SOCI 620: Quantitative Methods 2

Location	Onliine
Time	Winter 2022, Tue and Thu 4:05–5:25pm
Instructor	Peter McMahan (peter.mcmahan@mcgill.ca
Office hours	Wednesdays 1:30–3:00pm Zoom
Lab sessions	TBA
Syllabus	https://soci620.netlify.com/

Description

As the second of two courses in the quantitative methods sequence, this class will introduce students to statistical techniques that extend and diverge from standard multivariate regression models. To this end, the course will have two main goals. First, students will become familiar with a range of quantitative methods common in social science research. Methodological topics will include generalized linear models for predicting categorical, ordered, and count data, multilevel/hierarchical linear models, and strategies for analyzing time-series and panel data. Students will learn to critically interpret these methods as they are used in the literature, and to utilize the methods for their own research.

The second goal of the course will be to provide students with an overarching framework to understand the methods we discuss as well as techniques they may encounter elsewhere. Instruction will therefore focus on a probabilistic interpretation of statistical models through a Bayesian lens. In addition to fostering a strong understanding of statistical dependence and parameter covariance, this approach will unify the methods we cover and enable students to build interpretable, theory-driven models of their own.

Requirements

Students are expected to be familiar with the readings, engage in discussions (in-class and/or online), complete assignments, and prepare an independent research project.

Class

The scheduled classes for the course will be hybrid lectures, discussions, and lab sessions. It is vital that students attend class regularly, having completed the readings and prepared to engage with the topics covered. If attending class is difficult or impractical for you (due to, e.g., time zone differences or Internet access), please get in touch with me as soon as possible to arrange accommodations.

Classes will be held online through Microsoft Teams. Slides and code will be made available before class, and lectures will be recorded and posted shortly after class.

Labs

In addition to the scheduled classes, we will have a weekly lab session (location to be determined). These sessions will be led by the T.A. as a space to practice techniques, raise questions and concerns, and discuss course content with one another. Attendance at these sessions is optional but *strongly* encouraged.

Equipment and software

We will be working with data and learning analysis and visualization in-class, so students must bring a laptop computer with them. Mobile devices such as tablets and phones, even with an external keyboard, will not be sufficient. If you do not have access to a laptop please talk to me as soon as possible so we can work out a way for you to participate.

We will be using the R statistical language and software for data processing, statistical estimation, and visualization in this class. It is recommended that students install the RStudio graphical interface, which I will be using for demonstration in class.

Readings

We will use the second edition of the textbook *Statistical Rethinking* by Richard McElreath for the course (McElreath 2020). I have put the book on reserve at the McLennan library building, and it is available for free online and as an ebook through the library website.

Assignments

There will be a short assignment due on most of the Thursdays of the semester (see the schedule for specific dates). These are intended to help you learn to use the methods we discuss in R and to give you practice in interpreting statistical models. Assignments should be submitted online through the Microsoft Teams site. Students can work together and consult one another on assignments, but each student should create their own, unique write-up for submission.

Independent research project

Each student will finish an independent research project by the end of the semester. These projects will be empirical, scholarly analyses, including a source of data, a well formed research question, a motivated statistical analysis, and a thorough interpretation of the results. Ideally, the projects will be related to work students are doing outside of class. Projects that represent a piece of a student's broader research agenda are encouraged. The projects will be graded on the basis of four required assessments:

1. Precis (Due March 9): This will be a short (no more than one page) description of the research project. It should include a specific research

question, a brief description of the data that will be used, and an outline of the analytical strategy that will be employed. The purpose of the precis is to motivate the project and to establish its feasibility, not to perform any analyses or to answer any research questions.

2. Proposal (Due March 30): Based on the feedback received from the precis, the project proposal will give a more detailed account of the research project. A good proposal will give a thorough account of the data that is being used, including some preliminary summaries and analyses. It will also articulate the research question in terms of statistical models and will specify those models formally.
3. ~~Presentation (In class, April 8 and 13): Each student will give a brief, PechaKucha-style presentation of their final project in class, consisting of twenty slides that will automatically advance ever twenty seconds. The presentation should describe your research question succinctly, give a clear account of the statistical model(s) used, and briefly interpret the results in light of the research question.~~ Due to the shortened semester, the presentations are canceled.
4. Project write-up (Due April 21): The writeup for the final project will take the form of a formal scholarly paper. This should go into careful detail about the project, including a full description of the data, exposition and motivation of the statistical models used, a summary of the estimation of the model parameters, and a careful, thorough interpretation of the results. It should include tables and figures to illustrate your analysis.

Each student should arrange a brief meeting with me early in the term to discuss ideas for their research project and the appropriateness for the course.

Evaluation

The evaluation components and due dates for this course are strict. If outside circumstances will make it difficult to meet a requirement *please raise the issue with me as soon as possible so we can find a solution*. Regular absences will affect your ability to do well on assignments and the final project.

Item	Due	Weight
Assignments	*due dates listed on schedule*	50% of final grade
Project precis	March 9	7.5% of final grade
Project Proposal	March 30	12.5% of final grade
Project writeup	April 21	30% of final grade

Accessibility

I strive to make the classroom as accessible as possible and to accommodate the particular needs of individual students. Students with disabilities in need of accommodation please contact the Office for Students with Disabilities

(http://www.mcgill.ca/osd/, phone 514-398-6009) to work out a plan for meeting the course requirements. Students are encouraged to contact me with any further issues they may have attending class or completing the work.

