PL SC 503: "Multivariate Analysis for Political Research"

Spring 2017

Professor Christopher Zorn
Department of Political Science
Pennsylvania State University
E-mail: zorn@psu.edu
Tuesday - Thursday, 11:15-12:30 p.m.
Mateer Building, Room 110

Course Description

This is the second (full) course in quantitative methods in Penn State's political science Ph.D. program. The course introduces students to linear regression models for the analysis of quantitative data, and provides a basis of knowledge for more advanced statistical methods. It will also have a substantial programming/computation focus. The course assumes basic math literacy, including familiarity with probability theory, properties of estimators, rudimentary calculus, and linear algebra, as well as mastery of the basic statistics taught in PLSC 502. The bulk of the course will focus on regression models for continuous response variables, and will include discussions of the mathematical bases for such models, their estimation and interpretation, model assumptions and techniques for addressing violations of those assumptions, model diagnostics, and topics related to model specification and functional forms. We'll conclude with an introduction to the idea of maximum likelihood, including a brief overview of generalized linear models (logit, probit, etc.).

Note that all course materials (including this syllabus, slides, notes, data, computer code, homework exercises, etc.) will be available on a dedicated Github repo, which can be found at https://github.com/PrisonRodeo/PLSC503-Spring-2017-git. Throughout this syllabus, hot links are in Penn State Blue.

Texts

Required:

Weisberg, Sanford. 2013. *Applied Linear Regression*, 4th Ed. New York: Wiley. (*ALR*'s Wiley page.)

Faraway, Julian J. 2006. *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression*. London: Chapman & Hall.

Additional readings as necessary, all of which will be available via JSTOR^{TM}or on the course github repo.

The Weisberg text will be the primary text for the course. It's expensive, but a good reference, and is available in an e-book version. Faraway is for the latter part of the course, and (among other things) will also be used in PLSC 504. There is a second edition in print; either version will work for this course.

Recommended:

Kennedy, Peter. 2003. *A Guide to Econometrics*, 5th Ed. Cambridge: MIT Press. "Cliff's notes" for linear regression. There will be some readings assigned from Kennedy below, but these will be made available on the github repo.

Fox, John. 2008. *Applied Regression Analysis and Generalized Linear Models*, Second Edition. Thousand Oaks, CA: Sage Publications. Nice to have if you can get it cheaply; previous versions of this course used this as its main text.

Fox, John, and Sanford Weisberg. 2011. *An R and S-Plus Companion to Applied Regression*, Second Edition. Thousand Oaks, CA: Sage Publications. A companion to the Fox text, for S-Plus^{TM} and R users.

Nagler, Jonathan. 1996. "Coding Style and Good Computing Practices." *The Political Methodologist* 6(2):2-8. Contains words to live by.

Other Good Regression Texts:

Chatterjee, Samprit, and Ali S. Hadi. 2006. *Regression Analysis by Example*, 4th Ed. New York: Wiley.

Cohen, Jacob, Patricia Cohen, Stephen G. West, and Leona S. Aiken. 2002. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 3rd Ed. Lawrence Erlbaum.

Gelman, Andrew, and Jennifer Hill. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press.

Montgomery, Douglas C., Elizabeth A. Peck, and G. Geoffrey Vining. 2006. *Introduction to Linear Regression Analysis*, 4th Ed. New York: Wiley.

Seber, George A.F., and Alan J. Lee. 2003. *Linear Regression Analysis*, 2nd Ed. New York: Wiley.

A Few Other Useful References:

Chang, Winston. 2013. *The R Graphics Cookbook*. Sebastopol, CA: O'Reilly Media.

Crawley, Michael J. 2012. *The R Book*, 2nd Ed. New York: Wiley.

Gardener, Mark. 2012. *The Essential R Reference*. New York: Wiley.

Teetor, Paul. 2011. *The R Cookbook*. Sebastopol, CA: O'Reilly Media.

Some "Econometrics" Texts (can generally be ignored):

Dougherty, Christopher. 2007. *Introduction to Econometrics*, 3rd Ed. New York: Oxford University Press.

Greene, William. 2008. *Econometric Analysis*, 6th Ed. New York: Prentice-Hall.

Gujarati, Damodar. 2003. *Basic Econometrics*, 4th. Ed. New York: McGraw-Hill.

Hill, R. Carter, William E. Griffiths, and Guay C. Lim. 2007. *Principles of Econometrics*, 3rd Ed. New York: Wiley.

Kmenta, Jan. 1997. *Elements of Econometrics*, 2nd Ed. Ann Arbor, MI: University of Michigan Press.

