1 Course Description

Although the ordinary least squares regression is the workhorse of political science research, it is often an inappropriate model choice, particularly when the outcome variable of interest is not a normally distributed, interval-level random variable. However, generalizations of the linear model allow us to exploit the logic of OLS when analyzing limited dependent variables. Through the power of maximum likelihood estimation, we can solve these generalized linear model equations and produce useful empirical analyses of our concepts of interest.

In this course, we will begin by introducing the logic of generalized linear models. We will review the available link functions through which we relate the linear model to the dependent variable. During the first part of the course, we will focus on cross-sectional, single-level models of limited dependent variables. The second half of the course will introduce the general concept of multi-level modeling. This approach allows for flexible accommodation of non-independent observations. Throughout, we will learn to analyze data using computer software (Stata and R). Additionally, we will learn how to communicate these findings in professional, print-ready documents using \LaTeX{} and rmarkdown.

In this course, we will be reading (and re-reading) a lot. We will be engaging some of the mathematical foundations of the linear regression model. To avoid this is to resort to relying upon "magical thinking," something that is not among the accepted tools of social science. Although you will generally not be required to derive your own mathematical proofs of the concepts presented here, you will need to understand how these models work on the basis of more than just intuition. Of course, reading about regression models is not enough. Instead, we need to learn how to do this kind of research. As such, this course also requires a good deal of hands-on work. This is based upon the assumption that learning by doing is the best way to build deep understanding and sustainable mastery of these techniques.

This course builds on the content of the prerequisite courses, PSC 701 and PSC 702. It assumes basic competency in math, including an understanding of linear algebra, matrix notation, rudimentary calculus, probability theory, linear regression, etc. I also assume basic competency in computer coding, especially the use of Stata. Competency in any of variety of coding languages (e.g., HTML, C++, S, Java, Python) is very helpful, as well; it will make learning R, \LaTeX{}, and rmarkdown much easier. If these areas are not your strong suit, I highly recommend brushing up on these skills.
2 Objectives

By the end of this course, the successful student will be able to

1. describe the logic of generalized linear models, including the mathematical bases and appropriate uses of the various common link functions.
2. explain the concept of multilevel modeling as a way to accommodate complex data structures.
3. critically assess the use of advanced modeling techniques for limited dependent variables in the scholarly literature.
4. analyze existing datasets with limited dependent variables and complex data structures in Stata and R.
5. appropriately diagnose models and engage in empirically-based model selection and robustness tests.
6. appropriately document research to maintain code readability, research transparency, and replicability.
7. produce publication-ready documents using TEX and markdown.
8. design and implement an original research project using limited dependent variables, and complex data structures.

3 Course Materials

3.1 Required Resources


Additional readings will be available on WebCampus.

3.2 Recommended Resources


3.3 Software Packages

For this course, we will be using software programs to conduct quantitative analyses. Our main computer package will be Stata. Stata is relatively user-friendly, and many students are familiar
with it from their PSC 701 course. This software will be pre-installed on the lab computers we’ll be using in class. Outside of class, you will need access to a copy of Stata 13 or 14 (IC, SE, or MP). You can purchase access to this software by visiting: http://www.stata.com/order/new/edu/gradplans/student-pricing/. We will also be doing some of our analyses in \texttt{R}. \texttt{R} is a very powerful and flexible program, and it’s also free. It is a little bit trickier to learn, but the effort is well worth it in the long run. We will be using \texttt{R} with \texttt{RStudio}, a graphical user interface. UNLV graduate students have access to powerful tutorials to get you up and running in \texttt{R} and \texttt{RStudio} through the Lynda portal (oit.unlv.edu/lynda). I also highly recommend the DataCamp tutorials, especially the free “Introduction to \texttt{R}” course (www.datacamp.com/courses/free-introduction-to-r).

We will also be familiarizing ourselves with the \texttt{\LaTeX} and \texttt{rmarkdown} programming languages for the production of documents. At the beginning of the semester, I will send all students an invitation to Share\texttt{\LaTeX}, which is a free, web-based \texttt{\LaTeX} editor. This language is especially useful for those students using \texttt{R}, since you can use the \texttt{knitr} package to conduct your analyses right in your document. If you’re familiar with any of the main programming languages, \texttt{\LaTeX} won’t be too terribly difficult to learn. An even easier option is the \texttt{rmarkdown} language, which allows you to incorporate \texttt{R} code, \texttt{\LaTeX} equations, and formatted text into an original source document to be rendered into html, pdf, or Word documents. You can even make interactive websites!

