POL 419 H1S: Quantitative Methods and Data Analysis

University of Toronto
Winter 2014

Kenichi Ariga
Email: kenichi.ariga@utoronto.ca

Time: Friday, 10:00AM – 12:00PM
Location: FE 36
Office Hours: Thursday, 3:00PM – 5:00PM
Office: SS 3047
Prerequisite: POL242, POL322, or equivalent

Additional Office Hours by
Dr. Hanil Chang (hanil.chang@utoronto.ca)
Tuesday, 1:00PM – 3:00PM
at SS3045

Overview
This is an advanced-level class on quantitative empirical research methods for political science for those who have taken POL242, POL322, or equivalent, and have basic understanding of statistical inference and linear regression model.

The primary goals of this class are:
1. To learn the basics of statistical computing using statistical software called R, a freely available, increasingly popular, and flexible and powerful statistical computing language; and
2. To be able to interpret and apply statistical models for discrete dependent variables, such as binary variables, ordered categorical variables, nominal categorical variables, and event counts, that are widely used in quantitative empirical political science research.

1. Statistical Computing
Quantitative social science research requires the use of computers. In the past two decades, the development of affordable yet very powerful personal computers has revolutionized the use of statistical analyses in social sciences. Various statistical models have been developed and made readily available for researchers. The number of social scientists who have been trained for applying those models has increased, and the volume of applied empirical work using quantitative data and methods has skyrocketed. One of the key driving forces behind these advancements is the wide adaptation of statistical software among social scientists.

In this class you will learn the basics of how to use R to conduct statistical analyses in political science research. This class being designated as a political science course rather than a

computer programming course, the emphasis will be placed on using the program, applying models, and interpreting results rather than on learning how to program.

2. Statistical Models for Discrete Dependent Variables
While a basic linear regression model is a canonical tool in political science research, there are many important research questions and subjects that require more advanced statistical models. Among these advanced statistical models, we will focus on the models for *categorical dependent variables* in this class.

A linear regression model is appropriate if a dependent variable is continuous or can be considered continuous. Many political phenomena of our interest, however, may not necessarily be expressed in a continuous variable. For example, we may be interested in a dichotomous outcome, such as approving or disapproving the current prime minister, winning or losing a seat, engaging in a war or avoiding one, and complying with international law or breaching one. We may also be interested in nominal multiple outcomes, such as a vote choice of the Liberals, the Conservatives, and the New Democrats. Some multiple outcomes also involve a natural order, such as an individual's choice on the Likert scale from strongly disagree to strongly agree. We may also be interested in the number counts of certain events, such as terrorist attacks, incidents of ethnic violence, and amendments to bills in the parliament. A linear regression model is not appropriate for these types of dependent variables.

Accordingly, there have been many research articles in political science that use statistical models suitable for categorical dependent variables. In this class, you will learn how to estimate and interpret major categorical dependent variable models.

In addition, in order to understand statistical computing and discrete dependent variable models, we will cover more theoretical foundations of statistical analysis than was covered in POL242 and POL322.

Readings
Required readings will be made available through the class Blackboard site. In addition, the following book will be used as reference on how to use the R statistical computing language and made available for purchase at the UofT Bookstore.

John Fox and Sanford Weisberg, *An R Companion to Applied Regression, Second Edition* (SAGE
        Publications, 2011).

Blackboard / Learning Portal
The class Blackboard site (or the Learning Portal: https://portal.utoronto.ca/) will be the primary means through which class announcements and assignments will be distributed. Readings and datasets will be made available in the class Blackboard site as well. Its Discussion Board will be the primary medium by which you will ask questions about the course materials and get them answered (more on this below). It will be your responsibility to obtain access to the class Blackboard site and regularly check it. There will be an important update to the class Blackboard site at least once a week.

Course Requirements
Your grade will be determined by three components: 1) homework assignments, 2) a paper assignment, and 3) class participation.

1) Homework assignments: 60% (Each homework assignment will be weighted equally.)
2) Paper assignment: 30%
3) Class Participation: 10%

1) Homework Assignments: There will be almost-weekly assignments, which would take different formats depending on the class materials covered each week. Many of them will involve data analysis and interpretation using R. Other assignments may include a mathematical problem set and a short methodological critique to academic journal articles. For some assignments, you will be required to work and submit the assignment individually, but for many others, you will be allowed and encouraged to work and submit the assignments in a small group (the maximum number of group members will be determined by the instructor later). If you submit your work as a group, everyone in the group will receive the same mark. Please note that the group and individual submissions will be treated equally.

