



MODULE DESCRIPTION

Module title	Module code
Statistical data analysis	

Lecturer(s)	Department where the module is delivered
Coordinator: dr. Rūta Levulienė	Department of Mathematical Statistics Faculty of Mathematics and Informatics Vilnius University
Other lecturers:	

Cycle	Type of the module
First	Optional

Mode of delivery	Semester or period when the module is delivered	Language of instruction
Face-to-face	5 th or 7 th semester	Lithuanian

Prerequisites
Prerequisites: Mathematics for Software Engineering I, II, III

Number of credits allocated	Student's workload	Contact hours	Self-study hours
5	130	68	62

Purpose of the module: programme competences to be developed		
Purpose of the module – to give knowledge of statistical data analysis (concepts and methods), to develop capabilities for choosing statistical methods, to do analysis using statistical packages (R, SAS), to interpret results.		
Generic competences: <ul style="list-style-type: none"> • Communication and collaboration (<i>GK1</i>). • Life-long learning (<i>GK2</i>). 		
Specific competences: <ul style="list-style-type: none"> • Knowledge and skills of underlying conceptual basis (<i>SK4</i>). • Technological and methodological knowledge and skills, professional competence (<i>SK6</i>). 		
Learning outcomes of the module: students will be able to	Teaching and learning methods	Assessment methods
Collect and prepare statistical data for analysis. Know types of statistical data and correctly choose statistical techniques. Perform initial statistical data analysis: summary statistics, graphical methods. Apply methods of mathematical statistics: - Estimate parameters, test parametric hypotheses. - Test nonparametric goodness-of-fit independency and homogeneity hypotheses. - Do analysis of variance, regression and correlation analysis. - Perform cluster, discriminant and factor analysis. Analyse data using statistical packages: R, SAS.	Problem-oriented teaching, analysis of examples, information retrieval, individual reading, laboratory works	Exam (written). Laboratory works using statistical software (R, SAS), results presentation.

Content: breakdown of the topics	Contact hours						Self-study work: time and assignments		Assignments
	Lectures	Tutorials	Seminars	Practice	Laboratory work (LW)	Tutorial during LW	Contact hours	Self-study hours	
1. Basic concepts of statistical data analysis: problems and methods. Statistical packages: main features, comparisons.	2				6	8	8	6	Individual reading; problem solving, laboratory works
2. Special functions (probability, quantile, random number generation).	1				2		3	4	
3. Empirical characteristics and their properties.	3				4		7	4	
4. Methods of parameter estimation (point estimators, confidence intervals).	4				4		8	4	
5. Parametric hypotheses testing.	4				4		7	6	
6. Nonparametric hypotheses testing	5				4		10	8	
7. Analysis of variance	4				2		6	7	
8. Regression and correlation analysis	5				4		9	7	
9. Multivariate data analysis: cluster, discriminant and factor analysis.	4				2		6	6	
10. Tutorials before exam		2						2	
11. Preparing for the exam and taking the final exam (written).							2	10	Self-study of literature. Self-control tasks.
Total	32	2			32	8	68	62	

Assessment strategy	Weight %	Deadline	Assessment criteria
Laboratory work	40	During the semester	During the laboratory work in the computer class students have to perform statistical analysis using statistical software (R, SAS) and interpret results of computations. The maximum number of points is given for complete and correct analysis and interpretation of results. The partial implementation of the tasks proportionally reduces the assessment.
Exam (written)	60	Exam session	The exam consists of 12 exercises. The assessment of each is 0.5 points.

Author	Publishing year	Title	Number or volume	Publisher or URL
Required reading				
Čekanavičius V., Murauskas G.	2000, 2002	Statistics and its Applications	I, II	TEV, Vilnius
Kruopis J.	1993	Mathematical Statistics		Mokslas, Vilnius
Levulienė R.	2009	Applied Statistics using SAS		VUL, Vilnius
Recommended reading				
Everitt B.S., Hothorn	2010	A Handbook of Statistical Analyses using R		Chapman & Hall / CRC
Kleinman K., Horton N. J.	2010	SAS and R. Data Management, Statistical Analysis and Graphics		Chapman & Hall / CRC
Johnson R.A.	2002	Applied Multivariate Statistical Analysis		Prentice Hall, NJ