

## Political Science 799: Multivariate Analysis

Fall 2022

Tuesday 4–6 (2325 MH)

Professor: Walter R. Mebane, Jr.

Office: 7735 Haven Hall (607/592-0546); email [wmebane@umich.edu](mailto:wmebane@umich.edu)

Office hours: Tue 1–3 or other times by appointment.

GSI: Fabricio Vasselai; 7740 Haven Hall; email [vasselai@umich.edu](mailto:vasselai@umich.edu)

GSI Office hours: Wed 2:30-4:30 or other times by appointment.

Course web page: in Canvass; syllabus also at <http://www.umich.edu/~wmebane/ps799.html>

### Assignment Due Dates

due date	description	weight
TBA	problem sets	60%
Dec 15, 1:30pm–3:30pm	final paper presentations	10%
Dec 16	final paper	20%
—	participation	10%

See `fpaper.pdf` posted on the Canvas site for more information about the final paper and presentation.

### Reading Availability

Much of the course will refer to journal articles. I plan to follow or refer to a few chapters in the following books (others also appear below).

Cameron, A. Colin, and Pravin K. Trivedi. 2005. *Microeconometrics: Methods and Applications*. Cambridge UP.

Kenneth E. Train. 2003. *Discrete Choice Methods with Simulation*. Cambridge UP.

In the following listing, required reading is preceded by a bullet. Other items are recommended.

### Class meeting and reading schedule

#### 1. computing (Aug 30)

- John Chambers. 1999. Computing with Data: Concepts and Challenges. *The American Statistician* 53: (1, Feb.): 73–84.
- Bierlaire, M. 2018. BIOGEME: A free package for the estimation of discrete choice models, <http://biogeme.epfl.ch/>
- OpenBUGS. <http://www.openbugs.net/w/FrontPage>
- WinBUGS. <https://www.mrc-bsu.cam.ac.uk/software/bugs/the-bugs-project-winbugs/>
- Bob Carpenter, Andrew Gelman, Matthew D. Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. Stan: A probabilistic programming language. *Journal of Statistical Software* 76(1). DOI 10.18637/jss.v076.i01 (in file `v76i01.pdf`)

The Comprehensive R Archive Network. 2019. <https://cran.r-project.org/>

- Crawley, Michael R. 2007. *The R Book*. Wiley.
- Spector, Phil. 2008. *Data manipulation with R*. Springer.
- Albert, Jim. 2007. *Bayesian Computation with R*. Springer.
- Chambers, John M. 2008. *Software for Data Analysis*. Springer.
- Braun, W. John, and Duncan J. Murdoch. 2007. *A First Course in Statistical Programming with R*. Cambridge.
- Bierlaire, M. (2018). PandasBiogeme: a short introduction. Technical report TRANSP-OR 181219. Transport and Mobility Laboratory, ENAC, EPFL.  
<http://transp-or.epfl.ch/documents/technicalReports/Bier18.pdf>
- Plummer, Martyn. 2003. “JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling,” Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003), March 20–22, Vienna, Austria. ISSN 1609-395X.  
<http://www.ci.tuwien.ac.at/Conferences/DSC-2003/Proceedings/Plummer.pdf>
- JAGS. <http://mcmc-jags.sourceforge.net/>
- Pemstein, Daniel, Kevin M. Quinn and Andrew D. Martin. 2011. “The Scythe Statistical Library: An Open Source C++ Library for Statistical Computation.” *Journal of Statistical Software*. 42. DOI 10.18637/jss.v042.i12
- Martin, Andrew D., Kevin M. Quinn, and Jong Hee Park. 2011. “MCMCpack: Markov Chain Monte Carlo in R.” *Journal of Statistical Software*. 42. DOI 10.18637/jss.v042.i09

