Seminar on empirical processes, SP 4

Dozent/Art: Axel Munk/Seminar

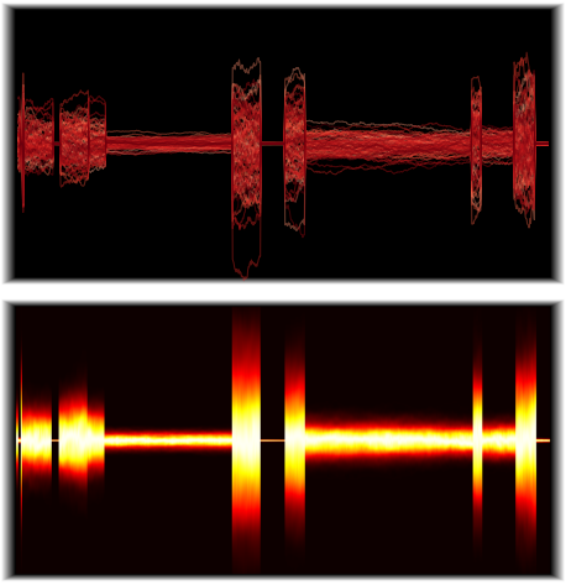
Zeit: Fridays, 10:15-11:45

Ort: Seminarraum 5.101, Institut für mathematische Stochastik, Goldschmidtstr. 5-7

Zielgruppe/Sprache: Advanced BSc and beginning MSc students. Language is English. Possible Modules: B.Mat.3444: Seminar on mathematical statistics, M.Mat.4844: Seminar on mathematical statistics, B.Mat.3447: Seminar on statistical foundations of data science, M.Mat.4847: Seminar on statistical foundations of data science.

Vorkenntnisse: Mandatory: Prob. & Measure Theory (B.Mat.1400). Useful in parts, but not mandatory: Stochastics (B.Mat.2410), Statistical data science (B.Mat.2420), and/or Statistical foundations of data science (B.Mat.3147).

Beschreibung



Visual representation of an M-process with discontinuities. Top: Sample trajectories, Bottom: Estimated probability intensity.

For many tasks of modern data analysis the data do not just occur as vectors, but rather as functions or even more generally, as random objects taking values in metric spaces. Such data have an inner structure which is to be respected for a sensible data analysis, e. g. each data point can be described as a metric graph or satisfies certain geometric properties. Many important statistical quantities can be written as functionals of such data. A simple but fundamental example is the empirical distribution function (on the real line) which leads (after centering and rescaling) to a stochastic process taking values in the space of bounded functions — the empirical process. The sample mean (after centering and rescaling) of real valued data then is a simple functional (an integral) of the empirical process. One possible path to studying the properties of such a statistic is to separately investigate the properties of the functional (mainly using functional analytic tools) and of the underlying empirical process. Classical Glivenko-Cantelli and Donsker theorems establish almost sure uniform and weak convergence of the empirical process, respectively, and from this properties of the sample mean and also of more complex functionals of the empirical process can be derived. The modern take of empirical process is to view the empirical process as a process indexed in indicator functions on intervals $(-\infty, t]$, which describes the complexity of this process. This view leads to remarkable generalizations to empirical processes indexed in function spaces. The theory developed to prove them provides powerful techniques that can be employed to understand the properties of modern statistical methods (e. g. bootstrap) in a broad range of scenarios. These techniques are fundamental to statistical and machine learning theory and have manifold applications nowadays.
In this seminar we cover the basic relevant mathematical concepts and main results of empirical process theory. Our focus will be on carefully understanding the basic principles rather than obtaining results

in most generality. Topics include: properties of sub-Gaussian random variables such as concentration of measure, maximal and log-Sobolev inequalities, Dudley's entropy integral, bracketing, Talagrand's chaining, inequalities for suprema of empirical processes, weak convergence in separable and non-separable metric spaces. As an application of the theory, properties of M-estimators are derived.

Application

To provide participants with the material to be presented at an early stage, we ask you to preregister for this seminar. To this end, please email Jan Victor Otte (`mail:janvictor.otte@stud.uni-goettingen.de`) and indicate your interest to give a seminar talk. Please include information about relevant courses you have taken in your email. Deadline for preregistration is Wednesday, 12 March 2025.
A preparatory virtual meeting, during which topics will be assigned to participating students, is scheduled for Monday, 17 March 2025 (11:00-12.30am).

Literatur

Topics for presentations will be assigned along the lines of

- Sen, B. (2018): Lecture Notes "A Gentle Introduction to Empirical Process Theory and Applications". Available for download via `http://www.stat.columbia.edu/~bodhi/Talks/Emp-Proc-Lecture-Notes.pdf`

References for further reading

- Gine, E., Nickl. R. (2016). Mathematical Foundations of Infinite-Dimensional Statistical Models. Cambridge Series in Statistical and Probabilistic Mathematics no. 40.
- Kosorok, M. R. (2008): Introduction to Empirical Processes and Semiparametric Inference. Springer.
- Pollard, D. (1990). Empirical Processes: Theory and Applications, IMS Hayward, CA.
- Shorack, G. R. and Wellner, J. A. (2009). Empirical Processes with Applications to Statistics. SIAM.
- van der Vaart, A. W. and Wellner, J. A. (2000). Weak Convergence and Empirical Processes: With Applications to Statistics. Springer.