

SYLLABUS

APPLIED BIOSTATISTICS IN R

University year 2025-2026

1. Information regarding the programme

1.1. Higher education institution	Babeş-Bolyai University
1.2. Faculty	Faculty of Biology and Geology
1.3. Department	Integrative Biology Doctoral School
1.4. Field of study	Biology
1.5. Study cycle	Doctorate
1.6. Study programme/Qualification	Doctoral training / PhD in Biology
1.7. Form of education	Full-time

2. Information regarding the discipline

2.1. Name of the discipline	Applied biostatistics in R			Discipline code	BDR1112		
2.2. Course coordinator	CSIII Dr. Fülöp Attila / Conf. Dr. László Zoltán						
2.3. Seminar coordinator	CSIII Dr. Fülöp Attila / Conf. Dr. László Zoltán						
2.4. Year of study	1	2.5. Semester	1	2.6. Type of evaluation	C	2.7. Discipline regime	O

3. Total estimated time (hours/semester of didactic activities)

3.1. Hours per week	4	of which: 3.2 course	2	3.3 seminar/laboratory	2
3.4. Total hours in the curriculum	48	of which: 3.5 course	24	3.6 seminar/laborator	24
Time allotment for individual study (ID) and self-study activities (SA)					hours
3.5.1. Learning using manual, course support, bibliography, course notes (SA)					48
3.5.2. Additional documentation (in libraries, on electronic platforms, field documentation)					48
3.5.3. Preparation for seminars/labs, homework, papers, portfolios and essays					12
3.5.4. Tutorship					12
3.5.5. Evaluations					4
3.5.6. Other activities					3
3.7. Total individual study hours					127
3.8. Total hours per semester					175
3.9. Number of ECTS credits					7

4. Prerequisites (if necessary)

4.1. curriculum	Not applicable
4.2. competencies	<ul style="list-style-type: none"> • User-level computing skills (medium-level) • Speaking and writing skills in English (medium-level) • Basic knowledge of mathematics

5. Conditions (if necessary)

5.1. for the course	Audio-video logistics, whiteboard, access to WiFi internet
5.2. for the seminar /lab activities	Audio-video logistics, whiteboard, access to WiFi internet

6.1. Specific competencies acquired

Professional/ essential competencies	<ul style="list-style-type: none"> • Competences in designing scientific studies, both on the field and in the laboratory • Competences in collecting, analyzing and interpreting scientific data • Programming, data analysis and data visualization competences in the R statistical environment
---	---

Transversal competencies	<ul style="list-style-type: none"> Competences in the statistical analysis of data, a requisite of scientific research in any field which involves a quantitative approach
---------------------------------	---

6.2. Learning outcomes

Knowledge	The student knows to perform a statistical analysis in R
Skills	The student is able to calculate, compute and program in R
Responsibility and autonomy:	The student has the ability to work independently for a research project or any field of biology which includes data analysis

7. Objectives of the discipline (outcome of the acquired competencies)

7.1 General objective of the discipline	<ul style="list-style-type: none"> To gain quantitative analytical skills of biological data in the R statistical environment
7.2 Specific objective of the discipline	<ul style="list-style-type: none"> To learn to design scientific studies and to collect data in biology and ecology, both in the laboratory and on field To learn to analyze data using the adequate statistical tests, as implemented in the R statistical environment To learn to visualize data using plotting tools implemented in R

8. Content

8.1 Course	Teaching methods	Remarks
1. Introduction: - The scientific method - Data, observations and variables - Probability distributions	Presentation, discussion, case studies	2 hours
2. Estimation: - Samples and populations - Common parameters and statistics - Standard errors and confidence intervals for the mean - Methods for estimating parameters	Presentation, discussion, case studies	2 hours
3. Graphical exploration of data: - Exploratory data analysis - Transforming data - Standardization - Outliers - Censored and missing data	Presentation, discussion, case studies	2 hours

