STA261: Probability and Statistics II

Shahriar Shams

Winter 2019

1 Course Description

STA261H is a rigorous introduction to the theory of statistical inference and to statistical practice. Some of the topics that will be covered are: Statistical models, parameters, and samples; Estimators for parameters, sampling distributions for estimators, and the properties of consistency, bias, and, variance; The likelihood function and the maximum likelihood estimator; Hypothesis tests and confidence regions; Examples illustrating statistical theory and its limitations; Introduction to the use of a computer environment for statistical analysis. This course builds on the material introduced in STA257H.

Contents, emphasis, etc. of the course is defined by means of the lecture materials - *not only the texts.* Table 1 shows the tentative lecture guide. Lecture slides will be uploaded every week. However, they are just rough, point-form notes, with no guarantee of completeness or accuracy. They should in no way be regarded as a substitute for attending the lectures and tutorials, or for doing the weekly non-credit assignments.

There will be some lecture examples using statistical software **R**. Students are encouraged to get familiar with this software through the course however *no previous exposure is expected.* Important announcements, problem sets, additional examples, and other course info will be posted on the course web page on Quercus. Check it regularly.

2 Course Schedule

- **Lectures:**
  - L0101- Room # MC-102, Mon (3 - 5pm) and Wed (3 - 4pm)
  - L5101- Room # MS-3153, Wed (7 - 10pm)
- **Instructor:** Shahriar Shams, PhD in Biostatistics candidate, Dalla Lana School of Public Health, University of Toronto.
- **email:** shahriar.shams@mail.utoronto.ca (Please write "STA261" in the subject of your email. PLEASE!)
- **Office hours:** Room # WE-74, Thursday 9.30 - 11.30am (Starting from Jan 17th)

3 Textbooks

1. Probability and Statistics: The Science of Uncertainty, Second Edition, by Michael J. Evans and Jeffrey S. Rosenthal
Available online (FREE!) on the web-page of Professors Evans and Rosenthal
http://www.utstat.toronto.edu/mikevans/jeffrosenthal/

2. Mathematical Statistics and Data Analysis, 3rd Edition, John A. Rice

4 Tutorials

The tutorials will start on the second week and run until the last week of class. Tutorials will cover topics taught in the previous week's lecture. In preparation for the tutorials, you should do weekly non-credit assignments. There will be short **quizzes every other week starting from week 3** based on previous 2 week's lectures and non-credit assignments. You have to write the quizzes in your assigned tutorial. Out of the 5 quizzes, your lowest one will be dropped. The remaining 4 will worth 12% (3% each). If you miss any quiz for any unforeseen circumstances you do not need to submit any paperwork. You will get zero for any missed quizzes and since the lowest one will be dropped it should not affect your grade in case you miss one.

5 Non-credit assignments

Every week after the lecture a set of exercises will be provided. These assignments are **not for credit**. They are only meant to give students opportunities to learn the materials and prepare themselves for the tutorials/quizzes/tests and exam. The instructor will help you solve few of the problems from these assignments during the weekly office hours. And the TAs will also solve few during the tutorials.

6 Assignment for credit

Sometime during the middle/end of the term, you will be doing an assignment which will help you to understand few of the theories that we will be using in this course. The assignment will be a bit tedious if you are doing the calculations by hand. But using a software (like R) will make it a really easy one to complete. It will worth 3% and clear instructions will be given on how to complete and submit your work.

7 Evaluation

- **Quizzes + Assignment:** 12% + 3% = 15%
- **Mid-term:** 35% (date and place to be announced later)
- **Final:** 50% (will cover everything taught in the course)

8 Missed test

There are **NO make-up tests**. If a test is missed for a valid reason, you must email the instructor within 24 hours, and submit appropriate documentations (doctor's notes, prescriptions, money receipts) to the course instructor within 2 business days of the test. Print your name, student number, course number, and date on the submitted documents. If documents are not received in time, your test mark will be ZERO. If a test is missed for a valid reason, its weight will be shifted to the final exam.

9 Calculators

Hand calculators will be needed for this course. **NO phone calculators** are allowed in tests/exams/quizzes. Any regular one with a square root and logarithmic function will do.

10 Computing

Statistical software **R** will be used occasionally. **No previous computing experience is assumed**. In no way students will be tested on their expertise on this software. Any code used in the lectures to demonstrate any example will be available on the course web-page for students to practice at their own time. Here are the links to download this open source software:

- **for windows:** `https://cran.r-project.org/bin/windows/base/`
- **for mac:** `https://cran.r-project.org/bin/macosx/`

After you have installed R, you can install R-studio (an IDE for R) by downloading it from `https://www.rstudio.com/products/rstudio/download/#download`

11 Communicating with your Instructor

Please do not email your instructor asking questions like "how to do problem 10.3.4?", "when is the midterm?", "how to submit the assignment?". Emails like these will be ignored. If you need help with any problem you can either talk to your instructor right after the class if you expect the answer to be brief or come during the office hour. If you have any question/concern that you don't want to ask in front of others you can email the instructor to set up a one-on-one meeting.

12 April 5th: last date for grade center correction

After every quiz, your TAs will enter your mark within a week. Check periodically to make sure that the marks entered are correct. **Contact your TA** right away if you see any discrepancies. After April 5th, no request for grade center correction will be accepted.

Table 1: STA261, Winter 2019 **tentative** lecture guide

Week	**Dates**	**Readings**	**Tutorial**	**Quiz**
1	Jan 07 - 11	Review, Some distributions and Intro to R	-	-
2	Jan 14 - 18	Statistical Models, Empirical CDF, Density Histogram, Boxplot	1	-
3	Jan 21 - 25	Likelihood function, Maximum Likelihood Estimation, Unbiasedness, MSE	2	1
4	Jan 28 - Feb 01	Consistency, Sufficiency, Factorization theorem, Efficiency, CRLB	3	-
5	Feb 04 - 08	Interval estimation, z, t, χ^2 intervals, MLE based confidence interval	4	2
6	Feb 11 - 15	Testing Hypothesis, Type I & Type II error, Likelihood Ratio Test	5	-
-	Feb 18 - 22	*Reading Week*	-	
7	Feb 25 - Mar 01	Comparing two samples: independent and paired	6	3
8	Mar 04 - 08	Probable week for the midterm	7	-
9	Mar 11 - 15	Goodness of Fit test, Model Checking, Normal Probability Plot	8	4
10	Mar 18 - 22	χ^2 test of independence and homogeneity	9	-
11	Mar 25 - 29	Regression: Both X & Y continuous	10	5
12	Apr 01 - 05	Regression: Continuous Y, categorical X	11	-