BIOTECH-06:	**BIOINFORMATICS AND BIOSTATISTICS**
GENERAL INFORMATION	
Course Coordinator(s)	Hrvoje Brkić, PhD, Assist. Prof. Dario Faj, PhD, Full Prof.
Associate(s)	-
Study Programme	Interdisciplinary Graduate Study Programme in English: Biotechnology
Course Status	Obligatory
Year of Study, Semester	1st Year / 1st Semester
Credits (ECTS)	**6**
Teaching Method (number of classes)	Lectures: 20 ; Seminars: 20 ; Exercises: 20
Expected Number of Students in the Course	25-30

COURSE DESCRIPTION

Course Aims

Teach students how to use bioinformatics tools, as well as search and process bioinformatics data bases. Introduce students with most common distributions in biosciences and theories of samples. Elaborate students' stochastic bond and mathematical modelling, application of statistical test and obtaining the results from the tests.

Prerequisites for Enrolment and the Entry Competencies Required for the Course

Completed undergraduate university study programme from the area of natural sciences (chemistry, biology) or biotechnical sciences, or biomedicine and healthcare.

Learning Outcomes at the Programme Level Contributed by the Course

BIOTECH-4; BIOTECH-9; INDBIOT-6; MEDBIOT-3

Learning Outcomes at the Course Level

After completing the course, the student will be able to:

1. Independently use bioinformatics tools and operational systems that are commonly used in bioinformatics.
2. Use shell scripts for data analyses.
3. Transform and compare different types of biological structures.
4. Use web based tools, as well as their shell versions of most common bioinformatics tools that have application in biology and chemistry.
5. Interpret parameters of statistical set, and based on them use appropriate theoretical distribution and test it.
6. Use appropriate statistical test (expectation test, expectation comparison (t-test), variance comparison (F-test), ANOVA, proportion test, proportion comparison, $\chi 2$-model fit test data, $\chi 2$-independence test, $\chi 2$-homogeneity test), with and without usage of computers, and interpret results correctly.
7. Suggest method of specimen sampling, determine parameters of population samples conduct appropriate statistical tests and interpret solutions.
8. Critically judge laboratory experiments, and statistical analyses that are present in professional papers

Course Content

Lectures. History and introduction to bioinformatics; Bioinformatics operating systems. Shell programming. Descriptive statistics. Statistical characteristics. Tabular and graphical presentation of data. Mean values. Variability measures. Location measures. Basics of Probability Theory. Probability Space. Probability setting. Conditional probability. Independence. Discrete and continuous random variables. Mathematical expectation and variance of a random variable. Binomial distribution. Hypergeometric distribution. Poisson distribution. Normal distribution. Statistical test. Errors of the first and second kind. Expectation test. Reliable intervals for direction

parameters. Prediction. Reliable intervals for the predicted value of the dependent variable and for its mean.

Seminars. Bioinformatics operating systems: getting to know and working on one of the Linux operating system distributions; Shell programming: basic commands and ways to execute them, create and run executable scripts. Analysis of primary DNA and protein sequences using bioinformatics tools. Descriptive statistics: Statistical characteristics. Testing statistical hypotheses and reliable intervals: A random sample. Point estimates of parameters (expectations and variance). Expectation comparison test (t-test). Variance comparison test (F-test). One-factor analysis of variance (ANOVA). Proportion parameter test. Reliable interval for the proportion parameter. Proportion Comparison Test. χ^2 - model fit test. χ^2 - independence test. χ^2 - homogeneity test. Linear regression model: Direction adjustment; least squares method.

Exercises. Bioinformatics tools, downloading of tools from repositories, their installation and usege. Performing background jobs. Practical examples of using bioinformatics tools in research and statistics, biological databases. Tools for transforming coordinate systems used in the file formats of biological molecules, comparing different types of formats, and searching for them. Expectation test; sample from a normally distributed population, and a large sample. Reliable interval for the expectation parameter; sample from a normally distributed population, and a large sample. Pearson's correlation coefficient. Correlation coefficient test. Testing direction parameter hypotheses. Nonparametric tests: Wilcoxon rank sum test, Mann Whitney test. Designing and conducting experiments.

Teaching Methods

Lectures; seminars; exercises

Students' Obligations

Attendance at all forms of classes is mandatory and the students are obligated to attend all knowledge tests. The students may be absent from 30% (full-time students) and 50% (part-time students) of each of the forms of classes, provided that the absence is justified. An exercise or a seminar which has not been completed must be made up through a midterm exam.

Monitoring the Activity of the Students *(Connecting Learning Outcomes, Teaching Methods, and Grading)*

Class-related activity	ECTS	Learning outcome	Student activity	Evaluation method	Grade points	
					Min.	Max.
Attending classes	1.25	1-7	Attendance at classes	Keeping records	10	25
Seminar work	1.25	1-7	Practical work	Seminar task	10	25
Knowledge tests (partial tests)	1.5	1-8	Studying for partial knowledge tests	Written exam	6	10
Final exam	2	1-8	Studying for the final exam	Written final exam	24	40
Total	**6**				**50**	**100**

Evaluation of the written part of the final exam

Percentage of correct answers (%)	Grade
>95.00	40
90.00-94.99	38
85.00-89.99	36
80.00-84.99	34
75.00-79.99	32
70.00-74.99	30

65.00-69.99	27
60.00-64.99	24

Forming the final grade:
The points granted for the final exam are added to the grade points awarded during class attendance. The grading process is conducted by absolute distribution, i.e. based on total achievements, and compared to the numerical system in the following manner:
A – Excellent (5): 90-100 grade points; B – Very Good (4): 80-89.99 grade points; C – Good (3): 65-79.99 grade points; D – sufficient (2): 50-64.99 grade points

Mandatory Literature (available in the library and via other media)

Title	Number of copies in the library	Availability via other media
Lesk A: Introduction to bioinformatics. Oxford University Press, 2014.		
Bernard R: Fundamentals of Biostatistics, 7th Ed., Brooks/Cole, 2011.		

Additional Literature

Scripts available on www.physics.mefos.hr

Quality Assurance Procedures Designed to Ensure the Acquisition of Outcomes and Competencies

Anonymous, quantitative, standardised student survey on the course and the teacher's work implemented by the Quality improvement office of the Faculty of Medicine Osijek Faculty and/or the Faculty of Food Technology Osijek.

Note

E-learning is not included in the class quota, but it is used in teaching and it contains links to various sites and video and audio materials available on websites.