STOR 557, FALL 2021
ADVANCED METHODS OF DATA ANALYSIS
Instructor: Richard L. Smith

Time of Class: This course is on **Tuesdays and Thursdays, 9:30-10:45 am**. The first class is on Thursday, August 19, and the last class is on Tuesday, November 30. There will be no class on Thursday October 21 because of Fall Break, or Thursday November 25 because of Thanksgiving. Class will not be canceled on September 7 (Rosh Hashana) or September 16 (Yom Kippur) or October 12 (University Day), but if you need accommodation on those days or other holidays that I am not yet aware of, please discuss it with me. The final exam is scheduled by the Registrar's office for **8:00 am** on **Tuesday, December 7.** Any changes to these arrangements will be discussed in advance with the class.

I *may* be out of town on Thursday, November 18, but if this should come to pass I will inform the class of alternative arrangements for that date. As of now, I am not aware of any other date that I will not be available myself.

Location: Unless a change in class status is announced by the university administration, all classes are scheduled to be **in-person** in **Gardner 105**.

Instructor: Richard L. Smith, **Hanes 303**. rls "at" email "dot" unc "dot" edu

Office hours (tentative): Mondays 1:00 am – noon; Tuesdays 2:00-3:00 pm, and Wednesdays 10:00-11:00 am. These office hours may be changed after consultation with the two classes I am teaching this semester; updates will be posted here and announced directly to the class. Students may attend office hours in person or remotely, but remote attendance is encouraged for those who would like to maintain social distancing and minimize personal contact. In-person office hours will take place in Hanes 303; the zoom link for remote office hours is:

https://unc.zoom.us/j/96580181311?pwd=YUlvMy9xZWI4eE9jeWV3MFljbUFnQT09
Meeting ID: 965 8018 1311
Passcode: STOR

You are also free to approach me at the end of class but please be aware that I have another class in Hanes Hall starting at 11:00 am; therefore, I will not be able to stay of more than a few minutes.

Instructional Assistant and Grader: The IA will be Haodong Wang (haodong "at" live "dot" unc "dot" edu) and the grader will be Kate Konrad (kskonrad "at" live "dot" unc "dot" edu). Haodong also has an office hour that will be Thursdays, 1:00-2:00 pm: zoom link https://unc.zoom.us/j/91062925146

Prerequisites: STOR 435, STOR 455. Students without those prerequisites *may* be admitted but please discuss all such requests with the instructor.

Class Attendance Policy

This is an in-person class and attendance at all class sessions is expected of all students, unless explicitly excused. That said, I recognize that we are still far from "back to normal" with respect to COVID-19, and I am willing to make further accommodations as necessary. Please do not attend class if you or a close contact of yours receive a positive test until the requisite quarantine period has passed. I would appreciate being informed if you expect to be absent so that I do not count it as an unexcused absence, but I will make every effort to be flexible in handling such requests. All classes will be video-recorded, and the videos as well as course handouts and slides posted online as soon as practicable after class. These materials will be made available to assist all students in following the course, but please do not interpret them as license to skip class without good reason. Other excuses that are not illness-related (e.g. attending job interviews) will be accepted so long as you give me adequate notice.

Face Masks

This semester, while we are in the midst of a global pandemic, all enrolled students are required to wear a mask covering your mouth and nose at all times in our classroom. This requirement is to protect our educational community – your classmates and me – as we learn together. If you choose not to wear a mask, or wear it improperly, I will ask you to leave the class, and I may submit a report to the Office of Student Conduct. At that point you will be disenrolled from this course for the protection of our educational community. Students who have an authorized accommodation from Accessibility Resources and Service have an exception. For additional information, see Carolina Together.

Assignments and Exams

The grading of the course will be split among homeworks (25%), midterm exam (35%) and final exam (40%). Homeworks will be given at weekly or bi-weekly intervals and will mostly consist of numerical exercises to be completed in R or RStudio. You may use R-Markdown (if you are familiar with it) for your assignments but this is not required; alternative formats are latex or Word (e.g. copy your figures and tables into a Word document and add explanations as appropriate). Assignments will be announced on the course sakai page and are to be handed in via gradescope through the gradescope link on sakai.

When submitting homeworks using gradescope, we ask you to:

1. Assign pages for each question (if you submit the whole assignment in a single file) or submit separate solutions for each question or part-question;
2. Highlight your final statement for each question or sub question (e.g. use the yellow highlighter feature in R-Markdown or Word) or write your final statement of each

question or sub question before all your code and explanations; above all, **always answer the question**;

3. When you have a question about grading, we suggest submitting a regrade request on gradescope stating your concerns instead of sending us email;
4. Check the deadline of each homework to ninimize the number of students who miss the deadline.

The midterm is planned to be a take-home exam and will also consist primarily of numerical exercises to be conducted in R or RStudio. *Very tentatively*, the exam will be posted online at 6:00 pm Thursday, October 7 and due (via sakai) at 6:00 pm Friday, October 8 (in which case, the October 7 class will either be cancelled or reorganized as an optional review session). However, this will only be finalized after the class begins; let me know if you have a specific conflict with those dates.

For the final exam, as of now I am planning this as an in-class final exam on the scheduled date of Tuesday, December 7, 8:00-11:00 am. However, I am considering making this a take-home exam as well and will keep the class informed as the plan is developed.

You are reminded that the university Honor Code is in effect for this course. For homework assignments, you are allowed to discuss the problems among yourselves, but the work you hand in must be your own; direct copying is not permitted. For exams, whether in-class or take-home, you are expected to work the problems entirely by yourselves and consultation of any kind is forbidden, unless it is with me or the Instructional Assistant.

