STA442: Methods of Applied Statistics
Fall 2023

Course information

- *Lecture*: Tuesdays, 2 –5 PM
- *Location*: OISE 2212
- *Quercus*: https://q.utoronto.ca/courses/317121/
- *Piazza*: https://piazza.com/utoronto.ca/fall2023/sta442h1flec0101/
  - Any questions about **contents, logistics, and homework** should be posted to Piazza.
- *Course email*: STA442@utoronto.ca
  - Use course email for **any communications related to administrative issues**.

Instruction team

- *Instructor*: **Jun Young Park**, PhD
  - Assistant Professor, Department of Statistical Sciences and Department of Psychology
  - How to call: either Professor Park (preferred) or Dr. Park
  - Email: junjy.park@utoronto.ca
    - Use this account only if it is necessary. In most cases (more than 99%), Quercus, Piazza, and course emails should suffice. The instructor won't reply to emails that can be addressed using existing resources.
    - Don't cc this account when sending emails to the course email.
- *Teaching Assistant*: **Michael Chong** and **Kevin Zhang**
  - Ph.D. students in Statistical Sciences

Office hours

Day	Time	Instructor/TA	Delivery	Note
Mondays	10-11AM	Michael Chong	Online	See Quercus for the link.
Thursdays	4-5PM	Jun Young Park	Online	See Quercus for the link.
Fridays	11AM-noon	Kevin Zhang	Online	See Quercus for the link.

Getting help with the course

- The **best way** to ask any questions about the course materials and administrative issues is to ask directly the instructor during (or after) the lecture or to use office hours.
- You may also use Piazza or course emails, but it may take some time to get a response (up to a few days) because it is not monitored every hour/day.
- You may email the instructor on personal issues that deserve protection of privacy. These will not be shared with anyone else.

Course description

This course is an upper-level course in applied statistics offered by the University of Toronto, Department of Statistical Sciences. The course will focus on using and interpreting advanced statistical methods with applications in several different areas. The course combines theory and applications and will include computing with **R**. The aims for this course are 3-folds:

- Revisit statistical methods covered in prerequisite(s) and integrate them into key concepts and perspectives.
- Learn how to apply statistical methods "appropriately" and see what could go wrong when applied inappropriately by using simulations.
- Be able to explain "why" a method is performing better than the others through statistical reasoning and simulations.

Prerequisite(s): STA303

Topics

- Introduction to computer-intensive methods
- Revisiting statistical inference
  - Revisiting key statistical concepts / the role of statisticians
  - Simulation studies, permutation tests, power analysis
- Revisiting linear models
- Correlated data analysis and random effects modeling
  - Longitudinal data analysis
- Understanding different models
  - Bias-variance tradeoff
  - Regularization methods
  - Bayesian methods
  - High-dimensional and semiparametric learning
- Multiple comparisons
- Research ethics in statistics

Textbooks

There is no single textbook that addresses all topics covered in this course. The course slides will be uploaded to the Quercus on weekends before the class. Topics will be selected from the textbooks below, but please note that the technical difficulties of the books are highly variable.

- *Applied Longitudinal Analysis* by Fitzmaurice et al.
- *Applied Linear Regression Models* by Kutner, Nachtsheim, and Neter
- *Bayesian and Frequentist Regression Methods* by Wakefield
- *Linear Mixed Models: A Practical Guide Using Statistical Software* by West et al.
- *Linear Models with R* by Faraway
- *Mathematical Statistics with Resampling and R* by Chihara and Hesterberg
- *Introduction to Statistical Learning with R* by James et al.
- *Statistical Computing with R* by Rizzo

Course resources

Lecture slides will be uploaded to Quercus before the lecture.

Important dates

- November 7: No class (Fall Reading Week)
- October 24: Midterm exam (during class hours at the same place)
- November 5: Midterm exam grades to be released (not guaranteed for the make-up exams)

Evaluations

- *Assignment (35%)*: There will be 4 assignments that ask you to internalize course materials and apply them to research problems. Discussions with peers are allowed and encouraged, but these are to be completed (written up) on your own. Assignment types include (i) short response, (ii) multiple choice and (iii) **R** programming.
  - Based on previous years, it is **strongly encouraged** to start working on the assignments.
- *Midterm exam (30%)*: It will evaluate your understanding of scientific research using data. It includes (i) short response, (ii) T/F, (iii) multiple choice (iv) case studies but not **R** programming.

- *Final exam (35%)*: Similar to the midterm exam, it will include (i) short response, (ii) T/F, (iii) multiple choice (iv) case studies but not **R** programming. In addition, there will be up to 3 research articles that students must read before the exam to answer questions related to the articles. The list of articles will be announced by October 31. The exam will be cumulative, although more emphasis is given to the topics not covered in the midterm exam.
- *Participation credits (up to 3%)*: Extra credits will be given based on participation in class. Participation is defined by one of the following:
  - Responses to questions asked by the instructor, or
  - Questions to course contents. These *exclude* clarification questions.

