

University of Toronto
Department of Statistical Sciences
STA304H1F LEC0101
Surveys, Sampling and Observational Data
Course Outline - Summer 2024

Lectures

TU 2:00 pm – 5:00 pm in MC 102
TH 2:00 pm – 5:00 pm in MC 102

Instructor Office Hours

TU 1:00 pm – 2:00 pm in MC 102
TH 1:00 pm – 2:00 pm in MC 102

Instructor: Dr. Luai Al Labadi

E-mail: luai.allabadi@utoronto.ca

TA Information: The TAs' contact information and office hours will be posted on Quercus.

COURSE OVERVIEW

- **Course Description:** Design of surveys, sources of bias, randomized response surveys. Techniques of sampling; stratification, clustering, unequal probability selection. Sampling inference, estimates of population mean and variances, ratio estimation. Observational data; correlation vs. causation, missing data, sources of bias.
- *Prerequisite:* ECO227Y1/ STA255H1/ STA261H1/ STA248H1/ STA238H1/ STAB57H3/ STA258H5/ STA260H5/ ECO227Y5
Exclusion: STAC50H3, STAC53H3, STA304H5
- Pre-requisites are **strictly enforced by the department, not the instructor**. Students who do not meet the equivalent pre-requisites will be automatically dropped from the course.
- **Learning Outcomes:** By the end of this course, students should be able to:
 1. Design a survey or sample that is appropriately gathering information of interest.

2. Carry out a variety of statistical analyses to make inference on the data collected from a survey/sample.
3. Identify and implement different sampling techniques and different study designs and the trade-offs involved in each.
4. Identify sources of bias within a study and comment on a study's design, including its weaknesses, strengths, and appropriate analyses.
5. Clearly communicate results of statistical analyses to technical and non-technical audiences.

COURSE MATERIALS

- **Course Content:** All lecture slides, materials, and important announcements will be posted on Quercus <https://q.utoronto.ca>. Please make sure to check it regularly so you do not miss anything.
- **Required Textbook**
 - ✓ *Elementary Survey Sampling, 7th edition*, by Scheaffer, Mendenhall, Ott & Gerow
 - ✓ Publisher: Cengage
 - ✓ ISBN-10: 0840053614 | ISBN-13: 9780840053619
 - ✓ Publication Date: 2012
 - ✓ **Available as an E-book**
- **Additional References**
 - ✓ *Sampling: Design and Analysis, 2nd edition*, by Sharon L. Lohr. Cengage, ISBN 9780495105275
 - ✓ *Sampling Theory and Practice* by Changbao and Thompson. Springer, ISBN: 978-3-030-44246-0
 - ✓ *Complex Surveys: A Guide to Analysis Using R*, by Lumley. Wiley, ISBN 978-0-470-28430-8
- **Calculators:** Handheld, **non-programmable** calculators may be used during tests/exam. Any calculator that has a logarithm, square root, and one memory button will suffice for this course, so there is no need to buy an expensive calculator.
- **Statistical Computing:**
 - ✓ This course uses the statistical package R/RStudio/Jupyterhub.
 - ✓ R is free statistical software and it can be downloaded from <http://cran.r-project.org/>.
 - ✓ JupyterHub (<https://datatools.utoronto.ca/>) allows you to work with this software without having to download anything to your computer.

COURSE COMPONENTS

- **Lectures:** Attending lectures, actively engaging with the topics, and seeking clarity on any questions are crucial aspects of succeeding in this advanced-level course. To effectively learn the material, students should make the most of the lecture time and regularly practice the concepts covered. Relying on last-minute cramming before tests/exams is unlikely to be beneficial.
- **Office Hours:** The instructor/TAs will hold office hours as described above or on Quercus. It is recommended that you visit office hours whenever you have a question about the material. It is very important to have material clarified as quickly as possible. Do not wait until the last minute to ask your questions. Any change in office hours will be communicated on Quercus or Piazza
- **Piazza:** This is for student-led discussion. All questions about course material should be posted here or asked during instructor/TAs office hours. The instructor and TAs will monitor the board and will help answer questions, but students are encouraged to answer posts and help their fellow classmates.

