

Special Topics Course

Statistics 700: Bayesian nonparametrics

Instructors: Jayaram Sethuraman and Long Nguyen
Department of Statistics
Fall 2014
2:00–4:00 pm Friday

Course description. This course provides an introduction to Bayesian nonparametrics (BNP). BNP is an area of statistical modeling and inference with highly flexible models whose complexity grows appropriately with the amount of data. It has a rich literature benefiting from developments in stochastic processes, computational algorithms, asymptotic theory, machine learning, and a variety of modern applications. Topics to be covered include Dirichlet and related processes, hierarchical modeling with nonparametric prior distributions, Markov Chain Monte Carlo and variational inference, Bayesian asymptotics, and applications in information retrieval and biostatistics.

Prerequisites. The prerequisites are previous courseworks in graduate-level probability (at the level of Stats 620 and 621), or the instructor’s approval (Contact Prof. Nguyen for details).

Structure/Evaluation. The class will meet once a week and will follow a regular lecture format. Students will be required to present a research paper in the final weeks in order to receive a grade for the course.

Course homepage. All announcements, reading materials and project information will be posted at the course’s Ctools site.

Textbook. There will be no textbook for the class. Lecture notes and relevant research papers will be provided. Nonetheless, a number of books may be useful for reference and learning purposes, and are highly recommended:

1. Bayesian nonparametrics. Edited by N. L. Hjort, C. Holmes, P. Mueller and S. Walker. Cambridge University Press, 2010. This book is a collection of extremely well-written survey papers providing a broad introduction to the field.
2. Prior processes and their applications. Eswar G. Phadia. Springer, 2013. A timely treatment with emphasis on statistical modeling.
3. Bayesian nonparametrics. J. K. Ghosh and R. V. Ramamoorthi, Springer 2002. Primary emphasis on statistical theory; the first four chapters serve as an elegant introduction to Dirichlet processes and issues arising in Bayesian asymptotics.

Tentative outline.

- Introduction to Dirichlet processes (Ferguson, Blackwell-MacQueen, Sethuraman), and extensions (e.g. Poisson-Dirichlet processes)
- Completely random measures and models for discrete structures (e.g., Beta processes, Indian Buffet process,...)
- Mixture and hierarchical modeling with Bayesian nonparametric priors.
- Algorithms for BNP models (Markov chain Monte Carlo, variational inference).
- Asymptotic theory and connection to optimal transportation.
- Topics in exchangeability, partial exchangeability (for matrices and graphs).
- Topics on random partitions, fragmentation-coagulation processes and applications.