# COURSE OUTLINE

## (1) GENERAL

| | | | |
|---|---|---|---|
| **SCHOOL** | SCHOOL OF SCIENCES | | |
| **DEPARTMENT** | DEPARTMENT OF CHEMISTRY | | |
| **LEVEL OF STUDIES** | ISCED level 6 – Bachelor's or equivalent level | | |
| **COURSE CODE** | EN4 | **SEMESTER** | 7th / 8th Semester |
| **COURSE TITLE** | Biostatistics | | |

| **TEACHING ACTIVITIES** *If the ECTS Credits are distributed in distinct parts of the course e.g. lectures, labs etc. If the ECTS Credits are awarded to the whole course, then please indicate the teaching hours per week and the corresponding ECTS Credits.* | **TEACHING HOURS PER WEEK** | **ECTS CREDITS** |
|---|---|---|
| | 3 | 3 |
| | | |
| | | |
| *Please, add lines if necessary. Teaching methods and organization of the course are described in section 4.* | | |

| | |
|---|---|
| **COURSE TYPE** *Background, General Knowledge, Scientific Area, Skill Development* | Skill Development |
| **PREREQUISITES:** | NO |
| **TEACHING & EXAMINATION LANGUAGE:** | GREEK |
| **COURSE OFFERED TO ERASMUS STUDENTS:** | NO |
| **COURSE URL:** | https://eclass2.emt.duth.gr |

## (2) LEARNING OUTCOMES

**Learning Outcomes**
*Please describe the learning outcomes of the course: Knowledge, skills and abilities acquired after the successful completion of the course.*

The aim of the course is to provide students with knowledge of basic and advanced statistical methods for problem-solving through data analysis in various scientific fields, such as Biology, Medicine, and Life and Health Sciences. The expected learning outcomes include acquiring knowledge of fundamental concepts in the scientific domain of Statistics and Probability (population, sample, discrete and continuous probability distributions, etc.), as well as basic methods of Descriptive and Inferential Statistics (parametric and non-parametric hypothesis tests). In addition, particular emphasis will be placed on well-known Biostatistical methods for knowledge discovery in multivariate data and modeling the relationships between a dependent variable and a set of independent variables using Regression Models (Linear Regression, Logistic Regression, etc.) and Survival Analysis. To facilitate the understanding of these methods and the extraction of meaningful conclusions, the course will focus on the detailed presentation of applications and case studies using real data from Biology, Medicine, and Life and Health Sciences. Students will also acquire practical skills in implementing these methodologies and techniques in the open-source programming language R, developing code within the RStudio Integrated Development Environment (IDE).

Upon successful completion of the course, students will be able to:

- Understand the fundamentals of Biostatistics and its applications in Biology, Life Sciences, and Health Sciences.
- Describe populations and samples, compute descriptive statistics, and perform exploratory data analysis and multivariate visualizations using R.

- Understand and apply probability distributions, including Normal, t-student, Binomial, and Poisson distributions, in Biostatistical contexts.
- Construct confidence intervals and perform hypothesis testing, including parametric and non-parametric tests, and interpret Type I and Type II errors.
- Apply parametric tests such as t-tests, ANOVA, and tests for categorical data, using R for practical case studies.
- Perform non-parametric tests, including Wilcoxon signed-rank, Mann-Whitney U, and Kruskal-Wallis tests, with applications in real datasets.
- Conduct correlation and regression analyses (simple and multiple linear regression), check assumptions, perform diagnostics, and interpret models in R.
- Apply logistic and Poisson regression for binary, categorical, and count data using R.
- Understand clinical study designs and apply statistical methods for analyzing data from experimental and observational studies.
- Use advanced methods including generalized linear models, mixed-effects models, hierarchical ANOVA, and repeated measures analysis in practical case studies.
- Perform survival analysis using Kaplan-Meier estimators, log-rank tests, and Cox proportional hazards models with R.
- Apply statistical methods in genomics, including GWAS data analysis, Hardy-Weinberg equilibrium, and SNP association studies using R.

**General Skills**

*Name the desirable general skills upon successful completion of the module*

| | |
|---|---|
| *Search, analysis and synthesis of data and information,* | *Project design and management* |
| *ICT Use* | *Equity and Inclusion* |
| *Adaptation to new situations* | *Respect for the natural environment* |
| *Decision making* | *Sustainability* |
| *Autonomous work* | *Demonstration of social, professional and moral responsibility and* |
| *Teamwork* | *sensitivity to gender issues* |
| *Working in an international environment* | *Critical thinking* |
| *Working in an interdisciplinary environment* | *Promoting free, creative and inductive reasoning* |
| *Production of new research ideas* | |

Search, analysis and synthesis of data and information, Adaptation to new situations, Decision making, Autonomous work, Teamwork, Working in an international environment,

## (3) COURSE CONTENT

Week 1: Introduction to Biostatistics: Importance and applications in the Life Sciences (Biology, Life and Health Sciences), observations, variables, and data types.

