STOR 590, SPRING 2020
ADVANCED LINEAR MODELS
Instructor: Richard L. Smith

This course is on Mondays, Wednesdays and Fridays, 9:05-9:55 am.
Location: Hanes 120
Instructor: Richard L. Smith, Hanes 303. rls *at* email.unc.edu
Office hours: Provisionally, my office hours will be 12:00-2:00, Mondays and Wednesdays. These are subject to change but I will announce any changes in class and in a revision of this syllabus.
Instructional Assistant: Taebin Kim
Grader: Xinjie Qian

Prerequisites: STOR 435, STOR 455.

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

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) 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 took STOR 556 in Fall 2019 are welcome to take STOR 590 as well, but STOR 556 is not a prerequisite for STOR 590. Students who took STOR 556 in Spring 2019 are not eligible to take STOR 590 in Spring 2020.

Course webpage: https://rls.sites.oasis.unc.edu/s590-2020/s590.html

It is likely that I will put most materials on the sakai page rather than a public webpage. However, you will find a link to last year's STOR 556 webpage and you are welcome to consult the materials posted there (please note, however, that I won't necessarily use exactly the same handouts and examples as I did for last year's class; whatever you find on the course sakai page, you can assume is up to date and intended for this year's class.

The first class is on Wednesday, January 8 and the last class is on Friday, April 24. There will be no classes on Monday January 20 (MLK Holiday), the week of March 9/11/13 (Spring Break) or on Friday April 10 (state holiday). The final exam is fixed by the University Registrar for

Monday, May 4 at 8:00 am, but this will most likely be replaced by a take-home exam (further details below).

The grading of the course will be split among homeworks (25%), midterm exam (25%) and final exam (50%). Homeworks will be given at roughly two-week intervals and will mostly consist of numerical exercises to be completed in R or RStudio. They will be announced on the course sakai page and are to be handed in via the Assignments tab on sakai.

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, Feb 27 and due (via sakai) at 6:00 pm Friday, Feb. 28. However, this is subject to change as I will consult the class about the date as well as the exact times for posting and collection.

Tentatively again, the final exam will be handled the same way, as a take-home with a 24-hour window for completion. At the moment I am proposing to use the reading day, Wednesday, April 29 as the main date for this. A possible schedule is that the exam will be posted online at 6:00 pm Tuesday, April 28 and due (via sakai) at 6:00 pm Wednesday, April 29. However, I can only do it this way if I can agree a schedule with the whole class. For the moment, please regard all this information as tentative but let me know at once if you know of a problem with this proposal.

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.

Class policies

1. Attendance in class is expected, and I may ask you to sign an attendance sheet from time to time. I will not normally penalize an occasional absence from class, but if I note you are frequently absent, that may impact your grade.
2. The class begins at 9:05 and ends at 9:55. Please do not expect me to end before 9:55.
3. If you have to leave the class early (or arrive late), I ask you to do so quietly and not disturb other students.
4. If you know in advance that you will not be present or will not attend the full class, I will appreciate receiving a note about it (personal email to me).
5. Please limit any use of cellphones or computers in class to tasks directly connected with the class.

List of Course Topics

The following is tentative, and based on the material I covered as STOR 556 last year; I may add or delete topics as I go. Chapter numbers are as in Faraway; however I plan to bring forward the material in Chapter 8 as this gives the general principles used in each of Chapters 2-7.

1. Quick 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 – if time permits).
12. Generalized linear mixed models (Chapter 13).
13. Introduction to nonparametric regression (Chapter 14).