

DOKTORANTŪROS STUDIJŲ DALYKO APRAŠAS

Dalyko pavadinimas	Mokslo kryptis (šaka) kodas	Fakultetas (institutas)	Katedra, skyrius
Statistinė duomenų analizė	Matematika	Matematikos ir informatikos	Ekonometrinės analizės katedra, Tikimybių teorijos ir statistikos skyrius
	01P		
Studijų būdas	Kreditų skaičius	Studijų būdas	Kreditų skaičius
Paskaitos	0	Konsultacijos	1
Individualus	4	Seminarai	0
Dalyko anotacija			
<p>Statistinės duomenų analizė plačiai taikoma mokslo tyrimuose, valstybės valdyme, versle. Sparčiai vystantis kompiuterinėms ir informacinėms technologijoms jos vaidmuo vis auga.</p> <p>Nors šis modulis yra skirtas matematikams, jo adaptuotas ir orientuotas į technologijas variantas bus naudingas ir kitų mokslo sričių doktorantams.</p> <p>Pagrindinės temos:</p> <ol style="list-style-type: none"> 1. Tikimybiniai skirstiniai ir statistiniai modeliai. Eksponentinė skirstinių šeima. 2. Imčių metodai. Imtys iš baigtinių populiacijų. Saviranka. Monte Karlo metodai ir modeliavimas su kompiuteriu. 3. Robastinės statistikos. 4. Bayeso statistika. Apriorinių skirstinių parinkimas. Empirinis Bayeso metodas. 5. Statistinė sprendimų teorija. 6. Klasikinės daugiamatės duomenų analizės modeliai ir metodai: regresija, klasifikavimas, klasterizacija, pagrindinių komponentų metodas, faktorinė analizė. 7. Tiriamaoji duomenų analizė ir vizualizacija. Glodinimas ir vidurkinimas. 8. Duomenų dimensijos sumažinimo problema, tikslingasis projektavimas. 9. Aproximacijų tikslumas ir statistinės paklaidos. Skleidiniai bazinėmis funkcijomis, splainai. 10. Branduolio metodai. Reproduktojančio branduolio Hilberto erdvės. 11. Neparimetriniai ir adaptyvūs metodai. 12. Modelio parinkimo problema. Informaciniai kriterijai. Kryžminio patikrinimo metodas. 13. Apibendrintieji tiesiniai ir mišrieji modeliai. Skirstinių mišiniai. Latentiniai kintamieji. 14. Kategorinių duomenų analizė. Logtiesiniai modeliai. 15. Modelių ir metodų jungimas ir kombinavimas, sudėtiniai modeliai ir metodai. 16. Grafiniai modeliai. Statistinė priežasčių analizė. 17. Duomenų tipai ir struktūros. 18. Duomenų analizės technologijos. Atraminų vektorių mašinos. Sprendimų ir regresiniai medžiai, atsitiktiniai miškai. MARS. Neuroniniai tinklai. AdaBoost, LASSO, asociacijos taisyklių paieška. 19. Duomenų analizės algoritmai ir programinė įranga. R paketas. Duomenų analizės taikymai. 			
Pagrindinė literatūra			
Hastie T., Tibshirani R., and Friedman J. <i>The Elements of Statistical Learning: Data Mining, Inference, and Prediction</i> , 2nd ed., New York: Springer, 2009. http://www-stat.stanford.edu/~tibs/ElemStatLearn/			
Wasserman L. <i>All of Statistics: A Concise Course in Statistical Inference</i> , Springer Texts in Statistics, 2004. ISBN: 978-1-4419-2322-6 (Print) 978-0-387-21736-9 (Online). http://www.bioinfo.org.cn/~wangchao/maa/w.statistic.pdf http://read.pudn.com/downloads158/ebook/702714/Larry%20Wasserman_ALL%20OF%20Statistics.pdf			
Zaki, M.J. and Wagner, M.JR. <i>Data mining and analysis</i> , Cambridge University Press, 2014. ISBN 978-0-521-76633-3 Hardback.			
Papildoma literatūra			
Bishop, Christopher. <i>Pattern Recognition and Machine Learning</i> , Singapore, Springer, 2006. http://users.isr.ist.utl.pt/~wurmd/Livros/school/Bishop%20-%20Pattern%20Recognition%20And%20Machine%20Learning%20-%20Springer%20%20202006.pdf			
Ledolter, Johannes. <i>Data mining and business analytics with R</i> , Hoboken, New Jersey: John Wiley & Sons, Inc.			
Čekanavičius V., Murauskas G. <i>Statistika ir jos taikymai</i> . I, II, III, Vilnius: TEV, 2000, 2002, 2009.			

