

UČNI NAČRT PREDMETA / COURSE SYLLABUS

Predmet:	Statistična analiza bioloških podatkov
Course title:	Statistical analysis of biological data

Študijski program in stopnja Study programme and level	Študijska smer Study field	Letnik Academic year	Semester Semester
Interdisciplinarni doktorski študijski program BIOZNANOSTI 3. stopnja	Biologija	1,2	1,2,3,4
Interdisciplinary Doctoral Study Programme in BIOSCIENCES 3rd cycle	Biology	1,2	1,2,3,4

Vrsta predmeta / Course type Izbirni predmet / Elective subject

Univerzitetna koda predmeta / University course code:

Predavanja Lectures	Seminar Seminar	Vaje Tutorial	Klinične vaje work	Druge oblike študija	Samost. delo Individ. work	ECTS
10	10	20	/	5	80	5

Nosilec predmeta / Lecturer: Nosilec: prof. dr. Andrej Blejec

Jeziki / Languages:	Predavanja / Lectures:	slovenski Slovene
	Vaje / Tutorial:	slovenski Slovene

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

splošni pogoji za vpis na doktorski študij in pridobljenih vsaj 3 do 5 KT iz osnov statistike na predhodno končanih študijih.

General requirements. In addition, At least 3 to 5 ECTS gained in basic statistics in previously completed studies.

Vsebina: **Content (Syllabus outline):**

1. Pregled osnovnih statističnih metod in njihova uporaba za analizo podatkov. Statistično preskušanje domnev. Metode proučevanja odvisnosti pojavov.

2. Osnove uporabe okolja za analizo podatkov »R«. Vrste podatkov, priprava in urejanje podatkov. Vnos in izpis podatkov, izmenjava podatkov z drugimi programskimi okolji. Grafično prikazovanje podatkov. Priprava lastnih funkcij. Statistične porazdelitve in simulacija podatkov. Analiza podatkov z R.

3. Pregled metod multivariatne analize. Osnovni pojmi linearne algebre za uporabo v statistiki večdimenzionalnih podatkov. Vektorska algebra, matrike in matrični račun, pojem lastnih vrednosti in lastnih vektorjev. Statistična in geometrijska interpretacija pojmov linearne algebre. Metoda glavnih komponent, diskriminacijska analiza, faktorska analiza, razvrščanje v skupine, vizualizacija podatkov.

4. Statistično ozadje analize mikromrež. Načrt poskusa, priprava podatkov, metode za odstranjevanje šuma ozadja, normalizacija podatkov, analiza diferencialne izraženosti, grafične predstavitve in vizualizacija rezultatov, analiza omrežij, povezovanje z bazami podatkov in ontologij na svetovnem spletu.

5. Izbrane metode za analizo podatkov. Izbor posebnih metod bomo prilagajali glede na usmeritev in področje dela prijavljenih študentov.

1. Review of basic statistical methods and their use for the analysis of data. Statistical testing of assumptions. Methods of studying the dependence of phenomena.

2. Basis of use of the environment for analysis of data »R«. Types of data, preparation and arrangement of data. Entry and extraction of data, exchange of data with other programme environments. Graphic presentation of data. Preparation of own functions. Statistical distribution and simulation of data. Analysis of data with R.

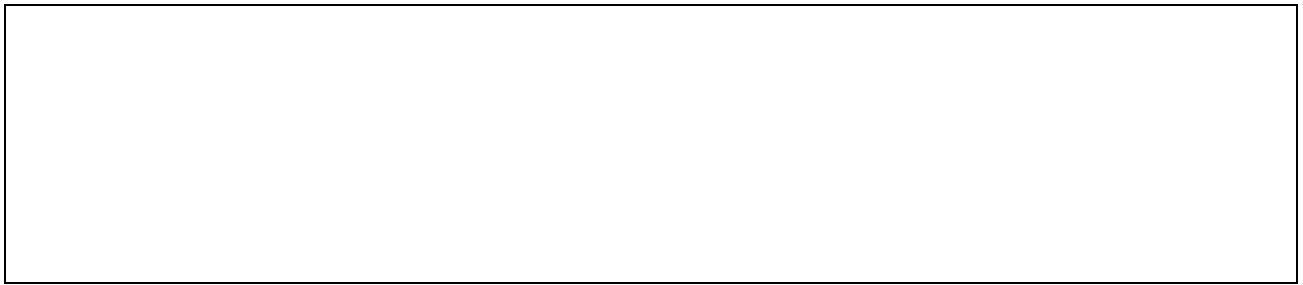
3. Review of methods of multivariate analysis. Basic concepts of linear algebra for use in statistics of multidimensional data. Vector algebra, matrices and matrix calculation, concept of own values and own vectors. Statistical and geometric interpretation of concepts of linear algebra. Method of main components, discrimination analysis, factorial analysis, classifying in groups, visualisation of data.

4. Statistical background to analysis of micronets. Plan of experiment, preparation of data, methods for removing background noise, normalisation of data, analysis of differential expression, graphic presentation and visualisation of results, analysis of networks, linkage with databases and ontologies on the internet.

5. Selected methods for data analysis. The selection of special methods will be adapted to the orientation and field of work of students.

