

STA 304F/1003F:
Samples, surveys and observational data/Sample survey theory

Time Wednesday 1-3 pm Friday 1-2 pm

Place MS 2158

Course description This course teaches mathematical and statistical reasoning behind sampling, aspects of inference from surveys, and the interplay with observational studies. In addition to the topics listed in the calendar description, I will include discussion of statistics and society, and interpreting studies reported in the news. The undergraduate calendar description is:

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: ECO220Y1/ECO227Y1/GGR270Y1 / PSY202H1/
SOC300Y1/STA221H1/STA255H1/261H1/248H1)

(Exclusion: STA322H1)

Grading There will be two midterm tests, on October 21, 2011 (25%) and November 25, 2011 (25%) and a final exam (50%). Test weights will be adjusted if tests are missed for valid medical reasons. The material covered on successive tests and exams is cumulative. Test solutions will be posted on the course web site, which is on Blackboard. Homework will be assigned and discussed in class, but not graded.

Text The text for this course is “Elementary Survey Sampling” by Scheaffer, Mendenhall and Ott. The seventh edition is current but the 6th edition is fine. I will make sure all the questions and readings from the text are compatible with the 6th edition.

The book “Sampling: Design and Analysis” by Sharon Lohr may be a useful reference.

Computing This course does not require extensive computing, but there will be some calculations needed. You are welcome to use a programmable calculator, or the statistical computing package of your choice. The textbook web site provides some Excel macros for the examples in the text, and Appendix B gives SAS macros as well. I use the R computing package, and will refer to this from time to time. There is a good introduction to R at <http://wiener.math.csi.cuny.edu/Statistics/R/simpleR/>, and this link has been added to the course web page. Additional help with R will be provided during Friday classes.

Wednesday and Friday

On Wednesdays I will generally lecture on material from the book, on issues arising in the practical implementation of surveys, and on material related to “Statistics in the News”. Fridays will be for review, homework questions, term tests, tutorials on using R, and for clarifying points from the Wednesday lecture, as needed.

Syllabus

We will cover most of Chapters 2 through 6, and selected parts of Chapters 7 through 11. A more detailed syllabus will be provided as the course progresses.

What’s on the test?

The tests will cover material in the textbook and material discussed in class. The tests will also be cumulative in terms of the material they cover. Detailed section numbers for the text will be given well in advance of each test. Each test, and the final exam, will have at least one question from the assigned homework.

What can I bring to the test?

A single 8 1/2 x 11 sheet of notes and a **non**-programmable calculator.

Course web page

The course web page is on Blackboard. Slides, handouts, homework solutions and grades will be posted here. The blog is for discussion of the course; questions that require an answer from the TAs or instructor should be posted on the Discussion Board.

Students in 1003

I will add you as a guest to STA304, so you can access notes, handouts and so on. Please send me an email if this is not working. Please also check the STA304 page for announcements. Your grades will be posted to the STA1003 page.

Contacts Nancy Reid: SS 6002A, reid@utstat.utoronto.ca, 978-5046
Office hours Thursday 1 to 3

Uyen Hoang: u.hoang@utoronto.ca

Rihong Hui: rihong@utstat.utoronto.ca