Statistics 200B: Introduction to Probability and Statistics at an Advanced Level

Soyeon Ahn <iamyeoni@berkeley.edu>
Department of Statistics

UC Berkeley, Spring 2019

Syllabus

T/Th 9:30-11am Lecture, Etcheverry 3107
W 10am-noon (101 Discussion Section), Evans 9

Instructor: Dr. Soyeon Ahn
Office: 307 Evans
Office Hours: T/Th 11am-noon
e-mail: iamyeoni@berkeley.edu

GSI: Omid Shams Solari
Office: Evans 434
Office Hours: W 12-3, Piazza
email: solari@berkeley.edu

Syllabus

About This Course: This is the second part of a sequence of courses on probability (200A) and statistics (200B) at the masters level. We will cover the fundamentals of statistical inference. Prerequisites: multivariable calculus; basic linear algebra (vectors and matrices); probability topics covered in 200A.

Textbooks

There is no required textbook for this course. But there are several recommended ones.

1. "All of Statistics" by Larry Wasserman. We'll cover the topics listed in Section II and parts of Section III. Even though we will follow the topics in this book, this book cannot serve as our textbook. The material content of the book is not detailed enough.
2. "Statistical Inference" by Casella and Berger. It is much more detailed and present "traditional" topics in more depth. But it lacks materials on modern topics.
3. "Theoretical Statistics: Topics for a Core Course" by Robert Keener. This book is detailed and has relatively more modern topics than the Casella and Berger book.

Homework and exam questions will be based on the lectures.

Grading

Your final grade will be maximum of [a weighted average of your average homework score (25%), two midterms (40%), and the final exam (35%)] and [a weighted average of your average homework score (25%), one midterm (20%), and the final exam (55%)].

- **Homeworks:** Problem sets will be assigned each Thursday, for a total of 11 assignments. You should download the assignments from the bCourses page for this class (https://bcourses.berkeley.edu). Each problem set is to be turned in **the following Thursday at the beginning of class**. No late assignments will be accepted. The lowest two homework grades will be droped when calculating your final score. Not all problems will be graded, and you should review the solutions carefully for those problems for which you don't get graded feedback. Students can discuss

homework assignments in groups of at most three. Each student must write up his/her own solutions individually, and must explicitly name any collaborators at the top of the homework. Any evidence of cheating will be subject to disciplinary action.

- **Midterm 1:** The first in-class midterm is tentatively scheduled on Tuesday, February 19. You are allowed to bring one double sided A4 page of handwritten notes to the exam.

- **Midterm 2:** The second in-class midterm is tentatively scheduled on Tuesday, April 9. You are allowed to bring one double sided A4 page of handwritten notes to the exam.

- **Final:** The final exam will be cumulative. The exam is scheduled on Wednesday, May 15, 2019 (11:30am - 2:30pm). You are allowed to bring two double sided A4 pages of handwritten notes to the exam.

Other logistics

Computing: The assignments will involve some computing. We will use `R`, which is available for free at http://cran.r-project.org/. Your GSI will teach you everything you need to know about `R` for this course.

Students with disabilities: If you need accommodations for any disabilities, please speak to me after class or during office hours so that we can make the necessary arrangements.

Academic Integrity: Any homework, test or report submitted by you and that bears your name is presumed to be your own original work. In all of your assignments, you may use words or ideas written by other individuals in publications, web sites, or other sources, but only with proper attribution.

UC Berkeley Honor Code: As a member of the UC Berkeley community, I act with honesty, integrity, and respect for others.

Tentative Schedule of topics, readings, and assignments

- **Week1** 1/22, 1/24 Intro, Probability review (HW1 assigned)
- **Week2** 1/29, 1/31 Intro to inference, Empirical CDF (HW2 assigned)
- **Week3** 2/5, 2/7 Bootstrap, Point estimation overview (HW3 assigned)
- **Week4** 2/12, 2/14 Properties of MLE, Delta method, Review
- **Week5** 2/19, 2/21 **Midterm 1**, Intro to hypothesis testing (HW4 assigned)
- **Week6** 2/26, 2/28 Likelihood ratio tests, p-values and testing history (HW5 assigned)

- **Week7** 3/5, 3/7 Multinomials, multiple testing, Bayesian inference (HW6 assigned)
- **Week8** 3/12, 3/14 Simulation in Bayesian models, Choice of priors (HW7 assigned)
- **Week9** 3/19, 3/21 Decision theory I, Decision theory II (HW8 assigned)
- **Week10** 3/26, 3/28 No class
- **Week11** 4/2, 4/4 Linear regression, Probability for Data Science, Review session
- **Week12** 4/9, 4/11 **Midterm 2**, Midterm Solution (no discussion session on Wed) (HW9 assigned)

- **Week13** 4/16, 4/18 Model selection, Generalized linear models (HW10 assigned)

- **Week14** 4/23, 4/25 Nonparametric curve estimation, Kernel density estimation (HW11 assigned)

- **Week15** 4/30, 5/2 Nonparametric regression, Classification

- **Week16** 5/7, 5/9 RRR