# Statistics

## Prof. Maria Chiara Zanarotti

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

The aim of the module is to teach students the statistical foundations for studying phenomena in a social, political and economic context. In particular, the course will present the main tools of data analysis, both descriptive and inferential. It will take a theoretical approach to the methodologies presented making several numerical examples for each of the various contexts in order to emphasise the transversal nature of the tools in question. By the end of the course, students will be able to interpret the main results of the statistical surveys conducted, with careful and critical awareness of potential errors and/or poor use of data. Students will also be able to use such tools independently, managing to identify the most suitable for each context.

***COURSE CONTENT***

1. *Data collection*.Introductory concepts and terminology; total surveys and sample-based surveys; descriptive statistics and statistical inference; an overview of sampling.

2. *Distribution of a variable and its representation*.Data matrix. Distributions of frequency and quantity and their graphic representation. Historical series and territorial series and their graphic representation.

3. *Synthetic indices - Means.*Averages of position (median, mode and percentiles). Analytical mean (arithmetic mean and its properties; other analytical means)

4. *Synthetic indices - Variability.* The variability of a distribution. Indices based on variance from a mean (variability and its properties; coefficients of variation). Other indicators of variability (field of variation; interquartile difference). Standardisation. Gini Index.

5. *Joint probability distribution of two variables.*Joint, marginal and conditional probability distributions; marginal and conditional means and variation. Graphical representations.

6. *Analysis of the association between two variables.*Dependence and independence. Association between two variables and how to measure it. Mean dependence and how to measure it. Measuring the interdependence between two quantitative variables.

7. *Interpolation and regression.*Functions of regression. Linear regression of quadratic minima. Measuring goodness of fit. Multivariate regression.

8. *From description to statistical inference*. Sampling, randomness and representativeness.

9. *Chance, probability and random variables*. Random experiment, revising set theory, defining probability. Conditional probability, independence and Bayes’ theorem. Discrete and continuous random variables. Binomial random variables. Normal and standardised normal random variables.

10. *Sampling statistics and their distribution*. Sample mean and proportion. Sampling variability and sampling error. Central limit theorem. De Moivre-Laplace theorem.

11. *Point estimation*. Estimators and their properties. Mean squared error.

12. *Interval estimation*. Confidence interval for mean and proportion (large samples).

13. *Verifying hypotheses*. Statistical hypotheses: null hypotheses and alternative hypotheses. First- and second-type errors Simple and complex hypotheses. Significance tests. Level of significance. Verifying hypotheses for mean and proportion. Verifying hypothesies for interdependence.

***READING LIST***

Slides and texts available on the Blackboard page.

One of the following volumes:

S. Borra-A. Di Ciaccio, *Statistica. Metodologie per le scienze economiche e sociali,* McGraw-Hill, Milan, 2004, chaps. 1, 2, 3, 4 (up to 4.3), 5 up to 5.3; 6 (except 6.7), 8-9 (except 9.7.1, 9.7.2, 9.7.4, 9.8.1, 9.8.4, 9.8.5 and 9.11); 11 (except 11.4, 11.5, 11.8 and 11.9); 12 (except 12.4 and 12.6); 13 (up to 13.5); 14 (paragraphs 14.1, 14.2.1, 14.2.3, 14.3, 14.6).

F. Mecatti, *Statistica di base,* McGraw-Hill, Milan, 2010, chaps. 1, 2, 3, 4, 5, 6 (only 6.1), 7-8 (only 8.1 and 8.2), 9-10 (except 10.4 and 10.5 and 10.6), 11, 12, 13 (except 13.8 and 13.9). 14, 15, 16, 17 (up to 17.6), 18 (except paragraphs 18.3 and 18.6), 19 (except paragraphs 19.4, 19.9, 19.10, 19.14 and 19.15)

B. Pacini-M. Raggi, *Statistica per l’analisi operativa dei dati,* Carocci, Rome, 2006 (chaps. 1, 2, 3, 4, 5, 6, 7, 8.2, 9, 10 (except 10.3.3), 11 (except 11.4 and 11.5) and 12 (except 12.2.3, 12.2.4, 12.2.5, 12.3, 12.4.2, 12.4.3 and 12.5).

K. Pelosi-T. Sandifer, *Introduzione alla Statistica,* McGraw-Hill, Milan, 2005 (see “*NOTES AND PREREQUISITES*” for more details).

D.M. Levine-T.C. Krehbiel-M.L. Berenson, *Statistica,* Apogeo, Milan, 2002 (see *NOTES AND PREREQUISITES* for more details).

***TEACHING METHOD***

Lectures in class: introduction to the main data analysis methods by means of examples from a maieutic perspective to promote understanding of their underlying logical processes. As well as theory-based lectures, there will also be practical activities to help students assimilate the data-analysis techniques and prepare for the exam.

***ASSESSMENT METHOD AND CRITERIA***

Students will be assessed on their learning of the methodological tools presented in the course. The exam is written with close-ended questions. During the written exam, students can use a calculator and will also have access to any materials they may wish to bring into the exam with them (texts, notes, formulas, etc.). This is because the focus of the exam is the students’ ability to use the statistical analysis tools independently (to check the accuracy of the formulae and properties learned). Each question of the test is worth a certain (variable) number of points. The final mark is the sum of the points obtained and will be then changed into a mark out of thirty. Attending students will sit two partial tests; the final mark will be the sum of the results of the two tests changed into a mark out of thirty. Students may request a supplementary oral exam.

***NOTES AND PREREQUISITES***

The texts on the reading list are to be considered alternatives. Students are advised to download the lecture slides from Blackboard, as they cover the course topics in detail. The last two texts on the reading list are also available in English. Students should refer to the syllabus and the above information (re. lecture slides) to identify the sections to be studied and those to be omitted.

Students will also find topics and exercises for the exam (some with answers) on Blackboard to help them prepare for the written test.

As the course is introductory in nature, there are no specific prerequisites in terms of content: basic secondary-school mathematics will suffice.

In case the current Covid-19 health emergency does not allow frontal teaching, remote teaching and assessment will be carried out following procedures that will be promptly notified to students.

The Blackboard platform will also be used to communicate with students. Other notifications and communications regarding the course, exams or office hours (e.g. postponement or cancellation) can be found on the lecturer's webpage at http://docenti.unicatt.it/web/searchByName.do?language=ENG, the Faculty notice board and/or the Università Cattolica website.