NATIONAL UNIVERSITY OF LIFE AND ENVIRONMENTAL SCIENCES OF UKRAINE

Department of Geoinformatics and Aerospace Research of the Earth

"CONFIRMED"
Dean of the Faculty of Land management
_____ T.O. Ievsiukov
«18» _May_ 2023

"APPROVED"
at the meeting of the department of Geoinformatics
and Aerospace Research of the Earth
Protocol № 11 from 14 April 2023
A.i. head of Department
_____ O. P. Drozdivskyi

"REVIEWED "
Program Coordinator
I.P. Kovalchuk.

PROGRAM OF THE COURSE

MATHEMATICAL STATISTICS FOR GEODATES

Specialization	193. Geodesy and land management
Educational program	"Geodesy and Land Management"
Faculty	Land management
Developers	Dr. Sci., Prof. Kokhan S.S.
	(посада, науковий ступінь, вчене звання)
	Dr. Sci., Associate prof. Koshel A.O.
	(посада, науковий ступінь, вчене звання)

Kyiv – 2023

1. Descreption of the course
MATHEMATICAL STATISTICS FOR GEODATES

Industry knowledge training direction, specialty, education level	
Educational qualification	Bachelor
Specialty	193 Geodesy and Land management
Educational program	Geodesy and Land management
Description of the course	
Type	Sample
Total number of hours	90
Number of credits ECTS	3
Number of content modules	2
Course project (work)	-
Form of Control	Final test
Descriptions of the course for full-time and distance learning	
	Full-time
Year of training	2
Semester	3
Lectures	15
Practical, seminars	-
Laboratory studies	30
Independent work	45
Individual tasks	-
Number of weekly hours for full-time study: classroom	3

2. Task and purpose of the discipline
MATHEMATICAL STATISTICS FOR GEODATES

Objective:

The discipline " Mathematical statistics for geodates " ensures that students master the fundamentals of using mathematical and statistical methods and applied processing of geodata and indicators regarding the state of land relations based on the use of modern computer technologies.

The aim of the discipline

The task of studying the discipline is the formation of a specialist in theoretical knowledge and practical skills of geostatistics, mathematical and statistical apparatus of geospatial data processing methods and their properties. In particular, the tasks use spectral properties of spatial objects, derived products (vegetation indices, leaf surface area) with corresponding cartographic projections, which require constant transformations and their further practical use on the basis of geoinformation technologies in the GIS analysis of data and indicators regarding the state of land relations and land cadastral data.

A study of the discipline the student must

know:

• basics of probability theory;

• main indicators and methods of basic statistical analysis of geodata.

be able to:

   • collect geodata and organize the necessary data;

   • perform the primary organization of geodata in the form of spreadsheets;

   • be able to use mathematical models for the analysis of geodata and indicators of the state of land relations;

   • have basic methods of multidimensional statistics in accordance with the analysis of geodata and indicators on the state of land relations;

• know the possibilities and be able to work with the Statistica package.

The discipline provides the formation of a number of competencies:

- *general competencies*:

• ZK01. Ability to learn and master modern knowledge;

• ZK02. Ability to apply knowledge in practical situations;

• ZK07. Ability to work autonomously;

• ZK12. The ability to exercise one's rights and responsibilities as a member of society; awareness of the value of a civil (free democratic) society and the need for its sustainable development, the rule of law, the rights and freedoms of a person and a citizen in Ukraine;

• ZK13. The ability to preserve and multiply moral, cultural, scientific values and achievements of society based on an understanding of history, the patterns of development of the subject area, its place in the general system of knowledge about nature and society, as well as in the development of society, technology and technologies, to use various types and forms of motor activities for recreation and leading a healthy lifestyle.

