Nonparametric theory

Phd Programme in Economics, Statistics and Data Science, DEMS, University of Milano-Bicocca

The Ph.D. in Economics, Statistics and Data Science and the Department of Economics, Management and Statistics of the University of Milano-Bicocca are pleased to announce the course Nonparametric theory held by Professor Szabó Botond, Bocconi University, during the period September 14 - October 24, 2022.

Schedule

- 14/9 10.30-12.30 (Room: U7-13) and 14.30-16.30 (Room: U7-13)
- 23/9 10.30-12.30 (Room: U7-13) and 14.30-16.30 (Room: U7-13)
- 10/10 10.30-12.30 (Room: U7-16) and 14.30-16.30 (Room: U7-16)
- 12/10 10.30-12.30 (Room: U6-26) and 14.30-16.30 (Room: U6-01e)
- 17/10 10.30-12.30 (Room: U7-16) and 14.30-16.30 (Room: U7-16)
- 19/10 10.30-12.30 (Room: U6-26) and 14.30-16.30 (Room: U6-01e)
- 24/10 10.30-12.30 (Room: U7-16) and 14.30-16.30 (Room: U7-16)

Lectures will be delivered in-class at the University of Milano-Bicocca, but there may be changes due to coronavirus.

Instructor

Botond Szabó is Associate Professor at Bocconi University, Milano. Previous to that, he was Associate Professor at VU Amsterdam and at University of Leiden. Before starting his position at Leiden, he was a postdoctoral fellow at CREST (Paris) and University of Amsterdam, and an assistant professor at Budapest University of Technology and Economics. In 2022 Szabó obtained a €1.5mln ERC Starting Grant for a project (*BigBayesUQ - The missing mathematical story of Bayesian uncertainty quantification for big data*) aimed at developing mathematical techniques capable of assessing the uncertainty inherent in the estimates derived from machine learning algorithms and, consequently, their reliability in context of statistical models. Currently he is associate editor of Annals of Statistics and Bayesian Analysis. `https://botondszabo.com/`

Enrollment

The course Nonparametric theory is part of the Statistical Learning course of the PhD in Economics, Statistics and Data Science, but it is open to all PhD students, researchers, academics, professionals and practitioners with strong quantitative background and motivation (see Prerequisites). Lectures will be delivered **in-class**.

Application period: September 1, 2022 - September 12, 2022. In order to apply you must fill in the following form:

`https://forms.gle/zQmvASAe3CUmP9T9A`

Admissions are conditional on place availability. **No fees are requested**. The ECOSTATDATA Administration Office will contact non-admitted candidates only.

Course objectives

Nonparametric theory.

The goal of the course is to provide a rigorous introduction to modern, nonparametric and high-dimensional statistical inference. After briefly recalling the results of parametric inference and introducing the Bayesian counterpart of the results we turn our attention to nonparametric inference. We introduce benchmark models and consider various standard estimators over them. We derive minimax theoretical lower bounds and show that well scaled methods achieve these optimal rates. Then we consider the Bayesian counterpart and derive posterior consistency and contraction rate results. Next we consider high-dimensional statistical models and investigate the standard frequentist and Bayesian methods over them. Finally we address the issue of scalability arising due to the ever increasing information amount and more complex models. We discuss two standard methods, the distributed and variational Bayes approaches and derive theoretical guarantees for them.

- Parametric inference:

  - Recalling asymptotic normality of MLE, and optimality guarantees of parametric models
  - Bayes in parametric models, Bernstein-von Mises theorem
  - Non admissibility of MLE, James-Stein's estimator.
- Nonparametric inference:
  - Benchmark models (Gaussian white noise, nonparametric regression and density estimation) and their connections
  - Haar wavelets, sequence estimators, adaptation with wavelet thresholding
  - Kernel estimators, adaptation with cross-validation
  - Minimax estimation lower bounds
- Nonparametric Bayes
  - Posterior consistency (Doob's and Schwartz's theorem)
  - Posterior contraction rates
  - Examples: Dirichlet Process mixtures and Gaussian Processes (the latter extensively)
  - Adaptation: empirical and hierarchical Bayes
- High-dimensional inference
  - LASSO, ridge regression in linear and logistic regression
  - Covariance and precision matrix estimation, matrix completion
  - Bayesian methods (spike and slab prior and horseshoe prior)
- Inference with scalable (Bayesian) statistical methods:
  - Variational Bayes: general contraction rate theorem, application in Gaussian process regression
  - Distributed Bayes: mixture versus product of experts

Lecture notes will be provided. Recommended reading:

- Bijma, Jonker, Van Der Vaart: Introduction to Mathematical Statistics
- Alexandre Tsybakov: Introduction to Nonparametric Estimation
- Ghoshal, van der Vaart: Fundamentals of Nonparametric Bayesian Inference (the chapter on consistency and contraction rate)

Prerequisites
The overall flavour of the course overall will be theoretical and methodological. The material should be understandable for anyone with a thorough knowledge of undergraduate Statistics, probability, real analysis and linear algebra.

Contacts
For more information:
Prof. Matteo Manera, Coordinator of ECOSTATDATA, e-mail: matteo.manera@unimib.it
Prof. Aldo Solari, ECOSTATDATA faculty member, e-mail: aldo.solari@unimib.it
Dr. Tommaso Rigon, ECOSTATDATA faculty member, e-mail: tommaso.rigon@unimib.it

For administrative issues:
Dr. Silvia Locatelli, ECOSTATDATA Administration Office, e-mail: phdecostat@unimib.it