Course Overview

The purpose of this course is to fully immerse graduate students into regression as an applied statistical tool. You - the student - will learn standard linear regression. What I refer to linear regression includes all parts of the general linear model (GLM) along with extensions (e.g., generalized linear models, generalized estimating equations, linear mixed effects models, path models, etc.). The first three-quarters of the semester focuses solely on the GLM while the final quarter shifts to the aforementioned extensions.

Due to time constraints along with a daunting work load, every student needs to devote no fewer than 10 hours each week to the course. Please consult students who completed the course with Jose Cortina and they will all tell you that my hour totals are probably underestimates. I specified a time because I want all students to be prepared. The huge course text requires your attention and every student must read the assigned readings prior to class. I say “must” because if you fail to read ahead of time, you will not be able to ask questions or understand in class assignment. Reading, therefore, takes place outside class and cannot be substituted by class attendance.

If you devote the prescribed time and energy to the course, you will learn how to conduct a thoughtful analysis with these tools. Additionally, you will be able to appreciate the strengths and limitations of linear models.

Required Texts

- Berry, W.D. (1993). Understanding regression assumptions. pp. 13-83. Newbury Park, CA: Sage. (Order it directly from Sage from order@sagepub.com. It is cheaper than going through the bookstore. This is #92 in their Quantitative Applications in the Social Sciences series.) (aka Berry)

Course Requirements

I expect all students to have a good grasp of basic linear regression. You ought to be able to run both a bivariate and multiple regression - not just run them but interpret the results.

My two-week review is fast and furious so if you have any reservations about your knowledge, please review your prior course material. Also, all of the assignments require a computer and a statistics program (more later). Lecture material will be available via youtube (or its equivalent) throughout the semester. If you do not own a computer with reliable internet access or have access to a statistical software program, I recommend you acquire them well before the course begins. All assignments must be completed and submitted electronically; please prepare accordingly and do not rely on other people to lend you these tools.

Statistical Software

Most psychologists use SPSS but I do not; I use R. You decide what tool you wish to use but please be aware that I am not able to help you with any software problems. Refer to the wealth of online resources to troubleshoot problems. I assume that you know how to use your computer and statistical software program according to the guidelines set forth in this class. For students who wish to learn R, I recommend you visit my online materials (see course website listing for more details). Additionally, those who are new to R and want to learn it, I require you to read Fox’s chapters 1 through 3. He provides an excellent overview of R for the course.

Course Structure

Each student will be required to read assignments, attend class, and contribute to the weekly group projects. Readings consist of book chapters from the required texts as well as additional, content-relevant articles posted on the course website. By attending class and contributing to class discussion, students get the opportunity to understand the assigned readings and discuss the implications and limitations of each topic. Finally, the weekly assignments will be conducted as groups but everyone in the group will be held accountable for the results. I will ask one person at random from your group to come demonstrate the results to the class. The random selection serves as your exam. Every student will be selected throughout the semester at least twice and at least three people will present each lecture day. Come prepared and you will have nothing to reason to worry.
Grading

All grades will be based upon student participation in weekly class discussions, weekly group projects, and individual randomly selected presentations. Every student gets two opportunities to present. Thus, your grade will be assigned according to the following criteria:

- Class Participation - 40%
- Weekly Group Projects - 40%
- Individual Presentation - 20%

I will assign a weight - ranging between 0 and 1.0 - according to my observations for the first two aspects of the course. Grades for the individual presentations will be assigned as pass/fail. I will assign final grades for the class according to the standard grading criteria for A (> 85%), B (> 75%), and C (< 75%) with pluses and minuses added in only questionable situations.

Academic Honesty

I must state for the record that cheating of any kind will be dealt with by rules set forth in the University Honor Code (see http://www.gmu.edu/catalog/apolicies/index.html). I prefer never to have any academic integrity problems arise during the semester. The aim of graduate education is to learn material that many others have not learned and master this material to ensure your future success. The degree you receive reflects the hard work you put into your courses. Please do not cheat yourself by misrepresenting your effort. Do the work or accept the consequences. Spend your effort learning the material and avoid being overly grade conscious. With a concerted effort to learn, you will not be tempted to cheat. Please note that academic dishonesty is not akin to studying with your classmates. I strongly encourage you to study together, exchange notes, and offer each other constructive feedback about your review preparation. My course is designed to eliminate any possibility of dishonesty. The only avenues to cheat yourself is by plagiarizing your group project but your classmates will control that more than any policy. So, let me repeat myself; please study with one another but do your work. I demand both.

Disability Accommodations

If you are a student with a disability and you need academic accommodations, please see me and contact the Disability Resource Center (DRC) at 703-993-2474. All academic accommodations must be arranged through that office.

Topics

Below lists selected topics with relevant readings. Note the abbreviations are included in the required text section above. Where I list a “c” before a number, you need to read the chapter. If I only provide a name for the reading, you must read the entire book.

1. Review of simple regression (CC c1 & c2, Fox c4)
   - (a) Purpose of regression
   - (b) Line of best fit
     i. Errors and least squares criterion
     ii. Calculation of b and a
   - (c) Standard error of estimate
   - (d) Meaning of b
   - (e) Standardized approach
   - (f) Example from Pedhazur
2. Other purposes of regression
   - (a) Contribution of multiple predictors
   - (b) Linear Interaction effects
   - (c) Simple Nonlinear relationships
   - (d) Categorical Predictors
   - (e) Nested data
   - (f) Nonlinear interaction effects
   - (g) Analysis of noncontinuous dependent variables
3. Assumptions of regression (Berry & CC c4)
4. General multiple regression (CC c3 & c5)
   - (a) Meaning of b
   - (b) Significance tests for R and B
   - (c) Multicollinearity
   - (d) Stepwise vs. Hierarchical
5. Moderated Multiple regression (CC c6 & c7)
   - (a) What is a moderator?
   - (b) Hierarchical test of significance for moderator
   - (c) Plotting interactions
   - (d) Power considerations
   - (e) Nonlinear effects
   - (f) Nonlinear interactions
6. Mediator relationships
   - (a) How do they differ from moderators?
   - (b) How are they tested?
   - (c) Preacher & Hayes, bootstrapping, etc
7. Regression with dummies (Alternative course title?: CC c8 & c9)
   - (a) How does one cope with categorical predictor variables in regression?
   - (b) How are the results interpreted?
8. Regression diagnostics (Fox c6)
   - (a) Using diagnostics to identify violated assumptions
   - (b) Using diagnostics to identify bizarre occurrences
9. Random Coefficient Modeling (CC c15, Fox c5)
10. Logistic regression (CC c13)
   (a) Regression with dichotomous criteria
   (b) Interpretation of output
11. Path Analysis (CC c12)
   (a) Differences from Multiple regression
   (b) Differences from structural equation modeling
   (c) Parameter estimation and model fit.

**Additional Readings**

Throughout the semester, I will post additional readings on the course website (see URL in the header). Please consult the course website every week for updates. I usually send out an email message alerting students to updates to the website but there are times when I simply forget to warn you. As a general rule, assume I will not alert you and, to avoid missing a reading, go to the course website every Sunday to check for readings. If you visit the website early enough in the week, you will be easily able to read the additional material.