STA303, Summer 2019: Methods of Data Analysis II

Instructor: Alex Stringer, Office: **EB103A**

You can call me… Alex. I don't have a title. Do not call me Dr. or Professor, this is incorrect.

Email: alex [dot] stringer [at] mail [dot] utoronto [dot] ca

Head TA: Robert Zimmerman, robert [dot] zimmerman [at] mail [dot] utoronto [dot] ca

- Head TA is the first point of contact for students' administrative requests (missed tests, etc). When in doubt of who to contact, contact **Rob**; he will forward to me if necessary.

Lectures: Mondays and Wednesdays 9:00AM - 12:00PM, **ES1050**

Office Hours: Mondays 12:00PM - 2:00PM, right after lecture. Give me 10 minutes to get there.

Course webpage: q.utoronto.ca

Marking Scheme:

- Test 1: 30%
- Test 2: 30%
- Final Exam: 40%

Test dates will be posted on Quercus. Final exam scheduled by faculty.

Course Description: Analysis of variance for one-and two-way layouts, logistic regression, loglinear models, longitudinal data, introduction to time series.

Prerequisite: STA302H1

Distribution Requirement: Science

Breadth Requirement: The Physical and Mathematical Universes (5)

Prerequisites will be *strictly* enforced for undergraduate students. Do not approach your instructor to ask to sign a form waiving prerequisites. Undergraduate students without the appropriate prerequisites will be removed from the course.

Textbook:
- Extending the Linear Model with R, 2nd edition (the 2016 one, that uses ggplot).
- Bayesian Regression Modelling with INLA (2018)
- Both books are available from the U of T library website.

The textbooks are mandatory. We will follow a detailed schedule of readings. Do all problems at the end of each section (there aren't that many).

Lectures

Lectures are mandatory, and are where the majority of the course material will be delivered. They will be a combination of verbal discussion, writing on the chalkboard, and interactive code examples in R. You have to come to every class; I will not be posting annotated slides or summaries of material covered. I will be posting the annotated code for the data analyses we do in class. The full lecture schedule is below. **Do all textbook readings before coming to lecture.** This will help you follow along.

Lecture #	Date	Topic	Textbook Reading/Notes
1	July 3	Review linear regression. Set up and use computing environment	Review. Required reading: ELMR Chapter 1
2	July 8	GLM: Exponential Family, Regression Model, IRWLS	ELMR chapter 6, sections 6.1, 6.2, 6.4
3	July 10	Binomial Regression + Examples	ELMR chapter 2 sections 2.1, 2.2, 2.3, 2.5, 2.10
4	July 15	Poisson Regression + Examples	ELMR chapter 3, sections 3.1 and 3.2
5	July 17	Penalized regression	https://web.stanford.edu/~hastie/glmnet/glmnet_alpha.html; up to and including linear regression
6	July 22	Dependent Data: LMEs, theory and fitting	ELMR chapter 8, sections 8.1, 8.2, 8.3
7	July 24	Dependent Data: LMEs, examples	ELMR chapter 9 sections 9.1, 9.2
8	July 29	Bayesian Inference: Introduction: posterior, conjugacy, Gaussians, Laplace Approximation	BRM Chapter 1 sections 1.1, 1.2, 1.3, 1.7
9	July 31	Bayesian Inference: Introduction: posterior, conjugacy, Gaussians, Laplace Approximation	BRM Chapter 1 sections 1.1, 1.2, 1.3, 1.7
10	August 7	Bayesian Inference: Linear Regression + Ridge and Lasso	BRM Chapter 3 sections 3.1, 3.2, 3.3, 3.7
11	August 12	Bayesian Inference: Hierarchical Models, Random Effects	BRM Chapter 5 sections 5.1, 5.2, 5.3
12	August 13	Exam Review	

Marking concerns

Any requests to have marked work re-evaluated must be made in writing within *one week* of the date the work was returned. You must email your head TA with any marking requests, **do not email your instructor with these**. Requests must include a detailed reason for the change that references **objective fact**, and must be made for **legitimate perceived errors only**. The following are unacceptable reasons for requesting a remark of any work:

- I feel my mark was unfair (the TA felt otherwise, and they know more than you)
- My friend got a better mark but they put the same thing as me (so should I lower your friend's mark then?)
- I need a bump to get my GPA over some threshold (I already bump final grades in accordance with recommendations outlined in the FAS instructor's handbook)

If you legitimately find an error, then I will happily change your mark, but in case of any ambiguity over the legitimacy of an error, **I will side with the TAs over you**.