Course Materials and Topics

Required Text: *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models, Second Edition* by Julian J. Faraway, Published 2016, Chapman and Hall/CRC Press.
https://www.crcpress.com/Extending-the-Linear-Model-with-R-Generalized-Linear-Mixed-Effects-and/Faraway/p/book/9781498720960

Copies will be available in the Student Stores: you are welcome to obtain your own copy from other sources if you like, but please make sure to get *Second Edition.*

Course Outline:

This course covers topics in linear models going beyond the material in STOR 455. The primary foci will be (a) Generalized Linear Models; (b) Random Effects; (c) Bayesian Statistics; (d) Nonparametric Methods (kernels, splines and related techniques). The course will be heavily computational, using the R statistical package (or RStudio), emphasizing the analysis of large datasets. However, you should expect to see some theoretical derivations as well where these are necessary to motivate the computational procedures; there will be no formal theorem/proof style mathematics. The material is distinct from that in STOR 556, which covers time series analysis; students who have taken or will take STOR 556 are welcome to take STOR 557 as well, but STOR 556 is not a prerequisite for STOR 557. Students who took STOR 556 from me in Spring

2019, or STOR 590 in Spring or Fall 2020, are not eligible to take STOR 557 in Fall 2021 (because it's the same course except for the number).

Course webpage:

http://rls.sites.oasis.unc.edu/s557-2021/s557.html

I plan to maintain a webpage separate from the course sakai site; please check both for course materials and assignments.

List of Course Topics

I reserve the right to modify or change the order of topics as the course proceeds, but as of the start of semester, the planned sequence of topics is as follows (all chapter references are to the Faraway course text):

1. Review of linear models and logistic regression (assumed to have been covered in STOR 455).
2. Binary response models (Chapter 2).
3. Generalized linear models: general theory and methods (Chapter 8).
4. Binomial and proportion models (Chapter 3).
5. Count regression (Chapter 5).
6. Contingency tables: two-way, three-way, matched pairs, ordinal variables (Chapter 6 – omit correspondence analysis).
7. Other GLMs: gamma, inverse gaussian, joint modeling of mean and dispersion, quasi-likelihood (omit Tweedie GLM).
8. Random effects: basic concepts, estimation and inference, prediction, diagnostics (first half of Chapter 10).
9. Examples of random effect models: block designs, split plots, nested effects, crossed effects, multilevel models (second half of Chapter 10).
10. Repeated measures (Chapter 11).
11. Bayesian methods (Chapter 12).
12. Generalized linear mixed models (Chapter 13).
13. Introduction to nonparametric regression (Chapter 14).

Accessibility Resources

The University of North Carolina at Chapel Hill facilitates the implementation of reasonable accommodations, including resources and services, for students with disabilities, chronic medical conditions, a temporary disability or pregnancy complications resulting in barriers to fully accessing University courses, programs and activities.

Accommodations are determined through the Office of Accessibility Resources and Service (ARS) for individuals with documented qualifying disabilities in accordance with applicable state and federal laws. See the ARS Website for contact information: https://ars.unc.edu or email ars@unc.edu. (source: https://ars.unc.edu/faculty-staff/syllabus-statement)

Counseling and Psychological Services (CAPS)

CAPS is strongly committed to addressing the mental health needs of a diverse student body through timely access to consultation and connection to clinically appropriate services, whether for short or long-term needs. Go to their website: https://caps.unc.edu/ or visit their facilities on the third floor of the Campus Health Services building for a walk-in evaluation to learn more. (source: Student Safety and Wellness Proposal for EPC, Sep 2018)

Title IX Resources

Any student who is impacted by discrimination, harassment, interpersonal (relationship) violence, sexual violence, sexual exploitation, or stalking is encouraged to seek resources on campus or in the community. Please contact the Director of Title IX Compliance (Adrienne Allison – Adrienne.allison@unc.edu), Report and Response Coordinators in the Equal Opportunity and Compliance Office (reportandresponse@unc.edu), Counseling and Psychological Services (confidential), or the Gender Violence Services Coordinators (gvsc@unc.edu; confidential) to discuss your specific needs. Additional resources are available at safe.unc.edu.

Honor Code: (For the complete honor code, please visit http://instrument.unc.edu/)

It shall be the responsibility of every student enrolled at the University of North Carolina to support the principles of academic integrity and to refrain from all forms of academic dishonesty, including but not limited to, the following:

1. Plagiarism in the form of deliberate or reckless representation of another's words, thoughts, or ideas as one's own without attribution in connection with submission of academic work, whether graded or otherwise.

2. Falsification, fabrication, or misrepresentation of data, other information, or citations in connection with an academic assignment, whether graded or otherwise.

3. Unauthorized assistance or unauthorized collaboration in connection with academic work, whether graded or otherwise.

4. Cheating on examinations or other academic assignments, whether graded or otherwise, including but not limited to the following:

(a) Using unauthorized materials and methods (notes, books, electronic information, telephonic or other forms of electronic communication, or other sources or methods);

(b) Violating or subverting requirements governing administration of examinations or other academic assignments;

(c) Compromising the security of examinations or academic assignments;

(d) Representing another's work as one's own; or

(e) Engaging in other actions that compromise the integrity of the grading or evaluation process.

5. Deliberately furnishing false information to members of the University community in connection with their efforts to prevent, investigate, or enforce University requirements regarding academic dishonesty.

6. Forging, falsifying, or misusing University documents, records, identification cards, computers, or other resources so as to violate requirements regarding academic dishonesty.

7. Violating other University policies that are designed to assure that academic work conforms to requirements relating to academic integrity.

8. Assisting or aiding another to engage in acts of academic dishonesty prohibited in the above items.

Administrative details

- All questions regarding course registration and waiting list should be directed at Ms. Christine Keat, crikeat@email.unc.edu.

- The instructor reserves to right to make changes to the syllabus.