Statistical Sciences
UNIVERSITY OF TORONTO

STA305H1S/STA1004HS - L0101 & L0201 Winter 2021
Design and Analysis of Experiments/ Introductory Experimental Design

Land Acknowledgement
We wish to acknowledge the land on which the University of Toronto operates. For thousands of years, it has been the traditional land of the Huron-Wendat, the Seneca, and most recently, the Mississaugas of the Credit River. Today, this meeting place is still the home to many Indigenous people from across Turtle Island and we are grateful to have the opportunity to work on this land.

Delivery: Online Synchronous. This requires online attendance at a specific time for some or all course activities and does not require attendance at a specific physical location for any activities or exams. Students are responsible for ensuring that they have proper equipment and reliable internet.

Course website: Available through the UofT Quercus (https://q.utoronto.ca)

Instructor	Synchronous lecture sessions
Dr. Shivon Sue-Chee (*she, her, hers*)	**Section L0101** • Mon 11:10-12:00 and Wed 11:10-13:00 **Section L0201** • Tues 15:10-16:00 and Thurs 15:10-17:00 Live class sessions will be held via Quercus in BbCollaborate or Zoom.

Teaching Assistants (TA): Jia, Sirui, Siyi, Xiaowen, Yuxuan, Amar, Hongda, Sebastian

Instructor and TA office times will be arranged with the class and posted later in our website.

Course Email: sta305@utoronto.ca

🖥 Course website
The course website is available through the Quercus Management Engine via https://q.utoronto.ca. It will be regularly updated with lecture notes, practice problems, assignments, announcements, readings and grades. Please turn on email notifications to receive announcements, and emails from me if urgent matters arise.

Lecture outlines will typically be posted by midnight on Sundays.

What to expect during live class sessions?

Live sessions will be conducted via Bb Collaborate or Zoom mainly by the instructor, with potential TA moderator support. **The main L0201 lecture will be recorded and will made available to students after the live classes.**

A typical 60-minute session, beginning at the `x'-th hour, will have the following program:

x:00- x:10	*(Re-)Entry and setup, Announcements, Community building intro polls*
x:10- x:45	*Main lecture (L0201's will be recorded)*
x:45- x:55	*Q&A discussions/ Piazza Live Q&A/ Small group discussion via breakout rooms*
x:55- x:60	*Wrap-up*

Course Content

This course will provide an introduction to the fundamental concepts of the design of scientific studies including the design of experiments and observational studies. Students will be become acquainted with statistical methods used to design and analyze experiments and observational studies. In particular, this course will cover: experiments versus observational studies, clinical trial design, comparing several groups using a completely randomized design, randomized blocks, Latin squares, incomplete block designs, factorial designs, causal inference in randomized and non-randomized studies, and adjusting for selection bias using propensity score methods.

The learning objectives of this course are:

- Understand the ideas, principles, and considerations that are common to the design and analysis of scientific studies including the statistical design of experiments and observational studies.
- Develop a statistical toolbox of methods for the design and analysis of experiments and observational studies.
- Identify appropriate uses and interpretations of experimental designs, and observational studies, including their strengths and limitations.

Topics

Experiments, observational studies, and causal inference
Experiments versus observational studies, and causal inference in randomized experiments.

Selection Bias in Observational Studies

Causal inference in randomized experiments versus observational studies. Introduction to the propensity score and three ways to use the propensity score to adjust for selection bias: matching; sub classification; direct regression adjustment.

Probability and Statistics

Mathematical statistics used in experimental design.

Comparing Several Groups

Comparing several groups in an experimental and observational setting and deciding whether differences that are found are likely to be real or due to chance.

Power and Sample Size

Power and sample size will be introduced for several designs. Applications will include the design and analysis of clinical trials with continuous or binary endpoints.

Blocking techniques

Blocked designs, Latin squares, randomized incomplete block designs.

Factorial Designs

Factorial, blocked factorial, and fractional factorial designs will be discussed.

Split plot designs

Split plot designs will be discussed as an example of restricted randomization in the design of experiments.

📖 Textbooks

Required Readings

The required readings are available via Professor Nathan Taback's text at https://scidesign.github.io/designbook/.

