence of students in two groups. The initial data are shown in the table.

Ar o		2. 2	N. K.	St	IQ s	cores		St. S	9. 00.	
1 group	112	105	109	<u>90</u>	130	117	117	125	134	109
2nd group	121	120	134	119	115	106	107	101	97	117

2. Using the Mann-Whitney U test, compare the duration of wakefulness in the first hour of life of infants born using the conventional and Leboyer methods. The initial data are shown in the table.

Childbirth according to the usual method	Childbirth using the Leboyer method
5.0 5.0	2.0
10.1	19.0
17.7	29.7
20.3	32.1
22.0	35.4
24.9	36.7
26.5	38.5
30.8	40.2
34.2	42.1
35.0	43.0
36.6	44.4
37.9	45.6
40.4	46.7
45.5	47.1
49.3	48.0
510 W (510 W)	49.0
53.1	50.9
55.0	51.2
56.7	52.5
58.0	53.3

3. Using the Mann-Whitney U test, compare systolic blood pressure (SBP) (in mm hg) in two homogeneous groups of healthy men:

• group 1 - persons with many years of work experience in conditions of disturbed sleep and wakefulness;

• group 2 - persons without disturbances in the circadian rhythm of sleep and wakefulness. The initial data are shown in the table.

· 60	F 1 9	KIN	2.0	90°. K	5	SBP	2.00	W. 11	, st		0
1 group	90	95	100	105	120	135	135	135	140	140	145
2nd group	110	115	115	122	122	125	130	150		12	, <u>v</u>

4. Using the Wilcoxon W test, check the effectiveness of a special diet to get rid of excess weight. The initial data are shown in the table.

OŃTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY



«Оңтүстік Қазақстан медицина академиясы» АҚ

Departments: "Medical Biophysics and Information Technologies",<br/>"Social health insurance and public health"№ 35-11(Б)-2024Guidelines for practical lessons of the course "Introduction to Scientific Research"№ 58-<br/>p.40 out of 68

No.	Weight (kg) before experiment	Weight (kg) after experiment
10	93.2	88.9
2	98.2	94.5
3	105.6	106.1
4	86.8	84.3
5	95.5	92.5

5. Using the Wilcoxon W test, check whether smoking affects platelet function. The initial data are shown in the table.

No.	Platele	t aggregation
3. 0	Before smoking	After smoking
1	25	27 6 27
2	25	29
3	27 5	37
4	A 44 1 5 . A	56 24 20 20
55	30. 1 6	46
6	67	82 82
5.7	53	14 3. 057 H S. KI
8	53	80
09	52	61
10	60	59. 20. 2
	28	43 8 1

6. Using the Wilcoxon W test, check whether a certain drug under study is a diuretic. The initial data are shown in the table.

No.	Daily	diuresis, ml
	Before taking the drug	After taking the drug
1	1490	1600
2	1300	1850
3	1400	1300
4	1410	1500
5	1350	1400
6	1000	1010

6. Assessment methods / technologies: Oral questioning. Practical work. (assessment using a checklist).

#### 7. Literature:

Main:

- 1. Biostatistics: electronic textbook / K. Zh. Kudabaev [et al.]. Shymkent: SKFFA, 2015.
- 2. Rosner Bernard Fundamentals of Biostatistics: Texbook/ B.Rosner 8th ed.: GENGAGE learning, 2016.
- 3. Armitage P. Encyclopedia of Biostatistics. Wiley, 2016. 6100 p
- 4. Le C.T. Introductory biostatistics. Wiley, 2013. 536 p
- 5. Newman S.C. Biostatistical Methods in Epidemiology. Wiley, 2017. 388p
- 6. http://www.biometrica.tomsk.ru

OŇTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY

«Оңтүстік Қазақстан медицина академиясы» АҚ

Departments: "Medical Biophysics and Information Technologies",<br/>"Social health insurance and public health"№ 35-11(Б)-2024<br/>№ 58- -2024<br/>p.41 out of 68Guidelines for practical lessons of the course "Introduction to Scientific Research"№ 10-2024<br/>№ 10-2024

#### 8. Control:

- 1. What is the difference between parametric and nonparametric statistical tests?
- 2. Why is the Mann-Whitney test called an analogue of the two-sample Student t-test?
- 3. What conditions must be met when using the Mann-Whitney test?
- 4. Why is the Wilcoxon test called an analogue of the paired Student t-test?
- 5. What conditions must be met when using the Wilcoxon test?

#### Lesson №6

1. The theme: Aanalysis of quality features.

**2. The aim:** formation of an idea of qualitative characteristics, contingency tables, familiarization with some of the capabilities of the STATISTICA program

#### 2. The learning objectives:

- understand the difference between qualitative and quantitative characteristics;
- learn to present qualitative data using contingency tables;
- understand in what cases Pearson's  $\chi 2$  test is used;
- learn to formulate null and alternative hypotheses;
- understand in what cases Pearson's  $\chi 2$  test is used for tables of size 2x2;
- find out why the Yates amendment is needed;
- understand in what cases McNemar's  $\chi 2$  test is used;
- master algorithms of criteria;
- learn to interpret the results;

#### 4. Key questions of the theme:

- 1. Quantitative and qualitative characteristics.
- 2. Contingency tables.
- 3. Conditions for applying the Pearson  $\chi^2$  test.
- 4. Scheme for applying the Pearson  $\chi^2$  test.
- 5. Contingency tables of size 2x2.
- 6. Scheme for applying the Pearson  $\chi^2$  test for tables of size 2x2.
- 7. Yates Amendment.
- 8. Scheme for applying the McNemar  $\chi^2$  test.

**5. Methods / technologies of teaching and learning:** Computer-based work / Solving situational tasks.

#### Tasks:

There are data on the number of observations and cases of mortality for four forms of acute purulent destruction of the lungs. Using Pearson's  $\chi^2$  test, it is necessary to assess the significance of the difference between groups in the number of deaths.

No onour	Earm diagona	Number o	of cases	Number of notionts
Nº group	Form diseases	Lethal outcomes	recovery	Number of patients
1.1	Purulent abscess	5	136	141
2	Gangrenous abscess	117	37	48
3	Gangrene of the lobes	CV VT S	.8	515
4	Total gangrene	6 9	5	14 10

#### Solution.

- 1) H₀: there is no connection between the signs
- $H_1$ : there is a connection between the signs
- 2) p=0.05 significance level

MEDISINA AKADEMIASY ңтүстік Қазақстан медицина академиясы» АҚ АО «Южно-Казахстанская мед	ицинская академия»
Departments: "Medical Biophysics and Information Technologies", "Social health insurance and public health"	№ 35-11(Б)-2024 № 582024
Guidelines for practical lessons of the course "Introduction to Scientific Research"	p.42 out of 68
$v_{ij}^* = v_{,i} \cdot \frac{v_{,j}}{v_{,i}}$ $\underbrace{\left(v_{ij} - v_{ij}^*\right)^2}_{v^*}$	edu. K. Sking
3.1) Calculate theoretical frequencies $V_{ij}$	Marg. er dui 1K2

SOUTH KAZAKHSTAN

160

SKMA

3) 
$$v_{ij}^* = v_{.i} \cdot \frac{v_{.j}}{v}$$

«C

$$\frac{\left(\nu_{ij}-\nu_{ij}^{*}\right)^{2}}{*}$$

OŃTÚSTIK-QAZAQSTAN

00	<b>B</b> 1	<b>B</b> 2	Total
A ₁	5	136	141
A ₂	011	37	48
A3	7	8	5 15
A4	6	5	11
Total	29	186	215

3.	800 JU. Kr 1 5.	Kuras. edu. K	
A C	e 211 1 1	string of du	Yr 5
- (v	$v_{ij} - v_{ij}^*$	1. 2411. 2. 00	
SF	V	1, 2, 24, 3,	
ies		1. H. 1. 3. Hr. 3	
A	B	B ₂	Total
<b>A</b> 1	29*141/215=19	186*141/215=122	141
A ₂	29*48/215=6,5	186*48/215==41,5	48
A ₃	29*15/215=2	186*15/215=13	15
A ₄	29*11/214=1,5	186*11/214=9,5	11
0	ev. 11. T. 1	14° - 20° - 1	4 3.
	· conduction	2411. 2. cor	Hr 1
	US. 60. "1'r.	1 241 20. 00	
St		Pr 3. (1, 8,	
1	(12( 122)	$\frac{B_2}{2(122-1)}$	
3	(136-122)	2/122=1.6	2.
3.1	(37-41.5)2	2/41.5=0.5	
-	(9 12)2/12	8-1.9	N. 21
1.10	(0-15)2/15		

3.2) Let's calculate the quantities skma.edu.kl

2	B ₁	B2	10, 60 m. 1
<b>A1</b> (5-	19)2/19=10.3	(136-122)2/122=1.6	and contraction
<b>A2</b> (1)	1-6.5)2/6.5=3.1	(37-41.5)2/41.5=0.5	H. US. 602 11
<b>A3</b> (7-	2)2/2=12.5	(8-13)2/13=1.9	24° 23° 200
44 (6-	1.5)2/1.5=13.5	(5-9.5)2/9.5=2.1	1. 241. 2. 00
culate	2 54 40 3. 8. 00 du	KI SKIIIA BOUND	U.K. KI SKAMAN

skma.edu.k2 skma.edu.k2

3.3) Let's calculate  

$$\chi^{2}_{pacy} = \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{\left(v_{ij} - v_{ij}^{*}\right)^{2}}{v_{ij}^{*}} = 10,3 + 3,1 + 12,5 + 13,5 + 1,6 + 0,5 + 2 + 2,1 = 45,5$$

4)  $\chi^2_{maon}(p, f)$ , where f =(r-1)(s-1) =(4-1)(2-1)=3 - number of degrees of freedom

 $\chi^2_{ma\delta,n}(0,05;3) = 7,8$  (see table 2)

