STA 304H1 F/1003H F, FALL 2019
Surveys, Sampling and Observational Data
Section: L0101

Class time and location: Wednesdays 13:10-14:00 and Fridays 13:10-15:00 in **ES 1050**

Course Website: https://q.utoronto.ca

Instructor: Dr. Fode Tounkara

E-mail: f.tounkara@utoronto.ca

Office hours: Wednesday 3-5 pm in **PHS 384**

Teaching Assistants::

Office hours: TBD in PHS Building, 3rdFloor(381)

Course Description

This course presents mathematical and statistical reasoning behind sampling, aspects of inference from surveys, and the interplay with observational studies.

Course Objectives

At the end of this course, students should be able to: 1) design and implement surveys with the following sampling designs: simple random, systematic, stratified, cluster and multistage; 2) estimate population mean and variance, population proportion, and ratio, 3) understand Observational data; correlation vs. causation, missing data, sources of bias.

Pre-requisite

Students should have the following statistics courses: ECO220Y1/ECO227Y1/GGR270Y1/PSY201H1/SOC300Y1

Exclusion: STA322H1

Resources

Course webpage

- Lectures notes are available through the learning portal at **https://q.utoronto.ca**.

Texbooks

1. *Elementary Survey Sampling*, **7th edition**, by Scheaffer, Mendenhall, Ott and Gerow.

(We will cover most of Chapters 1 through 5, and selected parts of Chapters 6 through 11.)

2. *Sampling: Design and Analysis*, **2nd edition** by Sharon Lohr. (Useful but not required)

Online Discussion Board

This term you will have the option to use Piazza for class discussion. If you decide not to use Piazza it will not disadvantage you in any way, and will not affect official University outcomes (e.g., grades and learning opportunities). If you choose not to opt-into Piazza then you can ask questions or discuss course material with the instructor or TAs during office hours. Be sure to read Piazzas Privacy Policy and Terms of Use carefully.

Take time to understand and be comfortable with what they say. They provide for substantial sharing and disclosure of your personal information held by Piazza, which affects your privacy. If you decide to participate in Piazza, only provide content that you are comfortable sharing under the terms of the Privacy Policy and Terms of Use. The Piazza system is highly catered to getting you course material help fast and efficiently from classmates, the TA, and the instructor. Rather than emailing questions, I encourage you to post your questions on Piazza.

To sign up for the discussion forum go to the link: http://piazza.com/utoronto.ca/fall2018/sta3041003

Evaluation

Assessment	Weight	Due data	Time	Location
Assignment 1	5 %	Thursday, September 26	11:59 pm	Submit online (Crowdmark)
Term Test 1	20 %	Wednesday, October 16	13:10-14:00	TBA
Assignment 2	5 %	Thursday, November 7	11:59 pm	Submit online (Crowdmark)
Term Test 2	20 %	Wednesday, November 20	13:10-14:00	TBA
Final Exam	50 %	Between December 7-20	(3h)	TBA

Assignments

- The assignments will each be of a practical nature, for which the use R will be required
- Must be submitted online into **Crowdmark** by 11:59pm on the due dates
- Late assignments will be accepted but subject to a 20% penalty per day late
- Late submissions will not be allowed beyond 48 hours of the due date

Term Tests

- The test will be written in locations to be announced (TBA).
- Term Test 1: One page, one-sided **handwritten 8-1/2 $\times$ 11** aid sheet allowed.
- Term Test 2: Two pages, one-sided (or one page, two-sided) **handwritten 8-1/2 $\times$ 11** aid sheet allowed.

Final exam

- Is cumulative
- Four pages, one-sided (or two pages, two-sided) **handwritten 8-1/2 $\times$ 11** aid sheet allowed.

Note: The formula sheet is cumulative, that is, every time you can add new pages; you cannot include any worked examples, only theoretical formulas and definitions, no properties; highlighting is allowed; hand-written only).

Note: Non-programmable, scientific calculators are permitted on the test and exam. Calculators on phones and other devices equipped with remote access will not be permitted during the tests or final exam.

