Quantitative Techniques in Sociology (SOC 506/CSSS 507)
COVID-19 Online Edition

Zack W. Almquist
Spring Quarter, 2020

Class Schedule

Lecture: TU & TH 3:00PM-4:20PM
Lecture URL: https://canvas.uw.edu/

Professor

Name:	Zack W. Almquist
Office:	Savery Hall 231
Office Hours:	11:00AM - 1:00PM (W)
Email:	zalmquist@uw.edu

Teaching Assistants

Name:	Connor Gilroy
Office:	
Office Hours:	2:00PM-3:00pm (TU & TH)
Email:	cgilroy@uw.edu

Course Description

SOC 506 covers statistical methods for analyzing and interpreting social data (e.g., individuals, groups and organizations) such as online behavioral data, survey data or experimental data. This course is designed for Sociology graduate students and assumes a background equivalent to SOC 505 and SOC 504, i.e., linear models, ANOVA and descriptive statistics. The class will be comprised primarily of introduction to modern statistical techniques such as categorical data analysis (e.g., logistic regression), time series analysis (e.g., event history analysis), and modern computational statistics (e.g., monte carlo tests). Students will receive both theoretical and technical training in statistics through the use of the R statistical programing language, and practical hands on data analysis experience with common

social science data sets such as the General Social Survey and US Census administrative data. Labs are organized to help students with the data analysis required to complete the exercises, develop the term paper, and to further training in statistical software used by social science researchers.

Prerequisites

SOC 504 and SOC 505 or equivalent; students are assumed to have an elementary understanding of the basic concepts of probability and statistics. English language proficiency appropriate to a graduate university class is assumed.

Course Requirements

Computers

This is an all online course and all students are expected to have computers and internet. If you do not or have limited access please let me know and I can share some resources with you.

Readings

Weekly readings are assigned on the course syllabus. All readings are assumed to be completed before each lecture/seminar. You are expected to read over the class notes each week and make sure you are familiar with the material as the course progress – questions are encouraged.

Participation: In-class Lab

Participation will be evaluated by completing the weekly lab and submitting an html file created from RMarkdown to canvas.

Labs will be assigned every week and will be due end of day (5:00pm) Friday. The lab component will consist of computing exercises in `R` and you are expected to write it up in R Markdown.

Homework

There will be 3 homeworks for the quarter. The homeworks are meant to provide experience with both the theoretical and practical side of the course material we will cover in class. You can submit Homework as either PDFs or HTML files.

A Note: Homework and lab assignments are meant to achieve three results: (1) provide practice with the computational and statistical programing language `R`, and (2) provide practice with the statistical concepts discussed in class and (3) provide a chance to demonstrate your mastery of material and highlight areas where more work is needed. You may work in a group, but all write-ups must be done independently. All collaborators should be appropriately cited in your write up and any `R` code from a given source should also be cited as you would a journal article.

Course project

Proposal

In the fifth week of class you are expected to submit a research proposal of no more than 2 pages. This proposal should consist of the following elements: (1) introduction to the problem, (2) hypothesis to be tested, (3) methods to use, and (4) expected results. I will provide four data sets for you to choose from.

Final report

The research report should mimic a standard empirical article or technical report and will be due the last day of finals week. This should be no more than 10 pages (including references and figures). Note that **maximum** is **not a minimum**, well written 5 page papers will receive higher marks than a poorly written 10 page paper. This should follow the same basic structure as the proposal, but be fully flushed out and with careful analysis and conclusions. Your report must come with an Appendix which includes all R code used in the analysis of this project and should be written up using RMarkdown. It can be submitted in either HTML or PDF.

Grading

Participation (Lab):	25%
Homework:	40%
Project Proposal:	15%
Project:	30%

Lectures, readings, labs, and homeworks are provided for each student's benefit. It is the responsibility of the student to take advantage of these opportunities to acquire and demonstrate mastery of course material, so as to achieve his or her desired grade.

Letter grade assignment

% Points Earned	Number Grade	Letter Grade
100-97	4.0-3.9	A
96-90	3.8-3.5	A-
87-89	3.4-3.2	B+
86-84	3.1-2.9	B
83-80	2.8-2.5	B-
79-77	2.4-2.2	C+
76-74	2.1-1.9	C
73-70	1.8-1.5	C-
69-67	1.4-1.2	D+
66-64	1.1-0.9	D
63-60	0.8-0.7	D-
59-0	0	F

Required Texts

Texts

JF John Fox. 2016. *Applied Regression Analysis and Generalized Linear Models*. 3rd ed. Sage.

MT Henrik Madsen and Poul Thyregod. 2010. *Introduction to general and generalized linear models*. CRC Press.

ALR Sanford Weisberg (2005). *Applied Linear Regression*, 3rd ed. Wiley.

SW John Fox and Sanford Weisberg. 2019. *An R Companion to Applied Regression* 3rd ed. Sage.

R Useful R Texts

GG Garrett Grolemund and Hadley Wickham (2017) *R for Data Science*. https://r4ds.had.co.nz/

HW Hadley Wickham (2016). *ggplot2: Elegant graphics for data analysis* (2nd ed., Use R!). Switzerland: Springer.

* Available free online via the UW Library https://www.lib.washington.edu/.

