

SURVMETH/SURV 625: Applied Sampling

Winter 2025

1 Overview

Applied Sampling is an applied statistical methods course concerned almost exclusively with the design of data collection. This course will concentrate on probability sampling. This course will cover the main techniques used in sampling practice: simple random sampling, stratification, systematic selection, cluster sampling, probability proportional to size sampling, and multistage sampling. The course will also cover sampling cost models, sampling error estimation techniques, non-sampling errors, missing data, and nonprobability samples.

- Instructor: **Yajuan Si**, PhD, Website: <http://www.umich.edu/~yajuan>;
Email: yajuan@umich.edu; Phone: 734-7646935; Office hours: By appointment.
- GSI: **Akari Oya**, Email: akarioya@umich.edu; Office hours: By appointment
- Dates: January 8, 2025 - April 23, 2025
- Lectures: 9:00 - 11:30 AM, Wednesdays
- Location: ISR-Perry G300 (UMICH); Lefrak 2208 (UMD)
- Lecture [Zoom link](https://umich.zoom.us/j/96603684256): <https://umich.zoom.us/j/96603684256>, password: 2025.

2 Competency

The student will learn (and will be tested on) the following competencies in this course:

1. Learn the meaning and application of expressions for sampling variance of means and proportions under a variety of finite population sampling techniques.

2. Understand the properties of (and how to apply) various sampling techniques, including simple random sampling, stratification and stratified random sampling, cluster sampling, systematic sampling, two-stage sampling, multistage sampling, probability proportionate to size sampling, and stratified multistage sampling.
3. Understand sampling variance estimation techniques for means and proportions for each of the sampling techniques in the course, including procedures for nonlinear statistics using the Taylor series expansion technique, as well as the balanced repeated replication and jackknife repeated replication techniques.
4. Learn how nonresponse can affect survey estimates, and what techniques can be used to reduce nonresponse and compensate for it through weighting and imputation.

3 Prerequisites

The course is presented at a moderate statistical level. While mathematical aspects of sampling theory are not covered, statistical notation and some algebraic derivations will be presented. Thus, a thorough understanding of statistical notation and principles will be needed.

- Basic terms of probability theory
 - probability, joint probability, conditional probability
 - expectation, variance, covariance
- Basic algebra and calculus
- Basic visualization techniques (R or Python)
 - histogram, density plot, scatter plot
- Basic programming skills. For learning R programming basics: [Garrett Grolemund, Hands-On Programming with R](#)

4 Textbook and readings

- Sampling: Design and Analysis, Third Edition. Sharon L. Lohr. 2022. See also [home page for the book](#)
- Recommended [readings](#)
- Papers on related topics will be assigned and available via Canvas

5 Examinations and Grading

Final grades will be a weighted composite of homework (25%), a midterm exam (25%), and a final project report (oral presentation: 15%; report: 30%; individual evaluation score: 5%). The final grades will not be based on any kind of curve, and will use a standard grading scale (95-100% = A+, 90-94% = A, 85-89% = A-, 80-84% = B+, 70-79% = B, 60-69% = B-).

There will be a two-hour, in-class, cumulative, open book / open notes midterm examination on Wednesday, March 26, 2025.

6 Homework

Five homework assignments are planned. Each are to be turned in by the beginning of the class session when due (see syllabus below), and each will receive equal weights for the final grades. Each homework assignment is to be submitted via the Canvas web site Assignment tool when due.

Homework assignments correspond with the course units and are designed to aid in skill development. Assignments will be graded check-plus (100 points), check (90), check-minus (80), and not submitted (0), and will ordinarily be marked and returned before the next class session. Students may request permission to submit homework late, but the request must be no later than one hour before the homework is due. Permission is not guaranteed, although typically granted. If late submission is granted, there will be an agreed upon date and time when the late assignment must be submitted; such assignments will be graded using the specified marking system. If homework is submitted late without prior permission, scores will be check-plus (70), check (60), and check-minus (50).