ASSESSMENTS AND DEADLINES

Type	Due Date	Weight
Test 1	May 21	25%
Test 2	May 30	25%
Final	TBA	50%

- **Term Tests Policies**
 - ✓ All the term tests start at 3:30 pm and end at 5:00 pm on the dates specified above.
 - ✓ There will be class from 2:00 pm -3:00 pm on the day of term test.
 - ✓ Your test may be in a different room. The location will be communicated on Quercus.
 - ✓ *Missed Term Work:*
 - If you are unable to attend a test due to a valid reason, please send an email to Luai in advance, if possible. Additionally, it is important to declare your absence on ACORN.
 - If you missed **one** term test, no makeup will be given for the missed term test. The mark of the missed term test will be substituted based on the final exam.

- If **both** tests are missed, a makeup test will be held on Thursday, June 13 from 5:30 pm to 7:00 pm. More information will be provided on Quercus. The makeup test will cover all course material and count for 25% of the final grade. Any unallocated weight from the missed tests will be shifted to the Final exam. **Not attending the makeup test will result in a score of zero on the test, with the weight of the final exam being 75% in that situation.**
 - **Re-mark Policy:** If you believe there is an issue with the marking, you can request a re-mark. The re-mark policy is in place to correct mistakes. Any request should clearly identify the error, such as an unmarked question or an incorrectly calculated total. Requests to correct such mistakes must be addressed by your instructor, not the TA. To be eligible for a re-marking request, you should either make the request upon collecting the paper (if the test is not marked using crowdmark) or by emailing Luai within 7 business days from when the graded test was initially available (if crowdmark is used). Ensure that the subject line of the email includes STA305, provide your full name and student ID number, and present specific, clear, and concise reasons for each request. Cite potential errors or omissions made by the marker. Re-marking requests lacking specific reasons will not be accepted. Note that your entire test may be remarked when submitting a re-marking request. Therefore, it is possible that your mark may go down if the regraded mark is lower than your original mark.
- **Final Exam**
 - ✓ Final exam is scheduled during the June examination period by the Office of Registrar. The final exam will cover the entire course. Final exam grades will not be posted in Quercus. Issues related to final exams (e.g. time conflict, remark requests/exam viewing) should be addressed to the Registrar's office. See
 - <https://www.artsci.utoronto.ca/current/faculty-registrar/exams-assessments/exam-conflicts>
 - <https://www.artsci.utoronto.ca/current/faculty-registrar/exams-assessments/exam-viewing>
 - ✓ *Missed Final Exam:* Students who cannot complete their final examination due to illness or other serious causes must file an online petition: <https://www.artsci.utoronto.ca/current/faculty-registrar/petitions/deferred-exam>. Students must also record their absence on ACORN on the day of the missed exam or by the day after at the latest. Please refer to <https://www.artsci.utoronto.ca/current/faculty-registrar/petitions/deferred-exams>

for more information on how to request a deferred exam, and deadlines.

USE OF GENERATIVE AI: In this course, students may use artificial intelligence tools, including generative AI, as learning aids. However, their use is strictly prohibited during term tests and final exam.

EMAIL POLICY: Your email must originate from your University of Toronto email account when you contact your instructor by email. The subject line should contain the course number and a relevant subject (indicating what the email is about). Be sure to include your full name and student ID number in the body of the message. Before you send an email, make sure that you are not asking for information that is already available from the course outline/website/announcements, or questions about the course material that are more appropriate for discussing during office hours or discussion board on Piazza. **In general, your instructor and TAs will not answer technical questions about the course material by email.**

INTELLECTUAL PROPERTY: Course materials provided on Quercus, such as lecture slides, assignments, tests and solutions are the intellectual property of your instructor and are for the use of students currently enrolled in this course only. Providing course materials to any person or company outside of the course is unauthorized use. This includes providing materials to predatory tutoring companies.

