SOCY-7111: <u>Data III</u>
Spring Semester 2014
Wednesdays, 7-9:50pm

Instructor: Ryan K. Masters
Office: Ketchum Hall 214
Office Hours: T & Th., 3:15-5:00 p.m.
Email: ryan.masters@colorado.edu

Course Overview

This course is intended to extend your understanding and application of statistical techniques for answering sociological questions. We will cover a wide array of statistical methods that move beyond basic OLS & MLE regression models to deal with:

(1) Categorical Dependent Variables

The first part of the course will introduce generalized linear models (GLMs), which are necessary for the analysis of non-continuous measured outcomes. No familiarity with this subject is assumed or required. We will begin by looking at logit models (i.e., logistic regression) as a subset of GLMs because of the often analyzed binary (1/0) outcomes in the social sciences.

(2) Generalized Linear Models (GLM) for Alternative Dependent Variables

We will continue our introduction to GLMs for analyzing other non-continuous outcomes (e.g., counts, nominal categories, and ordinal categories) and the various considerations in appropriate modeling.

(3) Multilevel/Hierarchical Models for Change

The second half of the course will introduce multilevel models for change (i.e., growth curve models), which are appropriate for the analysis of change in a continuous dependent variable over time. No familiarity with this subject is assumed or required. The major aspects of estimating and interpreting multilevel/hierarchical models will be reviewed in the context of growth curve modeling, which you should be able to extend to other forms of multilevel/clustered data structures (e.g., individuals nested within families, schools, states, etc.)

(4) Event History Analysis

The final component of this course will deal with continuous-time and discrete-time event history analysis (i.e., survival analysis, hazard rate models, etc.), which are techniques for modeling the transition from one status (or state) to another. Examples include life course transitions like marriage, birth, divorce, entry into the labor market, etc. We will focus on both discrete-time and continuous-time models that make few assumptions regarding time dependence of the hazard (e.g., piecewise constant exponential and the Cox proportional hazard model). We will focus on single transition models, although multi-state and competing risk methods may be discussed, time permitting. We will also consider frailty and shared frailty models as well as discrete- and continuous-time multilevel hazard models

Course Meetings

Class: Wednesdays, 7:00pm till 10:00pm

Course Material

Required Textbooks:

Generalized Linear Models and Extensions. 3rd Edition. By James W. Hardin and Joseph M. Hilbe. Referred to as "H&H" in the schedule.

Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence. By Judith D. Singer and John B. Willett. Referred to as "S&W" in the schedule.

Online Readings: I will post additional readings and Stata material to D2L.

Recommended Readings:

Statistical Methods for Categorical Data Analysis. 2nd Edition. By Dan Powers and Yu Xie.
Data Analysis Using Regression and Multilevel/Hierarchical Models. By Andrew Gelman and Jennifer Hill.
Regression Models for Categorical Dependent Variables Using Stata. 2nd Edition. By J. Scott Long.
Multilevel and Longitudinal Modeling Using Stata, Volume I: Continuous Responses. 3rd Edition. By Sophia Rabe-Hesketh and Anders Skrondal.

Multilevel and Longitudinal Modeling Using Stata, Volume II: Categorical Responses, Counts, and Survival. 3rd Edition. By Sophia Rabe-Hesketh and Anders Skrondal.
Flexible Parametric Survival Analysis Using Stata: Beyond the Cox. By Patrick Royston and Paul C. Lambert.

Course Requirements and Assessment

This is an applied course. We will learn most by applying the techniques covered in class to specific examples in multiple data sets (including your own). No previous programming experience is required, but familiarity with Stata and R will be a benefit.

Homework Assignments (5)	15% each
Final Homework Assignment (1)	25%

POLICIES & ACCOMMODATIONS

Accommodations

1. Appropriate academic accommodations will be provided to students with disabilities. Please contact the Disability Services office located in Center for Community as soon as possible to obtain documentation: N200 (303-492-8671) http://disabilityservices.colorado.edu/ . Guidelines for addressing temporary medical conditions and/or injuries can be found at http://disabilityservices.colorado.edu/general-information/temporary-injuries

2. The University of Colorado at Boulder has both legal and moral obligations to accommodate students who choose to abstain from classes and/or miss scheduled examinations in order to observe holidays. If you plan to be absent from class to observe a holiday, please notify me of any scheduling conflicts, in writing, by February 3.