## 2. maximum likelihood and numerical optimization (Sep 6)

- A. Colin Cameron and Pravin K. Trivedi. 2005. *Microeconometrics: Methods and Applications*. Cambridge. Chapter 5. (in file [CT5.pdf](#))
- Walter R. Mebane, Jr., and Jasjeet Sekhon. 2011. “Genetic Optimization Using Derivatives: The rgenoud Package for **R**.” *Journal of Statistical Software* 42(11). <https://www.jstatsoft.org/article/view/v042i11>
- Walter R. Mebane, Jr. 1999. “Congressional Campaign Contributions, District Service and Electoral Outcomes in the United States: Statistical Tests of a Formal Game Model with Nonlinear Dynamics” in Diana Richards, ed., *Political Complexity: Nonlinear Models of Politics*. Ann Arbor: University of Michigan Press. (in file [ghopf.pdf](#))
- Mebane, Walter R., Jr. 2000. “Coordination, Moderation, and Institutional Balancing in American Presidential and House Elections.” *American Political Science Review* 94 (March): 37–57. MS version available at <http://www.umich.edu/~wmebane/covote.pdf>

Philip E. Gill , Walter Murray and Margaret H. Wright. 1982. *Practical Optimization*. Classics in Applied Mathematics edition (2019). <https://doi.org/10.1137/1.9781611975604>

Stefan Theussl, Florian Schwendinger, Hans W. Borchers. 2021. “CRAN Task View: Optimization and Mathematical Programming.” <https://cran.r-project.org/web/views/Optimization.html>

### 3. generalized linear models and QMLE (Sep 13)

- John E. Jackson. 1983. “Election Night Reporting and Voter Turnout.” *American Journal of Political Science* 27(4): 615–635.
- McCullagh, Peter. 1983. Quasi-likelihood Functions. *Annals of Statistics* 11 (Mar.): 59–67.
- Leonard A. Stefanski and Dennis D. Boos, 2002. The Calculus of M-Estimation. *The American Statistician* 56 (1, Feb.): 29–38.
- White, Halbert. 1982. Maximum Likelihood Estimation of Misspecified Models. *Econometrica* 50 (1): 1–25.
- D. A. Freedman. 2006. On the so-called “Huber Sandwich Estimator” and “robust” standard errors. *The American Statistician* 60: 299–302.

Peter McCullagh and John A. Nelder. 1989. *Generalized Linear Models*. 2d ed. Chapman and Hall.

O. E. Barndorff-Nielsen. 1995. “Quasi Profile and Directed Likelihoods From Estimating Functions.” *Ann. Inst. Statist. Math.* 47(3), 461–464. [https://www.ism.ac.jp/editsec/aism/pdf/047\\_3\\_0461.pdf](https://www.ism.ac.jp/editsec/aism/pdf/047_3_0461.pdf)

S. A. Murphy and A. W. van der Vaart. 2000. “On Profile Likelihood.” *Journal of the American Statistical Association* 95 (450 Jun), 449–465. (in file `Murphy.Vaart.JASA2000.2669386.pdf`)

### 4. asymptotics, bootstrap and refinements (Sep 20)

- Bradley Efron. 1987. “Better Bootstrap Confidence Intervals (and discussion)” *Journal of the American Statistical Association* 82 (397): 171–200. (in file `efron.jasa1987.pdf` and nine other similarly named files)
- Gary W. Oehlert. 1992. A Note on the Delta Method. *The American Statistician* 46 (1, Feb.): 27–29. (in file `oehlert.amstat1992.pdf`)
- Thomas R. Fears, Jacques Benichou, Mitchell H. Gail. 1996. A Reminder of the Fallibility of the Wald Statistic. *The American Statistician* 50 (3, Aug.): 226–227. (in file `fears.benichou.gail.amstat1996.pdf`)
- Yudi Pawitan. 2000. A Reminder of the Fallibility of the Wald Statistic: Likelihood Explanation. *The American Statistician* 54 (1, Feb.): 54–56. (in file `pawitan.amstat2000.pdf`)
- Dennis D. Boos and Jacqueline M. Hughes–Oliver. 2000. How Large Does  $n$  Have to be for  $Z$  and  $t$  Intervals? *The American Statistician* 54 (2, May): 121–128. (in file `boos.hughesoliver.amstat2000.pdf`)
- J. Scott Long and Laurie H. Ervin. 2000. Using Heteroscedasticity Consistent Standard Errors in the Linear Regression Model. *The American Statistician* 54 (3, Aug.): 217–224. (in file `long.ervin.amstat2000.pdf`)
- John E. Jackson. 2019. “Corrected Standard Errors with Clustered Data.” *Political Analysis*, 28 (3, July): 318–339. (in file `cese_PA_final.pdf`)

Bradley Efron. 1979. “Bootstrap Methods: Another Look at the Jackknife,” *Annals of Statistics* 7 (1): 1–26. (in file `efron.aos1979.pdf`)

Barndorff-Nielsen, O. E., and D. R. Cox. 1984. “Bartlett Adjustments to the Likelihood Ratio Statistic and the Distribution of the Maximum Likelihood Estimator,” *Journal of the Royal Statistical Society. Series B (Methodological)* 46 (3): 483–495.