4 Course Requirements

4.1 Homework Assignments. 50 pts. each, up to 400 pts. total

Much of this course focuses on the practical application of statistical techniques. Because of this, it is absolutely imperative that you practice these skills as you learn them. To assist you in this case, there will be a series of twelve weekly assignments for you to complete. These assignments will require the use of Stata and/or \texttt{R}, along with \texttt{rmarkdown} or \texttt{\LaTeX}. Assignments are due before class via WebCampus. Details will be announced during the first class meeting. Please note that your score for the homework component is points-based. There are a total of 700 homework points available over the course of the term. Any points in excess of the 400 maximum will not be counted toward your final grade.

Please note that your homework assignments may be assessed, in part, by our course TA. These assessments will be according to a detailed rubric created by the professor. The TA will have only partial access to the WebCampus records, such that they will grade all of the assignments anonymously. As such, you will need to upload your homework assignments to WebCampus \textit{without} your name or other identifying information attached to the file.

4.2 Replicable Original Research. 400 pts.

As social scientists, much of our work consists of devising methodologically sound ways of testing research hypotheses. In this, our most advanced methods course, you’ll get a chance to put together a piece of original research. The aim is for you to produce a paper that is ready to present at a conference or to submit for peer-reviewed publication. This research must be replicable to the greatest extent possible. This means that you will need to submit a working dataset and the relevant scripts/codes to replicate your work.

To keep you on track, portions of this project will be due at various points throughout the semester. The final draft of your paper is due via the WebCampus link at 5 pm on December 5. You will present your paper in the style of a conference presentation during the final exam period, which is December 12. Although the exam period is from 6-8 pm, we will go ahead and meet at
the regular time (5:30 pm). Details will be announced during the first class meeting. Please note that this class may be held in a different room, and there may be an audience.

Please note that I follow the Academic Misconduct policy (found below in the University Policies section). I report every instance of plagiarism that I find. I reserve the right to exercise my discretion as to the severity of the punishment I suggest to the Office of Student Conduct, as per the rules laid out in the Student Academic Misconduct Policy. I take this very seriously. I interpret plagiarism more broadly than simply word-for-word adoption of another author’s text; I consider the unattributed use of ideas, literature review, or paper structure to be in violation of the Policy, as well. Given that you will be conducting replication studies, you must be very diligent about making sure that the resulting paper is your original work.

4.3 Seminar Contribution. 200 pts.

It is probably no surprise to you that this course will require a lot of reading. I have done my best to organize the readings to make the workload manageable. I have divided the readings into three different categories. The first is the required reading. The required readings are listed first and with the tag “Required” in Section 5.2. These readings must be read carefully before the class period. Although many of the required readings cover the same material, it is important to see this material presented in a few different ways in order to truly absorb it. The second category of readings is labeled “In Action” in Section 5.2. These readings provide you an opportunity to see some of the relevant concepts at work. You may choose to skim these readings, but make sure that you read the methodology section very carefully. You’ll need to learn how the methods are used and be exposed to the way the results are communicated. The final category of readings are “Recommended.” These readings provide you the information you need to delve deeper into the concepts. You may want to spend some time reading the recommended readings on topics that are particularly relevant to your own research interests.

To earn points for your contribution to the seminar, you need to come to all course meetings prepared to discuss the week’s readings. The course is a seminar, and it only works if students attend and are prepared to engage the course material. Attendance is required. I expect that all assigned reading will be completed prior to class. Repeated absences and/or failure to prepare for class will result in a significant reduction to the seminar contribution component of the grade. In addition, I expect that students will observe a baseline level of socially acceptable behavior in the seminar. Uncivil behavior will not be tolerated. Students are not allowed to record any portion of these seminars without express permission of the professor.

4.4 A Note on Grading

There is a norm in graduate school of grade inflation. Typically, grades for graduate seminars range from an A to a B+. Grades of B or below are rare, and are assigned only in cases of near total student dereliction. Many instructors of graduate courses give a grade of B- or C+ to all failing students, regardless how badly they have failed. However, in a course as critical to your graduate school success as this one, following this tradition may end up giving students a false sense of their performance and potential. As such, I will assign the full range of grades in this course. The cut-offs for final grades are as follows:
Table 1: Points Required for Letter Grades

<table>
<thead>
<tr>
<th></th>
<th>Passing Grades</th>
<th>Failing Grades</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>940-1000</td>
<td>800-839 D</td>
</tr>
<tr>
<td>A-</td>
<td>900-939</td>
<td>770-799 D-</td>
</tr>
<tr>
<td>B+</td>
<td>870-899</td>
<td>740-769 F</td>
</tr>
<tr>
<td>B</td>
<td>840-869</td>
<td>700-439 I [rare]</td>
</tr>
</tbody>
</table>

