Statistics lies at the foundation of data science, providing a unifying theoretical and methodological backbone for the diverse tasks encountered in this emerging field. This course rigorously develops the key notions and methods of statistics, with an emphasis on concepts rather than techniques.

Keywords
Data science, inference, likelihood, regression, regularisation, statistics.

Learning Prerequisites
Required courses
Real analysis, linear algebra, probability.

Recommended courses
A first course in statistics.

Important concepts to start the course
Students taking the course will need a solid grasp of notions from analysis (limits, sequences, series, continuity, differential/integral calculus) and linear algebra (linear subspaces, bases, dimension, eigendecompositions, etc). Though the course will cover a rapid review of probability, a first encounter with the subject is necessary (random variables, distributions/densities, independence, conditional probability). Familiarity with introductory level notions of statistics would be highly beneficial but not necessary.

Learning Outcomes
By the end of the course, the student must be able to:
• Derive properties of fundamental statistical procedures
• Estimate model parameters from empirical observations
• Test hypotheses related to the structural characteristics of a model
• Construct confidence bounds for model parameters and predictions
• Contrast competing models in terms of fit and parsimony

Assessment methods
Final exam.

Resources

Bibliography