HW2 Hadley Wickham (2019). *Advanced R*. https://adv-r.hadley.nz/

Readings

Be prepared to discuss all readings assigned at anytime in lecture/seminar.

Course Policies

Missing Class, etc.

It is expected that each member of the class will attend every lecture (and lab if taking the lab component of the course). If there is an appropriate reason to miss class it is expected that the individual will email or discuss in person with the instructor at least one week in advance. For any medical issues please see the UW website for university policies.

Cheating, etc.

All work is assumed to be your own and all individuals are expected to follow the university policy on cheating and misconduct. If you have any questions please consult the UW website for university policies and community standards.

Weekly Themes

Week 1 (03/31): Prep week

Week 2 (04/07): Review and R: Linear Models

Week 3 (04/14): Generalized Linear Models

Week 4 (04/21): Binary Data

Week 5 (04/28): Logistic and Probit Regression

Week 6 (05/05): Ordered Logit

Week 7 (05/12): Propensity Score Methods

Week 8 (05/19): Count Data Models

Week 9 (05/26): Mixed Effect Models

Week 10 (06/02): Bootstrap Methods

Week 11 (06/09): Finals Week

Course Calendar

Week 1 (03/31): Prep week

- *Readings (Text Books):*
  - SW Chapters: Preface
  - https://r4ds.had.co.nz/r-markdown.html
- *Readings (Articles):*
  -
- *Homework:*
  -

Week 2 (04/07): Review and R: Linear Models

- *Readings (Text Books):*
  - JF Chapters: 5-9,12-13
  - (Alternative JF) ALR Chapters: 1-11
  - SW Chapters: 1-4
  - GG Chapters: 9-12 and 27
- *Readings (Articles):*
  -
- *Homework:*
  - Lab 1 - Due Friday by 5:00pm
  - Homework 1 - Handed Out

Week 3 (04/14): Generalized Linear Models

- *Readings (Text Books):*
  - JF Chapters: 15.1 and 15.3.2
  - (Alternative JF) MT Chapters: 2-4
  - SW Chapters: 6.1-6.2
- *Readings (Articles):*

–

- *Homework:*
  - Lab 2 - Due Friday by 5:00p

Week 4 (04/21): Binary Data

- *Readings (Text Books):*
  - JF Chapters: 14.1
  - (Alternative JF) ALR Chapter: 12.1
  - SW Chapters: 6.3-6.4
- *Readings (Articles):*
  -
- *Homework:*
  - Lab 3 - Due Friday by 5:00pm
  - Homework 1 - Due Friday by 5:00pm
  - Meet about proposal (zoom) by Friday 5:00pm
  - Homework 2 - Handed Out

Week 5 (04/28): Logistic and Probit Regression

- *Readings (Text Books):*
  - JF Chapters: 14.1
  - (Alternative JF) ALR Chapter: 12.2
  - SW Chapters: 6.3-6.4
- *Readings (Articles):*
  -
- *Homework:*
  - Lab 4 - Due Friday by 5:00pm
  - Proposal due

Week 6 (05/05): Ordered Logit

- *Readings (Text Books):*
  - JF Chapters: 14.2
  - (Alternative JF) Almquist Notes on Ordered Logit for SOC 506
  - SW Chapters: 6.7
- *Readings (Articles):*
  -
- *Homework:*
  - Lab 5 - Due Friday by 5:00pm
  - Homework 2 - Due Friday by 5:00pm

Week 7 (05/12): Propensity Score Methods

- *Readings (Text Books):*
  -
- *Readings (Articles):*
  - Kosuke Imai and Marc Ratkovic. Covariate balancing propensity score. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 76.1 (2014): 243-263.
  - Jasjeet S.Sekhon. Multivariate and propensity score matching software with automated balance optimization: the matching package for R. Journal of Statistical Software (2008): https://www.jstatsoft.org/article/view/v042i07.
  - A useful list of software for propensity score methods: http://www.biostat.jhsph.edu/~estuart/propensityscoresoftware.html
- *Homework:*
  - Lab 6 - Due Friday by 5:00pm
  - Homework 3 - Handed Out

Week 8 (05/19): Count Data Models

- *Readings (Text Books):*
  - JF Chapters: 15.2
  - (Alternative JF) ALR Chapter: 12.3
  - SW Chapters: 6.4-6.5
- *Readings (Articles):*
  -
- *Homework:*
  - Lab 7 - Due Friday by 5:00pm

Week 9 (05/26): Mixed Effect Models

- *Readings (Text Books):*
  - JF Chapters: 23-24
  - (Alternative JF) MT Chapter: 5
  - SW Chapters: 7-8
- *Readings (Articles):*
  -
- *Homework:*
  - Lab 8 - Due Friday by 5:00pm

Week 10 (06/02): Bootstrap Methods

- *Readings (Text Books):*
  - JF Chapter: 21
  - (Alternative JF) Almquist Notes on Bootstrap for SOC 506
- *Readings (Articles):*
  -
- *Homework:*
  - Lab 9 - Due Friday by 5:00pm
  - Homework 3 - Due Friday by 5:00pm

Week 11 (06/09): Finals Week

- *Readings (Text Books):*
  - –
- *Readings (Articles):*
  - No reading
- *Homework:*
  - Final project due by Friday by 5:00pm