Advanced Topics in Maximum Likelihood Estimation:
Duration Analysis/Event History Analysis
Professor Brad Jones
University of California, Davis
Weeks 1-2

Contact Information: Professor Brad Jones
Department of Political Science
University of California, Davis
Davis, CA 95616
USA
E-mail: bsjjones@ucdavis.edu
Website: http://psfaculty.ucdavis.edu/bsjjones

Course Description
This section of the advanced MLE course will cover methods and models for duration data. Duration data record the length of time until some event occurs, for example, the termination of a cabinet government or the time until an unemployment spell ends. Because time-to-event occurrence is an important feature of these kinds of data, methods suitable to duration data are often referred to as event history analysis. In this course, we will consider a wide variety of event history modeling methods. Students will be asked to complete some problem sets that will involve estimating and interpreting event history models. In addition to consideration of duration models, we will take a “side-trip” and consider some non-traditional (i.e. not widely used) categorical models and consider their applicability to duration data. “Tutorials” will be available to students that describe some implementation issues pertinent to these models. Additionally, some lecture notes will be available as well. Any material necessary for downloading will be available at my website: http://psfaculty.ucdavis.edu/bsjjones. From here, you will be able to follow a link to this course to access tutorials, lecture notes, and article manuscripts.

Readings
The primary texts will be Box-Steffensmeier and Jones’ Event History Modeling: A Guide for Social Scientists (Cambridge University Press, 2004) and Kleinbaum and Klein’s Survival Analysis: A Self-Learning Text (Spring, 2005). In the daily itinerary of topics, several articles will be assigned (and will be posted on the website for the book). You will need to consult the website in conjunction with the date of the class to retrieve the readings.

Requirements
Students are expected to do the assigned readings and pay attention in class. There will be two or three short problem sets. The problem sets will entail estimation and interpretation of a variety of duration models. We will make use of both the R computing environment and Stata. Lecture notes and tutorials will be available on my website. Here, relevant code for both R and Stata can be found. Finally, students will be
asked to turn in a short (1-2 page) research prospectus that outlines a research question(s) and hypotheses that could be appropriately tested using duration modeling techniques.

**Itinerary**

My principal goal is to give you an introduction to the fundamental elements of duration modeling and then consider in some detail parametric, non-parametric (via the Cox model), and “discretized” duration models for single-event and multi-event duration data. I do not assume any prior knowledge of event history modeling, though I obviously will assume knowledge of the basic principals of maximum likelihood estimation as well as a thorough understanding of the classical linear model and traditional binary link models (like logit or probit).

The following gives you the day-by-day itinerary of topics. There are two “classes” of readings each day: core and application. It is important that the core readings be completed in their entirety. Several applications are listed for each day's topics. You will not have time to read each application; I recommend choosing a couple that may be of interest to you. Just about all of the application readings (as well as the core articles) are available from J-Stor (http://www.jstor.org). Applications are highly useful to read and I encourage you to read as many of these as you can. The readings below are primarily drawn from the social sciences.

---

**DAY 1: Monday, July 18**

**Preliminaries: Event History Data and the “Moving Parts” of Event History Analysis and an Introduction to Modeling Strategies**

Core Readings:
- Box-Steppensmeier and Jones, Chapters 1—2.
- Hosmer et al, Chapters 1-3.

**DAY 2: Tuesday, July 19**

**Describing and Modeling Duration Data: The Kaplan-Meier Estimator (and related estimators) and Parametric Models**

Core Readings:
- Box-Steppensmeier and Jones, Chapter 3.
- K and K: Chapters 1 and 2

**DAY 3: Wednesday, July 20**

**Estimation and Model Selection Issues in the Application of Parametric Duration Models**

Core Readings:
- Box-Steppensmeier and Jones, Chapter 3.
• K and K, Chapter 7

DAY 4: Thursday, July 21
The Cox Proportional Hazards Model

Core Readings:
• Box-Steensmeier and Jones, Chapter 4.
• K and K, Chapter 3

DAY 5: Friday, July 22.
The Proportional Hazards Property and other Cox Model Diagnostics and Other Issues

Core Readings:
• Box-Steensmeier and Jones, Chapter 8.
• K and K, Chapters 3-5.

DAY 6: Monday, July 25
“Discretized” Duration Data and Associated Models

Core Readings:
• Box-Steensmeier and Jones, Chapter 5.

DAY 7: Tuesday, July 26
Models for Competing Risks: Discrete and Cox

Core Readings:
• Box-Steensmeier and Jones, Chapter 10.
• K and K, Chapter 9

DAY 8: Wednesday, July 27
Repeatable Events

Core Readings:
• Box-Steensmeier and Jones, Chapter 10.
• K and K, Chapter 9

**DAY 9: Thursday, July 28**

**Frailty and Split-Population Models**

Core Readings:
• Box-Steppensmeier and Jones, Chapters 9.

**DAY 10: Friday, July 29**

**Special Topics in Duration Analysis**

Readings: TBA
Advanced Topics in Maximum Likelihood Estimation: Models for Clustered/Longitudinal Data  
Professor Rodolfo Espino (espino<at>asu.edu)  
Arizona State University  
Week 3-4 (August 1-12, 2011)

Course Description
This section of the advanced MLE course will cover methods for dealing with clustered and longitudinal data. While many of the methods for dealing with clustered data (e.g., students “nested” within schools for a single cross-section) and longitudinal data (e.g., countries “nested” within years across a span of time) are similar, we will generally approach the methods for each type of data separately as indicated in the syllabus itinerary. This will serve the purposes of 1) dividing the course over the different approaches and methods of interpretation researchers apply to each type of data/research question; and 2) utilizing a more accessible book (Goldstein) for the first week when we generally discuss approaches for dealing with hierarchical data for single cross-sections and then utilizing a more advanced book (Hsiao) in the final week when we generally discuss approaches to dealing with panel data.

Readings
The primary texts will be Harvey Goldstein’s *Multilevel Statistical Models* (1999) freely available on the internet and Cheng Hsiao *Analysis of Panel Data* (2003) available in the ICPSR library or for under $20 on Amazon. There will be supplementary required and application articles that are available through JSTOR or EBSCO.

Requirements
There will be 3-4 short problem sets assigned over this period of the course. We will make use of both Stata and the R computing language.

Monday (8/1): Introduction to Multilevel Models
Theoretical and statistical justifications for hierarchical models; clarifying concepts like fixed- and random- coefficients, fixed- and random- effects; and some examples of multilevel models

Tuesday (8/2): Interpreting the Two-level Model

Readings
Applications

**Wednesday (8/3): Generalizing to Multiple Levels and Estimation Issues**

**Required**

**Applications**

**Thursday (8/4): Inference and Model Assessment**

**Required**

**Applications**

**Friday (8/5): The Hierarchical Generalized Linear Model**

**Required**

**Applications**
Monday (8/8): Dealing with TSCS Data

Required

Applications

Tuesday (8/9): Dealing with Heterogeneity

Required

Applications

Wednesday (8/10): Dynamic Panel Models

Required

Applications
Thursday (8/11): Qualitative Dependent Variables in TSCS Models

Required


Applications