UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Probability and Statistics

2627-1-E3004Q005-E3004Q00501

Aims

The purpose of this class is to provide students with all the theory and practical tools to understand the fundamentals of measurement theory and uncertainty analysis, and to be able to perform simple data analysis on datasets at their disposal.

Contents

• Measurement theory and uncertainty analysis
• Estimators
• Error propagation
• Probability: definitions and usage
• Notable distributions (binomial, Poisson, Gaussian)
• Inference: (linear) regression
• Least squares method
• Regression examples using scikit-learn

Detailed program

Measurement theory and uncertainty analysis: general notions on the concept of experimental measurement, and on reducible and irreducible errors. Definitions of relative and absolute error. Definition of bias and statistical fluctuations. Introduction to statistical and systematic errors.
Estimators: distinction between error and uncertainty. Estimate of errors. Definition of the main estimators and their properties: variance and bias.
Error propagation: motivations and main techniques (analytical and numerical).

Probability, definition and usage: frequentist probability, its definitions, and its fundamental properties. Brief introduction to Bayesian probability.
Notable distributions: definitions and usage of some of the most used probability distributions, mainly binomial, Poissonian and Gaussian. Estimators evaluation for these distributions.
Inference and regression: introduction to statistical inference and regression on a dataset. Linear and quadratic regressions, and their properties. Fundamental concepts on fit quality evaluation (chi-squared, Fisher discriminant).
Least squared method: definition of loss function. Main advantages and disadvantages of the method, and possible alternatives.

Regression examples with scikit-learn: introduction to the python scikit-learn library and to its usage to solve regression problems. Simple practical examples.

Prerequisites

Fundamental mathematics notions. Introductory knowledge of python.

Teaching form

Formal introduction of statistical concepts, with mathematical derivation of the main properties of the tool introduced. Practical examples of error theory, solving simple problems.

The course is thus divided:

- 40 hours on the theoretical parts of the program
- 12 hours of laboratory on the practical aspects of data analysis using scikit-learn.

About 20 hours of remote teaching is foreseen, in synchronous mode.

Textbook and teaching resource

“An introduction to error analysis”, J. Taylor

Semester

Second semester

Assessment method

The assessment relies on an oral exam. During the examination, the instructor evaluates the student's learning level and the communication capabilities pertaining to the specific field... There will be no intermediate tests.

Office hours

Sustainable Development Goals