



Academic year	2019-20
Subject	10273 - Design of Experimental Studies and Statistics Applied to Nutrigenomics
Group	Group 1

**Subject**

<b>Subject / Group</b>	10273 - Design of Experimental Studies and Statistics Applied to Nutrigenomics / 1
<b>Degree</b>	Master's in Nutrigenomics and Personalised Nutrition
<b>Credits</b>	3
<b>Period</b>	1st semester
<b>Language of instruction</b>	Spanish

**Professors**

Lecturers	Office hours for students					
	Starting time	Finishing time	Day	Start date	End date	Office / Building
Arnau Mir Torres <a href="mailto:arnau.mir@uib.es">arnau.mir@uib.es</a>						You need to book a date with the professor in order to attend a tutoring session.
Jairo Enrique Rocha Cárdenas <a href="mailto:jairo@uib.es">jairo@uib.es</a>						You need to book a date with the professor in order to attend a tutoring session.

**Context**

'Design of Experimental Studies and Statistics Applied Nutrigenomics' is a compulsory subject in the Master in Nutrigenomics and Personalised Nutrition, it is located in Module 1, "Fundamentals of Nutrigenomics and Personalised Nutrition".

This course seeks to initiate students in the experimental design and in the statistical methodology in the context of nutrigenomics. The course is both theoretical and practical and is therefore intended that, on the one hand, students acquire a solid understanding of the process of experimental design and of the bases of the statistical tools but, on the other hand, it is intended that students learn to correctly apply these techniques (using specialized statistical software) and acquiring the ability to interpret statistical results with scientific rigor.

Learning Outcomes:

- \* Summarize the main types of designs used in clinical and basic research in the area of nutrigenomics
- \* Make proper and efficient use of statistics tools currently available to analyse data from omics techniques
- \* Correctly interpret the results of data analysis and execute different statistical tests in the context of nutrigenomics
- \* Properly discuss the results published in the scientific literature in the area of nutrigenomics

**Requirements**

**Recommended**

Knowledge of R software for statistical computing (<http://www.r-project.org>) and Rstudio (<https://www.rstudio.com/>)



## Syllabus

Basic knowledge of data management and analysis.

### Skills

#### Specific

- \* Ability to work proficiently in a professional environment related to nutrigenomics, personalized nutrition, molecular nutrition and functional foods
- \* Correctly apply the statistical analysis in the field of nutrigenomics

#### Generic

- \* Knowing the capabilities and potential of ICT (Information and communications technology) in the discipline area
- \* Ability to formulate hypotheses and design research studies
- \* Ability to analyze data and draw conclusions from research results
- \* Ability to search, organize and critically analyze the research literature related with the discipline

#### Basic

- \* You may consult the basic competencies students will have to achieve by the end of the Master's degree at the following address: [http://estudis.uib.cat/master/comp\\_basiques/](http://estudis.uib.cat/master/comp_basiques/)

### Content

The course will develop all the contents selected and listed below. These contents will be developed in lectures, in which the concepts and fundamentals of each topic will be introduced, and in practical classes (workshops) in which the statistics tools to solve specific problems in the field of Nutrigenomics will be applied.

#### Range of topics

##### Session 1. Introduction to the research applied to Nutrigenomics

- \* Steps of a study. Characteristics of clinical studies.
- \* Some types of clinical studies: cases and controls, cohorts, transverse, ecological, etc.
- \* Meta-analysis

##### Session 2. Probabilities

- \* Introduction, vocabulary and properties.
- \* Conditional probability, independence, Bayes formula.
- \* Odds.
- \* Diagnostic tests, ROC curves, risks

##### Session 3. Descriptive statistics

- \* Type of data, exploratory analysis of data, variables.
- \* Data structure.
- \* Qualitative variables, ordinal data, quantitative data, frequencies.
- \* Measures of central tendency.
- \* Measures of dispersion.
- \* Other measures: symmetry, kurtosis.
- \* Graphics: boxplots, dispersion.
- \* Multivariate data.