Konsultuojančiųjų dėstytojų vardas, pavardė	Mokslo laipsnis	Pedag. vardas	Svarbiausieji darbai mokslo kryptyje (šakoje) paskelbti per pastaruosius 5 metus
Gediminas Murauskas	Dr.	Doc.	<ol style="list-style-type: none"> 1. Kregždė, Arvydas; Murauskas, Gediminas. Analysis of Lithuanian credit default swaps // Journal of business, economics and management / Vilniaus Gedimino technikos universitetas. Vilnius, London: Technika, Taylor & Francis. ISSN: 1611-1699, eISSN: 2029-4433. 2015, Vol. 16, no. 5, p. 916-930. 2. Kregždė, Arvydas; Murauskas, Gediminas. Analysing sovereign Credit Default Swaps of Baltic countries (Baltijos šalių kredito apsikeitimo sandorių analizė) // Verslas : teorija ir praktika (Business: Theory and Practice) / Vilniaus Gedimino technikos universitetas. Vilnius: Technika. ISSN: 1648-0627, eISSN: 1822-4202. 2015, Vol. 16, No. 2, p. 121-131.
Marijus Radavičius	Dr. HP	Doc.	<ol style="list-style-type: none"> 1. Lapėnaitė-Gedvilė, Monika; Piaseckienė, Karolina; Radavičius, Marijus. Tekstų nehomogeniškumo tyrimas naudojant žymeklius = Analysis of text non-homogeneity using markers) // Lietuvos statistikos darbai. Lietuvos statistikos departamentas. 2015, 54 t., Nr. 1, 92–100 p. 2. Piaseckienė, K.; Radavičius, M. Empirical Bayes estimators of structural distribution of words in Lithuanian texts. // Nonlinear Analysis: Modelling and Control / ISSN: 1392-5113. 2014, Vol. 19, no. 4, p. 611-625. 3. Radavičius, M.; Samusenko, P. Goodness-of-fit tests for sparse nominal data based on grouping // Nonlinear Analysis: Modelling and Control / ISSN: 1392-5113. 2012, Vol. 17, no. 4, p. 489–501.
Rimantas Rudzkis	Habil. dr.	Prof.	<ol style="list-style-type: none"> 1. Rudzkis, R.; Valkavičienė, R. 2014. Econometric models of the impact of macroeconomic processes on the stock market in the baltic countries. <i>Technological and Economic Development of Economy</i>, 20(4), pp. 696–713, ISSN 2029–4913. 2. Bakshaev, A.; Rudzkis, R. 2014. Multivariate goodness of fit tests based on the kernel density estimators. <i>Nonlinear Analysis, Modelling and Control</i>, Vol. 20, No 4, pp. 585-602, ISSN 1392-5113. 3. Rudzkis, R.; Bakshaev, A. 2012. Probabilities of high excursions of Gaussian fields. <i>Lithuanian Mathematical Journal</i> 52(2): 196-213. ISSN 0363-1672 (print), ISSN 1573-8825 (online).

Patvirtinta Matematikos ir informatikos fakulteto taryboje 2017 m. kovo 14 d., protokolo Nr. 5

Fakulteto tarybos pirmininkas prof. habil. dr. Mindaugas Bloznelis