SOCY7111: Longitudinal Data Analysis
Spring Semester 2017
MF 3:00-4:15 PM, Ketchum 1B24

Instructor: Ryan K. Masters
Office: Ketchum 264
Office Hours: W 10:10am-12:00pm, F 10:10-11:00am, and by appointment.

Course Description

This is a course in statistical methodologies for analyzing longitudinal data structures. We will cover two main topics: (1) multivariate regression models for analyzing data collected on the same subjects (e.g., individual persons) over time (i.e., repeated occasions) and (2) methods for modeling event occurrence over time (i.e., survival analysis, event-history analysis, time-to-event data).

Topics

Multilevel/Hierarchical Models for Change.

The first half of the course will introduce and discuss multilevel models for change (i.e., growth curve models), which are appropriate for analyzing variation in continuous outcomes over time. No familiarity with this subject is assumed or required. The primary 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 data structures (e.g., individuals nested within families, neighborhoods, classrooms, states, etc.). We will also review grow curve models for categorical outcomes (e.g., counts) and nonlinear growth curve analyses.

Event History Analysis.

The second half of the course will introduce and discuss event history analysis (i.e., survival analysis, hazard rate models, time-to-event analysis, etc.), which is a technique for modeling the transition from one status (e.g., not a high school graduate) to another (e.g., a high school graduate) or one state (e.g., alive) to another (e.g., dead). Examples include classic life course transitions such as marriage, birth, divorce, entry and exit from the labor force, death, etc. We will focus on both discrete and continuous time models that make few assumptions regarding time dependence of the hazard, such as the Cox proportional hazard model – a continuous time model – and the piecewise constant exponential model – a discrete time model. We will focus entirely on single transition models. We will also consider frailty, shared frailty, and time-dependent modeling techniques.

Background

Students enrolled in SOCY7111 should have prior exposure to graduate level statistical methods. This should include a course in basic inferential statistics and a course in multivariate linear regression. Most of the material in the exercises and lectures is based on applied problems, as

opposed to theoretical (i.e., mathematical probability and estimation techniques). As this is a sociology course, we will cover examples that are mainly sociological and demographic in nature.

Course Requirements and Assessment

As 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 programming experience is required, but familiarity with Stata and/or R will be a benefit.

Homework Assignments

1. Working with Longitudinal Data and Trajectories
2. Fitting the Baseline Mean Model to the Unconditional Growth Model
3. Handling Various Aspects of Time
4. Nature of Event Occurrence and Discrete-Time Models
5. Continuous-time Models
6. Extending the Cox Model

Policies and Accommodations

Appropriate academic accommodations will be provided to students with disabilities. Please contact the Disability Services office located in the Center for Community (C4C Room N20) as soon as possible to obtain documentation (303-492-8671). Email: dsinfo@colorado.edu

Guidelines for addressing temporary medical conditions and/or injuries can be found at: http://www.colorado.edu/disabilityservices/students/temporary-medical-conditions

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

Expectations and Honor Code

1. I expect academic integrity and the university requires it. Although I encourage you to collaborate with one another on exercises and support each other's progress, you are expected to turn in original work. Students caught cheating will be reported to the Honor Code Council and will 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. Hard copies of homework assignments shall be turned in during the first five minutes of class on their scheduled due date. All assignments turned in after this time will be punished 10 percentage points per day late. Please type your homework assignments.
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 me and fellow students around you.

Laptops are not needed in class, but they are permitted. If you choose to bring a laptop to class, please be respectful to me and others in the classroom and use the device for class-related purposes only (e.g., following lecture slides, taking notes, running Stata and/or R).

Turn off all cell phones, iPods, and other noisy and distracting gadgets before class begins.

You and I both have the responsibility for maintaining a professional learning environment. Those who fail to adhere to a 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 identify and expression, age, and disability.

Please respect your classmates and me. 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.

Please know that the university provides me a class roster including a 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 email me early in the semester to indicate your preference(s).

Please include "7111" in the subject line of all emails related to the course, and please know that I return the majority of class-related emails during my office hours.

Readings

The bulk of the readings will be selected from two required texts. Some handouts and copies containing supplementary reading material will be posted to D2L.

Required Text

J.D. Singer and J.B. Willett (2003). *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence.* New York: Oxford University Press.
Referred to as "S&W" in the course schedule below.