A.C. Davison and D.V. Hinkley. 1997. *Bootstrap Methods and their Applications*. Cambridge.

Steven J. Sepanski. 1994. “Asymptotic for Multivariate  $t$ -Statistic and Hotelling’s  $T^2$ -Statistic under Infinite Second Moments via Bootstrapping,” *Journal of Multivariate Analysis* 49 (1): 41–54.

Cameron and Trivedi. Chapters 5, 11 and Appendix A.

Jeffrey M. Wooldridge. 2002. *Econometric Analysis of Cross Section and Panel Data*. MIT Press. Chapters 3, 12–14.

Russell Davidson and James G. MacKinnon. 1993. *Estimation and Inference in Econometrics*. Oxford UP. Chapters 4, 8–9.

Diogo Ferrari and John E. Jackson. 2019. “ceser R Package: Cluster Estimated Standard Error in R.” *Journal of Statistical Software*. (in file `ceser.pdf`)

## 5. text as data (Sep 27)

- Burt L. Monroe, Michael P. Colaresi and Kevin M. Quinn. 2008. “Fightin’ Words: Lexical Feature Selection and Evaluation for Identifying the Content of Political Conflict.” *Political Analysis* 16: 372–403. (in file `fightin_words_lexical_feature_selection_and_.pdf`)
- Justin Grimmer. 2010. A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases. *Political Analysis* 18 (No. 1, Winter): 1–35.  
(in file `bayesian_hierarchical_topic_model_for_political_texts_measuring_expressed_agendas_in_senate_press_releases.pdf`)
- Justin Grimmer and Brandon M. Stewart. 2013. “Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts.” *Political Analysis* 21: 267–297. (in file `text_as_data_the_promise_and_pitfalls_.pdf`)
- Margaret E. Roberts, Brandon M. Stewart, Dustin Tingley and Edoardo M. Airoldi. 2013. “The Structural Topic Model and Applied Social Science.” NIPS 2013 Workshop on Topic Models: Computation, Application, and Evaluation. (in file `stm_nips2013.pdf`)
- Patrick Y. Wu, Walter R. Mebane, Jr., Joseph Klaver, Logan Woods and Preston Due. 2019. “Partisan Associations of Twitter Users Based on Their Self-descriptions and Word Embeddings.” Presented at APSA 2019. (updated version in file `wepa.pdf`)

Matthew J. Denny and Arthur Spirling. 2018. “Text Preprocessing For Unsupervised Learning: Why It Matters, When It Misleads, And What To Do About It.” *Political Analysis* 26: 168–189. (in file `text_preprocessing_for_unsupervised_learning.pdf`)

David M. Blei, Andrew Y. Ng and Michael I. Jordan. 2003. Latent Dirichlet allocation. *Journal of Machine Learning Research* 3 (Jan.): 993–1022.  
<http://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf>

David M. Blei, Jon D. McAuliffe. 2007. Supervised topic models. *Neural Information Processing Systems* 21.  
<https://papers.nips.cc/paper/3328-supervised-topic-models.pdf>

Walter R. Mebane, Jr., Patrick Y. Wu, Logan Woods, Alejandro Pineda, Blake Miller, Joseph Klaver, Preston Due and Adam Rauh. 2020. “Diverse Election Experiences Reported without Bias: Observing Election Incidents in the United States via Twitter.” (in file `TEO.pdf`)

6. prediction, machine learning, LASSO, regularization (Oct 4, 11)