# ma.edu.kz skma.ed skma.edu.k table 2

степеней свободы	0,01	0,05	0,1	0,90	0,95	0,99
1	6,6	3,8	2,71	0,02	0,004	0,0002
$\sqrt{2}$	9.2	6.0	4,61	0,21	0,1 0	0,02
3 0	0H,3	7,8	6,25	0,58	0,35	0,12
4	13,3	9,5	7,78	1,06	0,71	0,30
55	15,1	11,1	9,24	0 1,61	1,15	0,55
66	16,8 🔿	12,6	10,6	2,20	1,64	0,87
. 7	18,5	014,1	12,0	2,83	2,17	1,24
8	20,1	15,5	13,4	3,49	2,73	1,65
9	521,7	16,9	14,7	4,17	3,33	2,09
10	23,2	18,30	16,0	4,87	3,94	2,56
	24,7	19,7	17,3	5,58	4,57	3,05
12	26,2 9	21,0	18,5	6,30	5,23	3,57
13	27,7	6 22,4	19,8	7,04	5,89	4,11
14 00	29,1	23,7	21,1	7,79	6,57	4,66
15	30,6	25.0	22.3	8.5	7.26	5.23

	Оңтүстік Қазақс	стан медицин	ŃTÚSTIK-QAZAC MEDI AKADEM на академиясы	STAN SINA IASY NAK	SOUTH KAZAKHS MEDICAL ACADEMY AO «Южно-Ка	STAN азахстанская ме	едицинская акаде	«RNM
)	Depa	artments: "Me	dical Biophysi	cs and Informat	ion Technologie	es",	№ 35-11(Б)	-2024
<u>n</u>	Guidelines f	"Socia" For practical le	al health insura ssons of the co	nce and public hourse "Introducti	nealth" on to Scientific	Research"	Nº 58- p.43 out o	-2024 of 68
	16 16	32.0	26.3	23.5	9.315	7.98	5.81	
	17	33.4	27.6	24,8	10,1	8.67	6.41	
	5 18	34,8	28,9	26,0	10,9	9,39	7,01	
	19	36,2	30,1	27,2	11,7	10,1	7,63	
	20	37,6	31,4	28,4	12,4	10,9	8,26	
	21	38,9	32,7	9 29,6	013,2	11,6	8,90	1.1
	22 5	40,3 (	33,9	30,6	14,0	12,63	9,54	
	23	41,6	35,2	32,0	14,8	13,1	10,2	
	24	43,0	36,4	33,2 9	15,7	13,8	6 10,9	
$\lambda$	25	44,3	37,7	34,4	16,5	14,6	11,5	)· (
	26	45,6	38,9	35,6	17,3	15,4	12,2	2.0
	27	47,0	40,1	36,7	218,1	16,2	12,99	
	28	48,3 9	41,3	37,9	18,9	16.9	13,6	
	29	49,6	6 42,6	39,1	19,8	17,7 0	14,3	
	30 0	.50,9	43,8	40,3	20,6	18,5	15,0	

5) If  $\chi^2_{pacy} > \chi^2_{ra\delta\pi}$  then "H0" is rejected, which means the differences between groups in the number of deaths are statistically significant.

#### An example of applying the Pearson $\chi^2$ criterion (2x2)

The relationship between maternal use of contraceptive pills and jaundice in breastfed children is being investigated. The data for the study are presented in the table.

Mother taking pills	There is jaundice	No jaundice	Total
Took pills	33	24	57
Didn't take pills	14	455	59
Total	47	69	116

#### Solution.

1) H₀: jaundice in children does not depend on the mother taking contraceptive pills

H₁: jaundice in children depends on the mother taking contraceptive pills

2)  $\alpha$ =0.05 – significance level

3) 
$$\chi^2_{pacy} = \frac{(ad-bc)^2 n}{(a+b)(c+d)(a+c)(b+d)} = \frac{(33\cdot45-24\cdot14)^2\cdot116}{57\cdot59\cdot47\cdot69} = 14,04$$

4) 
$$\chi^2_{maon}(0,05;1) = 3,8$$
 (see table 2)

5) Because  $\chi^2_{pacy} > \chi^2_{ma\delta n}$ , then the hypothesis of independence between jaundice and taking contraceptive pills is rejected, i.e. there is a dependency.

#### **Yates Amendment**

The above formula for  $\chi^2$  in the case of a 2x2 table gives overestimated values. In practice, this results in the null hypothesis being rejected too often. To compensate for this effect, the Yates correction is introduced into the formula:

$$\chi^2_{pacu} = \frac{n\left(ad - bc - \frac{n}{2}\right)^2}{(a+b)(c+d)(a+c)(b+d)}$$

For the example considered above, the calculated value of the criterion with the Yates correction

OŃTÚSTIK-QAZAQSTAN SOUTH KAZAKHSTAN MEDISINA SKM/ MEDICAL AKADEMIASY ACADEMY «Оңтүстік Қазақстан медицина академиясы» АҚ АО «Южно-Казахстанская медицинская академия» Departments: "Medical Biophysics and Information Technologies", № 35-11(Б)-2024 "Social health insurance and public health" Nº 58--2024Guidelines for practical lessons of the course "Introduction to Scientific Research' p.44 out of 68

$$\chi^{2}_{pacy} = \frac{n\left(ad - bc - \frac{n}{2}\right)^{2}}{(a+b)(c+d)(a+c)(b+d)} = \frac{116\left(33 \cdot 45 - 24 \cdot 14 - \frac{116}{2}\right)^{2}}{57 \cdot 59 \cdot 47 \cdot 69} = 12,66$$

#### Example of application of McNemar's $\chi 2$ test

**Example 3.** The effectiveness of a metabolite-type probiotic in complex therapy for complicated mixed respiratory viral infection and its effect on intestinal microbiocenosis are being investigated. 32 patients took part in the study. The data for the study are presented in the table.

2. du ki st d	After pro	biotic treatment
Before probiotic treatment	No dysbacteriosis	There is dysbacteriosis
There is dysbacteriosis	9 0 0 0	5.4 3. 00 .4
No dysbacteriosis	18	0 15 3. 30

#### Example of application of McNemar's $\chi 2$ test

#### Solution.

- 1) H₀: the incidence of dysbacteriosis after using the prebiotic did not change.
- H₁: the incidence of dysbacteriosis after using the prebiotic changed.
- 2)  $\alpha$ =0.05 significance level

3) 
$$\chi^2_{pacy} = \frac{(|a-d|-1)^2}{(a+d)} = \frac{(|9-0|-1)^2}{(9+0)} = 7,1$$

- 4)  $\chi^{2\text{keste}}(0.05;1)=3.8$  (see table 2)
- 5) Because  $\chi^2_{pacy} > \chi^2_{Ta6\pi}$ , then the hypothesis that the number of patients did not change after the use of the pribiotic is rejected.

**Example 4.** There are data on the number of observations and cases of mortality for four forms of acute purulent destruction of the lungs. Using Pearson's  $\chi^2$  test, it is necessary to assess the significance of the difference between groups in the number of deaths.

No anoun	Form diagona	Number of	cases	Number of notionts
Jve group	Form diseases	Lethal outcomes	recovery	Number of patients
1	Purulent abscess	5 5	136	141
2	Gangrenous abscess	511	37	48
3	Gangrene of the lobes	V Br do	8	150 0
4	Total gangrene	Nº 16 6 0	0 05	11 31

