STAT 135: Concepts of Statistics (Fall 2024)

Lectures: Mon, Wed, Fri 9-10am
Genetics & Plant Bio 100

Lab: Fri 11am-1pm, 342 Evans
Fri 1-3pm, 342 Evans
Fri 2-4pm, 330 Evans
Fri 4-6pm, 330 Evans

Bring your laptop. If you do not have access to a laptop, you can borrow one from the University library. See https://studenttech.berkeley.edu/hardware-lending for more details. The Student Technology Equity Program is another good resource. Feel free to contact the instructor if you have concerns about your access to needed technology.

Instructor: Sam Pimentel
Email: spi@berkeley.edu
Preferred pronouns: he/him
Office Hours:
Wed 11am-12pm, 429 Evans
Thu 12pm-1pm, 429 Evans

GSIs:

Email:
Preferred pronouns: she/her
Office Hours:
Remote: Zoom link. Password:

Email:
Preferred pronouns: they/them
Office Hours:
Remote: Zoom link. Password:

Topics:

A comprehensive survey course in statistical theory and methodology. Topics include parameter estimation, hypothesis testing, statistical tests (parametric and nonparametric) and linear regression (single and multiple). We will cover most of chapters 7, 8, 9, 11, 12, 13 and 14 in the textbook.

Textbook:

Mathematical Statistics and Data Analysis, John Rice, 3rd Edition. Previous editions are not acceptable.

This is a working draft of the syllabus and is subject to change.

Learning goals

By the end of the semester, you should be able to:

1. Interpret point estimates, confidence intervals, and hypothesis tests clearly for an audience without statistical training.
2. Construct common estimators, statistical tests and confidence interval procedures using probability theory.
3. Evaluate the relative strengths and limitations of several estimation or inference procedures for the same problem using mathematical concepts including unbiasedness, efficiency, and power.
4. Recommend an approach and carry out estimation and inference for canonical statistics problems including tests of association between two variables and fitting probability distributions to univariate data.

Prerequisites:

- **STAT 134 or an equivalent course in probability theory**.
- Multivariable calculus, especially Lagrange multipliers.
- Familiarity with moment-generating functions.
- Familiarity with linear algebra (matrix operations, inverses, and eigenvalues) for chapter 14.
- Familiarity with basic R concepts equivalent to the first ~6 weeks of Stat 133. Unless otherwise noted assignments involving computing must be completed in R.

Lecture:

Lectures will cover core theory and concepts, with supporting data analysis examples. To get the full benefit of lecture, it is best to read the supporting material ahead of time. I encourage active engagement and discussion during lectures, and I will frequently pose questions and call on students to answer. When slides or R code are shown in class, they will be posted online after class. However, many lectures will not have associated slides.

The current plan is to record lectures using a built-in classroom camera and post the videos for asynchronous viewing on bCourses. However, I cannot yet vouch for the quality of the video capture and recommend in-person attendance for at least the first few lectures. Masks are required for those attending in-person.

Lab:

Lab time will be spent working on practice problems, and occasionally for taking quizzes as discussed below. Since some problems will involve computing you should plan to bring your laptop. You may attend a lab for which you are not enrolled (physical space permitting). Labs will not be recorded so attendance is strongly recommended.

Assessment:

Weekly Assignments

We anticipate giving 12 homework assignments during the semester. Homework will be posted to bCourses, generally on Wednesdays, and will generally be due the following Friday. All homework should be submitted **via Gradescope** (linked through bCourses). Homework will be a combination of analytical and computational exercises done "by hand" and exercises using the computer.

Midterm and Final Exams

An **in-class** midterm and an in-person final exam will be given. The midterm exam will be held on **Friday October 11**. The final exam time and day are **Thursday December 19, 7:00 PM-10:00 PM**. I do not plan to offer makeup times for either exam so please confirm now that you can attend both exams.

Quizzes

There will be four 50-minute quizzes **in lab** to test your understanding of homework and lecture. The dates of the quizzes are:

- **Friday September 13**
- **Friday September 27**
- **Friday November 8**
- **Friday November 22**

Overall score

Your letter grade for the course will be based on the total points for all work in the semester, as follows:

- Homework (each assignment weighted equally, drop lowest): 15%
- Midterm (dropped if final exam score is higher): 25%
- Final exam: 40% or 65% (if higher than midterm score).
- Quizzes (drop lowest): 20%

Grades follow a curve under which roughly 30% of students receive letter grades in the A-range, roughly 30% receive letter grades in the B-range, and roughly 30% will receive letter grades in the C-range. Grade distributions for the quizzes and the midterm will be shared on bCourses when grading. Is complete.

Online Resources

bCourses

Homework assignments and material from lecture (where applicable) will be posted here. I will also make course announcements through bCourses.

Ed Discussion

I have created an Ed Discussion site for this course, which you can access through the link in bCourses. This is an online forum to ask questions to fellow students and course staff.

Gradescope

Homework assignments and regrade requests (see Policies section below) will be submitted through Gradescope, which you can also access through the link in bCourses. This is also where you can view your course grades.