2) Paper Assignment: You will be required to write a paper which employs statistical models covered in this class. You may use the data you have collected or obtained elsewhere, or the data to be provided by the instructor. The details of the format of the paper assignment will be determined in mid-February reflecting our progress and appropriate technical levels for the class. You are not expected to spend more than a month on this paper assignment. You are allowed and encouraged to work and submit the paper as a group (again, the maximum group size will be determined by the instructor). If you submit your work as a group, everyone in the group will receive the same mark. The group and individual submissions will be treated equally.

3) Class Participation: Your class participation grade will be determined by 1) whether you have fulfilled several participatory components of the class to be required throughout the semester, such as completing a survey and regularly attending the class, and 2) how well you participate in and contribute to the class learning, such as asking questions in lectures, responding to your classmates' questions on the class Discussion Board, and leading class discussions. Quality as well as quantity of your participation will be evaluated.

Group Work and Collaboration
Group work and collaboration is encouraged for this class. Given somewhat technical and advanced nature of the class materials, it is essential to have an opportunity to discuss with your classmates the concepts and methods you learn in class and how to apply them. Everyone has different strengths and weaknesses in their understanding of materials and learning style. Through working together, you are expected to facilitate learning for each other and deepen your understanding of the materials, which would be difficult if you worked alone. In fact, collaboration in a team of multiple scholars is a norm for contemporary social science research in general, and quantitative empirical political science research in particular. For many assignments of this class, collaboration in a team is encouraged so much so that even the submission by a

team is allowed and treated equally as the submission by an individual. Please note that there will be a cap on how many students may participate in one group, and the cap may change depending on assignments.

Turnitin
Normally, students will be required to submit their course essays to Turnitin.com for a review of textual similarity and detection of possible plagiarism. In doing so, students will allow their essays to be included as source documents in the Turnitin.com reference database, where they will be used solely for the purpose of detecting plagiarism. The terms that apply to the University's use of the Turnitin.com service are described on the Turnitin.com web site.

Students who wish to not use Turnitin.com may make an alternative arrangement with the instructor. They will need to let the instructor know well before the deadline of the assignment and ask for the alternative way to submit the essay. They will be required to save every version/draft of their essay electronically, and submit all of them at the time they submit the essay. They will also be required to hand in all notes, outlines, and bibliographic research at the same time.

Late Penalties
All work is late if submitted after the date and time specified as the due date. Any assignments handed in late will result in a penalty of 20 percentage points reduction per day (e.g., from 90% to 70%). Submitting assignments within 24 hours from the due date and time will be considered one day late; submitting after 24 hours but before 48 hours will be two days late, and so forth. Assignments handed in more than five calendar days late will receive a zero grade.

Extension
Extension may be granted only when there is a legitimate reason, such as an unforeseeable medical emergency, and when there is an acceptable official documentation, such as the UofT Medical Certificate, the College Registrar's Letter, and the Accessibility Services Letter. Students who know in advance they will need an extension for a legitimate reason should contact the instructor as early as possible before the deadline. Those who missed the deadline for a legitimate, unforeseeable reason should contact the instructor as soon as possible and no later than one week after returning to class.

Grade Appeals
Grade appeals must be received within one month from when the grade is assigned. When you appeal your grade, you are required to submit a documentation substantiating why you believe the grade is not appropriate.

Office Hours
You are welcome to visit during the instructor's office hours, which will be held during the time and date specified at the beginning of the syllabus, if you have any questions on the class materials.

There will also be additional office hours, held by Dr. Hanil Chang, a postdoctoral fellow at the Department of Political Science, during the time and date specified at the beginning of the syllabus. He is also available for appointments upon email request (hanil.chang@utoronto.ca).

