STAT 27400/37400: NONPARAMETRIC INFERENCE

Syllabus, Autumn 2018

Nonparametric Inference is an advanced undergraduate and beginning graduate level course in modern nonparametric methods for statistical estimation and inference.

Nonparametric inference is about statistical methods and models that make weak assumptions. A typical nonparametric approach estimates a nonlinear function from an infinite dimensional space, rather than a linear model from a finite dimensional space. This course gives an introduction to nonparametric statistics, with a focus on density estimation, regression, confidence sets, orthogonal functions, random processes, and kernels. The course treats nonparametric methodology and its use, together with theory that explains the statistical properties of the methods.

Schedule

LECTURES	Tuesday/Thursday 3:30-4:50 pm	Kent 101

Contact Information

Instructor:

Tingran Gao	Jones 316	`tingrangao@galton.uchicago.edu`
Office hours:	Tuesday, 2:00-3:00 pm, Jones 316	
	Thursday, 2:00-3:00 am, Jones 316	

Course Assistant:

Xialiang Dou	Jones 203/204	`xialiang@galton.uchicago.edu`
Office hours:	Monday, 9:00-10:00 am, Jones 304	
	Wednesday, 9:00-10:00 am, Jones 304	

Prerequisites

STAT 24400 is required; alternatively STAT 22400 and exposure to multivariate calculus and linear algebra.

Website

`https://canvas.uchicago.edu/courses/17607`

Course Structure and Grading

The course will have a standard lecture format. Assignments will be handed out and due every other week. Two short 15(-ish) minute quizzes will be given in class. These quizzes are not intended to be too challenging, but ways to reinforce some of the basic concepts in the class. An in-class midquarter exam will be given.

There are two sections of the course. The sections will be graded using separate scales. The graduate section, STAT 37400, will require a small project in addition to the final exam. This project will require analysis of a data set, to be chosen by the student. The analysis should include a comparison of parametric against nonparametric methods. A 4-6 page report should be submitted, describing the data, the statistical problem, the methods used, and the findings.

Each of these components will be weighted as follows to determine a final grade:

- Assignments: 40%
- In-class midquarter exam: 25%
- Final exam+project: 30%
- In-class quizzes: 5%

Policy on Assignments

Assignments will be posted on Thursdays, and are due fortnightly in class. Assignments may be handed in up to three days late, with a 20%-per-day penalty. That is, an assignment passed in on Friday will have a 20% penalty, on Saturday a 40% penalty, and on Sunday a 60% penalty. Students are encouraged, but not required, to use LaTeX to prepare solutions to the assignments.

Assignments will have at least one problem that involves computation. Students are required to use the R programming language for solutions. R is the language of choice for statistical computing; it is easy to learn, and the course does not require previous exposure to the language. Tutorial introductions can be organized as needed. Free downloads of the language, together with an extensive set of resources, can be found at `http://www.r-project.org`.

Collaboration on homework assignments with fellow students is encouraged. However, such collaboration should be clearly acknowledged, by listing the names of the students with whom you have had discussions concerning your solution. You may *not* share written work or code—after discussing a problem with others, the final solution must be prepared and written by yourself.

Course Calendar

The course calendar and other materials will be posted on the course Canvas site, `https://canvas.uchicago.edu/courses/17607`, and will be updated throughout the semester.

The schedule of topics, exams, and assignments follows.

Week	Date	Topic	Exams and Assignments
1	October 2	what is nonparametric inference? §§ 1.1–1.4	
	October 4	density estimation §§ 6.1–6.3	assn 1 out
2	October 9	nonparametric regression §§ 4.1, 4.5	
	October 11	nonparametric regression §§ 5.1–5.4, 5.6, 5.7	
3	October 16	CDFs and statistical functionals §§ 2.1–2.3	
	October 18	bootstrap §§ 3.2–3.4	assn 1 due, assn 2 out
4	October 23	nonparametric Bayes	quiz 1
	October 25	nonparametric Bayes	
5	October 30	normal means §§ 7.1–7.3	
	November 1	SURE and minimax risk §§ 7.4–7.6	assn 2 due, assn 3 out
6	November 6	—	midterm exam
	November 8	shape-constrained estimation	
7	November 13	splines and RKHS §§ 8.1–8.2	
	November 15	splines and RKHS	assn 3 due, assn 4 out
8	November 20	wavelets §§ 9.1–9.4	quiz 2
	November 22	—	thanksgiving break
9	November 27	wavelets	
	November 29	lower bounds on minimax risk § 7.5	assn 4 due
10	December 4	review	
	December 6	college reading period	STA37400 projects due
	December 13		final exam 4:00-6:00 pm

Textbook

The book used in the class is *All of Nonparametric Statistics*, by Larry Wasserman (Springer, 2006). The book will be complemented by extra notes. The book is succinct and technical in places. While we will cover much of the material in this text, it will be presented at a more elementary level.

Other books that will be used in part (but not required) are:

- Alexandre B. Tysbakov, "Introduction to Nonparametric Estimation," Springer Series in Statistics, 2009.
- L. Györfi, M. Kohler, A. Krzyżak, H. Walk, "A Distribution-Free Theory of Nonparametric Regression," Springer, 2002.
- T. Hastie, R. Tibshirani, J. Friedman, "The Elements of Statistical Learning: Data Mining, Inference, and Prediction," second edition, Springer, 2009.

The first two of these books are technical and theoretical; the third is strong on intuitive motivation and methods.