MATH 6262: Statistical Estimation – Spring 2024

Instructor: Dmitrii M. Ostrovskii – `ostrov@gatech.edu` – ostrodmit.github.io
TA: Shenduo Zhang – `szhang705@gatech.edu` – drop-in
Schedule: Mon/Wed 5-6:15pm – Skiles 171
Office hours: Wed/Thu 12:30-2pm – Skiles 263

Contact. The best way to contact me is via email (put 6262 in the subject field), or in-person during the office hours. Meetings outside of the office hours can be arranged via email (6262).

Disclaimer. This is a rigorous—and rather advanced—theorerical course in mathematical statistics, intended mainly for the graduate students whose current/anticipated research areas include *statistics, (foundations of) machine learning, optimization theory, information theory,* or a combination thereof. That said, any individual passionate about mathematics could benefit from taking this course. The goals are: (a) to understand the key principles behind statistical estimation, with the primary focus on the classical asymptotic theory but also going beyond it; (b) to learn how to apply these principles in concrete scenarios. Here is our tentative roadmap:[1]

0. Probability refresher. Basic tools from probability: probability spaces, random variables and vectors, expectations, independence, conditioning, moment generating functions.

1. Asymptotic theory. After a brief recap of the tools from probability, we shall review the basic techniques of statistical estimation, such as maximum likelihood estimation (MLE) and the method of moments, and the key notions such as consistency, sufficiency, and admissibility of an estimator. We shall then discuss the decision-theoretic viewpoint on statistical estimation, including the notions of minimax and Bayes estimators. We shall establish the local asymptotic normality (LAN) for MLE and Bayes estimators, and we shall also cover the asymptotic properties of some robust estimators.

2. Some finite-sample results. Here we shall first study linear regression in finite samples, establishing the sharp finite-sample counterpart of the corresponding LAN result. If time permits, we shall extend this to generalized linear models (GLMs), exploiting certain tools from analysis/optimization (such as self-concordant functions).

3. Convexity and information-theoretic methods. This block will be devoted to some applications of convex optimization and information theory in the context of statistical inference, with the following topics:

- KL-divergence and the geometry of MLE.
- Optimality of statistical testing via f-divergences; Pinsker's inequality; basic lower bound for estimation via Fano's method.
- Analysis of expectation-maximization (EM) algorithm for a log-concave density.

4. Optimality under restrictions We shall briefly overview the topics of minimum-variance unbiased estimators (MVUE), including the Rao-Blackwell and Lehmann-Scheffé theorems, and then establish the basic results for group-equivariant estimators.

[1] Please note that this syllabus is not a contract, and adjusting the selection of topics along the way is possible.

Prerequisites. MATH 4261; MATH 4262 or equivalent; MATH 6241.

Homework assignments. There will be 4-5 homework assignments with a few problems in each one. Some of the problems are rather comprehensive and divided into several steps through which you'll be guided. That said, as the main purpose of these problems is to help you develop your creativity and mathematical culture, I encourage you to ignore these hints at least occasionally. Some advice:

- Always start to work on homework assignments **as early as possible.** (Well, at least *look* in there, simply in order to assess the difficulty and get yourself properly stressed!)
- It's okay to work in a small group. However, always **try to solve the problem yourself first,** especially if you pursue/planning to pursue research in statistics/ML. If you are blocked, ask me a question, but meditate on it first (no answer is guaranteed otherwise).
- **Taking notes** during a discussion is a useful practice. (People forget things—I do all the time, for example—and they also tend to underestimate the extent of their forgetfulness.)
- Write solutions **on your own;** suspected plagiarism receives 0 credit and further sanctions.

Exams and project. The midterm will take place mid-March (before the Spring break), during a lecture in Skiles 171. The (comprehensive, 3-hour) final exam is on Friday, Apr 26 at 6:00pm, at Skiles 171. There is a possibility of taking a project instead of the final, or after it. More details on possible projects will be available after the Spring break; that said, the "comprehensive" homework problems are a good estimate of what you might have to deal with.

Grading. 50%/20%/25%/5% for homeworks/midterm/final/scribbling; might be curved later.

Attendance and scribing. I shall not keep track of attendance: I rely on your motivation. Scribblers shall be assigned in a round-robin fashion; you are responsible for being in class when it's your turn. Scribbling is to be done in LaTeX; I will give you feedback after your first turn.

Canvas. Link here. Will be used for homework submissions, file exchange and announcements.

Literature. We shall not follow a single textbook, but here is the list of some useful references.

- E. L. Lehmann, G. Casella. *Theory of Point Estimation, 2nd ed.*
- G. Casella, R. L. Berger. *Statistical Inference, 2nd ed.* (some parts).

Other references shall be provided along the way for specific topics. Also, you might consult my handwritten notes for MATH 541b, a not completely dissimilar course that I taught at USC.