# Statistics (Data analysis and probability)

## Professor Umberto Carugati

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

The aim of the course is to provide students with quantitative knowledge and abilities to construct and interpret synthetic indexes for data arising in economic, business and social phenomena. The course will provide also instruments to analyse the relationships between two or more variables, as well as basic elements of probability theory and random variables.

The following learning abilities are provided and expected to be achieved by participants at the end of the course:

1. Knowledge of concepts, terms and methods concerning descriptive statistics and probability.
2. Ability to correctly apply descriptive statistics and probability methods to problems arising in economics and management applications.
3. Quantitative thinking to make independent judgments driven by descriptive and probabilistic statements.
4. Ability to read, interpret and communicate data-driven results.
5. Mastery of tools useful for quantitative analyses in courses that the students will took later in the curriculum, as well as for simple quantitative analyses required in their careers involving management of data, rigorous reasoning and data-driven decision-making.

***COURSE CONTENT***

*Statistical methodology and scientific research.* The statistical analysis process: gathering and describing data. Frequency distributions. Graphical representations.

*Univariate analysis.* Location indices: mode, median and algebraic means. Indices to measures dispersion: indices of dissimilarity (mutability/heterogeneity) and variance. Remarks on Concentration*.* Symmetry/skewness indices.Box-plot

*Multivariate analysis.* Stochastic independence.Association, dependence and correlation.Analysis of conditional distributions.Measures of association. Measures of dependence. Study of the on-average relationship between two variables. The regression function. Decomposition of total variance and the correlation ratio.Measures of linear dependence. The simple linear regression model: assumptions, parameter estimation and the linear correlation coefficient. Brief introduction to multiple linear regression model.

*Elements of probability theory.* Definition of probability and basic theorems.Univariate probability models (Binomial, Hypergeometric, Normal).The Central Limit Theorem.

***READING LIST***

G. Cicchitelli. P. D’Urso, M. Minozzo  *Statistica: principi e metodi,* Pearson, Milano, 2018 (3a ed.). [The topics of the course are covered in chapters 1-14].

L. Deldossi-R. Paroli, Lezioni di Statistica, II Ed., Giappichelli, 2015

A similar English textbook can be suggested on demand.

***TEACHING METHOD***

Lectures and practical exercises.

***ASSESSMENT METHOD AND CRITERIA***

The assessment consists of a written exam to complete in 90 minutes and is made up of:

1. Four exercises. The first exercise is referred to univariate statistics and the study of the association between two or more variables; the second one concerns the concentration; the third one is related to the regression topic and the fourth one to the probability.
2. Eight multiple choice theoretical questions.

Students have the possibility to substitute the exam with two partial tests which equally contribute to the final evaluation. The midterm exam is planned during the class suspension period between the two semesters. More detailed information about the assessment process are available on the e-learning platform *Blackboard.*

The aim of the exam is to assess reasoning and analytic abilities on the course subjects. Properties of language and communication abilities are also assessed.