ESCI upf.
School of International Studies

TEACHING PLAN FOR

Biostatistics and Data Analysis

1. Basic description

Name of the course: Biostatistics and Data Analysis
Module: Mathematics and Statistics

Academic year: 2016-2017
Year: 2017
Term: Second
Degree / Course: First
Code: 51205
Number of credits: 4
Total number of hours committed: 200
Teaching language: English
Lecturer: Hafid Laayouni
Timetable: See official calendar

2. Presentation of the course

Biostatistics and Data Analysis provides an introduction to selected important topics in biostatistical concepts and reasoning. This course represents an introduction to the field and provides a survey of data and data types. Specific topics include tools for describing central tendency and variability in data; methods for performing inference on population means and statistical hypothesis testing; issues of power and sample size in study designs. The course will familiarize the students with the use of the R statistical package and give them the skills needed for effective data management, data manipulation, and data analysis at a basic level.

3. Competences to be worked in the course

I. General competences

- CB2. That the students know how to apply their knowledge to their work or vocation in a professional manner and have competencies typically demonstrated through devising and defending arguments and solving problems within their field of study.

- CB4. That the students can convey information, ideas, problems and solutions to both specialist and non-specialist audiences.

- CB5. That the students have developed those skills needed to undertake further studies with a high degree of autonomy

- CG1. That the students will acquire an intra- and interdisciplinary training in both computational and scientific subjects with a solid basic training in biology.

II. Specific competences

CE2. To manage and exploit all kinds of biological and biomedical information to transform it into knowledge.

CE9. To apply statistical and computational methods to solve problems in the fields of molecular biology, genomics and medical research and population genetics.

Learning outcomes

RA2.1. Visualize, manipulate and extract biological data.

RA5.2. Apply mathematical and statistical treatment to large amounts of biological data.

RA8.2. Quote valid sources of scientific information to support the state of the arts of a bioinformatic problem.

RA9.1. Identify and use appropriate statistical methods to each type of data.

4. Contents

Theme 1: Introduction. The nature and impact of variability in biological data. Observational studies and experiments. Random sampling. (1h)
Theme 2: Description of distributions. Frequency distributions, descriptive statistics, the concept of population versus sample. (1h)
Themes 3: Probability and the binomial distribution. The normal distribution. (1h)
Theme 4: Sampling distributions. (1h)
Theme 5: Confidence intervals for a single mean and for a difference in means. (1h)
Theme 6: Hypothesis testing. (2h)
Theme 7: The t test. (1h)
Theme 8: Relationships in categorical data, contingency tables. The chi-square goodness-of-fit test. (1h)
Theme 9: Analysis of variance. One-way ANOVA, multiple comparison procedures (4h)

Theme 10: Two-way ANOVA. (2h)
Theme 11: Correlation and regression. Descriptive and inferential aspects of correlation and simple linear regression. (3h)
Theme 12: Multiple regression analysis (2h)
Theme 13: Analysis of Covariance (3h)
Theme 14: Effect Size and a summary of inference methods. (2h)
Theme 15: Nonparametric Statistics (1h)

5. Assessment

The competencies of this subject will be evaluated by continuous assessment, including written, practical examinations and individual work.
Theory
Evaluation by two partial tests. The first part will weigh 20% -40% and the second part will weigh 30% -50%. The final test is aimed at students who have not passed the partial tests and the weight in the final grade is no more than 70%.

Seminars: evaluating problems section is carried out by
-Making Short problems in the classroom, these tests last approximately 25 minutes
- Delivery of answers before resolution in class. Delivery can be made by sending the work to module or in paper before class problem-solving of the first group (regardless of the group to which the student belongs).
This section represents represent 10%-15% of the final grade for the course.

Hands On: the evaluation of this section is carried out by a practical test at the computer lab in one to two hour session. In this exam session, the student must use the appropriate statistical program, enter the data of a study, propose an analysis of it and answer specific questions. The weight of this test is about 10% to 20% of the final grade of the course.

The average grade for the course is calculated by weighting coefficients for each section (theory, problems and practices). However, a minimum score of 4 in the second partial test or recovery test is required to calculate the average final grade.

Students who have passed the subject by partial assessments and would like to improve their final note, could take the recuperation test. Students who take part of recuperation exam to increase their marks renounce to the mark obtained in the partial exams.

Students who have not participated in the continuous assessment and delivery of work problems may be evaluated by the final test. Also, their final grade cannot exceed 70% of the maximum score.

It is necessary to obtain a final grade equal to or greater than 5 to pass the subject, by partial recovery or testing, in any case weighing with problems and practices work.

It is considered that a student has the qualification of "not evaluable" when the number of evaluation activities is less than 50% of those scheduled for the course.

Students who repeat the course must carry out all assessment activities including delivery of problems and evaluation of the practical test.
Copy in any exam or plagiarism in the essay implies failing the course.

Extraordinary exam will take place according to the schedule fixed by the Degree Coordination. Failure to attend this exam implies student will keep his initial score.

Assessment elements	Time period	Type of assessment		Assessment agent			Type of activity	Grouping		Weight (%)
		continuous	synthesis	Lecturer	Self-assess	Co-assess		Indiv	Group (#)	
Delivery of exercises	Approximately weekly sessions	x		x			Conceptual and pursuit of the subject	x		15
Delivery of practical works	After each practical session	x		x				x		15
Practical exam	By the end of practical sesions	x	x	x			Pursuit and synthesis			10
Evaluation I	Exam schedule	x	x	x				x		20
Evaluation II	Exam schedule	x	x	x				x		40

6. Bibliography and teaching resources

- Basic bibliography

  M.L. Samuel, J.A. Witmer, A. Shaffner. Statistics for the life Sciences.

- Supplementary bibliography

- Fundamentals of biostatistics. Rosner, Bernard . Thomson Higher Education, cop. 2006
- Sokal, R. R. y F. J. Rohlf. 1995. Biometry (3ª edición). W. H. Freeman and Co., Nueva York.

- Teaching resources
- https://www.otexts.org/book/biostat
- http://onlinestatbook.com/
- http://www.biostathandbook.com/

7. Methodology

It is strongly recommended that students follow the work plan outlaid in section 8 of this document. Work is divided into 3 types of activities: those developed in class (plenary, seminar and practical sessions), directed work (assignments) and autonomous work.

Class sessions (40 hours)

a) Plenary lessons: sessions where the teacher introduces and explains the course contents defined in the previous section. These sessions include definitions, introduction to the most relevant methodologies and examples of each model .

b) Seminar sessions: In the seminar sessions will take place control tests on the exercise lists. Correction of lists of exercises. Discussion of results and explanation. Is extremely important that student do the exercises before correction. Delivery of exercises (when asked for) is mandatory and will score 0 if exercises are not delivered on time without a justified cause.

c) Practical sessions

Hands on session will take place in parallel with theory and seminar sessions. Goal is to aquiere skills using appropriate software and to get familiar with the practice of statistical analysis. R software will be used in this part.

8. Scheduling activities

Week	Activity in the classroom Grouping/type of activity	Activity outside the classroom Grouping/type of activity
Week 1		
Week 2		
Week 3		
Week 4		
Week 5		
Week 6		
Week 7		
Week 8		
Week 9		
Week 10		
Week final exams		