Academic year	2019-20
Subject	10273 - Design of Experimental Studies and Statistics Applied to Nutrigenomics
Group	Group 1

### Session 4. Random variables.

- \* Definition. Events of a random variable. Independents random variables.
- \* Discrete variables. Mean, variance and typical deviation, quantiles.
- \* Continuous variables. Density. Mean, quantiles.
- \* Normal distribution. Properties, reference intervals and diagnostic criteria.
- \* Binomial, hypergeometric and Poisson distributions.

### Session 5. Point estimate and confidence intervals.

- \* Basic definitions. Inferential statistics.
- \* Typical deviation vs typical error.
- \* Sample mean and its distribution.
- \* Sample variance and its distribution.
- \* Sample proportion and its distribution.
- \* Introduction to confidence intervals.
- \* Confidence interval of the mean of a normal. Size of the sample.
- \* Confidence interval of the difference of two means.
- \* Confidence interval of proportions.
- \* "Finite" population.
- \* Confidence intervals of the quotient of proportions.

### Session 6. Hypothesis test

- \* Introduction
- \* Null hypothesis and alternative hypothesis. Scheme test. p-value. Types of errors.
- \* Example: t-test.
- \* Terminology: Rejection rules, statistical and clinical meaning. Power.
- \* Basic tests: t-test for two means, non parametric tests, chi square test for variances, approximate test for one proportion, tests for two independent and matched proportions.

### Session 7. Analysis of variance (ANOVA)

- \* Establishing the problem.
- \* Statistics, test assumptions, ANOVA table.
- \* Pairs comparison: t-test of Bonferroni.
- \* Verification of hypotheses: normality tests, QQ-plots, homocedasticity.
- \* Blocs ANOVA.
- \* Two ways ANOVA.

### Session 8. Test chi square

- \* Goodness of fit. Pearson's chi square test.
- \* Independence and homogeneity chi square test.
- \* Montecarlo's method.

### Session 9. Linear regression

- \* Basic problem.
- \* Simple linear regression. Least Squares, coefficient of determination, properties.
- \* Multiple linear regression. Properties and coefficient of determination.
- \* Linear regression and ANOVA.

## Teaching methodology

Learning and teaching activities:

Activities carried out by tutor with students:

- Lectures
- Seminars for student presentation and discussion
- Workshops with case studies



## Syllabus

Academic year	2019-20
Subject	10273 - Design of Experimental Studies and Statistics Applied to Nutrigenomics
Group	Group 1

- Computer-based analysis
- Tutorials

Activities to be carried out by students independently:

- Reading of recommended bibliography and study of lectures
- Literature searches
- Problem solving

Activities to be carried out by students in group:

- Reading and preparation of a critical comment (from a statistic point of view) of a scientific paper (or research work) chosen by the students
- Preparation of a seminar of the critical comment of a scientific paper
- Seminar presentation

In-class work activities (0.6 credits, 15 hours)

Modality	Name	Typ. Grp.	Description	Hours
Theory classes	Lectures	Large group (G)	<p><b>Purpose:</b> presentation and explanation of the main contents of the course.</p> <p><b>Methodology:</b> participatory lectures. To facilitate the understanding of concepts and the lectures follow-up, a power-point presentation will be projected, indicating the main topics of the subject and with diagrams and examples of the more difficult concepts. These slides will be uploaded in the specific Campus Extens space of the subject in order to facilitate that students come to class having reviewed and worked the material.</p>	6
Seminars and workshops	Seminar presentation	Medium group (M)	<p><b>Purpose:</b> To assess the degree of acquisition of basic knowledge of the subject, which includes managing of specific vocabulary, important concepts and current statistical tools, as well as the ability to synthesize and organize information from the literature. To stimulate the student to apply critical analysis and to develop speaking skills.</p> <p><b>Methodology:</b> Students in groups of 2-3 will present to the rest of the class, a seminar of a critical comment (from a statistic point of view) of a scientific paper (or research work) chosen by the students with teacher approval.</p>	3
Practical classes	Practical laboratory	Large group (G)	<p><b>Purpose:</b> To apply the knowledge acquired in lectures. These workshops include solving exercises and problems, with the teacher help, on data management (organization, storage and treatment) and the application of various statistical tests (t-test, Chi square, ANOVA, correlations and nonparametric test) with hypothetical data related to the field of nutrigenomics.</p> <p><b>Methodology:</b> Problem based learning. This activity will be held in the computer room in order to use specialized software for statistics (R). The teacher will guide the student to use the R program.</p>	4
Assessment	Synthesis exam	Large group (G)	<p><b>Purpose:</b> Perform a final exam of the synthesis of the subject.</p> <p><b>Methodology:</b> Each examination should be done individually</p>	2