7 Sampling project

The project is a team-based, semester-length exercise leading to 1) an oral presentation and 2) a 10-page technical memorandum describing the team's proposed approach to a survey sampling problem set by Dr. Si. The student projects will be assigned to teams of 3-4 students, depending on enrollment counts, and each team will submit a team project. The grading will consist of three parts: 1) a team oral presentation score; 2) a team project report score; and 3) an individual evaluation score. Each student will evaluate the contributions of each of the other team members to the oral presentation and the team project report. The student's teammate's evaluation scores will be averaged to generate the individual student project score. Each student must complete an evaluation for the other students on the team; students not submitting evaluations receive zero for the individual component score.

- Grading criteria for oral presentations: This is meant to be a team exercise, and all team members are expected to contribute equally to the final oral presentation. These details are described below.
1. The quality and appearance of the final oral presentation of the sampling plan (which can be done in PowerPoint or Beamer): Is the sampling process clearly described and justified with rationale? Are main points easy to understand from the slides? Are they structured in a logical way? Are there excessive typos and grammatical errors? [30 pts]
 2. The clarity of the technical information contained within the oral presentation: Is every key concept clearly described? Are figures and tables clearly labeled and easy to understand? Are formulas clearly labeled in terms of notation? Does it appear that calculations and formulas are correctly implemented? [30 pts]
 3. The quality of the oral presentation, whether it is provided by one team member or all team members: Did everyone clearly present the points? Was the presentation completed within the allotted time limit? Were the speakers clearly familiar with the content on the slides? [20 pts]
 4. The ability to answer audience questions, either from the instructors or the students: Were any questions from the audience handled professionally and clearly, and without hesitation? Were reasonable answers provided for the questions asked? [20 pts]

The instructor will score the oral presentations on each of the four criteria above. Each team will be expected to provide an overview of their analysis in no more than 15 minutes, and respond to questions for no more than 5 minutes. **The oral presentations will take place on April 16, 2025, and the final written projects will be due on April 23, 2025.** Scores and feedback on the final oral presentations will be provided during the week leading up to the submission of the final project.

8 Course attendance

Students are required to attend all live class sessions in person. A student who anticipates attending the class via Zoom or missing a class due to illness, emergency, or any other reason must notify the instructor via email at least 30 minutes before the session with the reason. University of Maryland students should examine University attendance policies at <https://faculty.umd.edu/teach/>. University of Michigan students should review <https://lsa.umich.edu/advising/policies-procedures/class-attendance.html>.

We follow the academic calendar at the University of Michigan and observe the Spring break at Michigan (March 1-9, 2025). We have one lecture scheduled during the University of Maryland Spring break (March 19, 2025). Students from the University of Maryland are permitted to attend the March 19 lecture via Zoom alive or watch Zoom recordings later.

9 Communication channels

- The primary communication channel is Canvas. Q & A are handled **ONLY** on Canvas (Piazza, etc.) and office hours, thus please avoid asking technical questions via email. Students are encouraged to help solving each other's problems on Canvas. Students are encouraged to share notes on Canvas. Bonus points (up to 5%) will be given according to students' involvement and contributions on Canvas.
- Don't ask technical questions via email or direct messages. By asking via common channels in the course chat, more eyes will see your question, it will get answered faster and it is likely that other students benefit from the answer.
- Please schedule with the GSI or the instructor for office hours if you have any questions that you do not want to post on Canvas discussions. Emails sent to the GSI or instructor about (a) contents that can be found on Canvas, or (b) technical questions including bugs in code, won't be replied.
- Proper collaboration is highly recommended. However, each student must submit individual homework. Refer to the Academic Conduct for policies of plagiarism. The instructor will follow the university honor code policy strictly. Lack of knowledge of the academic honesty policy is not a reasonable explanation for a violation. If you discuss with other students about homework, you are supposed to write down your own homework/code independently after the discussions. Under any circumstances, two exact copies of homework/code are considered to be plagiarism. Students are under their own risk of any type of penalty from the university in violating academic honesty.

10 Mental Health and Well-Being

University Students may experience stressors that can impact both their academic experience and their personal well-being. These may include academic pressures and challenges associated with relationships, mental health, alcohol or other drugs, identities, finances, etc. If you are experiencing concerns, seeking help is a courageous thing to do for yourself and those who care about you. If the source of your stressors is academic, please contact me so that we can find solutions together. For personal concerns, U-M offers a variety of resources, many which are listed on the Resources for Student Well-being webpage (<https://wellbeing.umich.edu/tools-resources/>). You can also search for additional well-being resources on that website.