1. Create a table in the Statistica program with a size of 2*215 and enter the initial data.

Departments: "Medical Biophysics and Information Technologies", "Social health insurance and public health" Guidelines for practical lessons of the course "Introduction to Scientific Research" Nº 58202 p.45 out of 68	Departments: "Medical Biophysics ar "Social health insurance a	d Information		медицинская академия»
"Social health insurance and public health"         № 58- p.45 out of 68           Guidelines for practical lessons of the course "Introduction to Scientific Research"         p.45 out of 68           1         1         1           2         1         1           3         1         1           3         1         1           4         1         1           5         1         1           6         1         1           6         1         1           9         1         1           11         1         1           12         1         1           13         1         1           14         1         1           15         1         1           16         1         1           17         1         1           18         1         1           19         1         1           19         1         1	"Social health insurance a	a mormation	n Technologies",	№ 35-11(Б)-202
Guidelines for practical lessons of the course "Introduction to Scientific Research"         p.45 out of 68           1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1         1	Social field in Surance e	und public hea	ilth"	Nº 58202
1         2           Форма         Исход           1         1           2         1           3         1           4         1           5         1           6         1           8         1           9         1           10         1           11         1           12         1           13         1           14         1           15         1           10         1           11         1           12         1           13         1           14         1           15         1           16         1           17         1           18         1           19         1           19         1           19         1           19         1	Guidelines for practical lessons of the course	"Introduction	to Scientific Research"	p.45 out of 68
1         2           1         1           2         1           3         1           4         1           5         1           6         1           8         1           9         1           11         1           12         1           13         1           14         1           15         1           10         1           11         1           12         1           13         1           13         1           14         1           15         1           16         1           17         1           18         1           19         1           19         1           19         1           19         1		7, 0	1 10 2.	V Kr 5. 1
Форма         Исход           1         1           2         1           3         1           4         1           5         1           6         1           8         1           9         1           10         1           8         1           9         1           8         1           9         1           11         1           12         1           13         1           13         1           14         1           15         1           16         1           17         1           18         1           19         1           18         1           19         1           19         1           19         1	A B. ON LE IS WIT AS			with at
1         1         петальный           2         1         летальный           3         1         летальный           4         1         летальный           5         1         летальный           6         1         выздоровлени           8         1         выздоровлени           9         1         выздоровлени           10         1         выздоровлени           11         1         выздоровлени           12         1         выздоровлени           13         1         выздоровлени           14         1         выздоровлени           15         1         выздоровлени           16         1         выздоровлени           17         1         выздоровлени           18         1         выздоровлени           19         1         выздоровлени           20         1         выздоровлени	WAR AN AN ALL ST AN	Форма	Исход	
2         1 петальный           3         1 летальный           4         1 летальный           5         1 летальный           6         1 выздоровлени           7         1 выздоровлени           9         1 выздоровлени           10         1 выздоровлени           11         1 выздоровлени           12         1 выздоровлени           13         1 выздоровлени           15         1 выздоровлени           16         1 выздоровлени           17         1 выздоровлени           18         1 выздоровлени           19         1 выздоровлени           20         1 выздоровлени		0 0 1	летальный	N. VI G
3         1         летальный           4         1         летальный           5         1         летальный           6         1         выздоровлени           7         1         выздоровлени           9         1         выздоровлени           10         1         выздоровлени           11         1         выздоровлени           12         1         выздоровлени           13         1         выздоровлени           14         1         выздоровлени           15         1         выздоровлени           16         1         выздоровлени           18         1         выздоровлени           19         1         выздоровлени           20         1         выздоровлени		1	летальный 🕤 📣	
4         1 летальный           5         1 летальный           6         1 выздоровлени           7         1 выздоровлени           9         1 выздоровлени           10         1 выздоровлени           11         1 выздоровлени           12         1 выздоровлени           13         1 выздоровлени           14         1 выздоровлени           15         1 выздоровлени           16         1 выздоровлени           18         1 выздоровлени           19         1 выздоровлени           20         1 выздоровлени	State Carlo		летальный	3. 00 Kr
5         1 летальный           6         1 выздоровлени           7         1 выздоровлени           8         1 выздоровлени           9         1 выздоровлени           10         1 выздоровлени           11         1 выздоровлени           12         1 выздоровлени           13         1 выздоровлени           14         1 выздоровлени           15         1 выздоровлени           16         1 выздоровлени           18         1 выздоровлени           19         1 выздоровлени           20         1 выздоровлени	4	$\Lambda$	летальный	
6         1         выздоровлени           7         1         выздоровлени           8         1         выздоровлени           9         1         выздоровлени           10         1         выздоровлени           11         1         выздоровлени           12         1         выздоровлени           13         1         выздоровлени           14         1         выздоровлени           15         1         выздоровлени           16         1         выздоровлени           17         1         выздоровлени           18         1         выздоровлени           19         1         выздоровлени           20         1         выздоровлени	5		летальный	
7         1         выздоровлени           8         1         выздоровлени           9         1         выздоровлени           10         1         выздоровлени           11         1         выздоровлени           12         1         выздоровлени           13         1         выздоровлени           14         1         выздоровлени           15         1         выздоровлени           16         1         выздоровлени           17         1         выздоровлени           18         1         выздоровлени           19         1         выздоровлени           20         1         выздоровлени	C C C C C		выздоровлени	
8         1 выздоровлени           9         1 выздоровлени           10         1 выздоровлени           11         1 выздоровлени           12         1 выздоровлени           13         1 выздоровлени           14         1 выздоровлени           15         1 выздоровлени           16         1 выздоровлени           17         1 выздоровлени           18         1 выздоровлени           19         1 выздоровлени           20         1 выздоровлени		-2, ( <b>1</b>	выздоровлени	
3         Павіздоровлени           10         1         выздоровлени           11         1         выздоровлени           12         1         выздоровлени           13         1         выздоровлени           14         1         выздоровлени           15         1         выздоровлени           16         1         выздоровлени           17         1         выздоровлени           18         1         выздоровлени           19         1         выздоровлени           20         1         выздоровлени			выздоровлени	·
10         1 выздоровлени           11         1         выздоровлени           12         1         выздоровлени           13         1         выздоровлени           14         1         выздоровлени           15         1         выздоровлени           16         1         выздоровлени           17         1         выздоровлени           18         1         выздоровлени           19         1         выздоровлени           20         1         выздоровлени			выздоровлени	
12         1         выздоровлени           13         1         выздоровлени           14         1         выздоровлени           15         1         выздоровлени           16         1         выздоровлени           17         1         выздоровлени           18         1         выздоровлени           19         1         выздоровлени           20         1         выздоровлени	Or the state of the	NU G	BLISTODOBREHL	
13 1 выздоровлени 14 1 выздоровлени 15 1 выздоровлени 16 1 выздоровлени 17 1 выздоровлени 18 1 выздоровлени 19 1 выздоровлени 20 1 выздоровлени	N' VI GN AN AN ON 12		выздоровлени	
14         1         выздоровлени           15         1         выздоровлени           16         1         выздоровлени           17         1         выздоровлени           18         1         выздоровлени           19         1         выздоровлени           20         1         выздоровлени	EL CALLANT A	1	выздоровлени	
15         1         выздоровлени           16         1         выздоровлени           17         1         выздоровлени           18         1         выздоровлени           19         1         выздоровлени           20         1         выздоровлени	1 AV VY S 1 (14		выздоровлени	
16         1         выздоровлени           17         1         выздоровлени           18         1         выздоровлени           19         1         выздоровлени           20         1         выздоровлени	0 15	1	выздоровлени	
17 1 выздоровлени 18 1 выздоровлени 19 1 выздоровлени 20 1 выздоровлени	3 0 16	1	выздоровлени	AN IN A
18 1 выздоровлени 19 1 выздоровлени 20 1 выздоровлени		0 1	выздоровлени	
19 1 выздоровлени 20 1 выздоровлени		12 - OT	выздоровлени	
СТ ЛУ СТ 1 Выздоровлени ЛУ	18			
	18 (n°	V 01	выздоровлени	

· 14. 3.	_01 ()	2
5.00	Форма	Исход
1 1 1	0' 001	летальный
2 7 12		летальный 🛇 🗌
3	<u> </u>	летальный
4	N AM	летальный
V 11/ 5		летальный
6		выздоровлени
2V V. 7	S A	выздоровлени
QY	1	выздоровлени
	<u></u>	выздоровлени
01		выздоровлени
A 11		выздоровлени
12		выздоровлени
13		выздоровлени
14		выздоровлени
0 15	20 11	выздоровлени
16	<u> </u>	выздоровлени
67	0.1	выздоровлени
18	$\wedge$	выздоровлени
5 19	V 01	выздоровлени
20	<u>` </u> ?' 1	выздоровлени
21		вызлоровлени

Tokna. Ma. edu. K. Marchander 2. Open the "Basic statistics and tables" module, select the "Tables and banners" tab. Make a contingency table. ma.edu.k2 skma.edu.k2 s





3. Options "Expected frequencies", "Pearson & ML Chi-square"

"Social health insurance and pu	iblic health"	№ 35-11(Б)-2 № 582
Guidelines for practical lessons of the course "Intro	oduction to Scientific Research"	p.47 out of
Crosstabulation lables Results: S	preadsheet1	J. 1 5
Quick Advanced Options		W. W.
Compute tables	Statistics for two-way tables	er fri vi
Highlight counts > 10	Pearson & M-L Chi-square	a. Co du.
	Fisher exact, Yates, McNemar (2	*2)
Residual frequencies	Phi (2x2 tables) & Cramér's V & (	Ma ev.
Percentages of total count	Kendall's tau-b & tau-c	st no.
Percentages of row counts	Gamma	St. no
Percentages of column counts	Spearman rank order correlation	
	Sommer's d	
a. ou the six alle a.	Uncertainty coefficients	
N. J. J. Kr. St. Mr. 6	W. I strate	

Форма	Исход выздоровление	Исход летальный	Row Totals	15
1 20	136	3.5	<u> </u>	
2 1	37	1 2 11	48	
3	8 6. 24	Ver Stray	15	
	5 1 5	P V 5 6		11
All Grps	186	29	215	30
10° 10°	Lit st do	2	14' 23'	

Summary Table: Expected Frequencies (Spreadsheet Marked cells have counts > 10 Pearson Chi-square: 45,6414, df=3, p=,000000 Исход Row Форма Исход Totals выздоровление летальный 9814 19.01860 141.0000 121 2 48,0000 41.5256 6.47442 2,02326 12,9767 15,0000 3 9,5163 1.48372 11,0000 All Grps 186,0000 29,00000 215,0000

Conclusion: P < 0.05, which means the hypothesis is rejected, that is, there is a relationship between the characteristics.

**Example 2**. Implementation of the calculated value of the characteristic with the Yates correction in the Statistica program for the above example.

1. Open the "Nonparametrics" module, select the "2x2 tables XI/VI/Phil, McNemar, Fisher exact" tab

2. Enter the data, click the "Summary" button

«Оңтүстік Қазақстан І	ОŃTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY медицина академиясы» АҚ	SOUTH KAZAKHSTAN SKMA -379- -379- -379- -379- -379- ACADEMY AO «Южно-Казахс	танская медицинская академия»
Departme	nts: "Medical Biophysics and I "Social health insurance and	nformation Technologies", public health"	№ 35-11(Б)-202 № 58202
Guidelines for pr	actical lessons of the course "In	troduction to Scientific Rese	earch" p.48 out of 68
I sknina.	Quick 33	Canc	

icy tabl

	2 x 2 Table	(Spreadshe	et1)
Kina. a. ea duite Ki skni	Column 1	Column 2	Row
Frequencies, row 1	33	0 1 24	57
Percent of total	28,448%	20,690%	49,138%
Frequencies, row 2	S	C 20 45	59
Percent of total	12,069%	38,793%	50,862%
Column totals	47	<u>69</u>	116
Percent of total	40,517%	59,483%	Hr 1
Chi-square (df=1)	14,04	p= ,0002	ON H
V-square (df=1)	13,92	p= ,0002	So. K
Yates corrected Chi-square	0 12,66	p= ,0004	x
Phi-square	,12105	Fr St d	
Fisher exact p, one-tailed	1	p= ,0002	de co
two-tailed		\ <b>`p= _0003</b>	A do
McNemar Chi-square (A/D)	6	p⇒ ,2130	St. JO
Chi-square (B/C)	/ 6 2,13	⊘p=,,1443	45 1

Conclusion: p<0.05, which means the hypothesis "H0" is rejected.

Tasks

1. The influence of the learning process on the results of a certain psychological test is being studied. Tests carried out on 100 schoolchildren revealed the results presented in the table below. Using the  $\chi 2$  criterion, it is required to investigate the presence of an influence of training on test results.

	Test results			
Age of schoolchildren	Low	Average	High	Total
Juniors	10	15	5	30
Average	6	16	8	30
Elder	67	13	20	40
Total	23	44	33	100

*Note.* Check your results in the STATISTICA program.

2. The table below provides information on newly diagnosed cases of primary and secondary syphilis, age and race rates per 100,000 population, USA,1989.Using the  $\chi^2$  test, it is required to

OŃTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY



«Оңтүстік Қазақстан медицина академиясы» АҚ

Departments: "Medical Biophysics and Information Technologies",<br/>"Social health insurance and public health"№ 35-11(Б)-2024<br/>№ 58- -2024<br/>p.49 out of 68Guidelines for practical lessons of the course "Introduction to Scientific Research"№ 36-11(Б)-2024<br/>№ 58- -2024

examine the presence of an effect of race on morbidity.

Age group	Kr. 1 3. KU.	Racial rate	per 100,000	10.1 5
(years)	White race	Black race	Other races	Total
< 14	0.0	2.4	0.8	0.4
15-19	2.4	131.5	51.0	24.3
20-24	5.8	323.0	139.2	55.9
25-29	5.4	270.9	117.9	44.1
30-34	4.7	256.6	83.2	38.8
35-44	2.9	135.0	47.8	19.0
45-54	1.70 00	76.7	29.6	10.5
>55	0.5	19.4	10.4	2.45
Total	2.2	115.8	45.8	Nº 17.7 St

Note. Check your results in the STATISTICA program.