At the beginning of the semester a schedule of the subject will be made available to students through the UIBdigital platform. The schedule shall at least include the dates when the continuing assessment tests will

be conducted and the hand-in dates for the assignments. In addition, the lecturer shall inform students as to whether the subject work plan will be carried out through the schedule or through another way included in the Aula Digital platform.

### Distance education tasks (2.4 credits, 60 hours)

Modality	Name	Description	Hours
Individual self-study	Problem solving	<p><b>Purpose:</b> To assess the degree of knowledge of the student and the capacity to apply statistical tools correctly with a set of proposed problems related to studies in the field of nutrigenomics</p> <p><b>Methodology:</b> Problem based learning in an individual self-study way. To develop this activity the use of any statistical software, for solving some of the problems, will be required (R is available in all computer rooms of the University).</p>	40
Group self-study	Seminar preparation	<p><b>Purpose:</b> To stimulate critical thought and the ability to synthesize and organize relevant information and to interpret statistical results with scientific rigor. It is also intended to develop teamwork skills.</p> <p><b>Methodology:</b> Students in groups of 2-3 will have to prepare a seminar of a critical comment (from a statistic point of view) of a scientific paper (or research work) chosen by the students with teacher approval.</p>	20

### Specific risks and protective measures

The learning activities of this course do not entail specific health or safety risks for the students and therefore no special protective measures are needed.

### Student learning assessment

To assess the knowledge acquired by students will be taken into account the acquisition of theoretical knowledge but also practical (including the ability to properly apply each statistical technique, what are the data requirements and how to interpret correctly the results of data related to Nutrigenomics) by solving a set of questions, problems and exercises, it is also intended to assess the acquisition of communication skills, critical analysis and the capacity to connect properly the various aspects of the subject by presenting a workshop. Raises two routes, the pathway A (continuous assessment) in which the assistance and class participation and the resolution of the final exam will be evaluated in addition to the other tools (problem solving and workshop). And the pathway B (final examination), which only take into account the other criteria of assessment (problem solving, workshop and final exam).

### Frau en elements d'avaluació

In accordance with article 33 of Regulation of academic studies, "regardless of the disciplinary procedure that may be followed against the offending student, the demonstrably fraudulent performance of any of the evaluation elements included in the teaching guides of the subjects will lead, at the discretion of the teacher, a undervaluation in the qualification that may involve the qualification of "suspense 0" in the annual evaluation of the subject".

Academic year	2019-20
Subject	10273 - Design of Experimental Studies and Statistics Applied to Nutrigenomics
Group	Group 1

### Seminar presentation

Modality	Seminars and workshops
Technique	Attitude scales ( <b>non-retrievable</b> )
Description	Purpose: To assess the degree of acquisition of basic knowledge of the subject, which includes managing of specific vocabulary, important concepts and current statistical tools, as well as the ability to synthesize and organize information from the literature. To stimulate the student to apply critical analysis and to develop speaking skills. Methodology: Students in groups of 2-3 will present to the rest of the class, a seminar of a critical comment (from a statistic point of view) of a scientific paper (or research work) chosen by the students with teacher approval.
Assessment criteria	It will be assessed the conciseness, accuracy and clarity of presentation, the oral communication, the use of terms learned during the course and the ability to relate the concepts learned. Also attendance and active involvement in the discussion of the classmate's seminars will be evaluated.