  The "participation count" is defined by the number of classes you participated in discussion points made by the instructor, or asked insightful questions in class. Clarification questions (e.g. *could you explain it one more time?* or *I don't get that*.) do not count toward the count. Students with participation counts ≥6 will get an extra 3%, and counts ≥4 will get an extra 2%, and counts ≥2 will get an extra 1% credit.

Software

We will use **R** throughout this course, a statistical software publicly available for free at https://mirror.csclub.uwaterloo.ca/CRAN/. Installing RStudio (https://posit.co/downloads/) for an interactive programming environment is also highly recommended. This course assumes basic knowledge of **R**, and some basics will be covered in class, including

- Basic functions for statistics including t.test(), lm() and glm().
- Built-in functions related to probability distributions, such as rnorm(), dnorm(), qnorm() as well as analogous functions for other probability distributions.
- for() loop and ifelse() statements.
- Writing basic functions.

Example codes relevant to course materials will be presented during the lectures. Note that homework assignments done with different programming languages (e.g., Python, SPSS and Stata) are not accepted.

Lecture recordings

The lectures will not be recorded nor distributed unless advised by the university policies.

Intellectual property statement (adapted from Dr. Liza Bolton's statement)

Course material that has been created by your instructor (i.e., lecture slides, assignments, solutions, and any other course material and resources made available to you) is the intellectual property of your instructor (or the credited holder of the copyright) and is made available to you for your personal use in this course. Sharing, posting, selling or using this material outside of your personal use in this course is not permitted under any circumstances and is considered an infringement of intellectual property rights. If you would like to record any course activities in this course, you **MUST** ask permission from your instructor in advance. According to intellectual property laws, not asking permission constitutes stealing.

Absence of declaration / Verification of illness

If you become ill and it affects your ability to do your academic work, consult the instructor right away. Normally, documentation in support of your specific medical circumstances is needed. It can be an Absence Declaration (via ACORN) or the University's Verification of Student Illness or Injury (VOI) form. The VOI indicates the impact and severity of the illness, while protecting your privacy about the details of the nature of the illness. You can submit a different form (like a letter from a doctor), as long as it is an original document, and it contains the same information

as the VOI. For more information on the VOI, please see http://www.illnessverification.utoronto.ca/ for the most recent university policy. For information on Absence Declaration Tool for A&S students, please see https://www.artsci.utoronto.ca/absence. If you get a concussion, break your hand, or suffer from some other acute injury, you should register with Accessibility Services as soon as possible.

Accommodations

If you have a disability or health consideration that may require accommodations, please contact Accessibility Services at https://studentlife.utoronto.ca/as or (416) 978-8060.

- For students being supported by Accessibility Services, it is recommended (though not required) to keep the instructor updated so that individualized assistance or accommodations (beyond the ones offered by AS) can be applied if needed. All information related to privacy or health conditions must not be shared.

Late submission of homework assignments

There will be a 1% deduction in the final course grade every 24-hour period past the deadline. Valid forms for requesting extensions without penalties will be Absence Declaration, VOI, or equivalent. Once these forms are submitted, the instructor will determine an appropriate extension.

Regrading policy

Students can ask for regrading of their homework problems and the midterm exam. It is expected that students clearly illustrate rationales for regrading. Note that regrading may result in the deduction of points, so it is encouraged to ask for regrading only if they strongly believe that the grading is wrong.

Missing the midterm exam

Students must take the term exam on the scheduled date to qualify for grades. In case of (medical or family-related) emergencies, students must declare the absence or verify illness immediately. For valid requests, a make-up exam will be scheduled by the instructor at the earliest date possible, and it will not be postponed until the end of the final exam.

Academic integrity

Academic integrity is essential to the pursuit of learning and scholarship in a university, and to ensuring that a degree from the University of Toronto is a strong signal of each student's individual academic achievement. As a result, the University treats cases of cheating and plagiarism very seriously. The University of Toronto's Code of Behaviour on Academic Matters (https://governingcouncil.utoronto.ca/secretariat/policies/code-behaviour-academic-matters-july-1-2019) outlines the behaviours that constitute academic dishonesty and the processes for addressing academic offences. Potential offences include, but are not limited to:

- Using someone else's ideas or words without appropriate acknowledgement
- Submitting your own work in more than one course without the permission of the instructor
- Making up sources or facts.
- Obtaining or providing unauthorized assistance on any assignment
- Misrepresenting your identity on exams

All suspected cases of academic dishonesty will be investigated following procedures outlined in the Code of Behaviour on Academic Matters. If students have questions or concerns about what constitutes appropriate academic behaviour or appropriate research and citation methods, they are expected to seek out additional information on academic integrity from their instructors or from other institutional resources.

Tutoring companies/services

The instructor is not associated with any private tutoring companies/services in any way in preparing course materials (lecture notes, assignments, exam questions, etc.) and delivering lectures.

Generative AI (e.g., ChatGPT, GPT-4, Bing, etc)

Students may use artificial intelligence tools for creating an outline for an assignment, but the final submitted assignment must be original work produced by the individual student alone.