3. Scientists compared the effectiveness of three antibiotics for recurrent urinary tract infections in girls 3-16 years old. After a short course of one of the antibacterial drugs (assigned at random), repeated urine cultures were taken over the course of a year. When bacteriuria was detected, a relapse was noted. The research results are presented in the table below. Using the  $\chi^2$  criterion, it is necessary to study the effectiveness of drugs.

All and Admin County 1	Relapse		
A drug	Eat	No	
Ampicillin	20	1 1 75 de	
Trimethoprim/sulfamethoxazole	5 24	XV 21 St	
Cephalexin	94	51 2	

*Note.* Check your results in the STATISTICA program.

4. The effectiveness of vaccination against typhus is being studied. The data is presented in the table below. Using the  $\chi^2$  criterion, draw conclusions about the effectiveness of vaccinations.

Observed values	Number of cases	Number of non-sick people	Total
With vaccination	72	7988	8060
Without vaccination	303	9322	9625
Total	375	17310	17685

*Note.* Check your results in the STATISTICA program.

5. The table below shows the classification data1000 people based on color blindness. Using the  $\chi^2$  test, check, is there a relationship between the presence of color blindness and a person's gender.

Observed values	Men	Women	Total
Colorblind	38	6	44
Not colorblind	442	514	956
Total	480	520	1000

Note. Check your results in the STATISTICA program.

6. It has been suggested that the skin reaction to dinitrochlorobenzene can serve as an assessment of the integrity of the patient's immunity. To test this assumption, we decided to compare the reaction to dinitrochlorobenzene with the reaction to croton oil, which has only a local irritant effect. To do this, a group of patients were applied to adjacent areas of clean skin with the agents being compared, and the reactions were recorded. The results are presented in the table below.

Reaction to di	nitrochlorobenzene	Total
Eat	No	Total
5 0	48	53
	Reaction to dir Eat 5	Reaction to dinitrochlorobenzeneEatNo548

мерізіла АКАРЕМІАЅҮ «Оңтүстік Қазақстан медицина академиясы» АҚ АО «Южно-Казахстанская мед	ицинская академия»
Departments: "Medical Biophysics and Information Technologies",	№ 35-11(Б)-2024
"Social health insurance and public health"	№ 582024
Guidelines for practical lessons of the course "Introduction to Scientific Research"	p.50 out of 68
	Kr S. Mi

SOUTH KAZAKHSTAN

No	23	10	33 5
Total	28	58	86

*Note.* Check your results in the STATISTICA program.

OŃTÚSTIK-QAZAQSTAN

7. There are data on mortality among patients suffering from open pulmonary tuberculosis during the first year after the disease.

It is required to establish whether the difference in mortality between men and women is significant.

Observed values	Alive	Died
Men	53	2.8
Women		10

Note. Check your results in the STATISTICA program.

8. There is evidence of the sociability of soldiers recruited from cities and rural areas. Test the hypothesis that city dwellers are more sociable than rural dwellers.

Observed values	Sociable	Unsociable
City	.10	5 14 0
Countryside	6	16

Note. Check your results in the STATISTICA program.

9. The effectiveness of high-frequency nerve stimulation as a pain reliever during tooth extraction is being studied. All patients were connected to the device, but in some cases it worked, in others it was turned off. Neither the dentist nor the patient knew whether the device was turned on. Do the following data suggest that high-frequency nerve stimulation is an effective analgesic agent?

Observed values	The device is turned on	The device is turned off
No pain	20	K 3.5 0 K
There is pain	16	12

Note. Check your results in the STATISTICA program.

## 6. Assessment methods / technologies: Oral questioning. Practical work. (assessment using a checklist).

#### 7. Literature:

Main:

- 1. Biostatistics: electronic textbook / K. Zh. Kudabaev [et al.]. Shymkent: SKFFA, 2015.
- 2. Rosner Bernard Fundamentals of Biostatistics: Texbook/ B.Rosner 8th ed.: GENGAGE learning, 2016.
- 3. Armitage P. Encyclopedia of Biostatistics. Wiley, 2016. 6100 p
- 4. Le C.T. Introductory biostatistics. Wiley, 2013. 536 p
- 5. Newman S.C. Biostatistical Methods in Epidemiology. Wiley, 2017. 388p
- 6. <u>http://www.biometrica.tomsk.ru</u>

#### 8. Control:

- 1. What is special about the analysis of qualitative features?
- 2. What is a contingency table of size mxn?
- 3. What conditions must be met when applying Pearson's  $\chi^2$  test?
- 4. Why is the Yates correction calculated?
- 5. To what cases does McNemar's criterion apply?
- 6. In what ways can it be implemented criterion 2-Pearson in a programme "STATISTICA"?
- 7. How is the information contained in the summary table interpreted?

MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ	SKMA -1979- , 1/, АО «Южно-Казахстанская ме	дицинская академия»
Departments: "Medical Biophysics and	Information Technologies",	№ 35-11(Б)-2024
"Social health insurance and	d public health"	№ 582024
Guidelines for practical lessons of the course "I	Introduction to Scientific Research"	p.51 out of 68

SOUTH KAZAKHSTAN

#### Lesson №7

#### 1. The theme: Correlation analysis

**2. The aim:** studying the principles of correlation analysis and the reliability of the obtained rank correlation coefficient when assessing risk factors for developing the disease

3. The learning objectives: find and study information on the following issues:

ONTUSTIK-QAZAOSTAN

- types of correlation dependence;
- Pearson pair correlation coefficient;
- assessment of the reliability of the correlation coefficient.
- conducting correlation analysis;
- calculate Spearman's rank correlation coefficient

#### 4. Key questions of the theme:

- 1. What is Correlation Analysis?
- 2. How is it located?Pearson's pairwise correlation coefficient?
- 3. How is the correlation coefficient determined?
- 4. What is the reference correlation coefficient used to estimate the relationship?
- 5. How p is calculatedSpearman's ang correlation coefficient?
- 6. What data will be included in the final table of correlation and regression analysis in Statistica?

**5. Methods / technologies of teaching and learning:** Computer-based work / Solving situational tasks.

#### Tasks:

**Example 1.**For the following data, calculate the linear Pearson correlation coefficient:

Incidence of acute respiratory infections per 1000 population, x	352 228	340	300	196	258	237
Incidence of pneumonia per 1000 population, in	64 60	52	48	46	41	32

#### Solution:

1) Create a calculation table:

No.	X	U	$x-\overline{x}$	$y - \overline{y}$	$(x-\overline{x})\cdot(y-\overline{y})$	$(x-\overline{x})^2$	$(y-\overline{y})^2$
0°1	352	64	79	15	1185	6241	225
2	228	60	-45	011	-495	2025	121
3	340	52	67	3	201	4489	9
4	300	48	27	-1	-27	729	OV Kr 1
55	196	46	<b>1</b> -77	-3	231	5929	9
6	258	41	-15	-8	120	225	64
475	237	32	-36	-17	612	1296	289
Sum	1911	343	0	0	1827	20934	718
Average	273	49		J. 1	St. Ja. 6	0-11.1.1	Kur S.

#### 2) Calculate the correlation coefficient:

$$r_{xy} = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}} = \frac{1827}{\sqrt{20934 \cdot 718}} = 0.47$$

ОŃTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ	SOUTH KAZAKHSTAN SKMA -597- -597- -597- -597- -597- ACADEMY AO «Южно-Казахстанская ме	едицинская академия»
Departments: "Medical Biophysics and In	nformation Technologies",	№ 35-11(Б)-2024
"Social health insurance and	public health"	№ 582024
Guidelines for practical lessons of the course "In	troduction to Scientific Research"	p.52 out of 68

3) Analyze the result obtained: the relationship between the characteristics under consideration is direct and moderate.

4) Calculate the average error of the correlation coefficient:

$$m_r = \pm \frac{1 - r_{xy}^2}{\sqrt{n}} = \pm \frac{1 - 0.47^2}{\sqrt{7}} = 0.3,$$

The correlation coefficient is not reliable, because does not exceed its average error by three times.

**Example 2.** In a certain locality, a chronic epidemic of Flexner's dysentery has been reported. Laboratory studies have shown that in the drinking water of the water supply network there are frequent "slips" of non-standard samples according to bacteriological indicators (risk factor). It is necessary to check whether there is a connection between these two signs.

Month	Number of patients with dysentery (x)	Proportion of non-standard water samples (y)
January	10	
February	K 3. 00 9k 5 K	0.5 (C)
March	2 4 3. 02 4 5 10	THE AND A PARTY OF ANY
April	5 10 × 70 × 5	
May	5 10 26 20 12 5	1.8 24 20 20
June	Ver St Will Con Wind	2.9
July	26 0 0	6.7
August	32	4.5
September	46	8.7
October	38	7.1
November	8 14 3.	3.2
December		

#### Solution:

1) Create a calculation table:

No.	X	at	r _x	$r_y$	$r_x - r_y$	$(r_x - r_y)^2$
F1 9	2	0	7	1.5	5.5	30.25
2	29	0.5	6	3	5 3	900
03	2	1.1	2. 10	4	5-3	9
4	7	2.0	4	6	-2	4
0.5 0	6	1.8	3	5	-2	5 4
6	11	2.9	8	. 7	80° (1	St. Mi
2 7 2	26	6.7	9	10	-1	J 5 1
8	32	4.5	10	9	2. 01	\$1.
9	46	8.7	11	12	· · -1	V K 19
10	38	7.1	12	11	1 N	du. Kh
	8	3.2	5	8	-3	9
12	5	0	2	1.5	0.5	0.25
Sum	11	Kr.	2. 0	32. K	1 5 1	70.5

e	ОНТОТІК-ОАДАСІАН МЕДІЗІНА АКАДЕМІАSY «Оңтүстік Қазақстан медицина академиясы» АҚ	-1979-	SOUTH KAZARHSTAN MEDICAL ACADEMY AO «Южно-Казахстанская ме	дицинская академия»
0	Departments: "Medical Biophysics and In	format	tion Technologies",	№ 35-11(Б)-2024
1	"Social health insurance and p	oublic 1	health"	№ 582024
0	Guidelines for practical lessons of the course "Intr	roduct	ion to Scientific Research"	p.53 out of 68

2) Calculate the correlation coefficient:

$$\rho = 1 - \frac{6}{n^3 - n} \sum_{i=1}^n (r_{x_i} - r_{y_i})^2 = 1 - \frac{6}{12^3 - 12} \cdot 70.5 \approx 0.75.$$

3) Analyze the result obtained: the connection between the characteristics under consideration is direct and high.