Recommended Reading

1. Statistics for Experimenters: Design, Innovation, and Discovery 2nd ed., Box, G.E.P., Hunter, J.S., Hunter, W.G. Wiley 2005
2. Design and Analysis of Experiments. Dean, A., and Voss, D. Springer. 1999. UofT library link: http://go.utlib.ca/cat/2573215
3. Design of Observational Studies. Rosenbaum, P. R. Springer 2010. UofT library link: http://go.utlib.ca/cat/7890274
4. Experiments: planning, analysis, and optimization 2nd ed., Wu, C.F.J., Hamada, M.S. Wiley, 2009. UofT library link: http://go.utlib.ca/cat/8598479
5. Causal inference for statistics, social, and biomedical sciences. Imbens and Rubin. Cambridge University Press, 2015. UofT library link: http://go.utlib.ca/cat/10127748

6. Experimental Design: Procedures for Behavioral Sciences, 4th ed., Kirk, R. E. Sage Publications 2013. UofT library link: http://go.utlib.ca/cat/10059810

▤ Evaluation

Students will be evaluated according to the following marking scheme.

	Weight (Undergrads)	Weight (Grads)	Date	Time
Course Surveys	2%	2%	Feb. 13, Apr. 9	by 22:00
Assignment 1	10%	10%	Jan. 30	by 22:00
Assignment 2	15%	15%	Feb. 13	by 22:00
Midterm Test	20%	10%	[L01] Mar. 3 [L02] Mar. 4	[L01] 11:10-12:40 [L02] 15:10-16:40
Assignment 3	18%	18%	Phase I: Apr. 3 Phase II: Apr. 9	by 22:00
Final project	35%	45%	April 16	by 22:00

Graduate students will be evaluated at the graduate level according to the University Assessment and Grading Practices Policy.

Assignment Due Date and Time
-   Assignments are due by 22:00 on the due dates and must be submitted electronically into Quercus or peerScholar.
-   Late assignments will be accepted but subject to a 1% penalty of the total assignment marks per hour late. Late submissions will not be allowed beyond 48 hours of the due date.
-   Students who would like additional accommodations should email the instructor at least 48 hours before the assignment is due.

Term test
The midterm test will be a Quercus Quiz and will be available during class time on the test day.
-   The test will have multiple choice and/or short-answer questions. Short-answers may require that you upload R output, hand-written answers and/or short videos of yourself. More details on the test coverage will be posted in our website.
-   Students who are eligible for special test accommodations will be facilitated through the university's Accommodated Testing Services.

Missed Test Guidelines

- If the test is missed, for any reason, you must send an email to sta305@utoronto.ca as soon as possible and no later than 5 business days after the test date. You must use the following sentences: "I affirm that I am experiencing an illness or personal emergency and I understand that to falsely claim so is an offence under the Code of Behaviour on Academic Matters. I will be available to do a makeup alternative after … (*please specify the date here*)."
- The makeup alternative will typically be held within 7 to 10 days after the original test date. **It may be an oral test or a test based on the cumulative course material up to the new test date or a combination of the two.**

Marking concerns

- Any requests to have marked work re-evaluated must be made in writing within <u>five (5) days</u> of the date the work was returned. Further, the request must contain a justification and the additional marks you think you deserve.
- Students should review the grading rubrics and solution guide, where available before submitting a regrade request.
- Requests must be sent to sta305@utoronto.ca. Include your official name and student number for identification purposes.
- Regrading requests will be processed by the teaching team within two weeks of the request date. Any request for a partial review can result in a full review of the entire work. **Unsuccessful disputes will result in a deduction of one assessment mark.**

Computing

We will use R/RStudio for all examples. R is freely available for download at http://cran.r-project.org for Windows, Mac, and Linux operating systems. For the test, you will need to know how to generate and interpret output from R. We will support the use of R to complete the assignments/project.

I recommend using R Studio as an integrated development environment to R. It is freely available at https://www.rstudio.com/products/rstudio/. RStudio can be downloaded onto your personal computer or used via our university's Jupyter hub or in the RStudio Cloud.

I am assuming that students have used R/RStudio before. However, I will provide you with the R syntax for all examples in lecture, which should be sufficient for you to do your assignments/projects.