- Trevor Hastie, Robert Tibshirani and Jerome Friedman. 2009. *The Elements of Statistical Learning*. Springer. Chapters 2, 5, 7 and 11. (in files `ESL2.pdf`, `ESL5.pdf`, `ESL7.pdf` and `ESL11.pdf`)
- Trevor Hastie, Robert Tibshirani and Martin Wainwright. 2015. *Statistical Learning with Sparsity: The Lasso and Generalizations*. CRC Press. Chapters 2, 3 and 11. (in files `SLS2.pdf`, `SLS3.pdf` and `SLS11.pdf`)
- Blake Miller, Fridolin Linder and Walter R. Mebane, Jr. 2020. “Active Learning Approaches for Labeling Text: Review and Assessment of the Performance of Active Learning Approaches.” *Political Analysis*, 28 (4, October): 532–551. DOI: <https://doi.org/10.1017/pan.2020.4> (in file `active_learning_approaches_for_labeling_text_review_and_assessment_of_the_performance_of_`

7. choice models (Oct 25)

- Daniel McFadden and Kenneth Train. 2000. Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics* 15 (5, Sep–Oct): 447–470. (in file `mcfadden.train.japplecon2000.pdf`)
- Kenneth E. Train. 2009. *Discrete Choice Methods with Simulation*. 2d ed. Cambridge UP. <http://elsa.berkeley.edu/books/choice2.html>
- John E. Jackson, Bogdan W. Mach and Radoslaw Markowski. 2010. “Party Strategies and Electoral Competition in Post-Communist Countries: Evidence from Poland.” *Electoral Studies* 29 (2): 199–209. (in file `jackson.mach.markowski.elecstud2010.pdf`)
- John E. Jackson, Bogdan W. Mach and Radoslaw Markowski. 2010. “Party Strategies and Electoral Competition in Post-Communist Countries: Evidence from Poland. Appendix A: Methodological Appendix.” (in file `jelsmethapp.docx`)
- Walter R. Mebane, Jr., John E. Jackson and Jonathan Wall. 2015. “Choice Function Heterogeneities in Models of Electoral Behavior.” Working paper (in file `mw14.pdf`).
- Garrett Glasgow. 2001. Mixed Logit Models for Multiparty Elections. *Political Analysis* 9 (1): 116–136. (in file `glasgow.pa2001.pdf`)

Hensher, David A., and William H. Greene. 2003. The mixed logit model: the state of practice. *Transportation* 30 (2): 133–176.

McFadden, Daniel. 1974. “Conditional logit analysis of qualitative choice behavior.” In P Zarembka, ed., *Frontiers of Econometrics*, New York: Academic Press. pages 105–142. <http://emlab.berkeley.edu/reprints/mcfadden/zarembka.pdf>

McFadden, Daniel. 1981. “Structural Discrete Probability Models Derived from Theories of Choice.” In Charles F. Manski and Daniel L. McFadden, eds, *Structural Analysis of Discrete Data and Econometric Applications*, Cambridge, MA: MIT Press, chapter 5, pp. 198–272. <http://emlab.berkeley.edu/discrete/ch5.pdf>

Mauricio Sarrias and Ricardo A. Daziano. 2017. “Multinomial Logit Models with Continuous and Discrete Individual Heterogeneity in **R**: The `gmn1` Package.” *Journal of Statistical Software* 79 (2). doi: 10.18637/jss.v079.i02

8. observational studies, RD and causal inference (Nov 1)

- Angrist, Joshua D., Guido Imbens and Donald B. Rubin. 1996. “Identification of Causal Effects Using Instrumental Variables.” *Journal of the American Statistical Association* 91(June): 444–455. (in file `angrist.imbens.rubin.jasa1996.pdf`)
- Holland, Paul. 1986, “Statistics and Causal Inference.” *Journal of the American Statistical Association* 81: 945–961. (in file `holland.jasa1986.pdf`)
- Lee, D. S. 2008. “Randomized Experiments from Non-random Selection in U.S. House Elections.” *Journal of Econometrics* 142:675-697. (in file `lee.jeconometrics2008.pdf`)
- Jasjeet Sekhon and Rocio Titiunik. 2017. “On Interpreting the Regression Discontinuity Design as a Local Experiment.” *Regression Discontinuity Designs (Advances in Econometrics, Vol. 38)*. Emerald Publishing Limited, pp. 1-28. <https://doi.org/10.1108/S0731-905320170000038001> . (in file `SekhonTitiunik-RD-2016.pdf`)
- Sebastian Calonico, Matias D. Cattaneo and Rocio Titiunik. 2014. “Robust Non-parametric Confidence Intervals for Regression-Discontinuity Designs.” *Econometrica* 82(6):2295–2326. (in file `Calonico-Cattaneo-Titiunik_2014_ECMA.pdf`)
- Sebastian Calonico, Matias D. Cattaneo and Rocio Titiunik. 2015. “Optimal Data-Driven Regression Discontinuity Plots.” *Journal of the American Statistical Association* 110(512):1753–1769. (in file `Calonico-Cattaneo-Titiunik_2016_JASA.pdf`)

