University of Toronto
Department of Statistical Sciences

STA304H1F LEC0101
Surveys, Sampling and Observational Data

Course Outline - Summer 2022

Lectures	**Instructor Office Hours**
TU 2:00 pm – 5:00 pm in BA 1160	MO 6:00 pm – 7:00 pm in Zoom
TH 2:00 pm – 5:00 pm in BA 1160	WE 6:00 pm – 7:00 pm in Zoom
	By appointment

Instructor: Dr. Luai Al Labadi

E-mail: luai.allabadi@utoronto.ca

Teaching Assistants: TBA 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**. If you do not have the equivalent pre-requisites, you will be un-enrolled 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 don't 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 quizzes and tests. 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:**

  o This course uses the statistical package R/RStudio/Jupyterhub.
  o R is free statistical software and it can be downloaded from http//cran.r-project.org/.
  o Jupyterhub (https://jupyter.utoronto.ca/) allows you to work with this software without having to download anything to your computer.

COURSE COMPONENTS

- **Lectures:** It is important to attend the lectures, follow the topics, and not leave any questions behind. This is a higher-level course, and students will need all the lecture time and regular practices to learn the material well. Crunching before tests/exams is unlikely to be helpful.

- **Office Hours:** The instructo will hold office hours through Zoom. The link will be posted 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. Don't wait until the last minute to ask your questions.

- **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 24	25%
Test 2	June 2	30%
Final	TBA	45%

- **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 2:00-3:00 on the day of term test.
  - Your test may be in a different room. The location will be communicated on Quercus.
  - *Missed Term Work:*
    - ✓ Students must record their absence in ACORN on the day of the missed test or by the day after at the latest. 0% will be recorded if they do not record your absence on ACORN.
    - ✓ If you missed **ONE** term test, no makeups will be given for the missed term test. The mark of the missed term test will be substituted based on the final exam.
    - ✓ Students are allowed to miss only one term test. If <u>both</u> tests are missed, then 0% will be recorded for missed Test 1. The marks of missed Test 2 will be substituted based on the final exam.

- **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.

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.

ACCESSIBILITY NEEDS

The University of Toronto offers academic accommodations 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, 2.3, 2.8, 2.9, 2.11, 2.12, 2.14, 2.15, 2.22, 2.25, 2.26, 2.27, 2.28, 2.29, 2.30
- **Ch. 3:** 3.1-3.8, 3.13, 3.14.
- **Ch. 4:** 4.1, 4.2, 4.4, 4.8, 4.11, 4.14, 4.15-4.28, 4.45, 4.47, 4.48.
- **Ch. 5:** 5.1, 5.2, 5.5, 5.13, 5.14, 5.25, 5.26, 5.29, 5.31, 5.36.
- **Ch. 6:** 6.1-6.5, 6.7, 6.9, 6.10, 6.11, 6.23.
- **Ch. 7:** 7.1, 7.2, 7.4, 7.5, 7.11, 7.12, 7.16, 7.18.
- **Ch. 8:** 8.1, 8.8, 8.9, 8,18, 8.20, 8.23.
- **Ch. 9:** 9.1, 9.2, 9.4, 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 10	Introduction	Ch. 1
		Technical terms	2.2
		How to select a sample: The design of the sample survey	2.3
2	May 12	Sources of error	2.4
		Questionnaires & planning	2.5-2.6
		Infinite/finite population summary statistics	3.1-3.3
		Sampling distributions, correlation & estimation	3.4-3.6
3	May 17	Simple random sampling	4.1-4.2
		Estimation of a population mean/total	4.3
4	May 19	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 24	Selecting & allocating sample sizes; **Test 1**	5.4-5.5
6	May 26	Estimating a proportion, and selecting & allocating sample sizes for proportions	5.6-5.7
		Optimal stratification	5.8-5.9
7	May 31	Ratio estimation	6.1-6.3
		Selecting sample size	6.4
8	June 2	Regression & difference ratio estimation; **Test 2**	6.5-6.7
9	June 7	Comparing Estimates	6.8

		Systematic sampling & estimation of a mean/total	7.1-7.3
10	June 9	Estimating a proportion, and selecting sample size	7.4-7.5
		Repeated systematic sampling	7.6
		Equal cluster sizes	8.4
11	June 14	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 16	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