Policies

Possibility of revisions to course policies

All course policies, including assessment, are subject to change during the course of the semester in response to unforeseen events including but not limited to developments in the COVID-19 pandemic, power outages, forest fires, and medical emergencies among members of the course staff.

Late Assignments

All students will have 5 late days that they may use for turning in homework after the due date. This will take the place of any extensions due to sickness or conflicts so use them wisely. To use a late day, **you must submit a Google Form at this link requesting a late day before the homework is due**, or you risk receiving a large penalty or a zero. Late day requests by email will not be answered. Late days cannot be used for the group project or the midterm.

Regrade requests

Regrade requests on an assignment are **due within one week of the release of the graded assignments and the solutions** (if applicable). Regrade requests should be submitted through

Gradescope. In writing a regrade request, please be specific about the nature and exact location of the error you feel the grader has made, with reference to the solutions if available.

Scheduling conflicts for exams and quizzes

All quiz and exam dates are provided on this syllabus and class members are expected to be present for each quiz and exam. I do not generally provide alternate exam times (except for students with DSP letters requiring this). If you have an unavoidable conflict, it is your responsibility to discuss it with me **during the first two weeks of class**.

Academic Honesty Policy

The student community at UC Berkeley has adopted the following Honor Code: *"As a member of the UC Berkeley community, I act with honesty, integrity, and respect for others."* My expectation is that you will adhere to this code. Beyond the importance of respecting your fellow students, acting with integrity in completing course assignments helps ensure that they achieve their purpose, which is to help you learn and develop valuable statistical understanding and skills.

- Homework must be done independently. If you get stuck or want to explore alternative approaches, feel free to discuss issues with students or course staff (including on Ed Discussion); however, you may not do the homework jointly, nor may you ask for or share complete code or solutions. Sharing solutions or obtaining and/or using solutions from previous years or from the Internet, if such are available, is considered cheating.
- On homework, you should include a section listing all the sources you drew on in producing your answers; on the homework, you should also list the names of other students with whom you consulted.
- During exams and quizzes, you must not consult with any other person besides the course staff or refer to any written resource (besides one double-sided page of notes on the midterm and final exam only).

Anyone caught cheating will be given a score of zero (0) on the assignment/exam and will be reported to the University's Office of Student Conduct.

Email

1) If you wish for your email to make it into my inbox, the subject of your email must contain the text "**135**."
2) Neither I (nor the GSIs) explain course material over email and will not respond to emails with such requests. Please use Ed Discussion, office hours, discussion section, or GSI's office hours (or schedule another time to meet if you have irreconcilable conflicts with the office hours).
3) I respond to email regarding the class roughly once a day, and rarely during the weekend.

Inclusivity and Accommodation

My hope is to establish a learning environment in this course that welcomes diversity of thought, perspective, and experience, and to be respectful of your individual identity as a student. I am happy to use your preferred name and/or personal pronoun. If you feel uncomfortable as a result of anything that is said in class, or if you feel that your performance in the course is being impacted by experiences outside of class, please do not hesitate to reach out to me about your concerns.

In addition, if you need accommodations for any physical, psychological, or learning disability, please speak to me after class or during office hours. Please note that you must make arrangements in a timely manner through DSP so that I can make the appropriate accommodations.

Acknowledgments

Most of the materials used in this course, including this syllabus, are close adaptations from materials originally created or compiled by Prof. Adam Lucas and generously provided for the current semester. In writing this syllabus I also adapted content from Prof. Chris Paciorek and from Prof. Monica Linden of Brown University.

Anticipated Course Schedule

Week	Topics	Assignments Due & Exams	Assigned reading from Rice
Aug 26	Introduction, sampling		7.3.1-7.3.3 (Fri)
Sep 2	Fitting distributions to data (no lecture Monday)	HW1 due Friday	8.1-8.4
Sep 9	Method of moments and delta method	Quiz 1 Friday HW2 due Friday	8.4, 4.6
Sep 16	Maximum likelihood estimation	HW3 due Friday	8.5
Sep 23	Theory for maximum likelihood	Quiz 2 Friday HW4 due Friday	8.7, 8.8
Sep 30	Hypothesis testing	HW5 due Friday	9.1 (first 3 paragraphs) 9.2
Oct 7	Review and midterm	Midterm	-
Oct 14	Testing/confidence interval duality, composite hypotheses	HW6 due Friday	9.3-9.4 11.1 (Fri)
Oct 21	Two-sample tests, goodness-of-fit for multinomial	HW7 due Friday	11.2, 9.5
Oct 28	Tests of independence	HW8 due Friday	13.3-13.4
Nov 4	Paired tests, rank tests, ANOVA	Quiz 3 Friday HW9 due Friday	11.3, 12.2
Nov 11	Simple linear regression	HW10 due Friday	14.2
Nov 18	Multiple linear regression	Quiz 4 Friday HW11 due Friday	14.5
Nov 25	Bayesian statistics (no lecture/lab on Wed or Fri)	-	8.6
Dec 2	Review	HW12 due Friday	-
Dec 9	RRR week		
Dec 16	Finals week	Final exam: Thu Dec 19 7:00-10:00 PM PST.	