Patrick Royston and Paul C. Lambert (2011). *Flexible Parametric Survival Analysis Using Stata: Beyond the Cox Model.* College Station, Texas: Stata Press.
Referred to as "R&L" in the course schedule below.

Recommended Readings on Specific Subjects

Sophia Rabe-Hesketh and Anders Skrondal (2012). *Multileevl and Longitudinal Modeling Using Stata, Third Edition*. *Volume I: Continuous Responses. Volume II: Categorical Responses, Counts, and Survival.* College Station, TX: Stata Press.

Paul D. Allison (1995). *Survival Analysis Using SAS: A Practical Guide.* Cary, N.C.: SAS Publishing.

Andy Gelman and Jennifer Hill (2005). *Data Analysis Using Regression and Multilevel/Hierarchical Models.* New York: Cambridge University Press.

David W. Hosmer, Stanley Lemeshow, and Susanne May (2008). *Applied Survival Analysis: Regression Modeling of Time-to-Event Data.* 2nd Edition. Hoboken, N.J.: Wiley-Interscience.

Dan Powers and Yu Xie (2008). *Statistical Methods for Categorical Data Analysis.* 2nd Edition. London: Emerald Press.

Mario Cleves, William W. Gould, and Yulia V. Marchenko (2016). *An Introduction to Survival Analysis Using Stata, Revised Third Edition.* College Station, TX: Stata Press.

Spring 2017 7111 Schedule

Week 1. Introduction to Investigating Change. S&W Ch 1

Frameworks for investigating change over time, distinguishing between types of questions about change.

Week 2. Introduction to Longitudinal Data on Change. S&W Ch 2-3

Person-level vs. person-period data (i.e., wide vs. long data structures), individual trajectories, improving OLS models.

Week 3. Multilevel Models for Change. S&W Ch 4

Individual change (i.e., within-person variation), interindividual differences in change (i.e., between-person variation), multilevel/hierarchical models, fixed and random effects.

Week. 4. Treatments of Time, Time-varying Covariates, and Centering. S&W Ch 5

Time doesn't take care of itself, so don't leave time alone.
Homework #1 Due Friday, 2/10

Week 5. Nonlinearity and Discontinuous Change. S&W Ch 6

Spline functions, transformations, nonlinear growth.

Week 6. Error Covariance Structures of Multilevel Models. S&W Ch 7

Understanding various components of error in multilevel models, GEEs, latent trajectories
Homework #2 Due Friday, 2/24

Week 7. Multilevel Models of Categorical Outcomes. R-H&S Ch 6

Individual change in non-continuous outcomes.
Homework #3 Due Friday, 3/3.

Week 8. Events and Event Occurrence. S&W Ch 9-10. R&L Ch 1-3

Introduction to the nature of event occurrence (i.e., event history) data, censoring and truncation, describing data using life table techniques.

Week 9. Discrete-time Hazard Models and Extensions. S&W Ch 11-12; R-H&S Ch 8

Split-episode data, time-varying covariates, Clog-Log, non-linear effects.
Homework #4 Due Friday, 3/17.

Week 10. Piecewise Constant Hazard Rate Models. R&L Ch 4.

Extending the exponential model using grouped and individual-level data, proportional hazard models, proportional hazard models with nonproportional (i.e., time-varying) effects.

Week 11. ******** SPRING BREAK **********

Week 12. Continuous-time Hazard Models. S&W Ch 13

Distributional aspects of time (*T*), survivor function, hazard function, exposure, nonparametric techniques for describing continuous-time survival.

Week 13. The Cox Proportion Hazards Model. S&W Ch 14

Partial likelihood estimation, fitting the Cox model, implications of Cox's method, inference about effects and interpretation of results.

Week 14. Extending the Cox Model. S&W Ch 16

Time-varying covariates, assessing and altering the proportionality assumption (i.e., time-dependent effects), model diagnostics, competing risks.
Homework #5 Due Friday, 4/21

Week 15. Unobserved Heterogeneity in Hazard Rate Models. R-H&S Ch 8-9

Frailty models, multilevel models for event occurrence, clustered data.

Week 16. Flexible Parametric Survival Analysis. R&L Ch 5-7

Royston-Parmar Models, model selection criteria, time-dependent effects.
Homework #6 Due the FOLLOWING Friday, 5/12