Final grade percentage: 20% for pathway A

Final grade percentage: 20% for pathway B

### Practical laboratory

Modality	Practical classes
Technique	Real or simulated task performance tests ( <b>non-retrievable</b> )
Description	Purpose: To apply the knowledge acquired in lectures. These workshops include solving exercises and problems, with the teacher help, on data management (organization, storage and treatment) and the application of various statistical tests (t-test, Chi square, ANOVA, correlations and nonparametric test) with hypothetical data related to the field of nutrigenomics. Methodology: Problem based learning. This activity will be held in the computer room in order to use specialized software for statistics (R). The teacher will guide the student to use the R program.
Assessment criteria	Resolution and discussion of the exercises proposed in class. It will be assessed by assistance, active participation, accomplishment of the tasks and quality of the answers.

Final grade percentage: 20% for pathway A

Final grade percentage: 0% for pathway B

### Synthesis exam

Modality	Assessment
Technique	Objective tests ( <b>retrievable</b> )
Description	Purpose: Perform a final exam of the synthesis of the subject. Methodology: Each examination should be done individually
Assessment criteria	Resolution of theoretical questions and problems individually. The statistical techniques used will be evaluated.  A minimum grade of 4 must be obtained in order to be able to balance with the other activities.

Final grade percentage: 40% for pathway A with a minimum grade of 3.5

Final grade percentage: 60% for pathway B with a minimum grade of 3.5

### Problem solving

Modality	Individual self-study
Technique	Extended-response, discursive examinations ( <b>retrievable</b> )
Description	Purpose: To assess the degree of knowledge of the student and the capacity to apply statistical tools correctly with a set of proposed problems related to studies in the field of nutrigenomics Methodology: Problem

Academic year	2019-20
Subject	10273 - Design of Experimental Studies and Statistics Applied to Nutrigenomics
Group	Group 1

Assessment criteria	based learning in an individual self-study way. To develop this activity the use of any statistical software, for solving some of the problems, will be required (R is available in all computer rooms of the University). Evaluation of all the contents of the subject by a set of proposed exercises. The questions will try to show that the student has understood the concepts and is able to connect properly the various aspects of the subject. Assessment: accomplishment of the tasks, delivery on time and quality of the answers.
Final grade percentage: 20% for pathway A	
Final grade percentage: 20% for pathway B	

## Resources, bibliography and additional documentation

### Basic bibliography

- 1) Course notes available on Campus Extens
- 2) Dalgaard, P. Introductory Statistics with R. Springer, 2002.
- 3) Altman, D.G. (1991). Practical statistics for medical research. Boca Raton: Chapman and Hall.

### Complementary bibliography

- 1) Ahrens, W. y Pigeot, I. (2005). Handbook of epidemiology. Berlin: Springer
- 2) Alberich, R., Mir, A., Rosselló, F. ActivaR: Introducció a R. Col·lecció Materials Didàctics 161.
- 3) Chongsuvivatwong, V. Analysis of epidemiological data using R and Epicalc. Recurso en the web: [https://cran.r-project.org/doc/contrib/Epicalc\\_Book.pdf](https://cran.r-project.org/doc/contrib/Epicalc_Book.pdf)
- 4) Peña, D. (2002). Análisis de datos multivariantes. McGraw-Hill.
- 5) Catena, A.; Ramos, M. y Trujillo, H. (2003). Análisis multivariado. Un manual para investigadores. Madrid: Biblioteca Nueva.
- 6) Cobo, E.; Muñoz, P. y Gonzalez, J.A. (2007). Bioestadística para no estadísticos. Bases para interpretar artículos científicos. Barcelona: Elsevier Masson.
- 7) Johnson, R. y Wichern, D. (2002). Applied Multivariate Statistical Analysis (5th edition). Upper Saddle River, NJ: Pearson Education.
- 8) Montgomery, D.C. (2001). Design and analysis of experiments. New York: John Wiley & Sons.
- 9) Rial, A. y Varela, J. (2008). Estadística Práctica para la investigación en Ciencias de la Salud. La Coruña: Netbiblo.
- 10) Tinsley, H. y Brown, S. (2000). Handbook of Applied Multivariate Statistics and Mathematical Modeling. San Diego: Academic Press