- *special competencies*:

• SK01. The ability to apply fundamental knowledge to analyze phenomena of natural and man-made origin when performing professional tasks in the field of geodesy and land management;

• SK02. Ability to apply theories, principles, methods of physical and mathematical, natural, socio-economic, engineering sciences when performing tasks of geodesy and land management;

• SK03. Ability to apply regulatory and legal acts, regulatory and technical documents, reference materials in professional activity;

• SK04. Ability to choose and use effective methods, technologies and equipment for carrying out professional activities in the field of geodesy and land management;

• SK05. The ability to use modern information, technical and technological support to solve complex issues of geodesy and land management;

• SK06. The ability to perform remote, ground, field and camera research, engineering calculations for the processing of research results, form research results, prepare reports when solving geodesy and land management tasks;

• SK07. The ability to collect, update, process, critically evaluate, interpret, store, publish and use geospatial data and metadata about objects of natural and man-made origin.

Program results:

• PH2. Organize and manage the professional development of individuals and groups;

• PH3. Convey information, ideas, problems, solutions, own experience and arguments to specialists and non-specialists;

• PH4. Know and apply in professional activity regulatory and legal acts, regulatory and technical documents, reference materials in the field of geodesy and land management and related fields;

• PH5. Apply conceptual knowledge of natural and socio-economic sciences when performing tasks of geodesy and land management;

• PH9. Collect, evaluate, interpret and use geospatial data, metadata about objects of natural and man-made origin, apply statistical methods of their analysis to solve specialized problems in the field of geodesy and land management.

3. Program and structure of the discipline
MATHEMATICAL STATISTICS FOR GEODATES

full-time study form:

Titles content modules and themes	Hours						
	weeks	total	including				
			l	p	lab	ind	i.w.
1	2	3	4	5	6	7	8
Content module 1. Fundamentals of the probability theory and basics statistical indexes							
Topic 1. The main provisions of probability theory and mathematical statistics	1-2	6	2		4		
Topic 2. Elementary analysis of geodata and indicators on the state of land relations	3-4	6	2		4		
Topic 3. Basic statistical models. Basic statistics of data and indicators on the state of land relations	4-5	6	2		4		
Total by Content module		18	6		12		
Content module 2. Fundamentals of land management geodata statistical analysis							
Topic 4. Methods of statistical evaluation of averages	6-7	14	2		2		10
Topic 5. Statistical models of geodata and indicators on the state of land relations based on samples	8-9	16	2		4		10
Topic 6. Fundamentals of statistical analysis of sample geodata	10-13	16	2		4		10
Topic 7. Multidimensional models of geodata	14-15	26	3		8		15
Total by Content module		72	9		18		45
Total		**90**	**15**		**30**		**45**

4. Themes of seminars

№	Name of theme	Hours

5. Topics of practical classes

№	Name of theme	Hours

6. Topic of laboratory classes

№	Topic	Hours
Content module 1. Introduction. Basic concepts of computer science		
1	Statistical and variation series of indicators of landowners and land users in MS Excel	2
2	Statistical and Variation Series for landowners and land users in STATISTICA	2
3	Basic statistics of land management indicators	2
4	Laws of the distribution of variables of random variables on an example of land valuation	2
5	Formation of sampling of land-use geodata from general collections	2
6	Criterion of reliability of estimates of cadastral geodata	2
7	Test hypotheses about distribution laws	2
8	Correlation analysis of land management geodata	2
9	Regression analysis of land management geodata	2
10	ANOVA of land management geodata	2
11	MANOVA of land management geodata	2
12	Cluster analysis of land management geodata	2
13	Discriminant analysis of land management geodata	2
14	Factor analysis of land management data and Principal component analysis	2
15	Multidimensional scaling of land management geodata	2

7. Topics of independent work

№	Topic name	Number of hours
1	The basic concept of geostatistics in the field of land relations	9
2	Basic aspects of using the R language for statistical analysis	10
3	Using a decision tree for geodata processing	8
4	ArcGIS software product for statistical analysis of geospatial data	8
5	Mathematical and statistical methods of geospatial data processing	10

8. Individual tasks

1. Get the task with geodata of the land resources characteristics in the Kyiv region

2. Sort and range the geodata for performance by years.

3. Compute the geodata frequencies.

4. Build a distribution bar chart. Conduct comparative analysis of the distribution of frequencies as compared by the years.

