# Probability and Statistics

## Prof. Alfredo Marzocchi

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

The main aim of the course is to introduce students to probability theory and inferential statistics, an essential tool for obtaining accurate information from variously distributed data. At the end of the course, students should know the main concepts of probabilistic and statistical methods and their applications. From the ability to apply knowledge and understanding perspective, students will be able to develop a statistical test suitable for a certain situation, motivating the choice with respect to any alternatives, and interpreting the possible results by drawing appropriate conclusions.

***COURSE CONTENT***

Recalls of probability theory. Random variables: discrete and continuous. Expectation, variance, and conditional expectation. Main univariate probability distributions. The change-of-variable formula. Random vectors and multivariate distributions. Expectation vector, variance-covariance matrix, correlation matrix. Multivariate normal and multinomial distributions. Convergence of random variables, law of large numbers, central limit theorem.

- Multivariate random variables. Marginal distributions. Independence and correlation. Conditional distributions. Conditional expectation. Examples.

- Functions of random variables: the general case and special examples.

- Sampling and sample variables. Sample space. Simple random sampling. Likelihood function.

- Statistics. Sample mean and sample variance. Exact and approximate distributions of the sample moments. Rao-Cramèr inequality. Sufficiency. Subordinate statistics and equivalent statistics. Minimal sufficiency. Complete statistics. Sufficiency and exponential family.

- Estimation. Estimator and estimate. Comparison of estimators.

- Methods of estimation. Method of moments. Maximum likelihood method. Maximum likelihood and exponential family. Asymptotic Optimality of the maximum likelihood method.

- Confidence intervals. Exact confidence intervals for the mean and the variance of a normal distribution~~s~~. Asymptotic confidence intervals: the discrete and continuous cases. Examples.

- Theory of statistical hypothesis testing. Neyman-Pearson lemma. Exact parametric tests for the mean and variance in the normal case. Chi Square test.

- Introduction to linear models.

- Introduction to non-parametric estimation methods.

***READING LIST***

- A.M. Mood, F.A. Graybill, D.C. Boes, *Introduction to the theory of Statistics,* Mc Graw-Hill, 1973.

Additional notes will be provided during the course.

***TEACHING METHOD***

Frontal, in-presence recorded lectures with ample opportunity for interaction with students, accompanied by tutorials with examples of the topics covered. Possible variations may be due to the evolution of the pandemic.

***ASSESSMENT METHOD AND CRITERIA***

Oral, written and laboratory. The oral exam consists of a short interview at the blackboard in which the written exam is critically examined and the main concepts contained in it reviewed. The laboratory part consists of the discussion of a script written in Python or R languages. Methods of assessment may vary in relation with the evolution of the pandemic.

The final assessment will focus not only on the written exam and laboratory, but also on relevance of students' answers, their appropriate use of specific terminology, well-argued and coherent structuring of their argumentation, and their ability to identify conceptual links and respond to open questions.

***NOTES AND PREREQUISITES***

No prerequisites, apart from high-school mathematics, are required. A pre-course to refresh the main mathematical tools needed in the Course will be offered in September. 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.