STA261: Probability and Statistics II (Summer 2020)

Instructor:	Rob Zimmerman	**Lectures:**	Prerecorded
Email:	robert.zimmerman@mail.utoronto.ca	**Tutorials:**	M,W 15:00 - 16:00
Course Website:	https://q.utoronto.ca/courses/155113	**Location:**	Online

Course Description: This course is a mathematically-rigorous introduction to statistical inference using the theory built up in STA257, roughly split into five modules. In Module 1, we will focus data reduction through the notion of sufficiency and other related properties of estimators, including ancillarity and completeness. In Module 2, we will study methods of point estimation and learn about how to assess the quality of various point estimators. In Modules 3 and 4, we will switch our attention to the the dual concepts of hypothesis testing and interval estimators, putting familiar concepts such as confidence intervals and p-values on a rigorous footing. In Module 5, we will learn about the properties of consistency and efficiency for sequences of estimators, and use them to derive some asymptotic extensions of earlier concepts. Additional topics related to statistical inference will be introduced as time permits.

Prerequisites: The official prerequisites are STA257H1/STAB52H3/STA256H5. Please don't email me about waiving these prerequisites – such matters are outside of my jurisdiction and can be addressed by contacting the department at ug.statistics@utstat.utoronto.ca. Unofficially, you should have a strong calculus background, and you should come into the course with a good amount of *mathematical maturity*, since you will be expected to write and understand coherent mathematical proofs.

Office Hours: Both instructor and TA office hours will take place online using Bb Collaborate; the exact weekly schedule will be posted on Quercus. The Bb Collaborate session for these will be "open" throughout the term, which means that you can enter at any time and discuss the course with your fellow students, regardless of whether there is a TA/instructor present.

Head TA: Elena Wiegelmann (elena.wiegelmann@mail.utoronto.ca). Elena is the first point of contact for all administrative requests (missed tests, remark requests, etc.). If you're not sure about who to contact, start with Elena, who will forward to me if necessary.

Piazza: We have a Piazza page for the course at http://piazza.com/utoronto.ca/summer2020/sta261h1s. While the TAs and I will try to monitor the page, it's no substitute for office hours and there's no guarantee that we'll be able to answer every question that gets posted. You are highly encouraged to participate and answer the questions of your fellow students (we will try to endorse good answers).

Lectures: Lectures will be used to go through theory and examples. The lectures have been prerecorded, and the recordings will be made available on Quercus according to the schedule on Page 3. They will mostly be between two and three hours long (several may go slightly beyond three hours). Once they are released the lectures will be available for viewing anytime, which gives you some freedom to manage your time effectively. For easier viewing, the mathematical notes that I jot down during the lectures will be available to view/download separately; however, these notes are provided only for your convenience and they are absolutely *not a substitute* for the lectures themselves. There will be no lecture on August 3 (the Civic Holiday).

Tutorials: Tutorials will be used to solve selected problems from the practice problems and demonstrate techniques in R that you might find helpful for the Term Project (see below). There will be no tutorials on July 6 (the first day of lectures), August 3 (the Civic Holiday) or on the dates of Quizzes #1, #2, and #3 below. Tutorials will be run by TAs and held online via Bb Collaborate.

Practice Problems: Each module will include a (large) set of practice problems, most (or all) of which which will be selected from the end-of-chapter exercises in the course textbook. Solutions will not be posted, except for those presented in tutorials (at the TA's discretion). You are encouraged to work on the problems with your fellow students, and you are *highly* encouraged to attempt the tutorial problems before tutorials.

Marking Scheme: Final grades will be calculated according to the following scheme:

Quiz #1 (15%) ..... July 15, 14:00 - 15:30
Quiz #2 (15%) ..... July 27, 14:00 - 15:30
Quiz #3 (15%) ..... August 10, 14:00 - 15:30
Term Project (25%) ..... Due August 17, 23:59
Final Assessment (30%) ..... TBA (During Final Assessment Period)

- The Quizzes and Final Assessment will be open-book and consist of problems similar to the examples shown in lecture and the practice problems (they will be cumulative, although the focus will be on material not covered in prior assessments). They will be conducted through Quercus. The questions for each will be released at the starting times listed above, and you'll be required to upload *clear, legible* photos of your handwritten answers by the ending time. You'll be able to re-submit each answer as many times as you like prior to the ending time, but you'll still need to plan your time carefully – it's best to think of Quizzes #1, #2, and #3 as 85-minutes long, with an additional 5-minute allowance to upload your answers. Similarly, the Final Assessment should be thought of as 175-minutes long.
- The Term Project will require you to submit a paper demonstrating several course concepts in R; you will be permitted to work in groups of (up to) three. Details and a rubric will be provided in a separate document.