4) Calculate the average error of the correlation coefficient:

$$m_r = \pm \frac{1 - r_{xy}^2}{\sqrt{n}} = \pm \frac{1 - 0.75^2}{\sqrt{12}} \approx 0.12,$$

The correlation coefficient is reliable because exceeds its average error by more than three times.

**Example 3.** In one locality, a chronic epidemic of Flexner's dysentery was registered. Preliminary analysis and laboratory studies have shown that in the drinking water of the water supply network there are frequent "slips" of non-standard samples according to bacteriological indicators (risk factor). Calculate Spearman's rank correlation coefficient. The initial data is presented in the table.

Month	Number of patients with dysentery (n)	Proportion of non-standard water samples (x)
January	24° 03° 010 14° 1 3° 160	
February	200 10 5	(10.5 St. 10.5 St. 10
March	2 32 32 KV S	
April	A S MI 7 S SO A	2.0
May	0. Hr 5. 6. 2. Kr. K	5 6 1.8 1 2 2
June	and the still and show	2.9
July		6.7
August	32 20 20	4.5
September	46	8.7
October	38 1 26 23	7.1
November	24 . 0. 8 . 4 . 1 . 12	3.2
December	24 3. C5 14 3. 16	

*1.* Select Statistics  $\rightarrow$  Nonparametrics  $\rightarrow$  Correlations (Spearman, Kendall tau, gamma) (Figure 1), click the "OK" button.



Figure 1. Selecting the "Correlations (Spearman, Kendall tau, gamma)" procedure

ОŃTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ ЭЗЭЭ- ССССАL АСАDEMY АО «Южно-Казахстанская ме;	дицинская академия»
Departments: "Medical Biophysics and Information Technologies",	№ 35-11(Б)-2024
"Social health insurance and public health"	Nº 582024
Guidelines for practical lessons of the course "Introduction to Scientific Research"	p.54 out of 68

2. Set the variables "Variables" (Figure 2), click the "OK" button.



*3.* Press the "Spearman rank R" button.

The result of calculating the Spearman rank correlation coefficient is the following table (Figure 3).

The value of the rank correlation coefficient of 0.760071 indicates that the relationship between the characteristics is direct and high.

na.e.edu.u.k. A skrin	Spearman Rank Order Correlations (Spreadsheet1) MD pairwise deleted Marked correlations are significant at p <,05000			
Variable	число больных дизентерией	доля нестандартных проб воды	SKII AR B. B.	
число больных дизентерией	1,000000	0,760071	2. 2. 10.	
доля нестандартных проб воды	0,760071	1,000000		

Figure 3. Calculation result of Spearman's rank correlation coefficient

#### Tasks

1. The relationship between the incidence of myocardial infarction by month of the year and the average monthly air temperature is studied. The initial data are presented in the table.

Month	Incidence of myocardial infarction by month (per 10,000 thousand inhabitants)	Average monthly air temperature
January	1.6	-7.1
February	1.23	-7.7
March 9	1.14	-5.8
April	5	4.1
May	1.12	3. 0° +13 ¹ 5° (1)
June	1.02	+14.9
July	0.91	+18.8
August	0.82	+15.6
September	1.06	+9.0

OŃTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY



Эңтүстік қазақстан медицина академиясы» Ақ			
Departments: "Medical Biophysics and Information Technologies",	№ 35-11(Б)-2024		
"Social health insurance and public health"	№ 582024		
Guidelines for practical lessons of the course "Introduction to Scientific Research"	p.55 out of 68		

October	1.22	+6.0
November	1.33	-1.0
December	1.4	-7.7° XV

Calculate the Pearson correlation coefficient, determine the direction, strength and reliability of the relationship between indicators. Draw a conclusion.

2. The relationship between systolic blood pressure (SBP) and weight in women aged 20 to 30 years is studied. The initial data is presented in the table.

SBP, at	110	125	80	120	115	140	120	110	85
Weight, (kg)x	53	60	58	55	68	70	64	55	55

Calculate the Pearson correlation coefficient, determine the direction, strength and reliability of the relationship between indicators. Draw a conclusion.

3. The relationship between the duration of smoking and the number of respiratory tract diseases is studied. The initial data is presented in the table.

Smoking experience (years), x	2	4	5	21	3	2	24	5	6
Number of diseases	05	6	4	1	2	5	55	6	4

Calculate the Spearman correlation coefficient, determine the direction, strength and reliability of the relationship between indicators. Draw a conclusion.

4. The relationship between the thickness of the skin scar and the time of its freezing for cryodestruction is being studied. The initial data is presented in the table.

	5					AU.		
Skin scar thickness (mm), x	3	5	8	9	12	14	17	20
Freezing of the skin scar during cryodestruction (min.), y	0.6		1.6	1.5	1.7	1.6	2.4	3

Calculate the Spearman correlation coefficient, determine the direction, strength and reliability of the relationship between indicators. Draw a conclusion.

The relationship between the duration of contact of workers with industrial poisons and the 5. incidence of toxic hepatitis is studied. The data is presented in the table.

Duration of work (years), x	Up to 1	2	3	4	5	6
Morbidity (‰), y	2.0	8	7	11	10	13

Calculate the Spearman correlation coefficient, determine the direction, strength and reliability of the relationship between indicators. Draw a conclusion.

6. In one of the districts of the West Kazakhstan region, cases of anicteric leptospirosis (water fever) appeared. A study was carried out on the relationship between the number of diseases and the amount of precipitation. The data is presented in the table.

Number of diseases, x	0	19	4	F1_	2	68	131	14	115	2
Precipitation amount (mm), y	54	101	185	85	30	128	143	74	28	132

Calculate Spearman's rank correlation coefficient, determine the direction, strength and reliability of

АКАДЕМІАЅҮ АСАДЕМҮ «Оңтүстік Қазақстан медицина академиясы» АҚ АО «Южно-Казахстанская медиц	цинская академия»
Departments: "Medical Biophysics and Information Technologies",	№ 35-11(Б)-2024
"Social health insurance and public health"	№ 582024
Guidelines for practical lessons of the course "Introduction to Scientific Research"	p.56 out of 68

(SKMA)

SOUTH KAZAKHSTAN

MEDICAL

the relationship between indicators. Draw a conclusion.

ΟΝΤΊΣΤΙΚ-ΟΑΖΑΟSTAN

MEDISINA

7. The relationship between the thickness of the coal seam and the incidence of hypertension among miners is studied. The data is presented in the table.

Seam thickness (m), x	0.6	0.8	1.0	1.2	1.4	1.5	1.6
Incidence of hypertension (per 1000 miners), in	3.5	4.2	5.6	6.3	7.4	8.9	10

Calculate Spearman's rank correlation coefficient, determine the direction, strength and reliability of the relationship between indicators. Draw a conclusion.

6. Assessment methods / technologies: Oral questioning. Practical work. (assessment using a checklist).

#### 7. Literature:

Main:

1. Biostatistics: electronic textbook / K. Zh. Kudabaev [et al.]. - Shymkent: SKFFA, 2015.

2. Rosner Bernard Fundamentals of Biostatistics: Texbook/ B.Rosner - 8th ed.: GENGAGE learning, 2016.

- 3. Armitage P. Encyclopedia of Biostatistics. Wiley, 2016. 6100 p
- 4. Le C.T. Introductory biostatistics. Wiley, 2013. 536 p
- 5. Newman S.C. Biostatistical Methods in Epidemiology. Wiley, 2017. 388p
- 6. <u>http://www.biometrica.tomsk.ru</u>

#### 8. Control:

- 1. How is linear correlation coefficient calculated?
- 2. What types of correlation do you know?
- 3. How is the reliability of the correlation coefficient determined?
- 4. What is correlation analysis?
- 5. What is the reference correlation coefficient used to estimate the relationship?
- 6. How is p calculatedSpearman's ang correlation coefficient?
- 7. What data will be included in the final table of correlation and regression analysis in Statistica?

#### Lesson №8

#### 1. The theme: Regression analysis

**2.** The aim:Formation of skills in drawing up a linear regression equation using the least squares method, checking the significance of regression coefficients and finding the coefficient of determination.

#### 3. The learning objectives:

- learn to create a linear regression equation using the least squares method;

-check the significance of regression coefficients, check the significance of the regression equation; -developing the ability to teach how to find the coefficient of determination.

#### 4. Key questions of the theme:

- 1. What is the main meaning of regression analysis?
- 2. What is regression and what are its types?
- 3. What types of pairwise regression equations do you know?
- 4. What is the main meaning of least squares?
- 5. What formula is used to determine the coefficients of a paired regression equation?

### **5. Methods / technologies of teaching and learning:** Computer-based work / Solving situational tasks.

#### Tasks:

**Example 1.** Based on the following data, construct and analyze a regression equation:

OŃTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY обо SKMA MEDICAL ACADEMY AO «Южно-Казахстанская медицинская академия»

«Оңтүстік Қазақстан медицина академиясы» АҚ

		ł
Departments: "Medical Biophysics and Information Technologies",	№ 35-11(Б)-2024	
"Social health insurance and public health"	№ 582024	
Guidelines for practical lessons of the course "Introduction to Scientific Research"	p.57 out of 68	
		1

Incidence of influenza per 1000 population, x	352	228	340	300	196	258	237
Incidence of pneumonia per 1000 population, in	64	60	52	48	46	41	32

#### Solution.

1) Calculate the correlation coefficient:  $r_{xy} = 0,47$ . The relationship between the signs is direct and moderate.

- 2) Construct a paired linear regression equation.
  - 2.1) Create a calculation table.