Academic integrity

McGill University values academic integrity. Therefore, all students must understand the meaning and consequences of cheating, plagiarism and other academic offenses under the Code of Student Conduct and Disciplinary Procedures (see http://www.mcgill.ca/students/srr/honest/ for more information). (approved by Senate on 29 January 2003)

L'université McGill attache une haute importance à l'honnêteté académique. Il incombe par conséquent à tous les étudiants de comprendre ce que l'on entend par tricherie, plagiat et autres infractions académiques, ainsi que les conséquences que peuvent avoir de telles actions, selon le Code de conduite de l'étudiant et des procédures disciplinaires (pour de plus amples renseignements, veuillez consulter le site http://www.mcgill.ca/students/srr/honest/).

Language of evaluation

In accord with McGill University's Charter of Students' Rights, students in this course have the right to submit in English or in French any written work that is to be graded. (approved by Senate on 21 January 2009)

Conformément à la Charte des droits de l'étudiant de l'Université McGill, chaque étudiant a le droit de soumettre en français ou en anglais tout travail écrit devant être noté (sauf dans le cas des cours dont l'un des objets est la maîtrise d'une langue).

Schedule

Background: probability and Bayesian statistics

Thu, Jan 7

Lecture: Introductions, course structure, syllabus

Lab:
Installing and testing software

Required:
⋮ (McElreath 2020, Ch. 1)

Tue, Jan 12

Lecture: Probability models of social processes

Lab:
Approximating simple posteriors

Required:
⋮ (McElreath 2020, Ch. 2)

Thu, Jan 14

Lecture: Probability distributions and random samples

Lab:
Working with random samples

Required:
⋮ (McElreath 2020, Ch. 3)

Tue, Jan 19

Lecture: Multi-parameter posteriors

Lab:
Maximum a posteriori estimation

Due:
⋮ HW 1

Linear models and model checking

Thu, Jan 21

Lecture: Linear regressions from a Bayesian perspective

Lab:
Estimating regressions in R with MAP

Required:
⋮ (McElreath 2020, Ch. 4)

Tue, Jan 26

Lecture: Covariates for causal analysis

Lab:
Creating indicators and transforming variables

Required:
⋮ (McElreath 2020, Ch. 5)

Thu, Jan 28

Lecture: Checking models and estimates

Lab:
Posterior predictive plots

Required:
⋮ (McElreath 2020, Ch. 6)

Tue, Feb 2

Lecture: Parsimony and overfitting

Lab:
Deviance and information criteria

Due:
⋮ HW 2

Generalized linear models

Thu, Feb 4

Lecture: Logistic regression: motivation

Lab:
Intercept-only logistic regression

Required:
- (McElreath 2020, Ch. 9 and Section 10.1)

Tue, Feb 9

Lecture: Logistic regression: methods and interpretation

Lab:
Prior-predictive simulation

Due:
- HW 3

Thu, Feb 11

Lecture: Counts and rates

Lab:
Poisson regression in R

Required:
- (McElreath 2020, Ch. 10)

Tue, Feb 16

Lecture: Expanding on Poisson regressions

Lab:
Overdispersed and zero-inflated Poisson regressions in R

Due:
- HW 4

Thu, Feb 18

Lecture: Categorical outcomes

Lab:
Multinomial regression in R

Required:
- (McElreath 2020, Ch. 11)

Tue, Feb 23

Lecture: Cumulative probability and ordinal outcomes

Lab:
Ordered logistic regression in R

Supplementary:
- Ordinal regressions (Michael Betancourt 2019)

Due:
- HW 5

Data complications

Thu, Feb 25

Lecture: Missing data

Lab:
Imputing missing data with brms

Required:
⋮ (McElreath 2020, Ch. 14)

Break

Tue, Mar 2

Spring break — no class

Thu, Mar 4

Spring break — no class

Data complications (continued)

Tue, Mar 9

Lecture: Non-uniform samples

Lab:
Weights and GLM in brms

Due:
⋮ Final project precis

Multilevel models

Thu, Mar 11

Lecture: Nested data and partial pooling

Lab:
Indexing by group in R

Required:
⋮ (McElreath 2020, Ch. 12)

Tue, Mar 16

Lecture: Random intercept models

Lab:
Random intercepts in R

Thu, Mar 18

Lecture: Estimation methods; introduction to random slopes

Required:
⋮ (McElreath 2020, Ch. 13)

Tue, Mar 23

Lecture: Covariance of coefficients

Due:
⋮ HW 6

Thu, Mar 25

Lecture: Two-level models in detail

Tue, Mar 30

Lecture: Multilevel GLM and R formula specification

Due:
⋮ Final project proposal

Building more complex models

Thu, Apr 1

Lecture: Time series: nesting time within individuals

Required:
⋮ (McElreath 2020, Ch. 14)

Tue, Apr 6

Lecture: Three-level and non-nested model

Presentations

Thu, Apr 8

Student presentations

Tue, Apr 13

Student presentations

References

McElreath, Richard. 2020. *Statistical Rethinking : A Bayesian Course with Examples in R and Stan*. Second. Boca Raton : Chapman & Hall/CRC,

Michael Betancourt. 2019. “Ordinal Regression.” May 2019. https://betanalpha.github.io/assets/case_studies/ordinal_regression.html.