Temeljni literatura in viri / Readings:

- Whitlock, Michael C.: The analysis of biological data / Michael C. Whitlock and Dolph Schluter. - Greenwood Village, Colorado : Roberts and Company Publishers, 2009. - ISBN 978-0-9815194-0-1
- Krzanowski WJ, Principles of Multivariate Analysis, Oxford Science Publications, 1988.
- Blejec, A: Introduction to R
<http://ablejec.nib.si/R/I2R/DOC/I2R.pdf>
- različni viri na svetovnem spletu.

**Cilji in kompetence:**

Študent nadgradi poznavanje statističnih metod z zahtevnejšimi metodami, ki jih bo potreboval pri raziskovalnem delu. Poudarek je na konceptualnem razumevanju metod, primernosti metod za različne probleme in samostojni analizi podatkov s pomočjo sodobne programske opreme (R).

Objectives and competences:

The student builds on understanding of statistical methods with more demanding methods required in research work. The stress is on conceptual understanding of methods, comparability of methods for various problems and independent analysis of data with the aid of up-to-date software (R).

Predvideni študijski rezultati:

Znanje in razumevanje:
Študent se usposobi za čim bolj samostojno izbiro ustrezne metode in analize problema, s katerim se vsebinsko ukvarja. Dosežena znanja mu pomagajo pri komunikaciji s strokovnjaki statističnih strok ter pri primerni vključitvi statističnih rezultatov v poročila in znanstvene članke

Intended learning outcomes:

Knowledge and understanding:
The student is trained for as independent as possible selection of suitable methods and analysis of problems with which he or she is dealing. The achieved knowledge will help him or her in communication with statistical experts and with suitable inclusion of statistical results in reports and scientific articles.

Metode poučevanja in učenja:

- predavanja,
- laboratorijske vaje,
- konzultacije,
- seminarske naloge

Learning and teaching methods:

- lectures
- work in computer lab
- consultations
- seminar

Načini ocenjevanja:

- ustni/pisni izpit
- praktično delo v računalnici

Delež (v %) /
Weight (in %)

Assessment:

- oral/written exam
- Practical work in the lab

Reference nosilca / izvajalcev / Lecturer's references:

1. FIŠER, Cene, BLEJEC, Andrej, TRONTELJ, Peter. Niche-based mechanisms operating within extreme habitats : a case study of subterranean amphipod communities. *Biology letters*, ISSN 1744-9561, 2012, vol. 8, no. 4, str. 578-581, doi: 10.1098/rsbl.2012.0125. [COBISS.SI-ID 2554447]
2. ROUYAR, Angela, PARTY, Virginie, PREŠERN, Janez, BLEJEC, Andrej, RENO, Michel. A general odorant background affects the coding of pheromone stimulus intermittency in specialist olfactory receptor neurones. *PloS one*, ISSN 1932-6203, 2011, vol. 6, no. 10, str. e26443-1-e26443-13. <http://www.plosone.org/article/info:doi/10.1371/journal.pone.0026443>, doi: 10.1371/journal.pone.0026443. [COBISS.SI-ID 1156703]
3. BAEBLER, Špela, STARE, Katja, KOVAČ, Maja, BLEJEC, Andrej, PREZELJ, Nina, STARE, Tjaša, KOGOVSĚK, Polona, POMPE NOVAK, Maruša, ROSAHL, S., RAVNIKAR, Maja, GRUDEN, Kristina. Dynamics of responses in compatible potato - potato virus Y interaction are modulated by salicylic acid. *PloS one*, ISSN 1932-6203, 2011, vol. 6, issue 12, str. 1-12. <http://dx.doi.org/10.1371/journal.pone.0029009>, doi: 10.1371/journal.pone.0029009. [COBISS.SI-ID 2492751]
4. ŠUŠTAR VOZLIČ, Jelka, ROSTOHAR, Katja, BLEJEC, Andrej, KOZJAK, Petra, ČERGAN, Zoran, MEGLIČ, Vladimir. Development of sampling approaches for the determination of the presence of genetically modified organisms at the field level. *Analytical and bioanalytical chemistry*, ISSN 1618-2642, 2010, vol. 396, iss. 6, str. 2031-2041, doi: 10.1007/s00216-009-3406-4. [COBISS.SI-ID 3210344]
5. PAPAŽOVA, Nina, ZHANG, David, GRUDEN, Kristina, VOJVODA, Jana, YANG, Litao, BUH GAŠPARIČ, Meti, BLEJEC, Andrej, FOUILLOUX, Stephane, DE LOOSE, Marc, TAVERNIERS, Isabel. Evaluation of the reliability of maize reference assays for GMO quantification. *Analytical and bioanalytical chemistry*, ISSN 1618-2642, 2010, vol. 396, no. 6, str. 2189-2201. <http://dx.doi.org/10.1007/s00216-009-3386-4>, doi: 10.1007/s00216-009-3386-4. [COBISS.SI-ID 2163791]
6. ROTTER, Ana, KRALJ NOVAK, Petra, BAEBLER, Špela, TOPLAK, Nataša, BLEJEC, Andrej, LAVRAČ, Nada, GRUDEN, Kristina. Gene expression data analysis using closed itemset mining for labeled data. *Omics*, ISSN 1536-2310, 2010, vol. 14, no. 2, str. 177-186. <http://dx.doi.org/10.1089=omi.2009.0126>, doi: 10.1089=omi.2009.0126. [COBISS.SI-ID 2210383]