👥 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 Piazza's [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 help fast and efficiently from classmates, and the TAs. Rather than emailing questions to the teaching staff, we encourage you to post your questions on Piazza. To sign up for the discussion forum click on the link:

piazza.com/utoronto.ca/winter2021/sta1004305

Note that moderation of the forum is subject to TA availability and further details of how and when the forum will be moderated, will be announced in our website. Please use the forum in accordance with its purpose. Inappropriate posts will not be tolerated and will be dealt with accordingly. **On the test day, the forum will be closed temporarily.**

📐 Academic integrity

The University supports acting in honesty, trust, fairness, respect, responsibility, and courage in all academic matters. You are responsible for knowing the content of the University of Toronto's Code of Behaviour on Academic Matters at https://governingcouncil.utoronto.ca/secretariat/policies/code-behaviour-academic-matters-july-1-2019. All suspected cases of academic dishonesty will be investigated following procedures outlined in the Code of Behaviour above. If you have questions or concerns about what constitutes appropriate academic behaviour or appropriate research and citation methods, you are expected to seek out additional information on academic integrity from your instructor or from other institutional resources (http://academicintegrity.utoronto.ca/). Here are a few guidelines that apply to this course:

- Students may consult course materials during tests, however, sharing or discussing questions and answers is an academic offence.
- Instructions for each assessment should be strictly followed.
- Do not personate another person, or have another person personate at any assessment.
- It is acceptable to get help with your assignments from someone outside the course, but the help must be limited to general discussion and examples that are

not the same as the assignments. As soon as you get some else to actually start working on one of your assignments, you have committed an academic offence!

- You must not copy someone else's answers or solutions provided to previous semesters of this course. Please read the UofT Policy on Cheating and Plagiarism, and don't plagiarize.

☞ Accessibility needs

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 Services as soon as possible: accessibility.services@utoronto.ca or http://accessibility.utoronto.ca.

For additional advice on course and degree planning, please contact your College Registrar, and, if you are in a StG Statistics program, the Statistics Undergraduate Office can answer your questions about STA courses and programs (see https://www.statistics.utoronto.ca/undergraduate-student-resources/office-hours).

Additional help

Need extra help with the coursework? Here are some options:
- For continued class discussion and questions outside of class, post and answer questions on our discussion forum.
- Visit your instructor or the teaching assistants during their office hours.
- Join an online recognized study group (RSG) or get advice on learning (outside of the classroom) from a learning strategist via Sidney Smith Commons at https://sidneysmithcommons.artsci.utoronto.ca/online-learning/.

Communication

E-mail is appropriate for emergencies or private matters. Use your *utoronto.ca account. You will not get a response if you email from other email addresses. Your email should contain an appropriate subject line, the addressee, your official name and UTORid for identification purposes. Please send course correspondence and regrade requests to sta305@utoronto.ca. This departmental course mailbox will be shared among the teaching team members to efficiently handle course correspondence. Please expect an answer to your e-mail within two business days.

For non-course related correspondence or strictly confidential emails, you can contact the instructor directly using the Quercus Inbox via the course website. I will generally answer e-mail within two business days.

The instructor and TAs are here to help you. **Ask questions and let me know promptly if there are any concerns.**

▦ Tentative Lecture Schedule and Important Dates

Week	Topics	Important Dates
1. Jan. 11-15	Introduction, review of mathematical statistics	
2. Jan 18-22	Comparing two groups via the randomization distribution	*Jan 24 – Last day to add S courses*
3. Jan 25-29	Power and sample size	Assignment 1: Jan 30, 22:00
4. Feb 1-5	Power via simulation Introduction to causal inference in randomized experiments	
5. Feb 8-12	Design of observational studies and propensity scores	Assignment 2: Feb. 13, 22:00 *Mid-Course Evaluation Survey*
Feb 15-19	**Reading Week – No classes**	
6. Feb 22-26	ANOVA - Comparing more than two groups	
7. Mar 1-5	**Review class, MIDTERM TEST**	[L01] Mar. 3, 11:10 – 12:40 [L02] Mar. 4, 15:10 – 16:40
8. Mar 8-12	ANOVA - Comparing more than two groups, multiple comparisons	
9. Mar 15-19	ANOVA - Comparing more than two groups, multiple comparisons	*Mar. 15- Last day to drop S course*
10. Mar 22-26	Factorial designs at two levels	
11. Mar 29-Apr 2	Randomized block designs, Latin Squares, Balanced Incomplete Block Designs	Assignment 3: Phase I: Apr. 3, 22:00 Phase II: Apr. 9, 22.00
12. Apr 5-Apr 9	Blocking in factorial designs, fractional factorial designs. Restricted randomization and split-plot designs.	*End-of-Course Evaluation Survey*
Apr 13- 23	Final assessment period	Final Project: April 16, 22:00