Stat 41521: Topics in Distribution-free Inference — Autumn 2025 Syllabus

Course description This course will focus on the recent field of distribution-free inference, which seeks to provide verifiable statistical guarantees without assumptions on the distribution of the data. Methods in this area include holdout set methods, cross-validation type methods, and conformal prediction. The course will cover theoretical advances and practical methodologies, theoretical hardness results, and open problems in the field.

Course info

- Instructor: Rina Barber, `rina@uchicago.edu`
- Class times: MW 1:30–2:50, Jones 226
- Office hours: Fri 11–12, Jones 214
- The main course information and materials will be on Ed Discussion. Problem sets will be turned in on Gradescope.
- Course textbook (available for free):

  Theoretical Foundations of Conformal Prediction, Angelopoulos, Barber, Bates 2025+

  We will also provide additional resources and references to go along with topics covered in the class. These resources will include papers, tutorials, videos of talks, etc, and will be posted on Ed Discussion.

Topics The topics covered in the course will depend on student interest. A tentative list:

- The distribution-free inference framework: goals, definitions, targets of inference
- Distribution-free predictive inference methods: holdout set methods for prediction; conformal prediction; jackknife+ and CV+
- Optimality results and hardness results for distribution-free predictive inference
- Weighted exchangeability; weighted versions of distribution-free inference methods, with applications to various problems in statistics
- Distribution-free prediction in different settings: streaming data ; relaxing exchangeability ; localized methods
- Distribution-free inference beyond the prediction problem: alternative measures of predictive risk; inference for regression; calibration

Grading All students taking the course for a quality grade are expected to (1) to contribute to one group presentation/project (details below), and (2) to submit substantial work for a majority of the homework exercises. Students taking the class Pass/Fail may choose to do either (1) or (2) but are not required to do both.

Homework exercises Homework exercises will be posted each week, due on Fridays. These exercises will vary in difficulty, and some will be at the level of current research problems in the field. It is not necessary to submit answers to every problem, but most problems should be attempted, and students are encouraged to submit partial solutions or ideas for problems even if the answer is not complete. The problems will mostly be open-ended, so there may be multiple possible solutions.

Presentations/projects Students will work in groups of size 2–3 to present an empirical exploration of a topic covered in the class, *or* may work on their own or in a group of size 2–3 to hand in a project exploring a topic in more depth. The goal for a presentation is to explore some interesting empirical aspects of methods covered in class (e.g., what types of data make a method succeed or fail, or, what types of empirical behavior of the method are or are not explained by the theory?). The presentation itself should be at most 15 minutes, but the group should plan to spend substantial time developing ideas, running simulations, etc, to prepare. If instead the group chooses to hand in a written project, the topic can be theoretical or empirical, and the group can consult with the instructor to develop a project idea.