STA2201H Methods of Applied Statistics II

Monica Alexander

Time and place:

- Wednesdays 2-5pm, starting 13 January
- Online synchronous via BBCollaborate

My contact details:

- monica.alexander@utoronto.ca. I do not check email after 5pm or on weekends.

Office hours:

- TBD, will be by appointment via Calendly

Course website:

https://github.com/MJAlexander/applied-stats-2021

Course description

This course covers a range of statistical methods, covering the theory, application and interpretation of models on a range of different datasets. Topics will include generalized linear models, Bayesian inference, generalized linear mixed models, generalized additive models involving nonparametric smoothing, model evaluation and selection. We will also cover some core statistical computing techniques.

A large focus of the outcomes on this course will also be on reproducible research, identifying and dealing with data and modeling issues, and model interpretation and communication.

Textbooks

There is no required textbook for this course. The following texts may be useful as reference:

- Dobson, A. J., & Barnett, A. G. An introduction to generalized linear models. Chapman and Hall/CRC, 2008.
- Gelman, Andrew, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. Bayesian data analysis. Chapman and Hall/CRC, 2013.
- Gelman, Andrew, and Jennifer Hill. Data analysis using regression and multilevel/hierarchical models. Cambridge university press, 2006.
- Hastie, T., Tibshirani, R., & Friedman, J. The elements of statistical learning: data mining, inference, and prediction. Springer Science & Business Media, 2009.
- Wickham, Hadley, and Garrett Grolemund. R for data science: import, tidy, transform, visualize, and model data. O'Reilly Media, Inc., 2016.

Mode of Delivery

This course will be held entirely online. Lectures and office hours will occur synchronously through BBCollaborate. Given the current circumstances, I understand it may be difficult to attend class for a number of reasons, and as such all lectures will be recorded so that they can be accessed at a later time.

Computing

Throughout the course we will be using R in all examples, labs and homework assignments. Exams will also require interpretation of R output. In particular, you will be learning and expected to code in the `tidyverse` style, and all labs and assignments will be written using R Markdown.

You will need to have R and RStudio installed on your computer:

- Download R here: https://www.r-project.org/
- Download RStudio (free version) here: https://www.rstudio.com/products/rstudio/download/

You will also need to have a GitHub account: https://github.com/

Assessment

Broadly, there are four components to the assessment for this course:

- Lab exercises (10 in total), 2% per week
- Two assignments worth 15% each
- Mid term exam (online), 15%
- Research project, 35%

All labs and assignments, and the written component of the research project will be completed in R Markdown. The '.Rmd', resulting pdf and any relevant files should be submitted, such that the file is reproducible and compiles to a pdf with no errors.

Lab Exercises

- Short hands-on exercises in R
- To be handed in via GitHub
- Will be due by 9am on the following Friday

Assignments

- A mix of short answer questions and applying methods to a dataset
- To be handed in via Quercus

Mid-term

- A mix of multiple choice and short answer questions
- Covers all material in first 5 weeks
- Will need to be able to write R 'pseudo code' and interpret R code and output
- Held online via Quercus

Research project

For the research project you will be required to investigate a research question of interest using a dataset of your choice and applying methods learned during the semester. The project consists of three main components:

- **Research proposal (7.5%)** describing your research question of interest, why it is interesting, any hypotheses you may have, the dataset and variables you are using, and what methods you intend to use. Also include some initial EDA. 1-2 pages excluding graphs and tables, due at the same time as Assignment 2. Submitted via Quercus.
- **Research paper (20%)** on your topic. It is expected that the report be written in the style of an academic paper, including the following sections: Abstract, Introduction, Data, Methods, Results, Discussion, and Conclusion. Must be written in R Markdown, and all files to reproduce the paper and analysis must be submitted via Quercus. Due at the end of semester.
- **Presentation (7.5%)** on you research project, including question, data, methods and results. Around 5 minutes. It is expected you will present with the aid of slides. To take place in Week 12, online via BBCollaborate.

Course policies

Attendance

You are expected to attend every class and arrive online in a punctual manner. However, as mentioned above, given the current circumstances, I understand it may be difficult to attend class for a number of reasons, and as such all lectures are being recorded.

Communication

The best way to ask questions about course material or assignments is in person during the lecture, lab or your instructor's office hours. The following are guidelines for email communication: please make sure that you have a legitimate need before you write and that you cannot resolve your question during the lecture, lab or office hours; email messages should state the course number and the purpose of the email clearly in the subject line. Please be respectful and clear when emailing.

