

STOR 557, FALL 2023
ADVANCED METHODS OF DATA ANALYSIS
Instructor: Richard L. Smith

Time of Class: This course is on **Tuesdays and Thursdays, 3:30-4:45 pm**. The first class is on Tuesday, August 22, and the last class is on Tuesday, December 5. There will be no class on Tuesday, September 5 (Well Being Day), Thursday October 19 (Fall Break) or Thursday November 23 (Thanksgiving). If you need accommodation on other holidays that I am not aware of, please discuss it with me. The final exam is scheduled by the Registrar's office for **4:00 pm** on **Saturday, December 9**. Any changes to these arrangements will be discussed in advance with the class.

Location: All classes will take place **in-person** in **Hanes 120**.

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

Office hours (shared with STOR 664): Mondays 1:00-3:00 pm; Tuesdays 2:00-3:00 pm, and Wednesdays 12:00-1:30 pm. There will be some weeks that I have to reschedule or cancel my office hours but I will make every effort to inform the class about this in advance. Students may attend office hours in person in Hanes 303, or remotely via <https://unc.zoom.us/j/8238989389>. You are also free to approach me at the end of class.

Instructional Assistant and Grader: The IA is Anna Myakushina (amyak "at" unc "dot" edu) and the grader is David Zhang (weiye "dot" zhang "at" unc "dot" edu). Feel free to contact them with queries about homeworks and grading. **Updated October 2:** Anna has an office hour on Mondays, 11:30-12:30, in person in Hanes B30 or by zoom at <https://unc.zoom.us/j/3184291031>.

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. If you will not attend class for some reason I appreciate being informed in advance; there is no need to seek a formal university excuse unless you expect to be away for several classes in succession. In particular, please do not attend class if you have a positive test for COVID or are experiencing COVID symptoms; I will make suitable accommodations in such cases. Although I have not yet made a final decision about recording classes, I expect to record a substantial portion of the class presentations. Excuses for other reason (e.g. religious holidays, attending job interviews) will be accepted so long as you give me adequate notice.

Face Masks and Vaccinations

There is no requirement that students at UNC must either wear a face mask or be vaccinated. Nevertheless, I urge you to be cautious and considerate of your fellow students (and me) on these matters, especially if you think you might have been exposed to one of the COVID variants. My recommendation is that you continue to wear a face mask in most indoor settings where you are with people who are not family members or close friends, but this is not a requirement.

If you have still not had a COVID vaccination, I urge you to consider doing so.

Assignments and Exams

The grading of the course will be split among homeworks (25%), two midterm exams (25% each) and final exam (25%). 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 canvas page and are to be handed in via gradescope through the gradescope link on canvas.

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 minimize the number of students who miss the deadline.

There will be two midterm exams, which I am scheduling as in-class exams though it is possible I will switch to a take-home format. The tentative dates for these exams are **Thursday** September 21 and Tuesday October 31; if you are aware of a conflict with either of those dates, please let me know a.s.a.p.

The final exam has been set by the registrar's office as Saturday, December 9, 4:00-7:00 pm. Again, it is possible I will switch to a different format (after consultation with the class), but please keep this date and time free.

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.

Course webpage:

<http://rls.sites.oasis.unc.edu/s557-F23/s557.html>

Most materials for the course will be posted on the course canvas site; please check it for course materials and assignments.

List of Course Topics

I may modify or change the order of topics as the course proceeds, but 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.