Paul R. Rosenbaum. 2002. *Observational Studies*. Springer.

Paul R. Rosenbaum. 2009. *Design of Observational Studies*. Springer.

Judea Pearl. 2009. *Causality: Models, Reasoning, and Inference*, 2d ed. Cambridge.

Joshua D. Angrist and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics*. Princeton.

Dunning, Thad. 2012. *Natural Experiments in the Social Sciences: A Design-Based Approach (Strategies for Social Inquiry)*. Cambridge.

Gerber, Alan S., and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. Norton.

9. causal identification norms, DAGs, interference (Nov 8)

- Thomas J. Rothenberg. 1971. Identification in Parametric Models. *Econometrica* 39 (May): 577–591.
- Judea Pearl. 1995. Causal Diagrams for Empirical Research. *Biometrika* 82 (4, Dec.): 669–688. (plus discussion, 688–710).
- Egami, Naoki and Imai, Kosuke, 2019. “Causal interaction in factorial experiments: Application to conjoint analysis.” *Journal of the American Statistical Association* 114 (526), 529-540.
- Aronow, Peter M. and Samii, Cyrus, 2017. “Estimating average causal effects under general interference, with application to a social network experiment.” *The Annals of Applied Statistics* 11 (4), 1912-1947.

- Hainmueller, Jens, Hopkins, Daniel J. and Yamamoto, Teppei, 2014. “Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments.” *Political Analysis* 22 (1), 1-30.
- Graeme Blair, Jasper Cooper, Alexander Coppock and Macartan Humphreys. 2019. “Declaring and Diagnosing Research Designs.” *American Political Science Review* 113 (3): 838–859.

Roger Bowden. 1973. The Theory of Parametric Identification. *Econometrica* 41 (Nov): 1069–1074.

Franklin Fisher. 1976. *The Identification Problem in Econometrics*. Krieger.

Roger Bowden and Darrell Turlington. 1984. *Instrumental Variables*. Cambridge UP.

Judea Pearl. 2009. *Causality: Models, Reasoning and Inference*, 2d ed. Cambridge UP. Chapters 1–5.

James M. Robins. 1999. Association, Causation, and Marginal Structural Models. *Synthese* 121: 151–179.

David A. Freedman and Jasjeet S. Sekhon. 2010. Endogeneity in Probit Response Models. *Political Analysis* 18 (2): 138–150.

James J. Heckman and Edward Vytlacil. 2005. Structural Equations, Treatment Effects, and Econometric Policy Evaluation. *Econometrica* 73 (3, May): 669–738.

Heckman, J. J. 1978. Dummy endogenous variables in a simultaneous equation system. *Econometrica* 46: 931–959.

Heckman, J. J. 1979. Sample selection bias as a specification error. *Econometrica* 47: 153–161.

#### 10. hierarchical models, MCMC (Nov 15)

- Simon Jackman. 2009. *Bayesian Analysis for the Social Sciences*. Wiley. Chapter 7.
- Andrew Gelman, John B. Carlin, Hal S. Stern and Donald B. Rubin. 2004. *Bayesian Data Analysis*, 2d ed. Chapman & Hall. Chapter 5. (Chapters 1–4 are probably necessary preparation.)
- Andrew Gelman and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge. Pages 109–117 251–265, 345–359, 366–371, 419–421.
- Brooks, S. P. 1998. Markov chain Monte Carlo method and its application. *The Statistician* 47: 69–100. (in file `brooks.statistician1998.pdf`)
- Brooks, S. P. and A. Gelman. 1998. Alternative methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* 7: 434–455. (in file `brooks.gelman.jcgs1998.pdf`)
- Spiegelhalter, D. J., N. G. Best, B. P. Carlin and A. van der Linde. 2002. Bayesian measures of model complexity and fit (with discussion). *J. Roy. Statist. Soc. B* 64: 583–640. (in file `spiegelhalter.jrssb2002.pdf`)

George Casella and Edward I. George. 1992. Explaining the Gibbs Sampler *The American Statistician* 46 (3, Aug.): 167–174.