By submitting a remark request, you are agreeing to have me (the instructor) remark your entire work, change the grade up, down, or not at all, and that the result of this represents your final mark on the work and is not to be contested further.

Missed Tests

If a test is missed for a valid reason, you must submit documentation to the **head TA**.

If a test is missed for a valid medical reason, submit a copy of the University of Toronto Verification of [Student Illness or Injury form](#) to your instructor within two weeks of the quiz. Please **scan and email the head TA** this form. The form is considered received when they email you back saying "Received; feel better!".

Important: The form must indicate that the degree of incapacitation on academic functioning is **moderate, serious, or severe** in order to be considered a valid medical reason for missing the quiz. If the form indicates that the degree of incapacitation on academic functioning is negligible or mild then this will *not* be considered a valid medical reason.

If a test is missed for a valid reason then the weight will be added to the final exam.

Computing

Using a computer is not a mandatory part of the course, but it is strongly recommended. We will be doing lots of examples of simulations and data analysis using R. You can download R from CRAN: https://www.r-project.org/

You can download RStudio from https://www.rstudio.com/products/rstudio/#Desktop

Get the open-source version, which is free and runs on Windows/Mac/Linux.

You will not be evaluated on *writing* R code in this course, but you will be evaluated on *reading* R output, and it would be a little bit silly to try to do this course without using R yourself. Do all exercises and recreate all in-class labs in R, by yourself.

Calculators

You will need a calculator. Any calculator that has logarithmic functions will be sufficient. Calculators on phones or other devices equipped to communicate with the outside world (for example, through the internet or cellular or satellite phone networks) will not be permitted during the term test and the final exam.

Online Discussion Board

This term you will have the option to use Piazza for class discussion. The link will be posted on Quercus. 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. They provide for substantial sharing and disclosure of your personal information. 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.

How to communicate with your instructor

I get a lot of email, so it is important to abide by my email policy in order to ensure you receive a prompt and efficient response. You'll get a good response from me if you chat to me after lecture or in office hours. If you feel the need to contact the course team, first email **your head TA**. This isn't because I don't want to talk to you, it's because the head TA is paid as part of their hours to respond quickly to student emails. I am extremely busy and can't guarantee good response times with hundreds of students; this is why we have a whole course team here to assist you. It's like if you're in a store buying shoes: first you ask the sales associate, then if the question is more complicated than usual, they get the manager. I'm the manager.

If you are following up on a conversation we have had in lecture or office hours, or if the matter is confidential and you don't want to share with the TA, then email me directly. But when in doubt over who to email first, email Rob. He'll just forward it to me if he can't answer you.

When communicating with *anyone* in any way, but especially by email, make sure you are professional: full sentences, no slang like "yo prof, I wanna get the lecture notes" (that's a quote from a real email I got), etc. This is good practice for your eventual transition into industry or grad school. *Make us want to reply to you.* We reserve the right to simply ignore any emails we don't like.

If you need to email us follow these steps:
- Put **STA303: Student Communication** in the subject line
- Start the email with your full name and student #, and "Hi Rob, …". First name is fine, we aren't faculty members.
- State the purpose of your email
- Say thank you or sincerely or something that indicates the email is over
- End with your name and student number

This policy is not designed to discourage student communication; it is designed to encourage *productive* student communication, by forcing you to communicate professionally. I can't overstate how much this will help you in your future career.

Example email:

"""

Hi Alex,

My name is <name>, student number <student number>, and I am a student in your STA303 class. I would like to follow up on our conversation after lecture yesterday.

<more content>

Thank you,

<name>, <student number>
"""

Academic integrity

You are responsible for knowing the content of the University of Toronto's Code of Behaviour on Academic Matters at www.governingcouncil.utoronto.ca/policies/behaveac.htm. If you have any questions about what is or is not permitted in this course, please do not hesitate to contact your instructor. **Don't use private tutoring services**, these are known for propagating mass academic fraud and if you use these services you are potentially liable for an academic offence. If you are paying a non U of T-affiliated business to receive course materials that I created, you are committing an academic offence.

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. Don't contact the course team about these things; while we're all certainly personally sympathetic to your various circumstances, we simply aren't trained to evaluate these situations. The people at accessibility services are. So contact them, and then if there is anything you need based on their documented recommendations, please reach out to me and I will happily accommodate.