

Subject code	ECTS credits
MAT5010	6

Course title in Lithuanian

DISKREČIOJI STATISTIKA

Course title in English

DISCRETE STATISTICS

Short course annotation in Lithuanian (up to 500 characters)

Šio kurso tikslas yra suteikti studentams teorinių ir praktinių žinių, reikalingų tiriant statistines priklausomybes tarp kategorinių kintamųjų. Kursas apima statistinių sprendimų teorijos pagrindus, klasifikavimo uždavinius, diskretaus pasirinkimo modelius, logtiesinę analizę, grafinius modelius ir jų taikymus.

Short course annotation in English (up to 500 characters)

Course objective – introduce basic theory and statistical methods for investigation of statistical dependences between categorical variables. The course covers the following main topics: basics of statistical decision theory; classification and clusterization; discrete choice modelling; loglinear analysis; graphical models and their application.

Prerequisites for entering the course

Probability Theory, Mathematical Statistics.

Course aim

Course aim is to provide deeper knowledge of decision theory, classification methods, and discrete statistical methods, develop students' skills in analytical thinking.

Links between study programme outcomes, course outcomes, criteria of learning achievement evaluation, study methods and methods of learning achievement assessment

Course outcomes	Criteria of learning achievement evaluation	Study methods	Methods of learning achievement assessment
1. Make decisions under uncertainty	Student demonstrates the ability to choose loss function and prior for Bayes risk, find optimal estimator and decision rule	Lectures, practical works, individual work,	Mid-term exam
2. Develop discrete choice model.	Student demonstrates the ability to estimate, validate and interpret discrete choice models	Lectures, practical works, individual work,	Mid-term exam
3. Apply statistical decision methods to classification problem	Student demonstrates the ability to identify risk function and choose the best classification rule	Lectures, practical works, individual work,	Exam
4. Identify and validate dependence structure between categorical variables and test their significance	Student demonstrates the ability to find relationships between categorical variables and test their significance, perform loglinear analysis	Lectures, practical works, individual work	Exam
5. Present report of performed study	Student demonstrates the ability to formulate task, present solution process, justify received results	Individual work, self-study of literature, discussions, consulting	Essay presentations

Links between study programme outcomes and course outcomes

Study programme outcomes	Running number of course outcome				
	1	2	3	4	5
1. Deepen and expand general knowledge of mathematics and apply it in a new non-standard environment	+	+	+	+	
3. Broaden and apply the knowledge of reliability analysis and statistical methods for data analysis	+	+	+	+	

4. Identify, select and understand the state-of-the-art literature of mathematics and apply the gained knowledge to specific scientific and practical tasks	+		+		
5. Develop mathematical models integrating the knowledge from various fields and different mathematical modelling techniques, and analyse the modelling results assessing the model adequacy and accuracy		+	+	+	+
7. Analyse, understand and use mathematical methods	+	+	+	+	+
11. Convey mathematical information to specialists of different fields orally and/or in written form, critically evaluate it		+		+	+
13. Take moral responsibility for the results of work					+

Content

No	Content (topics)
1.	Elements of statistical decision theory.
2.	Classification.
3.	Discriminant and cluster analysis.
4.	Discrete choice modelling in econometrics.
5.	Logistic regression.
6.	The generalized logit model.
7.	Loglinear analysis.
8.	Graphical models, application in linguistics and genetics.

Distribution of workload for students (contact and independent work hours)

Lectures	45 hours
Practical work	15 hours
Individual students work	100 hours
Total:	160 hours

Structure of cumulative score and value of its constituent parts

Final written exam (50%), mid-term written exam (25%), and assessments of homework work (25%).

Recommended reference materials

No	Publication year	Authors of publication and title	Publishing house	Number of copies in		
				University library	Self-study rooms	Other libraries
Basic materials						
1.	2002	V.Čekanavičius, G.Murauskas. <i>Statistika ir jos taikymai</i> . II dalis (Statistics and its application. II part.)	TEV	30	1	
2.	2009	V.Čekanavičius, G.Murauskas. <i>Statistika ir jos taikymai</i> . III dalis (Statistics and its application. III part.)	TEV	37	1	
Supplementary materials						
1.	2002	A.Agresti. <i>Categorical Data Analysis</i> .	Wiley & Sons	https://mathdept.iut.ac.ir/sites/mathdept.iut.ac.ir/files/AGRESTI.PDF		
2.	2004	M.A.T. Figueiredo. <i>Lecture Notes on Bayesian Estimation and Classification</i>		http://www.lx.it.pt/~mtf/learning/Bayes_lecture_notes.pdf		
3.	2008	M.J.Wainwright, M.I.Jordan. <i>Graphical models, exponential families, and variational inference</i> . Foundations and		https://people.eecs.berkeley.edu/~wainwrig/Papers/WaiJor08_FTML.pdf		

		Trends in Machine Learning. Vol. 1, Nos. 1–2, pp. 1–305		
--	--	---	--	--

Course programme designed by

Prof.dr. Marijus Radavičius