5 Course Schedule

5.1 Brief Overview of Course

Table 2: Tentative Schedule of Topics: Overview

<table>
<thead>
<tr>
<th>Session</th>
<th>Date</th>
<th>Topic</th>
<th>Highlights</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8/29</td>
<td>Introduction to GLM</td>
<td>HW1</td>
</tr>
<tr>
<td>2</td>
<td>9/12</td>
<td>The Logic of MLE</td>
<td>HW2</td>
</tr>
<tr>
<td>3</td>
<td>9/19</td>
<td>Binary Outcomes</td>
<td>HW3, Paper Topic</td>
</tr>
<tr>
<td>4</td>
<td>9/26</td>
<td>Ordinal Outcomes</td>
<td>HW4</td>
</tr>
<tr>
<td>5</td>
<td>10/3</td>
<td>Nominal Outcomes</td>
<td>HW5</td>
</tr>
<tr>
<td>6</td>
<td>10/10</td>
<td>Count Data</td>
<td>HW6</td>
</tr>
<tr>
<td>7</td>
<td>10/17</td>
<td>Duration Models</td>
<td>HW7, Data &amp; Methods</td>
</tr>
<tr>
<td>8</td>
<td>10/24</td>
<td>The Logic of MLM</td>
<td>HW8</td>
</tr>
<tr>
<td>9</td>
<td>10/31</td>
<td>Binary MLM</td>
<td>HW9</td>
</tr>
<tr>
<td>10</td>
<td>11/7</td>
<td>Ordinal MLM</td>
<td>HW10</td>
</tr>
<tr>
<td>11</td>
<td>11/14</td>
<td>Nominal MLM</td>
<td>HW11</td>
</tr>
<tr>
<td>12</td>
<td>11/21</td>
<td>Count MLM</td>
<td>HW12, Working Scripts</td>
</tr>
<tr>
<td>13</td>
<td>11/28</td>
<td>Duration MLM</td>
<td>HW13</td>
</tr>
<tr>
<td>14</td>
<td>12/5</td>
<td>Nested Models</td>
<td>HW14, Full Paper</td>
</tr>
<tr>
<td>15</td>
<td>12/12</td>
<td>Mini-Conference</td>
<td>Presentation; Discussant</td>
</tr>
</tbody>
</table>

5.2 Detailed Schedule of Topics and Readings

1. August 29: Introduction to Generalized Linear Models. During this session, we will review the basics of data organization and visualization. We will also become acquainted with Stata, R, \texttt{\LaTeX}, and \texttt{rmarkdown}. We will also spend some time talking about the generalization of the linear model (GLM), which forms the basis of all of the models we will learn about in the course. Homework 1 is due. We will complete this homework in class. Before class, read:

- **Required:** Long & Freese Ch. 1-2
- **Required:** The Syllabus
- **Recommended:** Fox & Weisberg Ch. 1-3
- **Recommended:** Fox Ch. 15
- **Recommended:** Liu Ch. 1
- **Recommended:** Smithson & Merkle Ch. 1.1-1.2
- **Recommended:** Relevant \texttt{\LaTeX}, R, and/or \texttt{rmarkdown} tutorials.
2. September 12: *The Logic of Maximum Likelihood Estimation*. In this session, we will discuss the differences between least squares estimation and maximum likelihood estimation (MLE). We will be relying heavily upon MLE to estimate the GLM models in this course. Homework 2 is due. Before class, read:

- **Required**: Long & Freese Ch. 3-4
- **Recommended**: Liu Ch. 2
- **Recommended**: Smithson & Merkle Ch. 1.3
- **Recommended**: Fox & Weisberg Ch. 5-5.2

3. September 19: *Binary Outcomes*. In this session, we will introduce the generalization of the linear model for binary dependent variables. We will cover the logic and interpretation of logit and probit. Homework 3 is due, as is your paper topic statement. Before class, read:

- **Required**: Long & Freese Ch. 5-6
- **Recommended**: Long Ch. 14.1
- **Recommended**: Liu Ch. 3
- **Recommended**: Smithson & Merkle Ch. 2
- **Recommended**: Fox & Weisberg Ch. 5.3, 6

4. September 26: *Ordinal Outcomes*. This session introduces the generalization of the logit and probit models to the case of an ordinal dependent variable. We will learn to assess the appropriateness of modeling ordered data this way. Homework 4 is due. Before class, read:

- **Required**: Long & Freese Ch. 7
- **Recommended**: Liu Ch. 4-8, 12.1
- **Recommended**: Long Ch. 14.2.3
- **Recommended**: Smithson & Merkle Ch. 4