5. Calculate the values of skewness and kurtosis and conclude that the nature of the variation curve.

6. Test the hypothesis of normality by Kolmogorov-Smirnov and Shapiro-Wilks tests. Make a conclusion about the normality.

7. Calculate the frequency distribution parameters for different distribution laws (with parameters mismatch normality).

8. Test the hypothesis match the actual frequency distribution of the calculated frequencies for different laws (by chi-square, «Observed versus expected X») and conclude that statistically significant distribution law.

9. Calculate basic statistical characteristics of market indicators earth. Conduct comparative analysis of basic statistics for years.

10. Test the hypothesis of equality of variances of geodata by years. Make a conclusion about the equality of variances.

11. Check for the hypothesis of equality of mean values and belonging to a samples of the population. (Note if the distribution does not match normality – used non-parametric criteria). Make a conclusion about where the geodata parameters for all the years to same population.

EXAMPLE OF TEST TASKS FROM THE DISCIPLINE "MATHEMATICAL STATISTICS FOR GEODATES"

Faculty: land management
Specialty: geodesy and land management
The form of study is full-time
Semester 3 Course 2
EQ "Bachelor"
Department of geoinformatics and aerospace research of the Earth
Discipline "Mathematical statistics for geodates"
Lecturer _____ A. Koshel
"Approve"
Chief of the department _____ S. Kokhan
_____2023 y.

Variant № 1

Question 1. «_____» – one of the general theories of the universe that generalizes the notion of randomness. In general, randomness is the result of the simultaneous action on an object or process of a set of partially independent factors.

Enter the correct answer in the form

Question 2. A set of relatively homogeneous but individually different observation units united within the framework of a research program is called:

1. Variation series
2. Multiple values
3. A statistical population
4. A set of numbers

Question 3. A characteristic property of a feature is "_____" of the value of the feature within certain limits.

Enter the correct answer in the form

Question 4. The distribution center around which all variants of the studied population are grouped.

1. Arithmetic average
2. Average harmonic
3. Mean squared
4. Geometric mean

Question 5. A probabilistic model explains the behavior of the "_____" set.

Enter the correct answer in the form

Question 6. Give the name of the coefficient calculated by the formula:

$$t_{As} = \frac{As}{S_{As}} = \frac{As}{\sqrt{6/(n+3)}}$$

Question 7. It is used when it is necessary to characterize the indicators of the area.

1. Arithmetic average
2. Average harmonic
3. Mean squared
4. Geometric mean

Question 8. The relative indicator of the deviation of the actual

Question 9. The indicator that is used when it is necessary to compare the characteristics of the variation of different aggregates, which can be expressed in different units:

1. Mean square deviation
2. Coefficient of variation
3. Variance
4. Standard deviation

Question 10. If a random variable can take any numerical value, it is called:

1. Standard
2. Discrete
3. Normal
4. Continuous

Question 11. An uncontroversial assumption about the distribution of a random variable is:

1. Statistical hypothesis
2. Variance
3. Sample
4. Event

Question 12. The "_____" distribution is a subset of the binomial distribution used to analyze rare random events.

Enter the correct answer in the form

Question 13. Arrange in the correct order

1. Quantiles	А. Quantities that divide the ranked variational series into ten equal parts
2. Quartiles	Б. Values that cut off a certain part of its members within a series.
3. Deciles	В. Quantities that divide the ranked variation series into 100 equal parts
4. Percentiles	Г. Three quantities (Q1, Q2, Q3) that divide the ranked variation series into four equal parts

Question 14. The average, relative to which the ranked variation series is divided into two equal parts, that is, the same number of variants is located on both sides.

1. Mode
2. Average harmonic
3. Median
4. Geometric mean

Question 15. With a complete coincidence of theoretical and actual frequencies χ2 =

1. 10
2. 1
3. 100
4. 0

Question 16. The set of conditions necessary to obtain data about the object's behavior is?

