California Institute of Technology
Department of Computing + Mathematical Sciences

ACM 270-1 Introduction to Statistical Inference

Winter 2016

Lectures:	MW 10:30-12:00 in 213 ANB
Instructor:	Konstantin (Kostia) Zuev
Office:	Annenberg 114
Email:	kostia@caltech.edu (please include "270-1" in the subject line)
Website:	http://www.its.caltech.edu/~zuev/
Office Hours:	TBA, or by appointment (please, send an email to schedule)
Discussions:	TBA
Teaching Assistant:	TBA (email, office)

Course Goals

Statistical Inference is a brunch of Mathematical Engineering that studies ways of extracting reliable information from limited data for learning, prediction, and decision making in the presence of uncertainty. The main goals of this course are:

- Develop statistical thinking and intuitive feel for the subject,
- Introduce the most fundamental ideas, concepts, and methods of Statistical Inference, and
- Explain how and why they work, and when they don't.

If you do well in the class, you should be able to read (and understand) most contemporary papers that use statistical inference and perform basic statistical analysis yourself.

Prerequisites

This is an introductory course on statistical inference. No prior knowledge of statistics is assumed. However, a solid understanding of Probability is required. Ma 3/103 or equivalent is a strong prerequisite. If you don't know Probability, you will suffer as someone who does not know Mandarin will suffer in the Chinese Poetry class. If you have not taken any Probability class, come to my "Introduction to Probability Models", and then come back in the next year. A key part of the course is homework, where you will get experience in using the learned methods and models in applications via simulations in MATLAB. So, some familiarity with MATLAB, e.g. ACM 11, and programming is desired, but this is a "soft" prerequisite.

Textbooks

This is a completely new course made "from scratch." Good news: I will provided a set of self-contained lecture notes which are heavily based on the following texts:

- G. Casella & R.L. Berger, *Statistical Inference,* 2002.
- L. Chihara & T. Hesterberg, *Mathematical Statistics with Resampling and R*, 2011.
- A.C. Davison, *Statistical Models*, 2003.
- D. Freedman, R. Pisani, & R. Purves, *Statistics,* 2007.
- M. Lavine, *Introduction to Statistical Thought*, 2013.
- S.L. Lohr, *Sampling: Design and Analysis*, 2010.
- D.C. Montgomery, E.A. Peck, & G.G. Vining, *Introduction to Linear Regression Analysis,* 2006.
- D. Nolan & T. Speed, *Stat Labs: Mathematical Statistics Through Applications*, 2000.
- L.A. Wasserman, *A Concise Course in Statistical Inference*, 2005.
- S. Weisberg, *Applied Linear Regression*, 2005.

Course Plan

The following is a tentative outline of the topics to be covered this term.

Week 1	Introduction, Summarizing Data, MATLAB crash course
Week 2	Classical Statistics: Fundamental of Survey Sampling
Week 3	Modeling and Inference: A Big Picture
Week 4	Statistical Functionals, Jackknife, and Bootstrap
Week 5	Method of Moments, Maximum Likelihood Estimation
Week 6	Hypothesis Testing: General Framework, p-Values
Week 7	The Wald, t-, Permutation, and Likelihood Ratio tests, Multiple Testing
Week 8	Regression Function, Scatterplots, Simple Linear Regression
Week 9	Ordinary Least Squares, Hypothesis Testing
Week 10	Interval Estimation, Prediction, Graphical Residual Analysis

Grading

Your final grade will be based on your total score. Your total score is a weighted average of Homework (60%), the Midterm exam (20%), and the Final exam (20%). There are no fixed thresholds for grades, but if your total score is 90% (80%, 70%, 60%), you are guaranteed at least "A" ("B", "C", "D").

Homework

There will be six homework assignments. Homework problems and due dates will be posted on the course website on TBA (for exact dates see "Important Dates" below). These problems will be collected TBA. Late homework will not be accepted for any reason, but the homework with the lowest score will be dropped and not counted toward your total score. Extensions may be granted for academic (with at least one week notice) or medical reasons.

Midterm Exam

There will be one (one hour) "in class" midterm exams: Monday, February 8. The exam will be given in regular class time. The place will be announced later. The midterm is open-book. Calculators are allowed.

Final Exam

The final exam will cover the material after the midterm. The exam is open-book. Calculators are allowed. The details regarding the time and place will be announced later.

Quizzes

There will be a random number of ~10 min quizzes given on random days. Please, consider quizzes not as a punishment (they will not affect your total score and final grade in any way!), but as an incentive to learn the material and prepare to the exams in advance. Think of a random quiz as a morning espresso shot.

Collaboration Policy

If you get stuck with a homework problem, I encourage you to discuss it with other students. Nevertheless, you will have to prepare and submit your homework by yourself. No collaboration is allowed on the midterm and final exams.

Important Dates

Quizzes:	random, mark your calendar
Homework:	Jan 15 (due 22), Jan 22 (29), Jan 29 (Feb 5), Feb 12 (19), Feb 19 (26), Feb 26 (Mar 4)
Midterm:	Monday, February 8
Final:	TBA

Websites

All lecture notes, announcements, homework assuagements, due dates, solutions, and other course related information and materials will be posted on the course website. You are expected to check it regularly.
http://www.its.caltech.edu/~zuev/teaching/2016Winter/ACM270-1.html
Class discussion will be managed via Piazza, which is designed such that you can get a quick help from your classmates, the TAs, and the instructor. Instead of emailing questions to the teaching staff, I encourage you to post your questions on Piazza because a) you will get the answers faster b) your classmates may also benefit from seeing the answers to your questions. Enroll here:
http://tba

Honor Code

You must conform to the honor code:
"No member of the Caltech community shall take unfair advantage of any other member of the Caltech community."
In-Class Behavior
Behavior that persistently or grossly interferes with classroom activities is considered disruptive behavior and may be subject to disciplinary action. Such behavior inhibits other students' ability to learn and an instructor's ability to teach. A student responsible for disruptive behavior may be required to leave class pending discussion and resolution of the problem and may be reported to the Undergraduate Board of Control or the Graduate Honor Council. In particular, the use of cell phones during class or conversation is disruptive behavior.
Academic Integrity
All students are responsible for maintaining standards of academic integrity. In particular, collaboration during midterms or the final are strictly prohibited.