Week 2: Population and Random Sample: Descriptive statistics, point estimators (measures of central tendency, measures of variability, measures of distribution shape), accuracy of mean estimation, standard error of the mean, exploratory analysis, multivariate data visualization techniques, case studies in the Life Sciences (Biology, Life and Health Sciences) using the R programming language.

Week 3: Probability Distributions: Random variables, discrete and continuous distributions, Normal distribution, t-student distribution, Binomial distribution, and Poisson distribution; applications of distributions in Biostatistics.

Week 4: Inferential Statistics: Confidence intervals for population parameters, hypothesis testing, goodness-of-fit tests, types of errors (Type I and Type II errors).

Week 5: Parametric Hypothesis Tests: t-tests (one-sample, two independent/dependent samples), analysis of variance (ANOVA), paired comparisons, hypothesis tests for categorical data, contingency tables, likelihood ratio, risk and relative risk, sensitivity and specificity; case studies in the Life Sciences using R.

Week 6: Non-parametric Hypothesis Tests: Wilcoxon signed-rank test, Mann-Whitney U test, Kruskal-Wallis test; case studies in the Life Sciences using R.

Week 7: Correlation & Regression Analysis: Simple linear regression, multiple linear regression, model assumptions and diagnostic checks, model interpretation; case studies in the Life Sciences using R.

Week 8: Logistic Regression: For binary response variables, multinomial logistic regression, regression for count data (Poisson regression); case studies in the Life Sciences using R.

Week 9: Clinical Studies: Types of clinical studies (experimental/observational, controlled/uncontrolled experimental, descriptive/analytical, cross-sectional/longitudinal, prospective/retrospective, factorial design).

Week 10: Generalized Linear Models and Mixed-Effects Models: Hierarchical ANOVA, repeated measures; case studies in the Life Sciences using R.

Week 11: Survival Analysis: Kaplan-Meier estimator, Kaplan-Meier curve, log-rank test, Cox proportional hazards model; case studies in the Life Sciences using R.

Week 12: Introduction to Statistical Methods for Genome-Wide Association Studies (GWAS): Data import, descriptive statistics, Hardy-Weinberg equilibrium, single nucleotide polymorphism (SNP) association analysis using R.

Week 13: Presentation of the Statistical Learning Life-cycle in real-life datasets.

## (4) LEARNING & TEACHING METHODS - EVALUATION

<table>
<tr><td>TEACHING METHOD<br>Face to face, Distance learning, etc.</td><td>Face to face</td></tr>
<tr><td>USE OF INFORMATION & COMMUNICATIONS TECHNOLOGY (ICT)<br>Use of ICT in Teaching, in Laboratory Education, in Communication with students</td><td>Use of ICT in Teaching<br>Use of ICT in Communication with students</td></tr>
<tr><td>TEACHING ORGANIZATION<br>The ways and methods of teaching are described in detail.<br>Lectures, Seminars, Laboratory Exercise, Field Exercise, Bibliographic research & analysis, Tutoring, Internship (Placement), Clinical Exercise, Art Workshop, Interactive learning, Study visits, Study / creation, project, creation, project. Etc.<br><br>The supervised and unsupervised workload per activity is indicated here, so that total workload per semester complies to ECTS standards.</td><td><table><tr><th>Activity</th><th>Workload/semester</th></tr><tr><td>Lectures</td><td>39</td></tr><tr><td>Study</td><td>33</td></tr><tr><td>Exams</td><td>3</td></tr><tr><td>Total</td><td>75</td></tr></table></td></tr>
<tr><td>STUDENT EVALUATION<br>Description of the evaluation process<br><br>Assessment Language, Assessment Methods, Formative or Concluding, Multiple Choice Test, Short Answer Questions, Essay Development Questions, Problem Solving, Written Assignment, Essay / Report, Oral Exam, Presentation in audience, Laboratory Report, Clinical examination of a patient, Artistic interpretation, Other/Others<br><br>Please indicate all relevant information about the course assessment and how students are informed</td><td>Written final exam that includes problem solving from different sections of the course.</td></tr>
</table>

## (5) SUGGESTED BIBLIOGRAPHY

- Βιοστατιστική με την R. Μία Εισαγωγή για τις Βιολογικές Επιστήμες. Leps J., Smilauer P. Μετάφραση: Κοκκόρης Γ., Δημητρακόπουλος. ΡΟΠΗ. Π. ISBN: 9786185289607.
- Βιοστατιστική. Τριχόπουλος Δ., Τζώνου Α., Κατσουγιάννη Κ.
- Βιοστατιστική και Εφαρμογές. Παπαγεωργίου Ε.. Εκδόσεις Νέων Τεχνολογιών, 2016