

Sampling Methodology in Finite Populations (Statistics 5201)

Zack W. Almquist
Fall Semester, 2014

Class Schedule

Lecture: M/W/F 1:25 – 2:15 pm 130 Ford Hall

URL: <http://moodle.umn.edu>

Note: Requires UMN login and registration in class to access.

Professor

Name: Zack W. Almquist
Office: 372 Ford Hall
Office Hours: 3:00-4:00PM W
Email: almquist@umn.edu
Telephone: 612-624-4300 (not recommended)

Teaching Assistant

Name: Ding Xiang
Office: 350 Ford Hall
Office Hours: 4:00-5:00pm Th
Email: xiang045@umn.edu

Course Objectives

This course will focus on traditional sampling methods employed on finite populations, including simple random, systematic, stratified, cluster, and unequal probability sampling. Further, it will expose the student to ratio, and model based estimation in the context of finite population sampling. Additional topics covered include Bayesian sampling and Network sampling.

Prerequisites

It is assumed that the student has taken an introduction sequence to statistics such as 3011/3021, introduction to statistical analysis and data analysis, or a mathematical statistics sequence equivalent to 4101/4102, 5021 or 50101/5102. This includes, but is not limited to, basic familiarity with probability distributions, hypothesis testing and linear models. While not required it is highly beneficial to have some familiarity with the R statistical platform.

Course Requirements

Computers

It is not required that students bring their computers/laptops to lecture and lab (if one is owned), but it is *highly* recommended since both lecture and lab will make extensive use of the computer software R. Computer labs are available on campus, please consult with the TA if you have trouble finding the various locations that computer labs reside on campus.

Readings

Weekly readings are assigned on the course syllabus. All readings are assumed to be completed before each lecture/seminar. You are expected to read over the class notes each week and make sure you are familiar with the material as the course progress – questions are encouraged.

Homework

Homework assignments will be administered on an approximately bi-weekly basis and will be due on every other Wednesday. Homework assignments are meant to achieve three results: (1) provide practice with the statistical concepts discussed in class and, and (2) provide practice with the computational and statistical programming language R and (3) provide a chance to demonstrate your mastery of material and highlight areas where more work is needed. You may work in a group, but all write-ups must be done independently. All collaborators should be appropriately cited in your write up and any detailed R code should also be provided.

Exams

To assess mastery of course material, two exams will be administered over the course of the semester. The first of these will be in-class exam, and the other will be a take-home exam provided in the last week of instruction. The examinations will cover all material

presented in lecture, in addition to assigned readings. Students are advised to keep up-to-date on reading assignments, and to attend lectures regularly, so as to be prepared for exams. Students must be present for the exam to obtain credit; exams are not rescheduled due to travel or other considerations. Note that dictionaries or other language aids may not be used when taking exam (see English proficiency, under Prerequisites).

No calculators or phones will be allowed. A cheat sheet will be allowed, but is limited to one 8 by 11 inch piece of paper (e.g., notebook paper or printer paper).

Participation

Individuals are expected to attend every course, to have completed every reading and to participate with questions and discussion on each topic as presented. If you plan on missing any class period you are responsible for all material and for contacting the instructor in a timely manner.

Grading

Participation:	5%
Labs:	10%
Homework:	25%
Exam 1:	30%
Exam 2:	30%

Lectures, readings, labs, and review sessions are provided for each student's benefit. It is the responsibility of the student to take advantage of these opportunities to acquire and demonstrate mastery of course material, so as to achieve his or her desired grade.

Letter grade assignment

A	93%+
A-	90-92.99%
B+	87-89.99%
B	83-86.99%
B-	80-82.99%
C+	77-79.99%
C	73-76.99%
C-	70-72.99%
D	60-69.99%
F	<59.99%

Required Texts

- Sharon Lohr (2010). *Sampling: Design and Analysis, 2nd Edition*, Duxbury Press.

Readings

Be prepared to discuss all readings assigned at anytime in lecture/seminar.

Required Software

We will be using the R statistical programming language. R can be downloaded at <http://www.r-project.org/>.

Recommended Software

RStudio IDE (Integrated Development Environment) is a software application which facilitates interaction with the R statistical programming language. It is often preferred to the GUI (Graphic User Interface) made available through CRAN. You can download it at <http://www.rstudio.com/>.

Course Policies

Missing Class, etc.

It is expected that each member of the class will attend every lecture/discussion. If there is an appropriate reason to miss class it is expected that the individual will email or discuss in person with the instructor at least one week in advance. For any medical issues please see the UMN website for university policies.

Cheating, etc.

All work is assumed to be your own and all individuals are expected to follow the university policy on cheating and misconduct. If you have any questions please consult the UMN website for university policies.

Assignments

Homework Assignments

Homework will be assigned on a biweekly basis starting on the second Wednesday of the Semester and will be due two weeks later at 5:00pm. There will be a total of six homework

assignments. Homework assignments will be graded out of 100 pts. Each assignment must be turned in through moodle, no late assignments will be accepted. Homework must be turned using sweave/latex and must include all R code. Your lowest score will be dropped in the final calculation of grades.

Lab Assignments

There will be occasional R lab assignments throughout the Semester. These assignments are to facilitate your experience and learning with the R statistical programming language. All lab assignments will be graded based on completion. Your lowest lab score will be dropped.

Exams

- *In Class Exam:*
 - 10/24/2014
- *Take Home Exam:*
 - Handed Out: 12/10/2014
 - Due: 12/15/2014 by 5:00pm

Reading Assignments

Week 1: Introduction and Review

- *Readings:*
 - Chapter 1

Week 2: Simple Random Sampling

- *Readings:*
 - Chapters 2.1-2.3

Week 3: SRS Weights, CI and Sample Size Estimation

- *Readings:*
 - Chapters 2.4-2.6

Week 4: Randomization Theory and Prediction

- *Readings:*
 - Chapters 2.8-2.9

Week 5: Stratified Sampling

- *Readings:*
 - Chapter 3.1-3.2

Week 6: Stratified Sampling

- *Readings:*
 - Chapter 3.3-3.2

Week 7: Stratified Sampling

- *Readings:*
 - Chapter 3.3-3.7

Week 8: Review and Exam

- *Readings:*
 - Review Chapters 1-3

Week 9: Ratio and Regression Estimation

- *Readings:*
 - Chapter 4.1-4.4

Week 10: Ratio and Regression Estimation

- *Readings:*
 - Chapter 4.5-4.6

Week 11: Cluster Sampling with Equal Probabilities

- *Readings:*
 - Chapter 5

Week 12: Cluster Sampling with Unequal Probabilities

- *Readings:*
 - Chapter 6

Week 13: Advanced Topics - Network Sampling

- *Readings:*
 - Handout on Network Sampling

Week 14: Advanced Topics - Bayesian Sampling

- *Readings:*
 - Handout on Bayesian Sampling

Week 15: Review

- *Readings:*
 - Chapters 1-6, Network Sampling, and Bayesian Sampling Handouts