STAT 525-01: Regression Analysis (Spring 2020)

Course Description:

Regression analysis answers questions about the dependence of a response variable on one or more predictors, including prediction of future values of a response, discovering which predictors are important, and estimating the impact of changing a predictor or a treatment on the value of the response. This course focuses on *linear* regression, which is the basis for many modern, advanced regression techniques, including those used by in statistics, machine learning, and data science. Specific topics include:

- Simple linear regression and statistical inferences
- Correlation analysis
- Diagnostics and remedial measures
- Matrix approach to simple linear regression analysis
- Multiple linear regression and statistical inferences
- Regression models for quantitative and qualitative predictors
- Model building (and variable selection) and validation
- Model diagnostics

A matrix formulation of the linear regression model is given partway through the course. This is for ease in presenting models and results and understanding some of the computational documentation, not for proving regression results using matrix theory. This is primarily an applied statistics course. While models and methods are written out carefully with some basic derivations, the primary focus of the course is on the understanding and presentation of regression models and associated methods, data analysis, interpretation of results, statistical computation, and model building.

Prerequisites: Stat 516 or equivalent previous coursework in probability and statistics, which includes knowledge of estimation, confidence intervals, and hypothesis testing and its use in at least one and two sample problems. You must be familiar with these statistical concepts beforehand. Stat 515 by itself is NOT a sufficient background for this course! Familiarity with basic matrix notation and operations is helpful.

Administrivia:

- Times: TTh 8:30AM-9:45AM
- Location: LGRT 141
- Website: `https://maryclare.github.io/stat525`
- Instructor: Maryclare Griffin
- Instructor Office: LGRT 1342
- Instructor Office Hours: Wednesday 9:00AM-11:00AM, or by appointment
- Instructor E-mail: `mgriffin@math.umass.edu`
- Teaching Assistant: Dongah Kim
- Teaching Assistant Office: LGRT 1435N
- Teaching Assistant Office Hours: TBD
- Teaching Assistant E-mail: `dkim@math.umass.edu`
- Textbook: Applied Linear Regression Models by Kutner, Nachsteim and Neter (4th edition) or, Applied Linear Statistical Models by Kutner, Nachtsteim, Neter and Li (5th edition). Both published by McGraw-Hill/Irwin.
  - Note: The first 14 chapters of Applied Linear Statistical Models (ALSM) are EXACTLY equivalent to the 14 chapters that make up Applied Linear Regression Models, 4th ed., with the same pagination. The second half of ALSM covers experimental design and the analysis of variance, and is used in our STAT 526. If you are going to take STAT 526, you should buy the Applied Linear Statistical Models (but it is a large book).

- The computing in this course will be conducted in R, freely available software available for multiple platforms at: `www.r-project.org`. We may have a few bring-your-laptop-to-class days, in which we will work through R exercises in class, and help everyone keep up with R. If you are completely new to R, there are many good online tutorials that can be helpful for getting started.

Schedule and Key Dates:

Week 1	1/21	1/23	
Week 2	1/28	**1/30**	(Optional) IE Step 0 Due Thursday
Week 3	2/4	**2/6**	IE Step 1 Due Thursday
Week 4	2/11	2/13	
Week 5	No class!	**2/20**	IE Step 2 Due Thursday
Week 6	2/25	2/27	
Week 7	3/3	**3/5**	Midterm Thursday in Class
Week 8	3/10	3/12	
Week 9	3/24	**3/26**	IE Step 3 Due Thursday
Week 10	3/31	4/2	
Week 11	4/7	**4/9**	IE Step 4 Due Thursday
Week 12	4/14	4/16	
Week 13	4/21	**4/23**	Last day of lecture Tuesday, presentations start Thursday
Week 14	**4/28**	No class!	Presentations continue Tuesday

- May 4: IE Final Papers Due
- May 6: 8:00-10:00: Final Exam (LGRT 141)

Grading:

- Homework 35%
- Project 25%
- Midterm exam 15%
- Final exam 20%
- Class participation 5%

Letter grades are typically as follows:

F	D	D+	C-	C	C+	B-	B	B+	A-	A
<60	60+	65+	70+	74+	77+	80+	84+	87+	90+	94+

Homework:
Homework will typically be due at the beginning of class on Thursday. For homework including code, please turn in your homework as a narrative, addressing the question in the homework. Rmarkdown or similar homework approaches that interweave pieces of text, code, and output are good, as long as you use the code, output, and plots to support the words and summary numbers, rather than expecting the reader to read what you've done or pick out the right numerical answers directly from raw output. Please include all code either with each question or at the end.

Project:
STAT 525 counts as an integrative experience (IE) course for undergraduate Math-Stat primary majors. The students in STAT 525 will complete a team project involving data analysis and interpretation in R, based on the principles introduced in class, as well as drawing on students' other learning and experience. Projects will be done in groups of about 4 students. Deliverables will include a project proposal, a presentation, and a final written report. Details of this project will be discussed in class.

Exams:
There will be a midterm and a final exam. The final exam will be cumulative, but mostly focused on later material, and earlier material as it relates to the later material.

Class Participation:
Full marks for class participation do not require asking a certain number of questions, or volunteering many answers, but do require attending class, answering when called on, paying attention in class, and participating in any in-class activities.

Missed Work:
I expect all students to take responsibility for keeping up with the material. Late or missed work will not be graded. The lowest homework grade will be dropped. In documented extenuating circumstances, I will reserve the right to make reasonable arrangements.

Grading Errors:
If you believe there was an error in grading your work, please return the work, along with a written description of the problem to me within 1 week of the return of that assignment, and I will re-evaluate.

Excused Absences:
If you miss class, it is your responsibility to find out what you missed, both class material and announcements. Because the lowest homework grade will be dropped, there will be no make-up homework assignments or quizzes, and missed assignments will be graded 0. If, due to extreme health or personal circumstances, you are unable to attend an exam you must notify me in advance to make alternate arrangements. In documented extenuating circumstances, I reserve the right to make reasonable arrangements.

Collaboration and Academic Honesty:
The goal of class is to facilitate learning. Most of your learning will come from struggling with the problems assigned in homework. Do not be afraid to struggle. Talking with classmates can also facilitate learning. Students are encouraged to work together and help each other on homework assignments. However, all submitted work must be your own. A good rule of thumb is that another student or online/print source might describe to you their approach to solving the problem, and even show you their work. You should think about their approach, and make sure you understand it. Then look away from their solution, and write up your own solution, or write your own code from your own understanding.

Pronouns:
If you would like me to refer to you using a pronoun other than the one on Spire, please let me know.