STA442: Methods of Applied Statistics

Instructor: Jun Young Park, PhD

E-mail:junjy.park@utoronto.ca

Teaching Assistants:
- (Head TA) David Decker, colin.decker@utoronto.ca
- Jiayue Huang, jiayue.huang@mail.utoronto.ca
- Ying Zhou, yingx.zhou@mail.utoronto.ca
- Sigeng Chen, sigeng.chen@mail.utoronto.ca

Notes:
- If you have inquiries about this course to be sent by email, please use STA442 - studentid - Subject (e.g. STA442 - 1234567 - Homework 1 grading)

Office hours: will be held in Zoom (the links will be provided on Quercus)
- Instructor: Tuesdays 9-10AM, Thursdays 9-10AM
- TA: Fridays 3-5PM

Course deliveries: Lectures will be delivered in person. The classroom is **MS3153**.

Lectures: Tuesdays 3-6PM Eastern time

Recorded lectures: Lectures may be recorded if there is a request by the university's accessibility services. It does not mean that the recorded lectures will be released to all enrolled students. If there are any circumstances that in-person lectures are unavailable, the lectures will be held virtually (online synchronously), recorded, and distributed to all enrolled students.

Prerequisites: Completion of **STA303**. There is no exception.

Course description: The course will focus the using and interpreting advanced statistical methods with applications in a number of different areas. The course is a mixture of theory and applications, and will include a number of assignments which will involve computing with R.

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
  - Linear algebra and distributional theory
  - Gauss-Markov theorem
  - Model checking and variable selection
- Correlated data analysis and random effects modeling
  - Longitudinal data analysis
- Bootstrapping
- Understanding different models
  - Bias-variance tradeoff
  - Regularization methods
  - Bayesian methods
  - High-dimensional and semiparametric learning

Piazza: We will use Piazza for discussing any questions on course materials, other than homework or exam questions: https://piazza.com/utoronto.ca/fall2022/sta442.

Textbooks: There is no single textbook that addresses all topics covered in this course. The course slides will be uploaded to the Quercus. I will select topics from the textbooks below, but please note that the technical difficulties of the textbooks are highly variable.

1. *Applied Longitudional Analysis* by Fitzmaurice et al.
2. *Applied Linear Regression Models* by Kutner, Nachtsheim, and Neter
3. *Bayesian and Frequentist Regression Methods* by Wakefield
4. *Linear Mixed Models: A Practical Guide Using Statistical Software* by West et al.
5. *Linear Models with R*, Faraway
6. *Mathematical Statistics with Resampling and R* by Chihara and Hesterberg
7. *Introduction to Statistical Learning* by James et al.
8. *Statistical Computing with R* by Rizzo.

Software: This course assumes that you are familiar with `R` programming, 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.

Evaluations: Attendance (5%), Homework (40%), 1 midterm exam (in-person, 30%), 1 final project (take-home, 25%)
- **Attendance**: Attendance will be checked in the following ways: A QR code for a Google Form will be shared in each class and students will be asked to fill out the form. Students will not be penalized, for any reason, for not attending the class up to **twice**.
- **Class participation**: Up to 3% of the final grade can be topped up based on the participation. The purpose of this portion is to encourage you to ask more questions and add your thoughts on the course material.
- **Homework**: There will be **four** homework sets, requiring a solid understanding of theories and computing. Homework is a **group work** consisting of 3-4 students, and students submit their responses as a team. Note that team members are chosen randomly in each assignment. Each homework is due sharply at the appointed date, and any late submission will be penalized.
- **Midterm exam**: In person exam is scheduled to be held in November –. It is to be done individually in person and evaluates technical skills learned in class.
- **Final project**: The final project is an individual work, and students will be asked to read and summarize a research paper and conduct simulation studies as needed.
- **Peer evaluations**: It is expected that every student contribute to the problems equally. Students will also be asked to rate other members' contributions to the work. For the homework sets, each student will be asked to submit the peer evaluation for others. It will be one of `exceptional contributions / equal contributions / low contributions`. To keep it anonymous, it will be reflected to the final grade only.
- **Submission guides**: Up to 20% of each homework and final project grades can be deducted based on how neatly it was prepared. Examples that would lead to deduction include, but not limited to,
  - Unorganized answers, including margins, spacing, etc.
  - Grammatical or typographical errors
  - Irreproducible R codes
  - Lengthy answers that are wrong or unrelated to the problems

Regrading policy: Students are allowed to ask for regrading of their homework problems and the midterm exam. It is expected that students clearly illustrate rationals for regrading. Note that regrading may result in deduction of points, so it is encouraged to ask for regrading only if they have a strong belief that the grading is clearly wrong.

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 students 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 (`link`) outlines the behaviours that constitute academic dishonesty and the processes for addressing academic offences. Potential offences include, but are not limited to:

On tests and exams:

1. Using or possessing unauthorized aids.
2. Looking at someone else's answers during an exam or test.
3. Misrepresenting your identity.

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.

Missing an examination: Students are required to take the midterm exam on the scheduled date to qualify for grades. In case of (medical or family-related) emergencies, students must declare the absence and send a proof to the instructor *at least 1 hour before the exam.* I will not consider any requests made afterwards. 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.

Accommodations for disability policy: If you have a disability or health consideration that may require accommodations, please feel free to approach me or Accessibility Services at [(416) 978-8060 or `https://studentlife.utoronto.ca/as`]. A copy of the letter of accommodation needs to be sent to the instructor 7 days before the exam date.

Intellectual property statement (adapted from Dr. Liza Bolton's statement): Course material that has been created by your instructor (i.e., lecture slides, questions/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.

Tutoring companies: I am not associated with any tutoring companies in any way in preparing course materials (lecture notes, homework sets, exam questions, etc) and delivering lectures.

Challenges: For any other challenges during your study, please visit Health and Wellness Centre or the Graduate Wellness Services.