Accessibility Services

It is the University of Toronto's goal to create a community that is inclusive of all persons and treats all members of the community in an equitable manner. In creating such a community, the University aims to foster a climate of understanding and mutual respect for the dignity and worth of all persons. Please see the University of Toronto Governing Council "Statement of Commitment Regarding Persons with Disabilities" at http://www.governingcouncil.utoronto.ca/Assets/Governing+Council+Digital+Assets/Policies/PDF/ppnov012004.pdf. In working toward this goal, the University will strive to provide support for, and facilitate the accommodation of individuals with disabilities so that all may share the same level of access to opportunities, participate in the full range of activities that the University offers, and achieve their full potential as members of the University community. We take seriously our obligation to make this course as welcoming and accessible as feasible for students with diverse needs. We also understand that disabilities can change over time and will do our best to accommodate you. Students seeking support must have an intake interview with a disability advisor to discuss their individual needs. In many instances it is easier to arrange certain accommodations with more advance notice, so we strongly encourage you to act as quickly as possible. To schedule a registration appointment with a disability advisor, please visit Accessibility Services at

http://www.studentlife.utoronto.ca/as, call at 416-978-8060, or email at: accessibility.services@utoronto.ca. The office is located at 455 Spadina Avenue, 4th Floor, Suite 400. Additional student resources for distressed or emergency situations can be located at distressedstudent.utoronto.ca; Health & Wellness Centre, 416-978-8030, http://www.studentlife.utoronto.ca/hwc, or Student Crisis Response, 416-946-7111.

Academic Integrity Clause

Copying, plagiarizing, falsifying medical certificates, or other forms of academic misconduct will not be tolerated. Any student caught engaging in such activities will be referred to the Dean's office for adjudication. Any student abetting or otherwise assisting in such misconduct will also be subject to academic penalties. Students are expected to cite sources in all written work and presentations. See this link for tips for how to use sources well: (http://www.writing.utoronto.ca/advice/using-sources/how-not-to-plagiarize). According to Section B.I.1.(e) of the Code of Behaviour on Academic Matters it is an offence "to submit, without the knowledge and approval of the instructor to whom it is submitted, any academic work for which credit has previously been obtained or is being sought in another course or program of study in the University or elsewhere." By enrolling in this course, you agree to abide by the university's rules regarding academic conduct, as outlined in the Calendar. You are expected to be familiar with the Code of Behaviour on Academic Matters (http://www.artsci.utoronto.ca/osai/The-rules/code/the-code-of-behaviour-on-academic-matters) and Code of Student Conduct (http://www.viceprovoststudents.utoronto.ca/publicationsandpolicies/codeofstudentconduct.htm) which spell out your rights, your duties and provide all the details on grading regulations and academic offences at the University of Toronto.

Equity and Diversity

The University of Toronto is committed to equity and respect for diversity. All members of the learning environment in this course should strive to create an atmosphere of mutual respect. As a course instructor, I will neither condone nor tolerate behaviour that undermines the dignity or self-esteem of any individual in this course and wish to be alerted to any attempt to create an intimidating or hostile environment. It is our collective responsibility to create a space that is inclusive and welcomes discussion. Discrimination, harassment and hate speech will not be tolerated. Additional information and reports on Equity and Diversity at the University of Toronto is available at http://equity.hrandequity.utoronto.ca.

Course outline

Planned content by week (note: subject to change)

- Week 1 (13/1/21): Introduction
- Week 2 (20/1/21): Generalized linear models recap
- Week 3 (27/1/21): Survival analysis
- Week 4 (3/2/21): Bayesian methods and inference (**Assignment 1 due**)
- Week 5 (10/2/21): Bayesian methods and inference II

READING WEEK (17/2/21)

- Week 6 (24/2/21): **Mid-term**
- Week 7 (3/3/21): Visualizing the Bayesian workflow and model checks
- Week 8 (10/3/21): Multilevel models I
- Week 9 (17/3/21): Multilevel models I
- Week 10 (24/3/21): Non-linear models I (**Assignment 2 and research proposal due**)
- Week 11 (31/3/21): Non-linear models II and extensions

- Week 12 (7/4/21): Recap and **Student presentations**
- **Research paper** due 16/4/21