Course Expectations and Honor Code

1. I expect academic integrity (and the university requires it!). While I encourage you to collaborate with one another on exercises, support each other in studying, and edit each other's work, you are expected to turn in original work and complete all exams on your own. Students caught cheating will be reported to the Honor Code Council, and will also have their course grade justly penalized. Information about the Honor Code can be found at http://www.colorado.edu/policies/student-honor-code-policy

2. Every homework assignment shall be turned in during the first five minutes of class on the scheduled due date. All assignments turned in after this time on the same day or emailed on the due date will be punished two points. Please type your homework assignments, exams, and papers.
3. Please know that the syllabus and course schedule are not set in stone. I reserve the right to change the basic course requirements, due dates, and overall content and schedule with adequate notice to students via D2L, class announcements, and/or email.

Classroom Etiquette
Please refrain from conversing with your neighbors during class. This can be quite disruptive to fellow students around you.

Laptops are not needed in class. If you choose to bring a laptop to class, please know that you're subject to a one-strike policy. If you are found using your laptop for an activity unrelated to class – a notepad/word processor and/or Excel – I will ban computers entirely from the classroom.

Turn off all cell phones before class begins.
You and I both have the responsibility for maintaining a professional learning environment. Those who fail to adhere to basic modicum of adult behavior may be subject to discipline. Please be courteous and sensitive to alternative perspectives, especially when dealing with topics pertaining to race, culture, religion, sexuality, political ideology, nationality, gender identity & expression, age, and disability.

Please know that the University provides me a class roster containing your picture and legal name. I will happily honor your request to remove your picture and/or address you by an alternative name if you like. Please notify me early in the semester.

Email Policy
Include "7111" in the subject line of all course-related emails.

Respect, Discrimination, and/or Harassment
Please respect your classmates. Topics discussed in class may be interpreted as contentious by some, and I would like everyone to feel comfortable enough to freely and openly participate. I will do my best to present the material and discuss the topics in an open and inclusive manner.

Spring 2014 Data III Schedule

January 15

Syllabus, Course Overview/Expectations, and Review of OLS/MLE

H&H   Chapter 1

January 22

Generalized Linear Models: Families (Distributions) and Links

H&H   Chapters 2, 4, and 6

January 29

Categorical Covariates, Contingency Tables, and Logit Regression

H&H   Chapter 10 through Chapter 12

February 5

Building and Interpreting a Logit Regression Model

Exercise #1 Due

February 12

Multinomial and Ordered Logit Regression

H&H   Chapter 16 and Chapter 17

February 19

Count Models: Poisson and Negative Binomial Regression

H&H   Chapter 13 through Chapter 15

Exercise #2 Due

February 26

GLM Extensions: Overdispersion, Rare Events, Clustered Data Structure

H&H   Chapter 18 through Chapter 20

March 5

Multilevel Models of Change: Data Formats & Conceptualizing Individual Trajectories

S&W   Chapter 1 through Chapter 3

Exercise #3 Due

March 12

Data Analysis with Multilevel Models of Change: Conditional Means to Slope Variation

S&W   Chapter 4 and Chapter 5

March 19

Importance of Error: Nonlinearity of Time/Trajectories and Error Covariance

S&W   Chapter 6 and Chapter 7

Exercise #4 Due

March 26

Spring Break

April 2

Event Occurrence

S&W   Chapter 9 and Chapter 10

April 8

Discrete-Time Survival Models

S&W   Chapter 11 and Chapter 12

Exercise #5 Due

April 15

Continuous-Time Survival Models & the Classics

S&W   Chapter 13

April 22

Cox Proportional Hazard Model

S&W   Chapter 14

April 29

Treating Time More Flexibly: Extending the Cox and the Royston-Parmar Model

S&W   Chapter 15

May 2

Unobserved Heterogeneity and Frailty Models

Final Exercise Due May 7