STAT 152: Sampling Surveys

Lectures: Tuesday, Thursday 12:30-2pm
20 Barrows

Lab: Friday 1-3pm
Friday 3-5pm
334 Evans

Instructor: Elizabeth Purdom
Office: 433 Evans Hall
Phone: 642-6154 (office)
Email: epurdom@stat.berkeley.edu
Office Hours: W 3:00-5:00

GSI: Inna Gerlovina
Email: innager@berkeley.edu
Office Hours: Th 2-4pm

Topics:

- Estimation in probability surveys
- Simple random samples
- Stratified samples
- Cluster samples
- Complex samples
- More advanced data analysis

Textbook:

Sampling, S. Lohr (2nd Edition)

Complex Surveys, Lumley (recommended)

Additional readings will be posted on bspace

Prerequisites:

STAT 134. Statistics 133 and 135 are recommended.

Lab:

Lab time will be spent working on practice problems, and how to analyze the data in R where applicable.

We assume familiarity with R. If you do not have experience with R, we will make online introductions to the programming language available (see resources on bspace).

Class Activities and Attendance:

Past student evaluations have said that attending lecture was critical for doing well in the class, and this is exactly what I expect. To get the full benefit of the class, you should read the material ahead of time, if possible.

I encourage active engagement in the class, and frequently will pose questions and call on people during class. I will also occasionally ask for you to answer questions on an index card in class and turn in the card or otherwise use them during class.

Handouts given in class deal with in-class activities and discussions and will generally **not** be posted online, unless they cover material not in the book.

Study Groups:

To aid in learning, I encourage students to form study groups of 5-6 people from the same section, who are willing to meet at the same time each week for at least 1 hour to review concepts from the class or study for exams; you can also take advantage of any time remaining in section for ad hoc sessions.

These are not intended to be sessions where you sit and do homework jointly – everyone should do their homework independently – but you can discuss issues that arise in doing your homework, swap R tips, and help each other in a constructive way. Remember that any increase you get on your homework grade by blindly following others will be drowned out by your inability to do well on the exam on your own!

Course Work:

Homework

Homework will be posted to bspace, and will generally be due 1-2 weeks later. All homework is due **in lecture**. Homeworks will be a combination of computational exercises and data analysis using the computer. Late homeworks will not be accepted except in exceptional circumstances.

The final homework score will be the sum of all homework grades, so 15 points in HW1 counts the same as 15 points in HW5, with points allocated to questions across the semester that aim to be compatible. This means the homework assignment as a block will not count equally, and it is always worth your while to turn in what you have done of your assignment, even if it is incomplete.

Exams

There will be two midterms, tentatively scheduled for the week of Feb 24th and April 14th, each of which will count equally.

Final Project

I am planning on creating a final project/exam covering the last portion of the class on complex surveys. You will have to analyze data from a complex survey. This is still in flux, and I may convert it to an in-class final exam.

Overall score

Your letter grade for the course will be based on the total points for all work in the semester.

- Homework: 20%
- Final Project/Exam (1): 35%
- Midterms (2): 45%

Policies

Late Assignments

Late assignments will not be accepted except with my personal approval, which will general result in a penalty except in extraordinary circumstances. I expect any requests to be done as timely as possible.

Email

1) If you wish for your email to make it into my inbox, the subject of your email must contain the text "152"
2) Neither I (nor the GSI) explain course material over email and will not respond to emails with such requests. Please come to 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 almost never in the evening nor weekend.

Academic Honesty Policy

- Homework must be done independently. If a homework problem is based on an experiment run in class or section, only the experimental results can be done jointly; answering HW questions about the experiment must be done independently. Obtaining and/or using solutions from previous years or from the internet, if such are available, is considered cheating. Within your study sessions, you may discuss issues about the homework, but you must **not** sit down and do the assignment jointly. Please note that while the homework is time-consuming, they are *en masse* 30% of your grade; the cumulative weight is large, but becoming desperate over single questions is **not** worth the risk of cheating!
- For projects students naturally will work with the students in their group, but no other students.
- For exams cheating includes, but is not limited to, bringing written or electronic materials into an exam or quiz, using written or electronic materials during an exam or quiz, copying off another person's exam or quiz, allowing someone to copy off of your exam or quiz, and having someone take an exam or quiz for you.

In fairness to students who put in an honest effort, cheaters will be harshly treated. Any evidence of cheating will result in a score of zero (0) on the **entire** assignment or examination. I will **always** report incidences of cheating to Student Judicial Affairs, which may administer additional punishment.

Disability

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

Syllabus

The following is only a guide, and there is likely to be slight variation as the semester progresses.

The reading described below is a guide to where the relevant material can be found in the book for the subjects described under 'Topic.' It is not a prescription to when **you** should actually read the material.

Week	Topics		Assigned Reading
Jan 20	Course Logistics Elements of a Sampling Survey		Lohr Ch 1, 2.1-2.3
Jan 27	Simple Random Sampling -- Estimation		Ch 2.4-2.8
Feb 3	Ratio and Regression Estimators with SRS		Ch 4.1-4.4
Feb 10	Sample Size		Ch 2.6
Feb 17	Bootstrap CI		Handout*
Feb 24		Midterm	
Mar 3	Stratified Sampling		Ch 3
Mar 10	Cluster Sampling		Ch 5
Mar 17	Complex Surveys		Ch 7
Mar 24	SPRING BREAK		
Mar 31	Variance in Complex Surveys Non-response Adjustment		Ch 9.1, 9.4, 8,1-8.2,8.4-8.5
Apr 7	Visualization with Complex Surveys		Lumley Ch 4
Apr 14	Regression	Midterm	Ch 11
Apr 21	Categorical Data		Ch 10
Apr 28	Estimating Population Size and Rare Population Estimation		Ch 14
May 5 (RR)			
	Final Exam/Project Due: May 16, 3-6pm		