ACADEMIC INTEGRITY: The University treats cases of plagiarism and cheating very seriously. It is the students' responsibility for knowing the content of the [University of Toronto's Code of Behaviour on Academic Matters](#). All suspected cases of academic dishonesty will be investigated following procedures outlined in the above document. If you have questions or concerns about what constitutes appropriate academic behaviour or appropriate research and citation methods, you are expected to seek out additional information on academic integrity from your instructor or from other institutional resources (see <http://academicintegrity.utoronto.ca/>). Offences include, but are not limited to:

On tests and exams:

1. Using or possessing unauthorized aids.
2. Looking at someone else's answers during an exam or test.

3. Obtaining or providing unauthorized assistance.
4. Misrepresenting your identity.

In academic work:

1. Falsifying institutional documents or grades.
2. Falsifying or altering any documentation required.

All suspected cases of academic dishonesty will be investigated following procedures outlined in the Code of Behaviour on Academic Matters. If you have questions or concerns about what constitutes appropriate academic behaviour or appropriate research and citation methods, you are expected to seek out additional information on academic integrity from your instructor or from other institutional resources.

ACCESSIBILITY NEEDS: The University of Toronto offers academic accommodation for students with disabilities. If you require accommodations, or have any accessibility concerns about the course, the classroom, or course materials, please contact Accessibility Services as soon as possible: accessibility.services@utoronto.ca or <http://accessibility.utoronto.ca>.

SUGGESTED PROBLEMS FROM THE TEXTBOOK: They will form the basis for the term tests and the final exam, although it does not imply that you will be tested on these problems. They will be essential for your understanding of the topics covered in class.

- **Ch. 2:** 2.2-2.5, 2.8, 2.9, 2.11, 2.12, 2.14, 2.15, 2.28, 2.29, 2.30.
- **Ch. 3:** 3.1-3.8, 3.13, 3.14.
- **Ch. 4:** 4.1, 4.2, 4.4, 4.5, 4.8, 4.11, 4.14, 4.15-4.28, 4.41, 4.42, 4.47, 4.48.
- **Ch. 5:** 5.1, 5.2, 5.5, 5.13, 5.14, 5.25, 5.26, 5.32, 5.36.
- **Ch. 6:** 6.1-6.5, 6.7, 6.9, 6.10, 6.11, 6.23, 6.31.
- **Ch. 7:** 7.1, 7.2, 7.4, 7.5, 7.11, 7.12, 7.16, 7.17, 7.18.
- **Ch. 8:** 8.1, 8.8, 8.9, 8.13, 8.18, 8.20, 8.23.
- **Ch. 9:** 9.1, 9.2, 9.4, 9.7, 9.8, 9.11, 9.12.
- **Ch. 10:** 10.1, 10.5, 10.8, 10.10, 10.11, 10.13.

APPROXIMATE LECTURE SCHEDULE: This schedule is subject to change. Updates will be announced in lectures and posted on Quercus.

Lecture #	Date	Topic	Text Reference
1	May 7	Introduction	Ch. 1
		Technical terms	2.2
		How to select a sample: The design of the sample survey	2.3
2	May 9	Sources of error	2.4
		Questionnaires & planning	2.5-2.6
		Ethics in research	-
		Infinite/finite population summary statistics	3.1-3.3
		Sampling distributions, correlation & estimation	3.4-3.6
3	May 14	Simple random sampling	4.1-4.2
		Estimation of a population mean/total	4.3
4	May 16	Sample size	4.4
		Estimation of a proportion & comparing estimates	4.4-4.6
		Stratified random sampling & estimation of a mean/total	5.1-5.3
5	May 21	Selecting & allocating sample sizes; Test 1	5.4-5.5
6	May 23	Estimating a proportion, and selecting & allocating sample sizes for proportions	5.6-5.7
		Optimal stratification	5.8-5.9
7	May 28	Ratio estimation	6.1-6.3
		Selecting sample size	6.4
8	May 30	Regression & difference ratio estimation	6.5-6.7

		Comparing Estimates; Test 2	6.8
9	June 4	Systematic sampling & estimation of a mean/total	7.1-7.3
10	June 6	Estimating a proportion, and selecting sample size	7.4-7.5
		Repeated systematic sampling	7.6
		Equal cluster sizes	8.4
11	June 11	Selecting sample size & estimating a proportion	8.5-8.7
		Two stage cluster sampling	9.1-9.2
		Estimation of a mean/total	9.3
12	June 13	Ratio estimation of a mean, estimating a proportion	9.4-9.5
		Equal sized clusters & probability proportional to size	9.6-9.7
		Estimating a population size	10.1-10.3