UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Strumenti Informatici per l'Analisi dei Dati

2526-1-ESM02Q030

Aims

The course introduces the main tools and methodology for analyzing, visualizing, and interpreting experimental data in scientific filed, with particular attention to examples and case studies from the course "Elementi di Metodo Sperimentale". Students will acquire practical skills in data management, the use of spreadsheets, and the application of fundamental statistical techniques.

Contents

1. Introduction to scientific data analysis
   Organisation of data files (CSV, TSV, TXT, etc.), overview of basic software for scientific data processing

2. Data management and cleaning
   Import/export, common formats, preprocessing.

3. Descriptive statistics
   Mean, standard deviation, histograms, xy graphs

4. Main statistical distributions
   Reconstruction of normal, binomial (with dice roll simulation) and Poisson distributions.

5. Analysis and interpretation
   Linear regression, estimation of slope and uncertainty, comparison with models (chi-square method)

Detailed program

1. Introduction to scientific data analysis
   Introduction to the concept of scientific data and its role in experimental activity. Organisation and structure of data files, with particular attention to the most common text formats (CSV, TSV, TXT). Overview of the main software tools used for scientific data processing and analysis and the typical workflow that leads from raw data to the interpretation of results.

2. Data management and cleaning
   Techniques for importing and exporting data to and from text files. Management of the most common formats and preparation of datasets for analysis. Identification and treatment of missing or invalid data, preprocessing and data transformation operations to make them suitable for statistical analysis.

3. Descriptive statistics
   Introduction to descriptive statistics for the analysis of experimental data. Calculation and interpretation of means and standard deviations. Construction and analysis of histograms and Cartesian (xy) graphs for data representation and visualisation of relationships between physical quantities.

4. Main statistical distributions
   Study of the main probability distributions of scientific interest. Reconstruction of the normal distribution from experimental data. Introduction to the binomial distribution, with examples and numerical simulations (e.g. dice throwing), and to the Poisson distribution for the description of rare events (e.g. description of the number of goals in a match). Comparison between theoretical distributions and observed data.

5. Data analysis and interpretation
   Analysis of relationships between physical quantities using simple models. Linear regression of experimental data and estimation of model parameters. Evaluation of uncertainties in slope and intercept. Comparison of experimental results with theoretical models using the chi-square method and critical interpretation of results.

Prerequisites

Basic mathematics (algebra, functions, elementary derivatives), no programming experience required.

Teaching form

Lectures with practical examples of statistical data processing.

Textbook and teaching resource

The spreadsheets generated during the lessons will be uploaded to the e-learning platform along with a brief guide to the main commands used in the lessons.

Semester

First and secon semester of the first year.

Assessment method

Students can earn the relevant credits (the exam is not graded, only pass/fail) in two different ways:

1. Complete and return all the exercises provided in the shared folder to the lecturer
2. During the laboratory part of the Elements of Experimental Method course, students must demonstrate, at the end of the various experiments, that they have correctly used the IT tools for processing experimental data

Office hours

By appointment with the lecturer: roberto.lorenzi@unimib.it

Sustainable Development Goals

QUALITY EDUCATION