No.	X	U	xy	x2	y _x	$(y_x - \overline{y})^2$	(y-y _x ) ²	$(x-\overline{x})^2$
1	352	64	22528	123904	55.89	47.54	65.70	6241
2	228	60	13680	51984	45.07	15.42	222.83	2025
93	340	52	17680	115600	54.85	34.19	8.11	4489
45	300	48	14400	90000	51.36	5.55	11.27	729
5	196	46	9016	38416	42.28	45.16	13.84	5929
6	258	41	10578	66564	47.69	1.71	44.77	225
07 V	237	32	7584	56169	45.86	9.87	192.05	1296
Sum	1911	343	95466	542637	343	159.45	558.55	20934
Average	273	49	13638	77519.6	49	22.78	79.79	2990.6

2.2) Calculate regression coefficients:

$$b = \frac{yx - \bar{y} \cdot \bar{x}}{\bar{x}^2 - \bar{x}^2} = \frac{13638 - 49 \cdot 273}{77519,6 - 273^2} = 0,087,$$
  
$$a = \bar{y} - b\bar{x} = 49 - 0.087 \cdot 273 = 25.17$$

Paired linear regression equation: yx=25.17+0.087x.

3) Find the theoretical values of "yx" by substituting the actual values of "x" into the regression equation.

4) Construct graphs of actual "y" and theoretical values of "yy" of the resultant characteristic:



ОЙТÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ Оңтүстік Қазақстан медицина академиясы» АҚ	цинская академия»
Departments: "Medical Biophysics and Information Technologies", "Social health insurance and public health"	№ 35-11(Б)-2024 № 582024
Guidelines for practical lessons of the course "Introduction to Scientific Research"	p.58 out of 68
5) Check the statistical significance of regression coefficients: 5.1) Calculate random errors: $m_{b} = \sqrt{\frac{\sum(y - y_{x})^{2}}{\sum(x - \overline{x})^{2}}} = \sqrt{\frac{558,55}{7-2}} \approx 0,073;$ $m_{a} = \sqrt{\frac{\sum(y - y_{x})^{2}}{n-2}} \cdot \frac{\sum x^{2}}{n\sum(x - \overline{x})^{2}} = \sqrt{\frac{79,8}{7-2}} \cdot \frac{542637}{7 \cdot 2990,6} \approx 20,34.$ 5.2)	edu. K. Skina. edu. K. Skina. e na. edu. K. Skina. e Skina. e s
$t_{b\ pacu} = \frac{b}{m_b} = \frac{0.087}{0.073} \approx 1.19$ , $t_{a\ pacu} = \frac{a}{m_a} = \frac{25.17}{20.34} \approx 1.24$ .	JU.KI SKN Ma. ed

- 5) Check the statistical significance of regression coefficients:
  - 5.1) Calculate random errors:

tical significance of regression coefficients:  
andom errors:  

$$m_{b} = \sqrt{\frac{\sum(y - y_{x})^{2}}{\sum(x - \bar{x})^{2}}} = \sqrt{\frac{558,55}{20934}} \approx 0,073;$$

$$m_{a} = \sqrt{\frac{\sum(y - y_{x})^{2}}{n - 2}} \cdot \frac{\sum x^{2}}{n\sum(x - \bar{x})^{2}}} = \sqrt{\frac{79,8}{7 - 2}} \cdot \frac{542637}{7 \cdot 2990,6}} \approx 20,34.$$

$$t_{b \ pacy} = \frac{b}{m_{b}} = \frac{0,087}{0,073} \approx 1,19, \qquad t_{a \ pacy} = \frac{a}{m_{a}} = \frac{25,17}{20,34} \approx 1,24.$$

$$5) = 2.57 \text{ (see table 1).}$$
Table 5.

$$t_{b \ pacu} = \frac{b}{m_b} = \frac{0.087}{0.073} \approx 1.19, \qquad t_{a \ pacu} = \frac{a}{m_a} = \frac{25.17}{20.34} \approx 1.24$$

I able o
----------

$m_{a} = \sqrt{\frac{\sum (y)}{n}}$ 5.2) $t_{b \ pacy} = \frac{b}{m_{b}} =$	$\frac{(-y_x)^2}{-2} \cdot \frac{\sum 5}{n \sum (x)}$ $= \frac{0,087}{0,073} \approx 1,19,$	$\frac{x^2}{(-\bar{x})^2} = \sqrt{\frac{79.8}{7-2}} \cdot \frac{4}{7}$ $t_a  pacy = \frac{4}{7}$	$\frac{542637}{2990,6} \approx 20,34.$ $\frac{a}{n_a} = \frac{25,17}{20,34} \approx 1,24.$	1.K2 SKINA
5.3) $t_{table} (0.05; 5) = 2.57$ (see the number of degrees of	table 1).	na.eachu.ku.ku	KA SKITTA	edu Ta
freedom	200, 4, 4, 1	significa	lice level, a	
C A A A	0,10	0,05	0,02	0,01
	6,31	12,7	31,82	63,7
	2,92	4,30	6,97	9,92
	2,35	3,18	4,54	5,84
St. 60 40	2,13	2,78	3,75	4,60
<u> </u>	2,01	2,57	3,37	4,03
6	1,94	2,45	3,14	3,71
2	1,89	2,30	3,00	3,50
	1,80	2,31	2,90	3,30
9	1,83	2,20	2,82	3,23
	1,01	2,23	2,70	3,17
	1,80	2,22	2,72	3,11
13	1,78	2,10	2,00	3,05
14	1,76	2,10	2,62	2.98
5 15 15	1,75	2.13	2,62	2,95
16	1,75	2,12	2,58	2,93
17	1,74	2,11	2.57	2.90
518	1,73	2,10	2,55	2.88
19 00 0	1.73	2,09	2,54	2.86
20	1,73	2,09	2,53	2,85
30	1,70	2,04	2,46	2,75
40 1 6	1,68	2,02	2,42	2,70
60 60	1,67	2,00	2,39	2,66
	1,66	1,98	2,36	2,62
120				

OŃTÚSTIK-QAZAQSTAN SOUTH KAZAKHSTAN MEDISINA SKM/ MEDICAL AKADEMIASY ACADEMY «Оңтүстік Қазақстан медицина академиясы» АҚ АО «Южно-Казахстанская медицинская академия» Departments: "Medical Biophysics and Information Technologies", № 35-11(Б)-2024 "Social health insurance and public health" Nº 58--2024Guidelines for practical lessons of the course "Introduction to Scientific Research' p.59 out of 68

5.4)  $t_{b calc} < t_{table}$ , which means the coefficient "b" is insignificant,

 $t_{a calc} < t_{table}$ , which means the coefficient "a" is insignificant.

6) Check the statistical significance of the regression equation:

(5.1) 
$$F_{calc} = \frac{\frac{\sum (y_x - \bar{y})^2}{k}}{\frac{\sum (y - y_x)^2}{n - k - 1}} = \frac{\frac{159,45}{1}}{\frac{558,55}{5}} = 1,43.$$

6.2)  $F_{table}(p;k; nk-1)=(0.05; 1; 5)=6.61$  (see table 2).

6.3)  $F_{calc} < F_{table}$ , then the regression equation is chosen incorrectly. This result can be explained by the low closeness of the relationship ( $r_{xy}=0.47$ ) and the small number of observations.

7) Calculate the coefficient of determination:  $R_2=(0.47)2=0.22$ , the constructed equation is of poor quality.

#### Tasks:

1. The relationship between the incidence of myocardial infarction by month of the year and the average monthly air temperature is studied. The initial data are presented in the table.

Month	Incidence of myocardial infarction by month (per 10,000 thousand inhabitants)	Average monthly air temperature
January	ST (1.6) 1.6	-7.1.
February	1.23	-7.7
March	1.14	-5.8
April	1.13	4.1 0 1
May	1.12	+13
June	1.02	+14.9
July	0.91	+18.8
August	0.82	+15.6
September	1.06	+9.0
October	1.22	+6.0
November	1.33	
December	9 9.4	-7.7

Construct and estimate a paired linear regression equation.

2. The relationship between systolic blood pressure (SBP) and weight in women aged 20 to 30 years is studied. The initial data is presented in the table.

SBP, at	110	125	80	120	115	140	120	110	85
Weight, (kg)x	53	60	58	55	68	70 ×	64	55	55

Construct and estimate a paired linear regression equation.

3. The relationship between the duration of smoking and the number of respiratory tract diseases is studied. The initial data is presented in the table.

Smoking experience (years), x	2	400	5	1	3	2	4	5	6
Number of diseases	5	6	4	1	2	5	5	6	4

Construct and estimate a paired linear regression equation.

устік Қазақст	ан меди	ОŃTÚS цина ака	TIK-QAZA MED AKADEN AGEMURCI	QSTAN ISINA MIASY ы» АҚ	SKMA -1979-	SOUTH K MEDICA ACADEJ AO «Ю»	AZAKHSTA NL MY KHO-Kasa	N	ая медиі	цинская а	кадемия»
Depart	tments: " So	Medical ocial hea	Biophys lth insura	sics and I ance and	nformati public h	on Techi ealth"	nologies"	edu.	J.K.	№ 35-1 № 58-	11(Б)-2024 -2024
Guidelines for	practica	l lessons	of the c	ourse "In	troductio	on to Sci	entific Re	esearch"		p.60	out of 68
ng.e.edu	N. KI	1-5×	Kino	0.00	200.1	XI A	SK.	<u> </u>	ead	S. K.	Table 2
k1 K1	60	2	3	40	<b>5</b> 00	6	8	5/12	24		Fr Si
241	161,4 18.51	199,5	215,7	224,6	230,2	234,0 19,33	238,9 19.37	243,9	249,0 19,45	254,3 19,50	JU. Y.KI

	1 1 1 1 1	0,24	0,59	0,39	0,20	0,10	0,04	(J),91 [	12,119	) J,0.
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,59	4,74	4,35	24,12	3,97	3,87	3,73	3,57	3,41	3,23
08	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
90	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11.0	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,1
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,52
60	4,00	3,15	2,76	2,52	2,37	2,25	2,10	1,92	1,70	1,39
120	3,92	3,07	2,68	2,45	2,29	2,17	2,02	1,83	1,61	1,25
00	3 84	2.99	2.60	2.37	2.21	2.09	1.94	1.75	1.52	1.00

4. The relationship between the thickness of the skin scar and the time of its freezing for cryodestruction is being studied. The initial data is presented in the table.

Skin scar thickness (mm), x	3	5	8	9	12	14	17	20
Freezing of the skin scar during cryodestruction (min.), y	0.6	1	1.6	1.5	1.7	1.6	2.4	33

Construct and estimate a paired linear regression equation.

«Оңтүстік

5. The relationship between the duration of contact of workers with industrial poisons and the incidence of toxic hepatitis is studied. The data is presented in the table.

OŇTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY



«Оңтүстік Қазақстан медицина академиясы» АҚ

Departments: "Medical Biophysics and Information Technologies",	№ 35-11(Б)-2024
"Social health insurance and public health"	№ 582024
Guidelines for practical lessons of the course "Introduction to Scientific Research"	p.61 out of 68

Duration of work (years), x	Up to 1	2	3	4	5	6
Morbidity (‰), y	$2^{2}$	8	75	11	10	13

Construct and estimate a paired linear regression equation.

6. In one of the districts of the West Kazakhstan region, cases of anicteric leptospirosis (water fever) appeared. A study was carried out on the relationship between the number of diseases and the amount of precipitation. The data is presented in the table.

Number of diseases, x	0	19	4	1	02	68	131	14	11	2
Precipitation amount (mm), y	54	101	185	85	30	128	143	-74	28	132

Construct and estimate a paired linear regression equation.

6. Assessment methods / technologies: Oral questioning. Practical work. (assessment using a checklist).

#### 7. Literature:

Main:

1. Biostatistics: electronic textbook / K. Zh. Kudabaev [et al.]. - Shymkent: SKFFA, 2015.

2. Rosner Bernard Fundamentals of Biostatistics: Texbook/ B.Rosner - 8th ed.: GENGAGE learning, 2016.

- 3. Armitage P. Encyclopedia of Biostatistics. Wiley, 2016. 6100 p
- 4. Le C.T. Introductory biostatistics. Wiley, 2013. 536 p
- 5. Newman S.C. Biostatistical Methods in Epidemiology. Wiley, 2017. 388p
- 6. <u>http://www.biometrica.tomsk.ru</u>

#### 8. Control:

- 1. What types of pairwise regression equations do you know?
- 2. What is the main significance of the least squares method?
- 3. What formula is used to determine the coefficients of a paired regression equation?

#### Lesson №9

- 1. The theme: Health systems in Kazakhstan. International cooperation in health care.
- 2. The aim: to familiarize students with the healthcare system in Kazakhstan.
- **3. The learning objectives:** 
  - The student must know and be able to:
  - •Basic principles and objectives of the state policy in the field of health care.
  - •The healthcare system in the Republic of Kazakhstan.
  - •International cooperation in the field of healthcare.

#### 4. Key questions of the theme:

- 1. State and non-state health sector.
- 2. Pros and cons of the healthcare system.
- 3. Prospects for the development of healthcare.
- 5. Methods / technologies of teaching and learning: Training cases, question and answer
- 6. Assessment methods / technologies Assessment using a check list
- 7. Literature: see Appendix 1.

#### 8. Control (questions):

- 1. What is included in the healthcare system?
- 2. What is the healthcare system in Kazakhstan?
- 3. At what level is medicine in Kazakhstan?
- 4. What problems exist in the healthcare system?
- 5. Name the priority directions of international cooperation in the field of health care.

онтизник-одагаозтан MEDISINA АКАДЕМІАЅҮ «Оңтүстік Қазақстан медицина академиясы» АҚ Оңтүстік Қазақстан медицина академиясы» АҚ	цинская академия»
Departments: "Medical Biophysics and Information Technologies",	№ 35-11(Б)-2024
"Social health insurance and public health"	Nº 582024
Guidelines for practical lessons of the course "Introduction to Scientific Research"	p.62 out of 68

#### Lesson №10

The theme: Methodology for calculating and analyzing medical and demographic indicators.
 The aim: to familiarize students with the medical and social aspects of demography and talk about the solution of demography is indicators.

the calculation and analysis of demographic indicators in a comprehensive health assessment.

#### 3. The learning objectives:

The student must know and be able to:

- Medico-social aspects of demography.
- The principle of calculation and analysis of demographic indicators.
- Comprehensive assessment of public health.

#### 4. Key questions of the theme:

- 1) What are the medical and social aspects of demography?
- 2) What demographics shape public health?
- 3) How are demographics calculated?
- 4) What methods are used in the analysis of demographic indicators?
- 5. Methods / technologies of teaching and learning: Training cases, case-study
- 6. Assessment methods / technologies Assessment interview using a checklist
- 7. Literature: see Appendix 1.

#### 8. Control (questions):

- 1. What current health questions can be answered by examining demographics?
- 2. What demographics shape public health?
- 3. What methods are used in the analysis of demographic indicators?
- 4. What indicators are needed to calculate a comprehensive assessment of public health?

#### Lesson №11

**1. The theme:** Modern medical and social problems, issues of strengthening the health of the population.

2. The aim: to familiarize students with the criteria for the incidence of the population.

#### 3. The learning objectives:

The student must know and be able to:

- morbidity of the population.
- Classification of the incidence of the population.
- Quantitative criteria for morbidity according to WHO.
- Methods for calculating the criteria for morbidity (indicator of the frequency of primary morbidity, the prevalence of general morbidity, the incidence rate).

#### 4. Key questions of the theme:

- 1) What do you understand by population morbidity?
- 2) On what grounds is the classification of the incidence of the population?
- 3) What morbidity criteria are quantitative according to WHO?
- 4) What calculation methods are intended for quantitative criteria of morbidity?
- 5) According to what principle is the incidence of the population divided into primary, secondary and tertiary?
- 5. Methods / technologies of teaching and learning: Educational cases, case-study
- 6. Assessment methods / technologies Assessment using a checklist
- 7. Literature: see Appendix 1.

 

 MEDISINA АКАДЕМІАЅҮ «Оңтүстік Қазақстан медицина академиясы» АҚ
 MEDICAL АСАДЕМҮ АО «Южно-Казахстанская медицинская академия»

 Departments: "Medical Biophysics and Information Technologies", "Social health insurance and public health"
 № 35-11(Б)-2024

 Guidelines for practical lessons of the course "Introduction to Scientific Research"
 № 35-11(Б)-2024

SOUTH KAZAKHSTAN

#### 8. Control (questions):

- 1. What morbidity rates do you know?
- 2. How does the ICD-10 classification affect the definition and calculation of incidence rates?
- 3. By what criteria does WHO classify incidence rates?
- 4. What methods of calculating incidence rates are most often used to determine the level of public health?

#### Lesson №12

1. The theme: Organization and conduct of medical and social expertise.

OŃTÚSTIK-QAZAQSTAN

2. The aim: to familiarize students with the activities of MSE.

#### 3. The learning objectives:

The student must know and be able to:

- Main functions of MSE.
- Criteria for determining disability groups.
- -The procedure for appealing the conclusion of MSE.

#### 4.Key questions of the theme:

- Grounds for conducting a medical and social examination.
- Inspection procedure.
- Criteria for establishing disability.
- 5. Methods / technologies of teaching and learning: Training cases, case-study
- 6. Assessment methods / technologies Assessment interview using a checklist
- 7. Literature: see Appendix 1.

#### 8. Control (questions):

- 1. What is a medical and social expert commission?
- 2. Who gives direction to MSE?
- 3. Who is part of MSE?
- 4. How to challenge the conclusion of MSE?

#### Lesson №13

1. The theme: Medical care and its types.

**2. The aim:** to familiarize students with the types of medical care provided and the forms of its provision.

#### 3. The learning objectives:

The student must know and be able to:

- Types of medical care.
- Forms of providing medical care.
- The principle of distribution of children and adolescents by health groups.
- The principle of distribution of adults by health groups.
- Distinctive features of the distribution of adults into 3 groups, and children and adolescents into 5 health groups.

#### 4. Key questions of the theme:

- 1) What types of medical care are provided by health facilities in the Republic of Kazakhstan?
- 2) What forms of medical care are used in the Republic of Kazakhstan?
- 3) What health groups do you know?
- 4) What is the difference between the distribution of adults and children by health groups?
- 5. Methods / technologies of teaching and learning: Educational cases, case study

ÍTÚSTIK-QAZAQSTAN
MEDISINA
AKADEMIASY
A1/

SOUTH KAZAKHSTAN **SKM** MEDICAL ACADEMY shi,

«Оңтүстік Қазақстан медицина академиясы» АҚ

O

АО «Южно-Казахстанская медицинская академия» Departments: "Medical Biophysics and Information Technologies", № 35-11(Б)-2024 "Social health insurance and public health' № 58--2024

Guidelines for practical lessons of the course "Introduction to Scientific Research" p.64 out of 68

- 6. Assessment methods / technologies Assessment using a checklist
- 7. Literature: see Appendix 1.

#### 8. Control (questions):

- 1. On what basis is medical care provided?
- 2. What are the general and distinctive features of the distribution of medical care into forms and types?
- 3. What is the reason for the distribution of health groups in the adult population into 3 groups, and in children - into 5 groups?
- What are the key distribution criteria for each adult and child health group?

#### Lesson №14

1. The theme: Medical secrecy.

2. The aim: to acquaint students with medical secrecy, the procedure for providing information and

the responsibility for disclosing it.

#### 3. The learning objectives:

- The student must know and be able to:
- •Definition of the term "confidentiality".
- •Hippocratic oath of medical secrecy.
- •Legal basis of medical secrecy.

#### 4. Key questions of the theme:

- 1. Medical secrecy: regulation.
- 2. Preservation of medical secrecy: medical documents.
- 3. Conditions for disclosure of medical confidentiality.
- 4. Punishment for disclosure of medical secrets.
- 5. Methods / technologies of teaching and learning: Training cases
- 6. Assessment methods / technologies Assessment using a checklist
- 7. Literature: see Appendix 1.

#### 8. Control (questions):

- 1. What is included in the concept of medical secrecy?
- 2. When is disclosure of medical confidentiality allowed?
- 3. What is the responsibility for disclosing medical confidentiality?

#### Lesson №15

1. The theme: The universality of the ethical norm and the uniqueness of the moral choice.

2. The aim: explain to the students the universality of the ethical norm and the uniqueness of the moral choice.

#### 3. The learning objectives:

The student must know and be able to:

- The principle of justice
- The concept of the universality of the ethical norm
- Moral choice and morality
- The uniqueness of moral choice in medicine

#### 4.Key questions of the theme:

- 1. How are the principle of justice and the universality of the ethical norm combined?
- 2. What is a moral choice?
- What is meant by the term "uniqueness"? 3.

OŃTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY

2962 SOUTH KAZAKHSTAN SKMA MEDICAL ACADEMY sh.

«Оңтүстік Қазақстан медицина академиясы» АҚ

АО «Южно-Казахстанская медицинская академия» Departments: "Medical Biophysics and Information Technologies", № 35-11(Б)-2024

"Social health insurance and public health" № 58--2024Guidelines for practical lessons of the course "Introduction to Scientific Research" p.65 out of 68

- 4. What is the uniqueness of moral choice in medicine?
- 5. Methods / technologies of teaching and learning: Educational cases, case study
- 6. Assessment methods / technologies Assessment using a checklist
- 7. Literature: see Appendix 1.

#### 8. Control (questions):

- 1. What are the main criteria for the principle of justice?
- 2. How do you understand the terms "universality" and "uniqueness"? Give examples.
- 3. What are the main differences between morality and moral choice?
- 4. How is the interaction of the principle of justice and the uniqueness of the moral choice?

#### Lesson №16

- 1. The theme: Ethical regulations.
- 2. The aim: familiarize students with normative-legal acts regulating medical activity.

#### 3. The learning objectives:

The student must know and be able to:

- the Constitution of the Republic of Kazakhstan.
- Code of honor for medical workers.
- Rights and obligations of persons in the field of healthcare.
- Code of the Republic of Kazakhstan "On the health of the people and the healthcare system"

#### 4.Key questions of the theme:

- The Constitution of the Republic of Kazakhstan.
- Ethical principles and requirements for medical workers.
- Principles of state policy in the field of healthcare.
- Criminal Code of the Republic of Kazakhstan.
- 5. Methods / technologies of teaching and learning: Educational cases, case study
- 6. Assessment methods / technologies Assessment using a checklist
- 7. Literature: see Appendix 1.

#### 8. Control (questions):

- 1. What legal acts regulate medical activity?
- 2. What normative acts include provisions on the protection of the health of citizens?
- 3. What is the fundamental legal document for doctors?

ортистик-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ АО «Южно-Казахстанская медиц	цинская академия»
Departments: "Medical Biophysics and Information Technologies",	№ 35-11(Б)-2024
"Social health insurance and public health"	№ 582024
Guidelines for practical lessons of the course "Introduction to Scientific Research"	p.66 out of 68

#### Appendix 1

#### Literature:

#### Main:

- Public health: textbook / A. A. Akanov [and others]. ; Approved and rec. committee for control in the field of education and science. Ministry of Education and Science of the Republic of Kazakhstan. - M. : "Litterra", 2017. - 496 p.
- Boleshov, M. A. Kogamdyk densaulyk zhane densaulykty saktau: okulyk / M. A. Boleshov. Almaty: Evero, 2015. - 244 bet p.
- Campbell, A. Medical Ethics / A. Campbell, G. Gillette, G. Jones; ed. Yu. M. Lopukhin. M. : GEOTAR - Media, 2014. - 368 bet. With.

#### Additional:

1. Rymanov D.M., Medic V. A. Public health and health care: hands. to practical exercises. - M.: GEOTAR - Media, 2012. - 400 p.

#### **Electronic resources:**

- Lisitsyn Yu. P. Public health and healthcare [Electronic resource]: textbook / Yu.P. Lisitsyn, G.E. Ulumbekova. 3rd ed., revised. and additional Electron. text data. (43.1Mb). M. : GEOTAR Media, 2017. el. opt.
- Medic V. A. Public health and healthcare [Electronic resource]: textbook / V.A. Medic, V. K. Yuryev. - Electron. text data. (47.6 Mb). - M. : GEOTAR - Media, 2013. - 608 p. email
- Lisitsyn Yu. P. Public health and healthcare [Electronic resource]: textbook / Yu. P. Lisitsyn, G. E. Ulumbekova. 3rd ed., revised. and additional Electron. text data. (40.9 Mb). M.: Ed. group "GEOTAR-Media", 2011. 544 el.
- Shchepin O. P. Public health and healthcare [Electronic resource]: textbook / O. P. Shchepin, V. A. Medic. - Electron. text data. (43.6 Mb). - M.: Ed. group "GEOTAR-Media", 2011. - 592 p. email opt. disc (CD-ROM).
- Medic V. A. Public health and healthcare [Electronic resource]: a textbook for honey. Schools and colleges / V. A. Medik., V. K. Yuriev. 3rd ed., revised. and additional Electron. text data. (37.2 MB). M.: Publishing group "GEOTAR-Media", 2011. 288 p. email opt. disk.

#### Electronic database:

No.	Name	Links
71 N	Digital library	http://lib.ukma.kz
25	Digital catalogue	http://10.10.202.52
KL.	- For internal users	http://89.218.155.74
· vi	- For external users	200
3	Intermediate republican higher educational	http://rmebrk.kz/
Y 11	institutions electronic library	and all the structure and
4	Electronic Library of the Medical School	http://www.studmedlib/ru
ð. 0	"Student Advisor"	T and a set with set and
5	Section "Paragraph" information system	https://online.zakon.kz/Medicine
SU.	"Medicine"	1 st no compt 1 st
6	Electronic source of legal information "law"	https://zan.kz
7	Scientific electronic library	https://elibrary/ru/
89	"BooksMed" electronic library	http://www.booksmed.com

О́́́ИТÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Оңтүстік Қазақстан медицина академиясы» АҚ	SOUTH KAZAKHSTAN MEDICAL ACADEMY AO «Южно-Казахстанская мег	цицинская академия»
Departments: "Medical Biophysics and Infor	mation Technologies",	№ 35-11(Б)-2024
"Social health insurance and pub	lic health"	№ 582024
Guidelines for practical lessons of the course "Introd	luction to Scientific Research"	p.67 out of 68
0 ("Web of asian as" (Thomson Deutero)	http://www.constaction	WY S. HU
9 Web of science" (Thomson Reuters)	http://apps.webotknowledge.com	
	https://www.sciencedirect.com	
10 "Science Direct" (Elsevier)	https://www.sciencedire	ect.com
10"Science Direct" (Elsevier)11"Scopus" (Elsevier)	https://www.sciencedire www.scopus.com	ect.com

Sudennes for practical ressons of the course introduction to Scientific Rescarein				
9 "Web of	science" (Thomson Reuters)	http://apps.webofknowl	edge.com	
10 "Science	Direct" (Elsevier)	https://www.sciencedirect.com		
11 "Scopus	(Elsevier)	www.scopus.com		
12 PubMed	13. 60 11 1 3 XIII.	https://www.ncbi.nlm.n	<u>ih.gov/pubmed</u>	
a.edu.KI	KIL SKINA. COLULIKI SI	KL SKINA. edu. KL SK	KL SKINA. BUNA	
skmared	edu, KI skina. eou	du.k. K. S. sknina.e. edu	du. K. K. St. St.	
JU.K. K. SKI	Na. C. COLOU.K. K. S. SKN	KM3.edu. Kr Sky	KIN3. et edu.	
ma.edu.uk	I Stakmana.edu.	KL SK SKMara.eu.K	L SKIMB.	
SKI SKIND. BU	edu. K. K. Skina.	200 edu. KL S. Skillena.	edu. Kr s	
du.K. K. St. St.	SKINB. Edu. KL KL SK	Kma. equilite Ki Ski	KIN3. Education	
KUUS EQUALITY	KI SKITTER SKITTER SCOUL	J.K. SK. SKMara.edu.	J. KI SKINO.	
KI SKMO.	a.edu. K. K. Skina.e	eac du.k. K. S. skina.e	edu. H	
source star	KMana.edu. V. K. Sk	SKM3. Eacout. K. SK	skina.edu	
13. c. edu, du, k	I.K. St. Skina. et. edu.	A Skina. eurou. k	A SKIII	
2 11 20		Wint st no e	N. W.	

OŃTÚSTIK-QAZAQSTAN MEDISINA AKADEMIASY «Онтустік Казакстан медицина академиясы» АК «Онтустік Казакстан медицина академиясы» АК	NUIINHCKad akanewida
Departments: "Medical Biophysics and Information Technologies",	№ 35-11(Б)-202
"Social health insurance and public health"	<u>№</u> 58202 p 68 out of 68
The second states and the se	KI ST ST
St. Mo. 200 Mr. D. St. as. 60 miler 1 2 Kr. 23.0	dr. Kr Sh
Structure Ser Write Structure Contraction of the Structure S	edu
A service service at the way control a shirt	3. equ. 14
i the second second in the second second in the	13. 60. 1.
and the second solution of the second solutio	41, 3. 60,
and the second second in the second second in	241, 33. 60
3. BUT THE STURE SECTION A SK MON ON WITH	1 ski al
and the structure with the structure of the	
the second states and second the spirition of the	J.T. L - 441 -
the second the structure with the structure on	11. 12 - 24
La Kara S. Con the second son the standard	er 10. 12
1 Star 3. On the Star and Star and the Star and	er 10. 11
1. M. B. Skill B. Shi M. Shi Kill Stranger All & Skill	No. CO WILL
With a star so of the service so and the st	na ev ju
SO MILLA SHILL S. ON MILL S. MILL S. C. ON. H.	
COMPART SKILLER OD WITH STATE STORY KI	St. no et
10. CO. 10. 1 2 41. 19. 60. 14. 1 2. 41. 10. 9. 10. 1	
100, 60, 10, 1, 1, 241, 13, 60, 1, 1, 1, 3, 40, 1, 3, 6, 90.	
34 Mon 60 M. 1 & 24 Mon 80 M. 1 2. All 3.	
St. Mor ov W. 1 - 34, War ov M. 1 2. All 3.	
1 St. Mon 60 M. 1 & Sh. W. S. 60 M. K. 1 & Sh. VS	. go. Kr
A St. Mar Co With A St. Mar Co With I a Star	S. On Kr
N. KI Ski Mar Er WI. I - Ski Mar Ear W. I - Ski	
dui ki ski ma en with i chi mai ca with i	Ku a. du
is due the second a second to the design second the	attin a. e
3. dr. Kr st. Ma ser with ski war con with	1 241, 20.
Mar du the state we we du the she was conn	
the second secon	W. W. KIN
3. Kur 3. du Kr St Kur 2. c Wr. A. St. Word	U. W. W. S
1 3. Hur s. du the structure of the structure	en 10. 11. 12
the stand and the standard with stand	0. 60 XN. L
With I start as about the start and so about the start	Mar CU XU.F
South 1 2. Hur 3. On the section 2. Co 40. At 3	
contraction of the state of the	
or contract a state of the second second of the	
us. compile 1 s. the vs. of the service unit	
, us. complete 1 setting a good the setting see the	VI SK M
Shi war con the I a the war go the set the set	du VI SKI