Enter the correct answer in the form

Question 17. If, when the conditions of the test are fulfilled, events cannot be known to be received, they are called

variant from the mean value or the measure of dispersion of the characteristic is "_____"
Enter the correct answer in the form

Question 18. The sum of the return values of an option divided by their number.
1. Arithmetic average
2. Average harmonic
3. Mean squared
4. Geometric mean

Question 19. Arrange in the correct order	
1. Van der Waerden's X-test	А. It is used to test the null hypothesis that two independent samples belong to the same population.
2. z-criterion of signs	Б. It is used for the analysis of dependent samples, when the values of their variants are connected in pairs by some common conditions.
3. Mann-Whitney U-test	В. It is used to test the null hypothesis when comparing independent samples.
4. Wilcoxon's t-test	Г. It is used if the research results have only two results, which can be expressed in the form of + and - signs.

Question 20. Give the name of the coefficient calculated by the formula:
$$t_{Ex} = \frac{Ex}{S_{Ex}} = \frac{Ex}{2\sqrt{6/(n+5)}}$$

Question 21. Most often, the Kolmogorov-Smirnov test is used to check the sample on
1. Significance
2. Admissibility
3. Normality
4. Volume

Question 22. If the result of the test can be only one and it is predictable - the events are called "_____".
Enter the correct answer in the form

Question 23. If there can be more than one outcome of an event and it is impossible to predict it a priori, the event is called "_____".
Enter the correct answer in the form

«_____».
Enter the correct answer in the form

Question 24. The value of the χ2 criterion is always:
1. Positive
2. Negative
3. Inconsequential
4. Continuous

Question 25. Variation series are:
1. Equal intervals
2. Unequally spaced
3. Intervalless
4. Arbitrary

Question 26. By comparing the ideal representation with the real one for any object, you can get a certain estimate of the real object, which is called:
1. Theory of probability
2. Statistics
3. Hypothesis
4. Event

Question 27. «_____»is some numerical measure of the objective possibility of the occurrence of an event in a single trial.
Enter the correct answer in the form

Question 28. A probability that can be determined before a test is called "_____".
Enter the correct answer in the form

Question 29. Each individual numerical value of the characteristic is called «_____» .
Enter the correct answer in the form

Question 30. The value that occurs most often in this population.
1. Mode
2. Structural average
3. Median
4. Geometric mean

9. Teaching Methods

In conducting lectures appropriate to use verbal teaching methods: explanation, narration, discussion, educational debate, with a combination of visual learning methods: illustration, showing.

In carrying out laboratory work should be used such as verbal learning method of instruction on the combination of visual learning methods of illustration and demonstration, the aspect of these studies is that they facilitate communication theory and practice, providing students acquiring skills using standard and specialized software , application of information technology to cadastral and form students' initial skills of research activities. Laboratory work in the laboratory are equipped computers.

10. Forms of control

Content module number	Division of discipline	The topic of the lecture	Topic of practical (laboratory) lesson	Form of knowledge control
I	Basics of probability theory and basic statistical indicators	Topic 1. Basic principles of probability theory and mathematical statistics	Statistical and variational series of indicators of areas of landowners and land users in MS Excel	Protection of laboratory work / assessment of modular control work
		Topic 2. Elementary analysis of geodata and indicators regarding the state of land relations	Statistical and variational series of indicators of areas of landowners and land users in STATISTICA	
			Basic statistics of the main indicators for geospatial data	
			Laws of distribution of variable random variables on the example of land valuation	
			Formation of samples of geodata and indicators regarding the state of land relations from general aggregates	
		Topic 3. Basic statistical models. Basic statistics of data and indicators regarding the state of land relations	Criteria for reliability of assessments of cadastral geodata	
			Testing hypotheses about distribution laws	
			Correlational analysis of geodata and indicators regarding the state of land relations	
II	Basics of statistical analysis of land management geodata	Topic 4. Methods of statistical estimation of average values	Regression analysis of geodata and indicators regarding the state of land relations	Protection of laboratory work / assessment of modular control work
		Topic 5. Statistical models of geodata and indicators regarding the state of land relations based on samples	One-factor dispersion analysis of geodata and indicators regarding the state of land relations	
			Multifactor dispersion analysis of geodata and indicators regarding the state of land relations	
		Topic 6. Fundamentals of statistical analysis of sample geodata	Cluster analysis of geodata and indicators regarding the state of land relations	
			Discriminant analysis of geodata and indicators regarding the state of land relations	
		Topic 7. Multidimensional geodata models	Factor analysis of cadastral geodata and the method of principal components	

			Multidimensional scaling of geodata and indicators regarding the state of land relations	

The main methods of control of knowledge and skills students have to study the subject are: oral examination, written and practical test, standardized control in the form of modular test papers, assessment for individual learning task, the final test.

