MATH-522 **Empirical processes**

Limnios Myrto

Cursus	Sem.	Type
Ing.-math	MA2, MA4	Opt.
Mathématicien	MA2	Opt.
Statistics	MA2, MA4	Opt.

Language of teaching	English
Credits	5
Session	Summer
Semester	Spring
Exam	Written
Workload	150h
Weeks	14
Hours	**4 weekly**
Lecture	2 weekly
Exercises	2 weekly
Number of positions	

Summary

From prototypical examples of estimators used by statisticians, to more complex nonparametric models, methods and theorems will be taught to study their (non)asymptotic behavior, when defined as mappings depending on an empirical measure based on a sample of random observations.

Content

This course provides central results and methods to understand and derive generalization guarantees of modern learning algorithms. The main topics are:

- Fundamental limit theorems for classes of empirical processes.
- Basic probabilistic inequalities: methods and results.
- Symmetrization and chaining methods.
- How to measure the complexity of classes of functions: from combinatorial and entropic considerations to Rademacher averages.
- Concentration-of-measure phenomenon: maximal inequalities and concentration inequalities of suprema of empirical processes.

The exposed methods and results will be illustrated with statistical applications related to machine learning and hypothesis testing throughout the lecture and exercise sessions.

Keywords

Empirical processes, limit theorems, concentration of measure, probabilistic inequalities, generalization properties, learning algorthims.

Learning Prerequisites

Important concepts to start the course

Previous courseworks in analysis, mathematical statistics and probability are highly required, with interests in statistical learning theory and algorithms.

Learning Outcomes

By the end of the course, the student must be able to:

- Formulate the fundamental framework related to empirical processes
- Manipulate probabilistic inequalities to empirical estimators

- Describe the concept of concentration-of-measure phenomenon for empirical processes
- Apply concentration inequalities to derive the performance of statistical learning procedures

Teaching methods

The semester will be composed of:

- Lectures using slides and blackboard (for some proofs and examples in particular).
- Exercise sessions.

Expected student activities

Student are expected to:

- Attend to all lectures.
- Participate to exercise weekly sessions.

Assessment methods

Written exams (midterm and final).

Resources

Bibliography

- Stéphane Boucheron, Gábor Lugosi, and Pascal Massart, *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, Oxford Academic, 2013.

- Sara van de Geer, *Empirical Processes in M-Estimation*. Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 2009.

- Aad W. Vaart , Jon A. Wellner, *Weak Convergence and Empirical Processes. With Applications to Statistics*. Springer Series in Statistics (SSS), Springer New York, 1996.

- Víctor H. Peña , Evarist Giné, *Decoupling. From Dependence to Independence*. Probability and Its Applications (PIA), Springer New York, 1999.

Moodle Link

- https://go.epfl.ch/MATH-522