5. October 3: *Nominal Outcomes*. Here, we will extend the logic of the logit and probit models to the multinomial case, including conditional logit. Homework 5 is due. Before class, read:

- **Required**: Long & Freese Ch. 8
- **Recommended**: Fox Ch. 14.2.1-14.2.2
- **Recommended**: Liu Ch. 12.2
- **Recommended**: Smithson & Merkle Ch. 3
- **Recommended**: Fox & Weisberg Ch. 5.6
6. **October 10: Count Data.** Here, we will look at ways to generalize the linear model to handle dependent variables that are count data. We will focus on the poisson and negative binomial regression techniques. Homework 6 is due. Before class, read:

- **Required:** Long & Freese Ch.9
- **Recommended:** Fox Ch. 15.2
- **Recommended:** Smithson & Merkle Ch. 5
- **Recommended:** Fox & Weisberg Ch. 5.5, 5.9

7. **October 17: Duration Models.** This session will cover the generalization of the linear model to dependent variables that represent event history or duration. Homework 7 is due, as is your data and methods section. Before class, read:

- **Required:** Box-Steefensmeier & Jones Ch. 1-6.
- **Recommended:** Smithson & Merkle Ch. 6-7
- **Recommended:** Fox & Weisberg Ch. 5.9, 5-10.4

8. **October 24: The Logic of Multilevel Modelling.** In this class, we will discuss various ways to accommodate systematic non-independence of observations. The focus will be on the logic of multilevel modeling and the different options for conceptualizing these interrelationships. Homework 8 is due. Before class, read:

- **Required:** Rabe-Hesketh & Skrondal Volume I preface
- **In Action:**
- **Recommended:** Fox Ch. 23
- **Recommended:** Smithson & Merkle Ch. 8-8.2
- **Recommended:** Liu Ch. 10-10.2

9. **October 31: Binary Response Multilevel Models.** In this session, we will apply to logic of multilevel models to binary dependent variables. We will focus on various specifications of multilevel logit. Homework 9 is due. Before class, read:

- **Required:** Rabe-Hesketh & Skrondal Ch. 10.


- **Recommended**: Liu Ch. 10.3
- **Recommended**: Smithson & Merkle Ch. 8.2.1

10. **November 7: Ordinal Response Multilevel Models.** Here, we will discuss how multilevel modeling works with ordinal dependent variables. We will concentrate on the proportional odds model and explore alternative link functions. Homework 10 is due. Before class, read:

    - **Required**: Rabe-Hesketh & Skrondal Ch. 11
    - **Recommended**: Liu Ch. 11-12.1

11. **November 14: Nominal Response Multilevel Models.** In this session, we will learn about the various models to accommodate categorical responses and discrete choice outcomes. We will focus on multinomial logit and conditional logit. Homework 11 is due. Before class, read:

    - **Required**: Rabe-Hesketh & Skrondal Ch. 12
    - **Recommended**:

12. **November 21: Count Data Multilevel Models.** This session will extend multilevel modeling techniques to count data. We will focus on multilevel Poisson and negative binomial models. Homework 12 is due, as is your working replication script(s) and datasets for your results section. Before class, read:

    - **Required**: Rabe-Hesketh & Skrondal Ch. 13

13. **November 28: Duration Data Multilevel Models.** In this session, we will cover multilevel models for duration data. Homework 13 is due. Before class, read:

    - **Required**: Rabe-Hesketh & Skrondal Ch. 14-15
    - **Required**: Box-Steffensmeier & Jones Ch. 7-11
    - **Recommended**:

14. **December 5: Models with Nested & Crossed Random Effects.** In this session, we will wrap up our discussion of multilevel models of limited dependent data. Here, we will deal with complex nested or crossed random effects specifications. Homework 14 is due, as is your final paper.

    - **Required**: Rabe-Hesketh & Skrondal Ch. 16


15. **December 12:** **Mini-Conference.** Your discussant comments will be due before class. You will present the results of your own research and provide professional commentary for a classmate. Please note that this class may be held in a different room, and there may be an audience.


**University Policies**

**Academic Misconduct** — Academic integrity is a legitimate concern for every member of the campus community; all share in upholding the fundamental values of honesty, trust, respect, fairness, responsibility and professionalism. By choosing to join the UNLV community, students accept the expectations of the Student Academic Misconduct Policy and are encouraged when faced with choices to always take the ethical path. Students enrolling in UNLV assume the obligation to conduct themselves in a manner compatible with UNLV’s function as an educational institution.

An example of academic misconduct is plagiarism. Plagiarism is using the words or ideas of another, from the Internet or any source, without proper citation of the sources. See the Student Academic Misconduct Policy (approved December 9, 2005) located at: https://www.unlv.edu/studentconduct/student-conduct.

