Syllabus for STAT 700, Fall 2024

Eric Slud 8/26/2024

This course introduces mathematical statistics at a graduate level, using tools of advanced calculus and basic analysis. The framework is to define families of probability models for observed-data structures and explain the sense in which functions of observed-data random variables can give a good idea of which of those probability models governed a particular dataset. The course objective is to treat diverse statistically interesting models for data in a conceptually unified way; to define mathematical properties that good procedures of statistical inference should have; and to prove that some common procedures have them. Aspects of the theoretical results are illustrated using demonstrations with statistical simulation. **The fall term (Stat 700) will primarily emphasize definitions, concepts and finite-sample optimality properties of estimators, while the spring term (Stat 701) emphasizes large-sample asymptotic theory and more modern topics.**

Course Coverage: STAT 700–701 divides roughly with definitions and properties of finite-sample statistics in the Fall (STAT 700), and large-sample limit theory in the Spring (STAT 701). The division is not quite complete, because we motivate many topics (Point Estimation, Confidence Intervals, identifiability) in terms of the Law of Large Numbers. The coverage in the Bickel & Doksum book for the Fall is roughly Chapters 1-4 along with related reading in Casella & Berger for special topics.

Outline of Topics:

We begin with an overview of statistical data structure, models and formal definition of statistics in Chapter 1 (Secs. 1.1.1-1.1.3). Succeeding lectures will review standard background material on probability and standard distributions (Appendix A, sections A.10-A.14) in order to set up later material on Exponential Families (Section 1.6). Additional review of types of stochastic convergence (Appendix B.7) and Limit Theorems (Appendix A.15). Brief review of basic statistical definitions will be done from the viewpoint of Decision Theory (Section 1.3). Introduction of the Bayesian viewpoint on statistical inference (Section 1.2) is naturally done in that context, including Bayesian mechanics. The other important material in chapter 1 concerns the notion of "sufficient statistics" and "prediction" versus "estimation" (Section 1.4).

Chapter 2 covers the main estimation techniques, (generalized) method of moments, maximum likelihood, and Estimating Equations as a way to unify these two different-seeming methods in a general framework. Computational topics (algorithms, including numerical maximization and EM) for the solution of Maximum Likelihood and Estimating Equation problems are also covered in Chapter 2.

Chapter 3 discusses notions of performance quality and optimality for statistical estimation procedures. In this connection we cover variance lower bounds (Cramér-Rao lower bound), minimum variance unbiased estimation and some decision-theoretic (Bates-optimality, minimax) optimality.

Chapter 4 introduces basic ideas and optimality principles related to hypothesis testing (Neyman-Pearson Lemma, UMP tests, MLR density families)

Readings in Casella and Berger will be occasional and topic-based. Some introductory Bayesian topics will be covered there, and basics on MCMC may also be discussed as part of Chapters 1 and 2 of Bickel and Doksum augmented by pdf handouts.

Assigned work and Grading Policy are covered on the Course Web-page `http://www.math.umd.edu/~evs/s700` **and the STAT 700 ELMS Home-page.**