

**Stat 669a :
Information and Statistics**

Instructors: Andrew Barron, Mokshay Madiman. Use `firstname.lastname@yale.edu` to contact either of us.

Class Time: 4:00–5:15 p.m., Tuesday and Thursday.

Class Location: Room 107, 24 Hillhouse Ave.

Grading: One problem will be assigned for each class meeting. Class participants are expected to participate in solving these problems at the beginning of the next class, since the problem will be relevant to the content being covered. [If you are taking the course for credit, you are expected to have worked the problem beforehand rather than making impromptu attempts on the board.]

You can optionally choose to present a review of a paper covering an important development in the field at the end of the semester; the choice of paper needs to be discussed with and approved by the instructors before November 16.

Course Description: Study of the pivotal role that information theory plays in illuminating modern statistics. Topics include the equivalence of data compression and statistical modeling (from the Shannon, universal coding, and Kolmogorov viewpoints), and its relationship to the minimum description length principle; and clean risk bounds for complex estimation scenarios based on an index of resolvability. Other possible topics include various aspects of hypothesis testing (fixed and sequential tests, error exponents, multiple testing); and the arbitrary sequence approach to on-line learning with applications to prediction, data compression and portfolio selection.

Prerequisites: Prior exposure to probability theory and basic statistical inference is assumed. Measure-theoretic probability might be occasionally useful, but is not essential. You do not need to know any information theory (although prior background will, of course, enrich your learning experience).

1 More about the course

Welcome to STAT 669, “Information and Statistics”. This is a research-oriented course primarily for graduate students, meant to be somewhat flexible and adapt somewhat to the interests and preferred pace of the participants. There is no textbook, and much of the material will be drawn from papers written in the last decade. The major theme of the course is the pivotal role that information theory plays in illuminating modern statistics.

Approximate outline:

Foundations: The equivalence of data compression and statistical modeling (from the Shannon, universal coding, and Kolmogorov viewpoints), and its relationship to the minimum description length (MDL) principle.

Flexible function fitting: Illustrated via several settings including projection pursuit regression, neural nets, MARS, MAPS, MART, and function aggregation. Various algorithms. Risk analysis, emphasizing the role of dimension and complexity, and bounds using the index of resolvability.

Minimax rates for function estimation: Via Fano’s inequality, and via total redundancy. If time permits, connections with distortion-rate theory.

MDL revisited: Density estimation formulation, Rissanen theory, minimum complexity density estimation, information projection, clean risk bounds, mixture models, greedy algorithms, information theoretic foundation of Bayes procedures.

Possible additional topics