Discussion Board
We will use the Discussion Board in the class Blackboard site as the main medium through which you can ask questions regarding class materials and get answers. Given the nature of the course materials, someone else may have the same question as yours and s/he would benefit from your posting the question and getting an answer through the Discussion Board. You are also encouraged to post an answer to the questions posted by your classmates so that we can maintain a mutually-supporting learning community from which all of you will benefit. As stated above, your response to your classmate's questions on the Discussion Board will be reflected on your class participation grade. The instructor will regularly check the Discussion Board (once on Mondays, Wednesdays and Fridays) and answer questions which have not been adequately addressed by peers. For more complex questions or those that would require an extensive treatment, you are best advised to visit the office hours.

Email Policy
If you have questions of personal nature (e.g., accessibility, deadline extension for legitimate reasons), you may email the instructor and expect a response within two working days. Please start the subject heading of your email with "POL419:..." I will not respond to, however, any questions over email that are of substantive nature concerning the class materials. You will need to post those questions on the Discussion Board.

Please note that I will not be able to answer emails or Discussion Board questions during weekends.

In the case of your questions of substantive nature on the Discussion Board or those of personal nature over email not answered within two working days (excluding weekends), send me an email to let me know they have not been addressed. Please include "POL419: Unanswered Question" in the subject heading of your email.

Accessibility
The University of Toronto is committed to accessibility. If you require accommodation for a disability, or have any accessibility concerns about the course, the classroom or course materials, please contact Accessibility Services at (416) 978-8060 or www.accessibility.utoronto.ca as soon as possible.

Academic Integrity
Academic integrity is fundamental to learning and scholarship at the University of Toronto. Participating honestly, respectfully, responsibly, and fairly in this academic community ensures that the U of T degree that you earn will be valued as a true indication of your individual academic achievement, and will continue to receive the respect and recognition it deserves.

You are expected to be familiar with the Code of Behaviour on Academic Matters (http://www.governingcouncil.utoronto.ca/policies/behaveac.htm), which is the rule book for academic behaviour at the U of T. Potential offenses include, but are not limited to, plagiarism, cheating on tests and exams, fraudulent medical documentation and improper collaboration on marked work.

The University of Toronto treats cases of academic misconduct very seriously. All suspected cases of academic dishonesty will be investigated following the procedures outlined in the Code. The consequences for academic misconduct can be severe, including a failure in the course and a notation on your transcript. If you have any questions about what is or is not permitted in this course, please do not hesitate to contact me.

Class Topics
The following is the list of topics to be covered throughout the semester. The dates for each topic are not given to allow some flexibility to adjust our course schedule to our needs and pace. More details on each class session and the reading assignments will be announced as we proceed.

Part 1: Basics of Statistical Computing and Statistical Analysis
(1) Basics of Statistical Computing in R
- Data structures
- Descriptive statistics, such as histograms, boxplots, and scatterplots
- Linear regression analysis
(2) Some Theoretical Foundations of Statistical Analysis
- *Conceptual* understanding of a little bit of math behind statistical analysis: functions, differentiation, integration, exponents, logarithms, vector, matrix, limits, etc.
- Probability and probability distribution
- Statistical inference using R
(3) Linear Regression Model and Research Design
- Reverse causality, confounding variable(s), and omitted variable bias
- Experiments and observational studies
- A few advanced topics on linear regressions, such as interaction terms, instrumental variable regressions, and robust standard errors

Part 2: Categorical Dependent Variable Models for Political Science Research
A list of actual models to be covered may be adjusted by the instructor as we go along taking into consideration our progress and appropriate technical levels for the class. Please note that we may not cover all the models listed here.

(1) Binary dependent variable model: Logit and probit models
- A dependent variable takes one of the two categories (i.e., either 1 or 0).
- The emphasis will be given to substantive interpretation and visualization of the model's implications.
- Examples of the application of these models

- o Benoit, Kenneth, and Michael Marsh. 2008. "The Campaign Value of Incumbency: A New Solution to the Puzzle of Less Effective Incumbent Spending." *American Journal of Political Science* 52 (4): 874-890. (http://onlinelibrary.wiley.com/doi/10.1111/j.1540-5907.2008.00348.x/abstract)
- o Miller, Michael K. 2012. "Economic Development, Violent Leader Removal, and Democratization." *American Journal of Political Science* 56 (4): 1002-1020. (http://onlinelibrary.wiley.com/doi/10.1111/j.1540-5907.2012.00595.x/abstract)
- o Stasavage, David. 2010. "When Distance Mattered: Geographic Scale and the Development of European Representative Assemblies." *American Political Science Review* 104 (4): 625-643. (http://dx.doi.org/10.1017/S0003055410000444)

