The class will introduce the students to formal statistical reasoning. Building on knowledge of probability and calculus, we will explore how limited noisy observations can be used to learn general characteristics of a population. We will study the basics of decision theory, including frequentist and Bayesian solutions to the “paradox of induction.” This is a mathematical class, with an emphasis of understanding the logic behind data analysis procedures, rather than introducing a menu of possible techniques or practical challenges of data analysis. These important goals are fulfilled by other courses like Stats101.

Topics

We will start by reviewing some probability concepts, with an emphasis on notable distributions and limit theorems (please note that this review is only intended as a refresher, not as a substitution for a proper probability course like Stats116). We will then move on to study sampling and the distributions of sample moments. At this point we will be ready to tackle point estimation and explore the properties of different approaches: methods of moment, maximum likelihood, Bayesian inference, etc. We will study hypothesis testing—a cornerstone of modern science—including the importance of hypothesis specification, the correct interpretation of p-values, and the implications of testing a large number of hypotheses simultaneously. Strong from our practice with test statistics, we will then tackle interval estimation. Finally, we will explore how the approaches we developed are applied in a variety of common inferential problems.

Textbook and references

The textbook for the class is Rice “Mathematical statistics and data analysis,” 3rd ed. The core material is in chapters 5-9. Knowledge of the subjects of chapters 1-4 is assumed as a prerequisite and topics in chapters 10-14 will be covered selectively. Lectures will complement and expand the textbook and attendance is required (we do not expect to distribute lecture notes).

There are several other textbooks that you might want to consult:

  Has more material and details than the typical book at this level.

Is an older textbook, but does a good job explaining motivations.

  Also older, with many worked out examples.

  A newer book, with a language and reference to materials that might make it easier to connect the topics in this course with other disciplines

  Another classical option.

### Teaching staff

<table>
<thead>
<tr>
<th>Name</th>
<th>Email (@stanford.edu)</th>
<th>Role</th>
<th>Office</th>
<th>Office Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theodor Misiakiewicz</td>
<td>TA</td>
<td>Office in Sequoia 207</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shuangning Li</td>
<td>TA</td>
<td>Tuesday 1:30–3:30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nan Bi</td>
<td>TA</td>
<td>Wednesday 3–5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhimei Ren</td>
<td>TA</td>
<td>Thursday 11–1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pete Mohanty</td>
<td>pmohanty</td>
<td>Sequoia 202</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chiara Sabatti</td>
<td>sabatti</td>
<td>Sequoia 230</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Getting Help

Each weekday (M-F) at least one member of the teaching team will hold office hours. You are strongly encouraged to make use of them and to consider them as the primary way to answer any of the questions you might have (aside from asking in class). The class has a wiki-style Piazza site ([https://piazza.com/stanford/winter2018/stats200/home](https://piazza.com/stanford/winter2018/stats200/home)) but please understand that, while we will do our best, we cannot guarantee timely online responses. In particular, do not expect answers to questions about the homework which are submitted after 5pm of the day before it is due. Please do not contact the teaching team using their individual e-mails; we all receive too many e-mails to monitor them effectively. Piazza, which displays equations and code nicely, is usually better than email anyway.
Requirements and policies

Homework

There will be weekly homework, due Fridays in class. Homework is intended primarily as a learning tool and it is important that you work on it in a timely fashion. Late homework will not be accepted under any circumstances. Note, however, your worst homework score will be dropped for grading purposes. You are welcome to discuss solutions with peers, but each student needs to hand in an individual write-up. Homework will include simple computing exercises asking you to perform small simulations, create histograms and plots, and analyze data. You may use any language (e.g. R, Python, Matlab) and will be graded only on your results, not on the quality of your code. Solutions will be provided in R.

Class attendance & participation

Lecture attendance and in-class participation contributes to the course grade. Lectures will include active learning components and so you may be asked to hand in short exercises and might be occasionally re-seated into small groups. In class, cell phone use is not allowed and we recommend against the use of tablets and laptops. Although some people prefer taking notes on a laptop instead of by hand, educational studies have shown that students using laptops tend to learn less effectively than those without laptops: there are too many distractions available on an open computer.

Grading

There will be a midterm exam during class time; the final exam is scheduled by the university for Friday March 23, 2018 8:30–11:30am. You may bring one double-sided 8.5 × 11” page of notes with you to the midterm and the final exam; other than that, exams are closed-book. Your lowest homework grade will be dropped when computing your average homework score. Your final grade will be the maximum of the following two weightings:

\[
\begin{align*}
5\% \times \text{ (participation)} + 30\% \times \text{ (average homework)} + 30\% \times \text{ (midterm)} + 35\% \times \text{ (final)} \\
5\% \times \text{ (participation)} + 30\% \times \text{ (average homework)} + 20\% \times \text{ (midterm)} + 45\% \times \text{ (final)}
\end{align*}
\]

We do not ‘curve’ (standardize) grades, and we will be thrilled to have everyone completing the course with an A.
The Honor Code

Violating the Honor Code is a serious offense, even when the violation is unintentional. The Honor Code is available at: http://studentaffairs.stanford.edu/communitystandards/honorcode. Students are responsible for understanding the University rules regarding academic integrity. In brief, conduct prohibited by the Honor Code includes all forms of academic dishonesty, among them copying from another’s exam, unpermitted collaboration, and representing as one’s own work the work of another.

Provost’s Statement Concerning Students with Disabilities

Students who have a disability which may necessitate an academic accommodation or the use of auxiliary aids and services in a class must initiate the request with the Office of Accessible Education’s Disability Resource Center (DRC). The DRC will evaluate the request with required documentation, recommend appropriate accommodations, and prepare a verification letter dated in the current academic term in which the request is being made. Please contact the DRC as soon as possible; timely notice is needed to arrange for appropriate accommodations (phone 723-1066; TDD 725-1067). FERPA: Student Record Privacy Policy: studentaffairs.stanford.edu/registrar/students/ferpa

Stats200 Website

You will find all relevant information for this course on the class webpage: http://statweb.stanford.edu/~sabatti/Stat200/index.html

The class webpage links to the CANVAS and PIAZZA sites that we will use to share documents with restricted access and for discussion of class-related materials.