Course Textbook: All readings, practice problems, and other course content will be available through Quercus in order to keep the course self-contained. That being said, our presentation will follow this textbook very closely, and you're welcome to acquire a copy of it if you'd like to follow along (I personally believe it belongs on every statistician's bookshelf!)

- George Casella and Roger L. Berger. *Statistical Inference*. Brooks/Cole Cengage, 2nd ed., 2002. (Chapters 6-10)

Additional References: These are mainly textbooks used by several previous STA261 offerings; you might find them useful as a source of additional practice problems (although you'll have plenty from Casella/Berger already!). Buyer beware: there's no guarantee that what you find in these will be perfectly consistent with our presentation of the course, and you shouldn't expect myself or the TAs to be completely familiar with their contents. Note that the first textbook is (officially) available for free on Mike Evans' website.

- Michael J. Evans and Jeffrey S. Rosenthal. *Probability and Statistics: the Science of Uncertainty*. Freeman, 2010.
- John A. Rice. *Mathematical Statistics and Data Analysis*. Duxbury Press, 3rd ed., 2006.
- Dennis Wackerly, William Mendenhall, et al. *Mathematical Statistics with Applications*. Duxbury Press, 7th ed., 2007.
- Richard J. Larson, Morris L. Marx. *An Introduction to Mathematical Statistics and Its Applications*. Pearson, 6th ed., 2017.

Tentative Lecture Schedule:

Subject	Lecture	Date	Topics
Data Reduction	1	July 6	• Sufficient Statistics • Minimal Sufficient Statistics
	2	July 8	• Ancillary Statistics • Sufficient, Ancillary, and Complete Statistics
Point Estimation	3	July 13	• Method of Moments • Maximum Likelihood Estimation • The EM Algorithm
	4	July 15	• Mean Squared Error • Minimum-Variance Unbiased Estimation • Sufficiency and Unbiasedness
Hypothesis Testing	5	July 20	• Likelihood Ratio Tests • Error Probabilities and the Power Function
	6	July 22	• Most Powerful Tests • p-values
Interval Estimation	7	July 27	• Inverting a Test Statistic • Pivotal Quantities • Pivoting the CDF
	8	July 29	• Size and Coverage Probabilities • Test-Related Optimality
Asymptotic Evaluations	9	August 5	• Consistency • Calculations and Comparisons
	10	August 10	• Robust Estimation • Bootstrapping
	11	August 12	• Hypothesis Testing • Interval Estimation
(TBD)	12	August 17	• (TBD)

Marking Concerns: Any requests to have marked work re-evaluated must be made in writing within *one week* of the date the work was returned. You'll have to email Elena with any marking requests (not me). Requests must include a detailed reason for the change that references *objective fact*, and must be made for *legitimate perceived errors only*. The following are examples of unacceptable reasons for requesting a remark:

- "I feel my mark was unfair" (the TA felt otherwise, and they know more than you)
- "My friend got a better mark but they put the same thing as me" (so should I lower your friend's mark then?)
- "I need a bump to get my GPA over some threshold" (final grades will already be adjusted in accordance with recommendations outlined in the FAS Academic Handbook for Instructors)

If you legitimately find an error, then we'll happily change your mark. However, in the case of any ambiguity over the legitimacy of an error, I'll side with the TAs over you. Please make sure you understand this:

By submitting a remark request, you are agreeing to have me (the instructor) remark your entire work, change the grade up, down, or not at all, and that the result of this represents your final mark on the work and will not be contested further.

