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 as 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 at 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.
Grading

As many of you know, I grade based upon demonstrated skill and knowledge via modules. You perform some pre-defined task or set of skills in front of me or one of my astute TA’s and you receive one of three grades - a “0” for failure, a “1” for good job, and a “2” for exceptional performance. The first two grades are easy. You get a “0” if you fail to demonstrate adequate mastery of the checklist items below - a “1” if you get a check for each item. The “2” gets assigned to students who show a mastery of the material by integrating optional reading into the module performance. Each grade depends upon your preparation and “game day” performance. If you rehearse your performance many times then you will do fine. Anything short of repeated rehearsals results in awkward negotiations between you and me. I suggest we avoid those negotiations and have you prepare well in advance for your performance.

Modules

As I mentioned above, your grades are based upon 15 minute modules. I require you to perform three (3) modules - at least one in front of me and the other two may be submitted via youtube or some other online video service. Students have given me great feedback about the modules and many like the video option while others prefer the standard, in person option. I leave it to you to decide which option you prefer but remember you must perform at least one (1) in front of me (yes ME and not one of my TA’s). Here is what I expect from each module:

Module 1: Standard Linear Model
- Explain the purpose
- Explain the parts of the model
- Conduct an analysis (SW)
- Explain ALL results
- Run diagnostics
- Discuss implications

Module 2: Advanced Linear Models
- Moderation and Mediation
  - Explain models
  - Explain model parts
  - Conduct both with continuous predictors (SW)
  - Conduct both with categorical predictors (SW)
  - Conduct a complex model with both mediation and moderation
- Explain ALL results
- Run diagnostics (SW)
- Discuss implications

Module 3: Generalized Linear Models
- Random Coefficient Models
  - Explain models
  - Contrast with standard linear models
  - Explain model parts
  - Conduct analysis (SW)
  - Explain ALL results
  - Run diagnostics (SW)
  - Discuss implications
- Non-normal distribution model
  - Explain your selected model
  - Contrast with standard linear models
  - Explain model parts
  - Conduct analysis (SW)
  - Explain ALL results
  - Run diagnostics (SW)
  - Discuss implications

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 chap-
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.