Bocconi University – PhD in Statistics - academic year 2021-2022

Bayesian Theory II

Instructor: Sonia Petrone

Course description.

The aim of this course is to discuss and dig further on the main fundamental principles of Bayesian statistics.

This means, first of all, having at least a glance on the foundations of Bayesian Theory - on the decisional-theoretic work to establish an axiomatic theory meant as normative formalization of agents' rational behavior towards decisions with incomplete information and risk; which leads to the subjective notions of utility and probability.

We then focus on the 'Theory of Prediction' of Bruno de Finetti. Needless to say, the central role of prediction, rather than inference on parameters, is very 'modern' in the ongoing debate in statistics and machine learning. Here we underline the role of exchangeability as a subjective judgment, that gives the basic scheme for Bayesian predictive learning; and present theoretical properties of the predictive distributions for exchangeable sequences, and implications in Bayesian statistics.

We will point out the breakthroughs in the 1950's, that further enhanced Bayesian procedures even from a classical perspective, to get to some powerful ideas in the theory of Bayesian statistics and pave the way to further study of Bayesian methods beyond exchangeability and modern challenges.

Syllabus

- **Part I: Foundations of Bayesian Theory: historical excursus and main principles**

  - Decisional foundations: historical background (*lecture notes and references therein*)
  - Bruno de Finetti. The role of prediction. (*de Finetti (1937); Parmigiani, Inue(2009)*)
  - Exhangeability. (*the most complete reference is Aldous, 1985. See lecture notes as a guideline*).
  - *Physical* or *subjective* exchangeability? Reasoning on Pólya urns and Bernoulli urns.
    A very current debate: is Bayesian Theory the study of frequentist properties of Bayesian procedures?
  - Properties of the predictive distributions for exchangeable sequences.
    - Weak convergence (*with proof*). A partially reverse implication: convergent predictive distributions and asymptotic exchangeability.
    - de Finetti representation theorem in terms of predictive distributions.
    - Higher order convergence. Central limit theorems for exchangeable sequences (*with proof: Aldous, pp. 16-17*). Stable convergence. Almost sure convergence of conditional probabilities.
  - What are Bayesian credible intervals? On the notion of frequentist coverage of Bayesian credible intervals.

- **1950's breakthroughs**.

  - Wald statistical decision functions. Minimax and expected utility (loss) principles.
    Bayesian decision rules in normal form and in extensive form.
    Admissibility and complete classes (*no proofs*) (*use lecture notes as a guideline. Parmigiani & Inue, Chapter 7). Of course Wald, 1949*
  - Stein's 'paradox' and shrinkage estimators. Empirical Bayes reformulation. (*Efron, 'Large-scale Inference', IMS Monographs, Ch 1. Of course Stein, 1956 and H. Robbins, 1956*)
  - Consistency and semiparametric inference (*Kiefer & Wolfowitz, 1956*).

- **Bayesian solutions and 'modern' developments..**

  - Bayesian mixture models. Inference on the latent distribution.
  - Bayesian shrinkage estimators (Gaussian).
  - "Empirical Bayes" choice of prior hyperparameters.
  - An introduction to state-space models. - (*if time permits*)

Exam: assignments and oral exam