# Statistics

## Prof. Giuseppe Arbia; Prof. Elisa Fusco

***Text under revision. Not yet approved by academic staff.***

***COURSE AIMS AND INTENDED LEARNING OUTCOMES***

The course aims to help students understand the logical aspects of statistics and, subsequently, to be able to plan a survey with statistical rigour, summarise the information, analyse the results in an inferential manner, and prepare the relative summary reports. The course also aims to equip students to carry out their own statistical analyses using the calculator.

At the end of the course, students will be able to:

1. plan a statistical survey and identify the sample size needed and the data collection criteria to be used;
2. read critically and comment on statistical reports;
3. perform statistical analyses on the computer using the R statistical language;
4. comment on the output of the calculations made.

***COURSE CONTENT***

The course first introduces the basic concepts of descriptive statistics needed for the analysis and summarising of data related to one or more variables, through the creation of tables and graphs and the calculation of appropriate statistical indicators. After this preliminary part, students will be taught the conceptual aspects of empirical-inductive knowledge, with particular reference to its implications in the era of Big Data. In order to introduce the statistical-inductive procedure, the course will then address concepts related to probability theory and the random variables needed to rigorously programme the selection of units in a sample survey, and to infer conclusions that may be extrapolated from the observed sample on the entire population. In particular, the course will deal with problems related to estimating any unknown characteristics of the population under investigation and to analysing their significance. The theory of statistical tests will also be introduced. The course also deals with problems during the collection of statistical information in sample surveys. Within this context, the various sampling criteria, determining an optimal sample size for a predetermined research objective, and the various survey data collection criteria will also be introduced, including a discussion of the validity of data acquired via the Internet in the era of Big Data.

***READING LIST***

The course reference textbook is:

D.M. Levine-T.C. Krehbiel-M.L. Berenson, Statistica, Pearson Italia, 7th edition, with MyLab e Etext, Milan, 2018. (Edition provided with access code to the MyMathLab online platform).

It is essential that students refer to the textbook which will also serve as an exercise book both for carrying out tutorials in traditional form and online on the *MyMathLab* platform associated with the book.

For an introduction to the logical aspects of statistical-inductive analysis, the following text is recommended:

* G. Arbia (2018) Statistica, società e nuovo empirismo nell’era dei Big Data, Nuova Cultura. Anche nella sua edizione rivista ed aggiornata in lingua inglese: G. Arbia (2021) Statistics, new empricism and society in the era of Big Data, Springerbrief in statistics, Springer Verlag.

### For the use of package R and for the exercises, the following texts are strongly recommended:

### Espa , G. e Micciolo, R. (2012) Analisi esplorativa dei dati con R,  Apogeo Education.

### Espa, G. Micciolo, R. e Canal, L. (2021) Probabilità e modelli. teoria e pratica con R, Maggioli editore.

Further material will be distributed by the teachers during the course.

***TEACHING METHOD***

Ex-cathedra lectures and tutorials, during which the concepts introduced during lectures will be studied in depth by conducting practical exercises and examining case studies. The tutorials will be conducted by Dr Ilaria Valentini. From the first weeks the tutorials will make use of the R statistical. To this end, students are required to come to class with the R software and the R Studio user interface/Integrated Development Environment (IDE) installed on their laptops**.** Students will be regularly assigned assessment exercises to be carried out individually outside of lecture hours, making use also of the online platform associated with the textbook. These exercises will provide the starting point for further investigation and clarification during the subsequent tutorials.

***ASSESSMENT METHOD AND CRITERIA***

The course offers those students attending lectures the possibility of an interim test, to be carried out on the computer using the statistical language R. This test will take place during the seventh week of the course (30 October – 3 November). Dates and methods that will be clarified at the beginning of the course. The interim test will contribute 50% towards the final mark. Anyone who successfully passes the interim test (with a mark greater than or equal to 18) may take the final exam with an exemption from the topics covered in the first test. The interim test can be taken throughout but only in the winter exam session.

As with the interim test, the final test also involves carrying out computer exercises using the statistical language R. The final test may include an oral interview at the end of the practical computer exercise.

* Students who prefer not to take the interim test may take the full exam in a single test on one of the exam dates. This test may include an oral interview at the end of the practical computer exercise. Students must come to class with the R language statistical software and **the R Studio user interface/Integrated Development Environment (IDE) (come sopra)** installed on their PCs. Anyone who has difficulties is invited to communicate these to the lecturer during the first lecture of the course.
* The interim and final tests will be held in the Faculty’s computer labs.
* Students undertake to take the interim test and the final exam in accordance with academic honesty. To this end, students should be present on the day of the exam with both their mobile phone. Anyone found to be in possession of a mobile phone or PC connected to the internet or an external source will have their exam cancelled and their conduct reported to the competent academic authorities.
* Attendance and active participation in all parts of the course (including the taking of the interim test) are not compulsory, although strongly recommended. However, it is possible, as an alternative, to write just the final exam, which will contribute 100% to the final mark.

PREREQUISITES

The topics typically taught within a basic course of mathematical analysis are a prerequisite. In particular: Basic elements of mathematical logic and set theory. Algebraic calculation. Powers, logarithms, exponentials. First degree equations and inequalities. Systems of equations and inequalities. The prerequisites are an integral part of the course syllabus and it is the students' responsibility to ascertain they have them before attending the course and undertaking individual study.

Further information can be found on the lecturer's webpage at http://docenti.unicatt.it/web/searchByName.do?language=ENG, or on the Faculty notice board