The total value of these methods is to make the best possible to ensure timely and comprehensive feedback between students and teachers, by which establishes how students perceive and learn the material.

The purpose determines the choice of control methods, it should be borne in mind that these methods can be applied in all kinds of control - only complete applications allows regularly and objectively identify the dynamics of the formation of knowledge and skills of students. Each control method has its advantages and disadvantages, scope of application, none of them can not be the only one able to diagnose all aspects of the learning process. So:

-        to control the absorption of lectures: oral questioning, written modular test papers, current testing score for an individual learning task, the final test.

-        for the monitoring and evaluation of laboratory work: practical test and evaluation of each laboratory work.

11. Ratio between national and ECTS values and rating from the discipline

The student's knowledge is assessed on a 100-point scale and translated into national assessments according to the table. 1 "Regulations on Examinations and Credits at NUBiP of Ukraine" (from April 26, 2023, protocol No. 10)

Student rating, points	The assessment is national for the results of the assembly	
	exams	credits
90-100	excellent	Enrolled
74-89	Good	Enrolled
60-73	Satisfactory	Enrolled
0-59	Unsatisfactorily	Not counted

To determine the student's (student's) rating for mastering the RDYS discipline (up to 100 points), the obtained rating from the certification (up to 30 points) is added to the student's (student's) rating for the RHP educational work (up to 70 points): R ДИС = R HP + R AT .

12. Methodological Support

1. Methodological guidelines "Statistical methods in land management". A.O. Koshel, O.I. Novikov. - Kyiv, 2018. - 44 p.

13. Suggested Reading

Basic

1. Viktor Barkovskyi, Nina Barkovskaya, Oleksiy Lopatin. Probability theory and mathematical statistics. Center for educational literature. 2019. 494 p.
2. Probability, Random Variables, Statistics, and Random Processes: Fundamentals & Applications. Wiles and Sons Inc. 2019. 416 p.
3. Matkovskyi C.O., Halkiv L.I., Hrynkevich O.S., Sorochak O.Z. Statistics: Study guide - Lviv.: "Novyi Svit", 2009. - 430 p.
4. Fotheringham, A.S., Brunsdon, C. and Charlton, M., 2003. Geographically weighted regression: the analysis of spatially varying relationships. John Wiley & Sons.
5. Rogerson, P.A., 2014. Statistical methods for geography: a student's guide. Sage.
6. Schabenberger, O. and Gotway, C.A., 2004. Statistical methods for spatial data analysis. CRC press.
7. Kang-Tsung Chang (2015) Introduction to Geographic Information Systems, Eighth Edition, McGraw Hill.
8. Deterministic and Geostatistical Interpolation Methods sections, Geospatial Analysis: Web site, M. J. de Smith, M. F. Goodchild, P. A. Longley. 2021. URL: http://www.spatialanalysisonline.com/HTML/index.html.

Support

9. Steven S. Skiena. The Data Science Design Manual. Computer Science Department Stony Brook University. 2017. 453 p.
10. Senyo P.S. Probability theory and mathematical statistics. - K.: Center of educational literature, 2004.
11. Adamov, V.E. Economics and statistics: Textbook. Ed. S.D. Ilyenkova. M.: Finances and Statistics, 2019. 287p.

13. Information Resources

12. Державна служба статистики України. URL: www.ukrstat.gov.ua
13. Державна служба України з питань геодезії, картографії та кадастру. URL: https://land.gov.ua/
14. Eurostat. URL: https://ec.europa.eu/eurostat