Note: I do not negotiate grates unless a mathematical error has been made on my part.

Re-grading Policy

Any requests to have marked work re-evaluated must be made in writing within one week of the date the work was returned to the class. The request must contain a justification for consideration. Be sure to include your official name, student number and/or paper number for identification purposes.

Missed Tests:

There are no make-up tests. For a missing test without a reason (U of T doctor's note) you receive a zero mark. With a valid reason your mark will be adjusted. If you miss the first test (or assignment), the second test (assignment) weight will be adjusted. If you miss the second test (assignment), its weight will be shifted to the final exam. (warning: difficulty increases from the first test to the final; final exam covers complete course).

Tutorials:

There are no tutorials, but you can come for help to Stat. Aid Centre, PHB 381 (3 rd floor), before tests: date and time TBA. Some extra office hours before the final will be available.

Software

The course includes a lot of numerical calculation. We will use **Statistical programming language and environment R** software to support the class and sampling from real populations (data sets provided by the textbook), but **for you using R is optional**. **Understanding of R outputs is required**, on the level explained in the class.

R is freely available for download at

- **R** : https://www.r-project.org/

Rstudio is a good integrated development environment to **R**. It is also freely available at:

- **Rstudio** : https://www.rstudio.org/

Good online reference

- https://cran.r-project.org/doc/contrib/Torfs+Brauer-Short-R-Intro.pdf

For assignment, you can use **Rmarkdown** to write your solution.

- **Good online reference for Rmarkdown** : https://www.rstudio.com/wp-content/uploads/2015/02/rmarkdown-cheatsheet.pdf

Course outline:

Almost all of the course material is covered by the textbook. Related to the basic level of the textbook, some theoretical results will be considered in more detail.

The following is a tentative schedule for the course:

1. Sampling problems and notions (Ch 2).
2. Basic concepts (Ch 3; 3.3 will be covered in Ch 8, not here).
3. Simple random sampling (Ch 4; 4.6 is optional).
4. Ratio, regression, and difference estimation (Ch 6; 6.5 will be covered with Ch 5, below).
5. Stratified random sampling (Ch 5; 5.9, 5.10, 5.11 are not covered).
6. Systematic sampling (Ch 7).
7. Cluster sampling (Ch 8; 8.8 is not covered).
8. Two-stage cluster sampling (Ch 9; part of 9.6, on pp 301, 302, is not covered).
9. Supplemental topics, nonsampling errors (Ch 11.1, 11.4, 11.8 are covered, and maybe more, if time allows).

How to communicate with your instructor

- Questions about course material such as:
  - How do I do question 3.7 in the textbook?
  - What is sample unit?
  - When is the midterm?

should be posted on the discussion forums on Piazza. Questions can be posted anonymously (so that the author is anonymous to other students but not to the instructors), if desired.

- For private communication, such as:
  - I missed the test because I was ill.

e-mail your instructor, and include your full name and student number..

Note: I will only respond to e-mails you send me if they come from your e-mail account *@utoronto.ca*.

University of Toronto academic integrity

You are responsible for knowing the content of the University of Toronto's Code of Behaviour on Academic integrity at http://www.governingcouncil.utoronto.ca/Governing_Council/policies.htm. If you have any question about what is or is not permitted in this course, please do not hesitate to contact your instructor.

Students with Disabilities

I am committed to teach every student in this course. The University of Toronto is committed to accessibility. If you require accommodations for a disability, or have any accessibility concerns about the course, the classroom, or course materials, please contact Accessibility Service as soon as possible at https://www.utoronto.ca/accessibility. Students who may need course adaptations because of disability are welcome to make an appointment to see me during office hours.

Your responsibility

The course is designed to actively engage you in the course material. We hope you'll find the subject of statistics interesting, challenging, and fun, and an excellent opportunity to truly learn the material. In order for these sessions to be effective, preparing by learning about the week's concepts through the notes is essential.