

UČNI NAČRT PREDMETA / COURSE SYLLABUS

Predmet: Analiza kategorialnih podatkov
Course title: Categorical data analysis

Študijski program in stopnja Study programme and level	Študijska smer Study field	Letnik Academic year	Semester Semester
Podatkovne znanosti, magistrski študijski program druge stopnje	-	Prvi	Drugi
The second cycle masters study programme Data Sciences	-	First	Second

Vrsta predmeta / Course type

Obvezni / Obligatory

Univerzitetna koda predmeta / University course code:

2-PZ-MAG-AKP-2024-01-31

Predavanja Lectures	Seminar Seminar	Vaje Tutorial	Klinične vaje work	Druge oblike študija	Samost. delo Individ. work	ECTS
30	-	30	-	-	90	5

Nosilec predmeta / Lecturer: Doc. dr. Nuša Erman

**Jeziki /
Languages:**

**Predavanja /
Lectures:** Slovenski, angleški / Slovene, English
Vaje / Tutorial: Slovenski, angleški / Slovene, English

**Pogoji za vključitev v delo oz. za
opravljanje študijskih obveznosti:**

Pogoj za vključitev v delo je absolvirano znanje pri predmetu Izbrana poglavja iz verjetnosti in statistike.

Pogoj za pristop k pisnemu izpitu je pravočasna oddaja in pozitivno ocenjena seminarska naloga.

Prerequisites:

Knowledge obtained at Selected topics in probability and statistics is required.

Student has to submit seminar work within the due time. If the seminar work is positively graded, he/she is allowed to write the exam.

Vsebina:

Content (Syllabus outline):

- Uvod v analizo kategorialnih podatkov: kategorialni podatki, verjetnostne porazdelitve kategorialnih podatkov, inferenčna statistika diskretnih podatkov.
- Kontingenčne tabele: verjetnostna struktura kontingenčnih tabel, primerjava deležev v 2x2 tabelah, testi neodvisnosti, natančno (eksaktno) sklepanje, razširitev na tri- in več-dimenzionalne tabele.
- Razmerja obetov v 2 x 2 in I x J tabelah
- Analiza 2^k tabel
- Loglinearni modeli: definicija, interpretacija, ocenjevanje.
- Enostavna korespondenčna analiza.
- Multipla korespondenčna analiza.

- Introduction to categorical data analysis: categorical data, probability distributions for categorical data, statistical inference for discrete data.
- Contingency tables: probability structure for contingency tables, comparing proportions with 2x2 tables, tests for independence, exact inference, extension to three-way and multi-dimensional tables.
- Odds ratios in 2 x 2 and I x J tables.
- Analysis of 2^k tables.
- Loglinear models: definition, interpretation, estimation.
- Simple correspondence analysis.
- Multiple correspondence analysis.

Temeljni literatura in viri / Readings:

- Agresti, A. (2019). *An Introduction to Categorical Data Analysis, 3rd Edition*. Hoboken, NJ: John Wiley & Sons.
- Rudas, T. (2018). *Lectures on Categorical Data Analysis*. New York: Springer.
- Greenacre, M.J. (2021). *Correspondence analysis in practice, 3rd Edition*. Boca Raton: CRC Press.
- Erman, N.: Analiza kategorialnih podatkov – prosojnice s predavanj in gradiva iz vaj, FIŠ, Moodle.

Cilji in kompetence:

Učna enota prispeva k razvoju naslednjih splošnih in predmetno specifičnih kompetenc:

Splošne kompetence:

- Sposobnost skrbeti za kakovost strokovnega dela skozi avtonomnost, samoiniciativnost, (samo)kritičnost, (samo)refleksivnost in (samo)evalviranje.
- Sposobnost interpretacije rezultatov podatkovne analize.
- Sposobnost analitičnega in algoritmičnega razmišljanja.
- Zmožnost artikulacije raziskovalnega problema in na tej podlagi sposobnost pridobivanja, selekcije, ocenjevanja in umeščanja novih informacij.
- Sposobnost fleksibilne uporabe znanja v praksi.

Objectives and competences:

The instructional unit contributes to the development of the following general and subject-specific competences:

General competences:

- The ability to manage quality of professional work through autonomy, initiative, as well as (self-)criticism, (self-)reflection and (self-)evaluation.
- The ability to interpret the results of data analysis.
- The ability of analytical and algorithmic thinking.
- The ability to articulate the research problem and correspondingly, obtain, select, evaluate and embed the new information.
- The ability of flexible usage of knowledge in practice.

Subject-specific competences:

Predmetno-specifične kompetence:

- obvladanje raziskovalnih metod, postopkov in procesov
- poznavanje osnovnih in zahtevnih metod analize kategorialnih podatkov

- competence in research methods, procedures and processes
- familiarity with basic and advanced methods for categorical data analysis

Predvideni študijski rezultati:

Znanje in razumevanje:

Sposobnost študenta/študentke bo:

- poznavanje osnovnih in zahtevnejših pristopov k analizi kategorialnih podatkov;
- poznavanje in kritično vrednotenje ustreznosti uporabljenih metod za reševanje praktičnih problemov v analizi kategorialnih podatkov;
- poznavanje oblikovanja ustreznih vsebinskih interpretacij rezultatov.

Intended learning outcomes:

Knowledge and understanding:

The ability of the student:

- knowledge of basic and advanced approaches to categorical data analysis;
- knowledge and critical evaluation of the appropriateness of the methods used to solve practical problems in the categorical data analysis;
- knowledge to draw relevant content interpretation of the results.

Metode poučevanja in učenja:

- *predavanja* z aktivno udeležbo študentov (razlaga, diskusija, vprašanja, primeri, reševanje problemov)
- *vaje*, kjer študentje na enostavnih primerih ponovijo temeljne koncepte in metode, predstavljene na predavanjih
- *laboratorijske vaje*, kjer se študenti seznanijo s programskimi orodji za zbiranje in analiziranje podatkov.

Learning and teaching methods:

- *lectures* with active students participation (explanations, discussion, questions, examples, problem solving);
- *tutorials* (students will recall, reinforce, and shed light on the concepts and methods taught on lectures);
- *lab work* (students will learn state of the art software for data collection and analysis).

Načini ocenjevanja:

Delež (v %) /

Weight (in %)

Assessment:

Način:

- pisni izpit
- seminarska naloga

60
40

Type:

- written exam
- seminar work

Reference nosilca / Lecturer's references:

- ERMAN, Nuša, GOLOB, Tea, JELOVAC, Dejan, RAKOVEC, Primož. The impact of internal dialogue on aggressive driving. *The social sciences*, ISSN 1993-6125. [Online ed.], 2020, vol. 15, iss. 3, str. 119-127
- ERMAN, Nuša, TODOROVSKI, Ljupčo. The effects of measurement error in case of scientific network analysis. *Scientometrics*, aug. 2015, vol. 104, iss. 2, str. 453-473.

- ERMAN, Nuša. Izbrani vidiki proučevanja znanstvenih omrežij : teorija in praksa. 1. izd. Ljubljana: Vega, 2015. 103 str.
- ERMAN, Nuša, KOROŠEC, Aleš, SUKLAN, Jana. Performance of selected agglomerative hierarchical clustering methods. *Innovative issues and approaches in social sciences*, Jan. 2015, vol. 8, no. 1, str. 180-204.
- ERMAN, Nuša, TODOROVSKI, Ljupčo, JEREB, Berta. Late somatic sequelae after treatment of childhood cancer in Slovenia. *BMC research notes*, May 2012, vol. 5, no. 254, str. [1-19].