(2) Ordinal dependent variable model: Ordered logit and ordered probit models
- A dependent variable takes multiple categories with natural order.
- Examples of the application of these models
  - o Kocher, Matthew Adam, Thomas B. Pepinsky, and Stathis N. Kalyvas. 2011. "Aerial Bombing and Counterinsurgency in the Vietnam War." *American Journal of Political Science* 55 (2): 201-218. (http://onlinelibrary.wiley.com/doi/10.1111/j.1540-5907.2010.00498.x/abstract)
  - o Croco, Sarah E. 2011. "The Decider's Dilemma: Leader Culpability, War Outcomes, and Domestic Punishment." *American Political Science Review* 105 (3): 457-477. (http://dx.doi.org/10.1017/S0003055411000219)
  - o Blinder, Scott, Robert Ford and Elisabeth Ivarsflaten. 2013. "The Better Angels of Our Nature: How the Antiprejudice Norm Affects Policy and Party Preferences in Great Britain and Germany." *American Journal of Political Science* 57 (4): 841-857. (http://onlinelibrary.wiley.com/doi/10.1111/ajps.12030/abstract)
  - o Singer, David Andrew. 2010. "Migrant Remittances and Exchange Rate Regimes in the Developing World." *American Political Science Review* 104 (2): 307-323. (http://dx.doi.org/10.1017/S0003055410000110)

(3) Nominal dependent variable model: Multinominal logit model
- A dependent variable takes multiple categories that do not have natural order.
- Examples of the application of these models
  - o Lacina, Bethany. Forthcoming. "How Governments Shape the Risk of Civil Violence: India's Federal Reorganization, 1950-56." *American Journal of Political Science*. (http://onlinelibrary.wiley.com/doi/10.1111/ajps.12074/abstract)
  - o Allee, Todd L. and Paul K. Huth. 2006. "Legitimizing Dispute Settlement: International Legal Rulings as Domestic Political Power." *American Political Science Review* 100 (2): 219-234. (http://dx.doi.org/10.1017/S0003055406062125)
  - o Whitten, Guy D. and Harvey D. Palmer. 1996. "Heightening Comparativists' Concern for Model Choice: Voting Behavior in Great Britain and the Netherlands." *American Journal of Political Science* 40 (1): 231-260. (http://www.jstor.org/stable/2111701)
  - o Dancey, Logan, and Geoffrey Sheagley. 2013. "Heuristics Behaving Badly: Party Cues and Voter Knowledge." *American Journal of Political Science* 57 (2): 312-325.

(http://onlinelibrary.wiley.com/doi/10.1111/j.1540-5907.2012.00621.x/abstract)

(4) Count model: Poisson and negative binomial models
- A dependent variable is a count of some events.
- Examples of the application of these models
  o Wilson, Matthew C. and James A. Piazza. 2013. "Autocracies and Terrorism: Conditioning Effects of Authoritarian Regime Type on Terrorist Attacks." *American Journal of Political Science* 57 (4): 941-955.
  (http://onlinelibrary.wiley.com/doi/10.1111/ajps.12028/abstract)
  o Weghorst, Keith R. and Staffan I. Lindberg. 2013. "What Drives the Swing Voter in Africa?" *American Journal of Political Science* 57 (3): 717-734.
  (http://onlinelibrary.wiley.com/doi/10.1111/ajps.12022/abstract)
  o Martin, Lanny W. and George Vanberg. 2005. "Coalition Policymaking and Legislative Review." *American Political Science Review* 99 (1): 93-106.
  (http://dx.doi.org/10.1017/S0003055405051518)
  o Prakash, Aseem and Matthew Potoski. 2006. "Racing to the Bottom? Trade, Environmental Governance, and ISO 14001." *American Journal of Political Science* 50 (2): 350-364.
  (http://onlinelibrary.wiley.com/doi/10.1111/j.1540-5907.2006.00188.x/abstract)

Part 3: Student Research Presentations
Toward the end of the semester, a couple of weeks will be spent on the presentations of students' research for the paper assignment.