Maddala, G. S. 2001. *Introduction to Econometrics*, 3rd Ed. New York: Wiley.

Stock, James S. and Mark W. Watson. 2011. *Introduction to Econometrics*, 3rd International Edition. New York: Pearson.

Wooldridge, Jeffrey. 2005. *Introductory Econometrics: A Modern Approach*, 3rd Ed. Mason, OH: South-Western College Publishing.

Most of these are generally similar to Fox (2008), though with more of an "econometric" flavor (more emphasis on proofs, less emphasis on visualization, etc.).

The Methods Preceptor

Nick Dietrich is the methods preceptor for PLSC 503. He is a third-year Ph.D. candidate, who studies international relations, human rights, and methods. He will serve as a "first line of defense" in the course: He can assist you with course material, software and programming issues, and other

matters related to the course work. He can be reached via e-mail at dietrich.nicholas [at] gmail.com.

Grading

Grading will be based on a total of 1000 points, divided as follows:

- Homework exercises: Ten worth 50 points each.
- A final paper/project, worth 500 points.

Details for the homework assignments and the final project will be announced in class.

Some Other Useful Resources

The **Inter-University Consortium for Political and Social Research** (ICPSR), at the University of Michigan, maintains an extensive archive of data in the social and behavioral sciences. Much of it is accessible via their homepage (`http://www.icpsr.umich.edu`).

The **Political Methodology Section** of the American Political Science Association was created to provide APSA members with an interest in political methodology with a forum in which to meet and discuss ideas. The section publishes a quarterly newsletter (*The Political Methodologist*), a quarterly journal on political methodology (*Political Analysis*), conducts a discussion list on topics relating to political methodology, and maintains an extensive electronic archive of papers, accessible via their homepage (at `http://polmeth.wustl.edu`).

The **Comprehensive R Archive Network** (CRAN) (`http://cran.r-project.org/`) is the place to go for downloads, packages, and documentation. Similarly, the **Stata**^{TM} homepage (`http://www.stata.com`) is a valuable resource for questions about Stata statistical software.

Obligatory Statement on Academic Dishonesty

The Department of Political Science, along with the College of the Liberal Arts and the University, takes violations of academic dishonesty seriously. Observing basic honesty in one's work, words, ideas, and actions is a principle to which all members of the community are required to subscribe.

All course work by students is to be done on an individual basis unless an instructor clearly states that an alternative is acceptable. Any reference materials used in the preparation of any assignment must be explicitly cited. Students uncertain about proper citation are responsible for checking with their instructor.

In an examination setting, unless the instructor gives explicit prior instructions to the contrary, whether the examination is in-class or take-home, violations of academic integrity shall consist but are not limited to any attempt to receive assistance from written or printed aids, or from any person

or papers or electronic devices, or of any attempt to give assistance, whether the one so doing has completed his or her own work or not.

Lying to the instructor or purposely misleading any Penn State administrator shall also constitute a violation of academic integrity.

In cases of any violation of academic integrity it is the policy of the Department of Political Science to follow procedures established by the College of the Liberal Arts. More information on academic integrity and procedures followed for violation can be found here.

Disabilities

The Pennsylvania State University encourages qualified people with disabilities to participate in its programs and activities and is committed to the policy that all people shall have equal access to programs, facilities, and admissions without regard to personal characteristics not related to ability, performance, or qualifications as determined by University policy or by state or federal authorities. If you anticipate needing any type of accommodation in this course or have questions about physical access, please tell the instructor as soon as possible. Reasonable accommodations will be made for all students with disabilities, but it is the student's responsibility to inform the instructor early in the term. Do not wait until just before an exam to decide you want to inform the instructor of a learning disability; any accommodations for disabilities must be arranged well in advance.

Course Schedule

Linear Regression: Basics

- **January 10**: *Course Introduction*
  Readings (for background):
  - Preface to the 4th Ed. of Weisberg.
  - Berk, Richard. 2010. "What You Can and Can't Properly Do with Regression." *Journal of Quantitative Criminology* 26(4):481-487.
- **January 12**: *No Class.*
- **January 17**: *Regression: A Conceptual Overview*
  Readings:
  - Weisberg, Chapter 1 and Appendix A.1 and A.2.

- **January 19**: *Bivariate Regression: A (Re)Introduction*
  Readings:
  - Weisberg, Chapter 2, pp. 21-30 and Appendix A.3.

