

STA 304F/1003F:

Samples, surveys and observational data/Sample survey theory

Time Wednesday 1-3 pm Friday 1-2 pm

Place SS 2117

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 16, 2009 (25%) and November 20, 2009 (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 "Sampling: Design and Analysis" by Sharon Lohr. The second edition will not be ready until mid-October so I recommend the first edition. I will make sure all the questions and readings from the text are compatible with the first edition. The version with the CD-ROM will be more convenient but this is not required. The book by Sheaffer et al., "Elementary Survey Sampling" may be a useful reference.

Computing This course does not require extensive computing, but there will be some. You are welcome to use the statistical computing package of your choice, but I will refer when needed to the R computing package. R is available on CQUEST and UTSTAT, but probably more convenient is to obtain the free download available at probability.ca/cran or www.r-project.org. There are many helpful introductions to R listed on the course webpage, and I will also provide needed details in lectures.

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