Missed Quizzes: If you've missed a quiz, reach out to Elena as soon as possible (no later than one week after the missed quiz). The remaining assessments will be reweighted; however, missing at least 50% of the assessments will result in an automatic fail. This means, for example, that if you miss any three of the elements in the marking scheme on Page 2, then you will fail the course. If you are taking another course which asks for your attendance during one of our quizzes, you will have to arrange alternative accommodations for the other course. This shouldn't be the case, though; STA261 is offered synchronously, and the quizzes are scheduled during the lecture/tutorial time slots specifically to avoid this kind of overlap. They will not be deferred.

Course Communication: I receive a lot of email, so it is important to abide by this email policy in order to ensure you receive a prompt response. You shouldn't email me about course concepts you're having trouble with – use office hours and Piazza for this. If you feel the need to contact the course team, first email Elena. This isn't because I don't want to talk to you, but rather because Head TAs are *paid* as part of their assigned hours to respond quickly to student emails. I'm extremely busy and can't guarantee quick response times with hundreds of students enrolled in the course; this is why we have a whole course team here to assist you. If you're following up on a conversation we have had in lecture or office hours, or if the matter is confidential and you don't want to share with the TA, then go ahead and email me directly. But when in doubt over who to email first, email Elena. She'll just forward it to me if she can't answer you.

Etiquette: When communicating with *anyone* in any way – but especially via email – make sure you courteous and respectful. This means using full sentences, not slang like "yo prof, I wanna get the lecture notes" (a real email received by a fellow instructor), etc. This is good practice for your eventual transition into industry or grad school. Make us *want* to reply to you. Importantly, *we reserve the right to simply ignore any emails that don't follow these guidelines.* If you need to email us, follow these steps (with "Rob" replaced by the name of whomever you're emailing):

- Put "STA261: Student Communication" in the subject line.
- Start the email with "Hi Rob, ...", followed by your full name and UofT Student Number.
- State the purpose of your email.
- End the email with a "Thank you", "Sincerely", or something that indicates that the email is over.

This policy may seem rigid, but it's not meant to discourage student communication; rather, it's designed to encourage *productive* student communication, by forcing you to communicate professionally. I can't overstate how much this will help you in your future career. Here's an example of a properly written and formatted email to me:

> Subject: STA261: Student Communication
>
> Hi Rob,
>
> My name is Bob Knobb (Student #1005551234), and I'm a student in your STA261 class. I would like to follow up on our conversation towards the end of yesterday's office hours, where we were discussing arithmetic functions. Since then, I was easily able to establish a tight bound on the L^1 distance between the second Chebyshev function and the identity: specifically, I found that $\left|x - \sum_{p^k \leq x} \log(p)\right| \leq \sqrt{x}\log^2(x)/8\pi$, at least for all $x > 73.2$, where $p \in \mathbb{P}$ and $k \in \mathbb{N}$. Does this seem like it could be useful?
>
> Thank you,
> Bob

Technology Requirements: A set of recommended technology requirements for online learning can be found on the Office of the Vice-Provost's website. For our course, the most important of these requirements are a *fast, reliable internet connection* and access to a scanner or camera for the Quizzes and Final Assessment. It is *your responsibility* to ensure your ability to satisfy these requirements. Shoddy internet connections or dead phone batteries are not valid excuses for missing a test. If you anticipate technology problems that could jeopardize your ability to write a test, you should reach out to to Accessibility Services (see below) as soon as possible.

Accessibility Services: Students with diverse learning styles and needs are welcome in this course. If you have a disability/health consideration that may require accommodations, please contact Accessibility Services at (416) 978-8060 or `http://studentlife.utoronto.ca/as`.

Academic Integrity: The University of Toronto's Code of Behaviour on Academic Matters outlines the behaviours that constitute academic misconduct, the processes for addressing academic offences, and the penalties that may be imposed. You are expected to be familiar with the contents of this document. Potential offences on tests include – but are not limited to – the following:

- Using unauthorized aids (internet resources, etc.)
- Contacting fellow students
- Misrepresenting your identity and/or using any kind of exam writing service
- Submitting altered tests for regrading
- Sending or receiving aid to/from anyone else

Course staff will be inspecting all submitted coursework for evidence of cheating, and will investigate any student suspected of an academic offense (which may include an interview with the student). If you refuse to participate in this investigation, or if we remain convinced that you have committed an academic offence, we will not hesitate to escalate the matter to the Department Chair.