11 Accommodations for Students with Disabilities

Students at University of Maryland: In order to receive services, contact the Accessibility & Disability Service (ADS) office. Students are responsible for scheduling a registration appoint-

ment with ADS well in advance of needed accommodation. Appointments can be scheduled by calling the ADS office at 301-314-7682. <https://www.counseling.umd.edu/ads/> Students at University of Michigan: If you think you need an accommodation for a disability, please let me know at your earliest convenience. Some aspects of this course, the assignments, the in-class activities, and the way the course is usually taught may be modified to facilitate your participation and progress. As soon as you make me aware of your needs, we can work with the Services for Students with Disabilities (SSD) office to help us determine appropriate academic accommodations. SSD (734-763-3000; <http://ssd.umich.edu>) typically recommends accommodations through a Verified Individualized Services and Accommodations form. Any information you provide is private and confidential and will be treated as such.

12 Academic Conduct

The course follows ethical standards at the respective campuses. The University of Maryland Honor Code (<https://president.umd.edu/administration/policies/section-iii-academic-affairs/iii-100a>) is administered by the Student Honor Council. The Student Honor Council statement I pledge on my honor that I have not given or received any unauthorized assistance on this assignment/examination ordinarily should be handwritten and signed on the front cover of all papers, projects, or other academic assignments submitted for evaluation in this course. The instructor waives this requirement for JPSM students, but not the other provisions of the Honor Code. The student name on a document is the student's pledge that this work is theirs and theirs alone. University of Michigan students are expected to follow academic integrity policies in the University of Michigan Standard Practice Guide (<https://spg.umich.edu/policy/303.03>). Violation of the Maryland Honor Code pledge or the University of Michigan policy is a serious matter and may lead to a failing mark on a submission or for the entire course.

13 Sexual Misconduct Policy

Title IX prohibits discrimination on the basis of sex, which includes sexual misconduct - including harassment, domestic and dating violence, sexual assault, and stalking. We understand that sexual violence can undermine students' academic success and we encourage anyone dealing with sexual misconduct to talk to someone about their experience, so they can get the support they need. Confidential support and academic advocacy can be found with the Sexual Assault Prevention and Awareness Center (SAPAC) on their 24-hour crisis line, 734-936-3333, and at sapac@umich.edu.

Alleged violations can be non-confidentially reported to the Equity, Civil Rights, and Title IX Office (ECRT) at ecrtoffice@umich.edu.

14 Course Schedule

All assigned readings need to be completed prior to the start of the indicated class, and topics from these readings will be discussed in class. The order of the topics and assigned readings are subject to change.

Week	Topic	Assigned readings/HW due dates	Project tasks
1. Jan 8	Course description (Syllabus); Review of basic statistics	Lohr 2.1-2.2, Appendix A	Introduction
2. Jan 15	Simple random sampling	Lohr 2.3-2.7; HW#1 released	Question description
3. Jan 22	Stratified sampling	Lohr 3.1-3.5	Sampling frame
4. Jan 29	Ratio and regression estimation	Lohr 4.1-4.2; HW 1 due ; HW 2 released	Project teams and target sample size
5. Feb 5	Equal sized cluster sampling	Lohr 5.1-5.2 ;	Subsample size, achieved precision.
6. Feb 12	Unequal sized cluster sampling	Lohr 5.3-5.4; HW 2 due ; HW 3 released	Sampling rates, domains
7. Feb 19	Systematic sampling	Lohr 2.8, 5.5	Stratification
8. Feb 26	Sampling probability proportional to (estimated) size	Reading ; HW 3 due ; HW 4 released	
9. Mar 5	No class , Happy Spring break		
10. Mar 12	Sampling with unequal probabilities	Lohr 6.1-6.5;	Second stage selection and subclasses
11. Mar 19	Complex surveys	Lohr 7.1-7.5; HW 4 due ; HW 5 released	
12. Mar 26	Midterm exam		
13. Apr 2	Variance estimation	Lohr 9.1-9.5	Estimation procedures
14. Apr 9	Nonresponse adjustment	Lohr 8.1-8.8; HW 5 due	
15. Apr 16	Review, Final project team presentation		
16. Apr 23	Final project team presentation, report		Report due at 5pm