Siddhartha Chib and Edward Greenberg. 1995. Understanding the Metropolis–Hastings Algorithm *The American Statistician* 49 (4, Nov.): 327–335.

Jeff Gill. 2002. *Bayesian Methods: A Social and Behavioral Approach*. Chapman & Hall.

11. latent variable models (Nov 22, 29)

- Simon Jackman and Shawn Treier. 2008. Democracy as a Latent Variable. *American Journal of Political Science* 52 (1): 201–217. (in file `jackman.treir.ajps2008.pdf`)
- Joshua Clinton, Simon Jackman and Douglas Rivers. 2004. The Statistical Analysis of Roll Call Data. *American Political Science Review* 98 (2, May): 355–370. (in file `clinton.jackman.rivers.apsr2004.pdf`)
- Kevin McAlister. 2020. Chapter II of “Essays on Latent Variable Models and Roll Call Scaling” (2020 Ph.D. dissertation) (in file `KevinMcAlisterUMDissFinal.pdf`)
- Quinn, Kevin M. 2004. “Bayesian Factor Analysis for Mixed Ordinal and Continuous Responses.” *Political Analysis* 12: 338–353. (in file `quinn.pa2004.pdf`)
- Karl G. Jöreskog. 1974. “Analyzing Psychological Data by Structural Analysis of Covariance Matrices.” In David H. Krantz, Richard C. Atkinson, R. Duncan Luce and Patrick Suppes, *Contemporary Developments in Mathematical Psychology*, Vol. II. W. H. Freeman and Company. (in file `Joreskog1974c.pdf`)
- Fariss, Christopher J. 2014. “Respect for Human Rights has Improved Over Time: Modeling the Changing Standard of Accountability.” *American Political Science Review* 108 (2, May): 297–318. (in file `Fariss2014APSR.pdf`)
- Walter R. Mebane, Jr., Diogo Ferrari, Kevin McAlister, and Patrick Y. Wu. 2021. “Measuring Elections Frauds.” (in file `measfrauds.pdf`)

Jian-Qing Shi and Sik-Yum Lee. 2000. Latent Variable Models with Mixed Continuous and Polytomous Data. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 62 (1): 77–87. (in file `shi.lee.jrssb2000.pdf`)

Michael A. Bailey. 2007. Comparable Preference Estimates across Time and Institutions for the Court, Congress, and Presidency. *American Journal of Political Science* 51 (3, Jul.): 433–448.

Sik-Yum Lee. 2007. *Structural Equation Modelling: A Bayesian Approach*. Wiley.

Sik-Yum Lee, Xin-Yuan Song, John C. K. Lee. 2003. Maximum Likelihood Estimation of Nonlinear Structural Equation Models with Ignorable Missing Data. *Journal of Educational and Behavioral Statistics* 28 (Summer): 111–134.

Sik-Yum Lee and Xin-Yuan Song. 2004. Maximum Likelihood Analysis of a General Latent Variable Model with Hierarchically Mixed Data. *Biometrics* 60 (Sep.): 624–636.

Sophia Rabe-Hesketh, Anders Skrondal and Andrew Pickles. 2004. *Generalized Latent Variable Modelling: Multilevel, Longitudinal and Structural Equation Models*. Chapman & Hall.