**Copyright** — The University requires all members of the University Community to familiarize themselves with and to follow copyright and fair use requirements. You are individually and solely responsible for violations of copyright and fair use laws. The university will neither protect nor defend you nor assume any responsibility for employee or student violations of fair use laws. Violations of copyright laws could subject you to federal and state civil penalties and criminal liability, as well as disciplinary action under University policies. Additional information can be found at: http://www.unlv.edu/provost/copyright.

**Disability Resource Center (DRC)** — Disability Resource Center (DRC)—The UNLV Disability Resource Center (SSC-A 143 http://drc.unlv.edu/, 702-895-0866) provides resources for students with disabilities. If you feel that you have a disability, please make an appointment with a Disabilities Specialist at the DRC to discuss what options may be available to you. If you are registered with the UNLV Disability Resource Center, bring your Academic Accommodation Plan from the DRC to the instructor during office hours so that you may work together to develop strategies for implementing the accommodations to meet both your needs and the requirements of the course. Any information you provide is private and will be treated as such. To maintain the confidentiality of your request, please do not approach the instructor in front of others to discuss your accommodation needs.

**Religious Holidays Policy** — Any student missing class quizzes, examinations, or any other class or lab work because of observance of religious holidays shall be given an opportunity during that semester to make up missed work. The make-up will apply to the religious holiday absence only. It shall be the responsibility of the student to notify the instructor within the first 1 calendar days of the course for fall and spring courses (excepting modular courses), or within the first 7 calendar days of the course for summer and modular courses, of his or her intention to participate in religious holidays which do not fall on state holidays or periods of class recess. For additional information, please visit: http://catalog.unlv.edu/content.php?catoid=6&navoid=531.

**Transparency in Learning and Teaching** — The University encourages application of the transparency
method of constructing assignments for student success. Please see these two links for further information:

https://www.unlv.edu/provost/teachingandlearning

https://www.unlv.edu/provost/transparency

Incomplete Grades—The grade of I—Incomplete—can be granted when a student has satisfactorily completed three-fourths of course work for that semester/session but for reason(s) beyond the student’s control, and acceptable to the instructor, cannot complete the last part of the course, and the instructor believes that the student can finish the course without repeating it. The incomplete work must be made up before the end of the following regular semester for undergraduate courses. Graduate students receiving “I” grades in 500-, 600-, or 700-level courses have up to one calendar year to complete the work, at the discretion of the instructor. If course requirements are not completed within the time indicated, a grade of F will be recorded and the GPA will be adjusted accordingly. Students who are fulfilling an Incomplete do not register for the course but make individual arrangements with the instructor who assigned the I grade.

Tutoring and Coaching—The Academic Success Center (ASC) provides tutoring, academic success coaching and other academic assistance for all UNLV undergraduate students. For information regarding tutoring subjects, tutoring times, and other ASC programs and services, visit http://www.unlv.edu/asc or call 702-895-3177. The ASC building is located across from the Student Services Complex (SSC). Academic success coaching is located on the second floor of the SSC (ASC Coaching Spot). Drop-in tutoring is located on the second floor of the Lied Library and College of Engineering TEB second floor.

UNLV Writing Center—One-on-one or small group assistance with writing is available free of charge to UNLV students at the Writing Center, located in CDC 3-301. Although walk-in consultations are sometimes available, students with appointments receive priority assistance. Appointments may be made in person or by calling 702-895-3908. The student’s Rebel ID Card, a copy of the assignment (if possible), and two copies of any writing to be reviewed are requested for the consultation. More information can be found at: http://writingcenter.unlv.edu/.

Rebelmail—By policy, faculty and staff should e-mail students’ Rebelmail accounts only. Rebelmail is UNLV’s official e-mail system for students. It is one of the primary ways students receive official university communication such as information about deadlines, major campus events, and announcements. All UNLV students receive a Rebelmail account after they have been admitted to the university. Students’ e-mail prefixes are listed on class rosters. The suffix is always @unlv.nevada.edu. Email within WebCampus is acceptable.

Final Examinations—The University requires that final exams given at the end of a course occur at the time and on the day specified in the final exam schedule. See the schedule at: http://www.unlv.edu/registrar/calendars.

Library statement:
Students may consult with a librarian on research needs. For this class, the Subject Librarian is (https://www.library.unlv.edu/contact/librarians_by_subject). UNLV Libraries provides resources to support students’ access to information. Discovery, access, and use of information are vital skills for academic work and for successful post-college life. Access library resources and ask questions at: https://www.library.unlv.edu/.