Postgraduate Course
Statistical Modelling (MSc)

Instructor Information

Pedro J. Zufiria

E-mail: pedro.zufiria@upm.es
Work Phone: +34 91 336 72 84

Course Information

Course Description

This course provides the foundations of the statistical analysis of data. Starting with an overview of the fundamentals of probability theory, the course then develops the basic concepts and tools related to classical statistical inference such as parameter estimation, confidence intervals and hypothesis tests. Finally, within this statistical framework, the course characterizes the regression, classification and density estimation problems, as a fundamental background towards the study of statistical learning techniques.

Prerequisites

Linear Algebra. Calculus of Several Variables.

Course Goal

To develop an understanding of the fundamental concepts and mathematical techniques required for constructing and employing statistical models.

Summary of intended course outcomes

The students will know how to define probability models for characterizing random phenomena, being aware of the assumed explicit or implicit inductive biases. In addition, they will be able to construct such type of models from data, as well as to test existing ones, assessing their quality. By the end of the course, students will be able to:

- Describe, model and analyze data.
- Develop and test generative and discriminative models for description and/or prediction of the behavior of systems with uncertainties.
- Be familiar with the foundations of statistical learning theory.

Syllabus

Introduction
Engineering and statistical modelling.
Foundations of probability.
Probability spaces.
Discrete and continuous random variables.
Joint distributions.

Descriptive statistics and data visualization
Random sampling: sample mean, median, range and variance.
Histograms, stem and leave diagrams, boxplots and basic time series graphs.

Sampling distributions and parameter estimation
Point estimation.
Sample distribution and Central Limit Theorem.
Unbiased estimators. Variance of a point estimator; mean square error.
Point estimation methods: Moments, Maximum likelihood, Bayesian.

Statistical intervals for a sample
Confidence intervals for mean and variance of a normal distribution.
Confidence interval for population proportion.
Tolerance and prediction intervals.

Hypothesis testing
Tests on the mean and variance of a normal distribution.
Tests on a population proportion.
Bayesian hypothesis testing.

Simple linear regression and correlation
Simple linear regression.
Correlation analysis.

Towards statistical learning
Generative and discriminative modelling.
Regression and classification: supervised learning.
Density estimation: unsupervised learning.

Bibliography

- Applied Statistics and Probability for Engineers. Fifth Edition. Douglas C. Montgomery and George C. Runger. Wiley & Sons, 2011.
- Statistical Inference. Second Edition. George Casella and Roger L. Berger. Thomson Learning, 2002.
- The Nature of Statistical Learning Theory. Vladimir N. Vapnik. Springer Verlag, 2000.

Student Assessment Criteria

Take-home assignments to be solved using analytical and computational tools (R) will the regularly provided along the course (30% of the final grade). In addition, the students will take an exam (30%) and will develop a final project (40%).