- **January 24**: *Bivariate Regression: Inference*
  Readings:
  - Weisberg, Chapter 2, pp. 30-38 and Appendix A.4.
- **January 26**: *Bivariate Regression: Model Fit*
  Readings:
  - Weisberg, Chapter 2, pp. 30-38 and Appendix A.4.
  - Lewis-Beck, Michael S. and Andrew Skalaban. 1990. "When to Use R-Squared." *The Political Methodologist* 3(2):11-12.
  - King, Gary. 1990. "When Not to Use R-Squared." *The Political Methodologist* 3(2):9-11.
  - Luskin, Robert C. 1991. "R-Squared Encore." *The Political Methodologist* 4(1):21-23.

  Homework One due.

- **January 31**: *Stupid Regression Tricks*
  Readings:
  - No readings assigned.
- **February 2**: *Special Topic: Bootstrapping And Other Delights*
  Readings:
  - No readings assigned.

  Homework Two due.

Multivariate Linear Regression

- **February 7**: *No Class.*

- **February 9**: *Multivariate Regression: Estimation*
  Readings:
  - Weisberg, Chapter 3, pp. 51-68 and Appendix A.8.

  Homework Three due.

- **February 14**: *Multivariate Regression: Inference*
  Readings:
  - Weisberg, Chapter 6, pp. 133-150.
- **February 16**: *Multivariate Regression: Dichotomous Covariates*
  Readings:
  - Weisberg, Chapter 5, pp. 98-123.

  Homework Four due.

- **February 21**: *(Non-)Linearity and Data Transformations*
  Readings:
  - Weisberg, Chapter 4, pp. 67-93; Chapter 8, pp. 185-199.
- **February 23**: *Variance Issues*
  Readings:
  - Weisberg, Chapter 7, pp. 156-179.

  Homework Five due.

- **February 28**: *Multivariate Regression: Collinearity, etc.*
  Readings:
  - Kennedy, Chapter 11, pp. 205-217.

Homework Four due.

- **March 2**: *Multivariate Regression: Residuals, Outliers, and Diagnostics*
  Readings:
  - Weisberg, Chapter 9, pp. 204-226.

- **March 7**: **No Class – Spring Break**
- **March 9**: **No Class – Spring Break**

- **March 14**: **No Class.**

- **March 16**: *Variable Selection*
  Readings:
  - Weisberg, Chapter 10, pp. 234-248.

  Homework Six due.

- **March 21**: *Specification Error, Random Regressors, and Simultaneity*
  Readings:
  - Kennedy, pp. 107-109; 180-191.
- **March 23**: *Multiplicative Interactions*
  Readings:
  - Friedrich, Robert J. 1982. "In Defense of Multiplicative Terms in Multiple Regression Equations." *American Journal of Political Science* 26(November):797-833.
  - Brambor, Thomas, William R. Clark, and Matt Golder. 2006. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14:63-82.

  Homework Seven due.

Beyond Linearity

- **March 28**: *Maximum Likelihood - Introduction*
  Readings:
  - Fox, Appendix D6, pp. 92-95.
  - Weisberg, Appendix A.11.
- **March 30**: *MLE: Estimation / Optimization*
  Readings:
  - No readings (but see above for some suggestions).

- **April 4**: *MLE: Inference and "Robust" Variance Estimators*
  Readings:
  - Buse, A. 1982. "The Likelihood Ratio, Wald, and Lagrange Multiplier Tests: An Expository Note." *The American Statistician* 36(3):153-57.
  - Freedman, D. A. 2006. "On the So-Called 'Huber Sandwich Estimator' and 'Robust' Standard Errors." *The American Statistician* 60:299-302.
  - King, Gary, and Margaret E. Roberts. 2014. "How Robust Standard Errors Expose Methodological Problems They Do Not Fix, and What To Do About It." *Political Analysis* 22:1-21.
- **April 6**: *Generalized Linear Models*
  Readings:
  - Weisberg, Chapter 12, pp. 279-285.
  - Gill, Jeff. 2000. *Generalized Linear Models: A Unified Approach*. Thousand Oaks, CA: Sage. pp. 51-61.

  Homework Eight due.

- **April 11**: *Binary Responses, I.*
  Readings:
  - Weisberg, Chapter 12, pp. 270-279.
  - Faraway, pp. 25-38.

- **April 13**: *Binary Responses, II*.
  Readings:
    - No readings assigned.

- **April 18**: *Nominal Responses*.
  Readings:
    - Faraway, pp. 97-103.
- **April 20**: *Ordinal Responses*.
  Readings:
    - Faraway, pp. 106-112.

  Homework Nine due.

- **April 25**: *Event Counts*.
  Readings:
    - Faraway, pp. 55-66.
- **April 27**: *Catch-Up, wrap-up, and review*.
  Readings:
    - No readings assigned.

  Homework Ten due.

- **May 3**: **Final Papers Due**.