4. Hypothesis testing: - Statistical hypothesis testing - Decision errors - Multiple testing - Critique of statistical hypothesis testing	Presentation, discussion, case studies	2 hours
5. Non-parametric statistics: - Comparing means - Comparing variances - Comparing frequencies	Presentation, discussion, case studies	2 hours
6. Correlation and regression: - Correlation analysis - Simple and multiple linear regression analysis	Presentation, discussion, case studies	2 hours
7. Design and power analysis: - Sampling, experimental design	Presentation, discussion, case studies	2 hours
8. Analysis of variance 1: - Single factor (one-way) design - Assumptions and diagnostics - Factor effects	Presentation, discussion, case studies	2 hours
9. Analysis of variance 2: - Multi-factor design - Repeated measures design	Presentation, discussion, case studies	2 hours
10. Analyses of covariance: - ANCOVA	Presentation, discussion, case studies	2 hours
11. Generalized linear models and logistic regression: - Logistic regression - Poisson regression	Presentation, discussion, case studies	2 hours
12. Multivariate analysis of variance - Principal components analysis (PCA) - Factor analysis (FA) - Discriminancy analysis (DA)	Presentation, discussion, case studies	2 hours
Bibliography (1) Quinn, G.P., Keough, M.K. (2002) Experimental design and data analysis for biologists. Cambridge Univ. Press, UK. (2) Field, A., Miles, J., Field, Z. (2012) Discovering statistics using R. SAGE, UK. (3) Crawley, M.J. (2013) The R Book, 2nd edition. John Wiley, UK.		
8.2 Seminar / laboratory	Teaching methods	Remarks
1. Introduction to R: - Setting up the R environment (R and R Studio) - Installing and loading packages - Help and troubleshooting - Introduction to the R language	Practical work. Problem-solving and discussion	2 hours
2. Data types and structures in R: - Operations with different data types and structures	Practical work. Problem-solving and discussion	2 hours
3. R programming: - Conditionals, loops, and functions - Debugging - Organizing and commenting code	Practical work. Problem-solving and discussion	2 hours
4. Data manipulation, exploration and cleaning (part 1): - Reading and writing data files - Sorting, filtering, merging, cleaning data - Preparing the data set for analyses	Practical work. Problem-solving and discussion	2 hours
5. Data manipulation, exploration and cleaning (part 2): - Reading and writing data files - Sorting, filtering, merging, cleaning data - Preparing the data set for analyses	Practical work. Problem-solving and discussion	2 hours
6. Descriptive statistics and exploratory analysis: - Distributions - Exploratory plots and summary statistics - Data transformation - Missing values - Handling outliers	Practical work. Problem-solving and discussion	2 hours
7. Non-parametric statistics:	Practical work. Problem-solving	2 hours

- t, Wilcoxon, Welch, sign, binomial, Kruskal-Wallis, chi-square, Fisher tests - Correlation analysis	and discussion	
8. Linear regression (part 1): - Simple and multiple linear regression - Model checking - Post-hoc comparisons - Analysis of variance	Practical work. Problem-solving and discussion	2 hours
9. Linear regression (part 2): - Simple and multiple linear regression - Model checking - Post-hoc comparisons - Analysis of variance	Practical work. Problem-solving and discussion	2 hours
10. Multivariate statistics: - PCA, FA, DA	Practical work. Problem-solving and discussion	2 hours
11. Plotting: - Preparing plots for publications	Practical work. Problem-solving and discussion	2 hours
12. Reproducible research, open science tools for R: - R Markdown - git, GitHub, GitLab, etc.	Practical work. Problem-solving and discussion	2 hours
Bibliography (1) Crawley, M.J. (2013) The R Book, 2nd edition. John Wiley, UK. (2) Ekstrøm, C.T. (2016) The R Primer, Second Edition. CRC Press, FL. (3) Chang, W. (2013) R Graphics Cookbook. O'Reilly Media, CA. (4) Gandrud, C. (2020) Reproducible Research with R and RStudio, Third Edition. CRC Press, FL. (5) Wickham, H., Grolemund, G. (2017) R for Data Science: Import, Tidy, Transform, Visualize, and Model Data. O'Reilly Media, CA.		

9. Corroborating the content of the discipline with the expectations of the epistemic community, professional associations and representative employers within the field of the program

<ul style="list-style-type: none"> The course has a similar content to courses from other European universities, and considers the level of training and abilities of doctoral students The content of the course is regularly updated and incorporates the most novel approaches from the field of statistics The course is fundamental for doctoral students, as quantitative skills are essential for scientific research activities
--

10. Evaluation

Activity type	10.1 Evaluation criteria	10.2 Evaluation methods	10.3 Percentage of final grade
10.4 Course	Assessment of knowledge	Colloquium (written)	50%
10.5 Seminar/laboratory	Activity during seminars	Individual projects	50%
10.6 Minimum standard of performance <ul style="list-style-type: none"> Knowledge of 50% of the information content of the course Knowledge of 50% of the information content of the laboratory work 			

11. Labels ODD (Sustainable Development Goals)

	General label for Sustainable Development						
							

Date: 04. 05. 2025

Signature of course coordinator

Signature of seminar coordinator

Conf. Dr. László Zoltán

Conf. Dr. László Zoltán

CSIII Dr. Fülöp Attila

CSIII Dr. Fülöp Attila

Date of approval:

...

Signature of the head of department

.....