12. hypothesis tests and model selection (Dec 6)

- Cameron and Trivedi. Chapter 7. (in file `CT7.pdf`)

- Benjamin, Daniel J, James O. Berger, et al. 2018. Redefine Statistical Significance. *Nature Human Behavior* 2, 6–10. <https://doi.org/10.1038/s41562-017-0189-z> (in file BenjaminEtAlRedefineStatisticalSignificance.pdf)
- Benjamini, Yoav and Yocef Hochberg. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society, Series B* 57 (1): 289–300.
- Benjamini, Yoav and Daniel Yekutieli. 2005. False Discovery Rate-Adjusted Multiple Confidence Intervals for Selected Parameters. *Journal of the American Statistical Association* 100 (Mar.): 71–81.
- Vuong, Quang H. 1989. “Likelihood-ratio Tests for Model Selection and Non-nested Hypotheses.” *Econometrica* 57 (2): 307–333.
- Imai, Kosuke and Tingley, Dustin, 2012. “A statistical method for empirical testing of competing theories.” *American Journal of Political Science* 56 (1) 218–236.

Kass, Robert E., and Adrian E. Raftery. 1995. “Bayes factors” *Journal of the American Statistical Association* 90 (430) : 773–795.

Chib, S., 2001. “Markov chain Monte Carlo methods: computation and inference.” In *Handbook of Econometrics* (Vol. 5, pp. 3569–3649). Elsevier.  
Section 10: MCMC methods in model choice problems. (in file `chib2001.pdf`)

Gelman, A. and Meng, X.L., 1998. “Simulating normalizing constants: From importance sampling to bridge sampling to path sampling.” *Statistical Science* 163–185.

### 13. partial identification and identification with missing covariates (Dec 6)

- Kei Kawai and Yasutora Watanabe. 2013. “Inferring Strategic Voting.” *American Economic Review* 103 (2): 624–662. (in file `kawai2013inferring.pdf`)
- Joel L. Horowitz and Charles F. Manski. 2000. Nonparametric Analysis of Randomized Experiments with Missing Covariate and Outcome Data. *Journal of the American Statistical Association* 95 (449, Mar): 77–84.
- Rosa L. Matzkin. 2007. Nonparametric Survey Response Errors. *International Economic Review* 48 (4): 1411–1427.
- Charles F. Manski. 1990. Nonparametric Bounds on Treatment Effects. *American Economic Review* 80 (2, Papers and Proceedings): 319–323.
- Francesca Molinari. 2010. Missing Treatments. *Journal of Business and Economic Statistics* 28 (1): 82–95.
- Walter R. Mebane, Jr. and Paul Poast. 2013. “Causal Inference without Ignorability: Identification with Nonrandom Assignment and Missing Treatment Data.” *Political Analysis* 21 (2): 233–251.

Charles F. Manski. 1995. *Identification in the Social Sciences*. Harvard UP.

Charles F. Manski. 2003. *Partial Identification of Probability Distributions*. Springer.

Rosa L. Matzkin. 2007. Nonparametric Identification. In James J. Heckman and Edward E. Leamer, eds., *Handbook of Econometrics* volume 6B. North-Holland. Pp. 5307–5368.

Charles F. Manski and Elie Tamer. 2002. Inference on Regressions with Interval Data on a Regressor or Outcome. *Econometrica* 70 (2, Mar): 519–546.

14. bounded influence estimation (Dec 6)

- Stefanski, Leonard A., Raymond J. Carroll, David Ruppert. 1986. Optimally Bounded Score Functions for Generalized Linear Models with Applications to Logistic Regression. *Biometrika* 73 (Aug): 413–424.
- Western, Bruce. 1995. Concepts and Suggestions for Robust Regression Analysis. *American Journal of Political Science* 39 (3): 786–817.
- Mebane, Walter R., Jr., and Jasjeet S. Sekhon. 2004. Robust Estimation and Outlier Detection for Overdispersed Multinomial Models of Count Data. *American Journal of Political Science* 48 (April): 392–411.
- Mebane, Walter R., Jr. 2010. Fraud in the 2009 Presidential Election in Iran? *Chance* 23 (Mar.): 6–15.

Hampel, Frank R. and Peter J. Rousseeuw and Elvezio Ronchetti. 1981. The Change-of-Variance Curve and Optimal Redescending M-Estimators. *Journal of the American Statistical Association* 76 (Sep): 643–648.

Croux, Christophe and Peter J. Rousseeuw and Ola Hossjer. 1994. Generalized S-Estimators. *Journal of the American Statistical Association* 89 (Dec): 1271–1281.

15. paper presentations (Dec 14, 1:30pm–3:30pm)