

# Applied Statistics 560: Visiting Scholar Statistical Research Seminar Spring 2009

Meeting Time: W 1430–1630, Siegle Hall L004

Visiting Statistician:	Ben B. Hansen, Ph. D.	Local:	Robert W. Walker, Ph. D.
Office:			Siegle Hall L085C
Phone:			(314) 935-5881
Email:	bbh@umich.edu		rww@wustl.edu
WWW:	www.stat.lsa.umich.edu/~bbh		rhubarb.wustl.edu

## 1 Opening sessions

- “Attributing Effects to A Cluster Randomized Get-Out-The-Vote Campaign” (Hansen and Bowers, 2009). (*To be presented in the Jan. 15 lecture*, which will not assume the audience has read it. It’s available at <http://www.jakebowers.org/PAPERS/bowershansen2008JASA.pdf>.)
- “The (mis)estimation of neighborhood effects: causal inference for a practicable social epidemiology” Oakes (2004), particularly pp. 1929–34, 38–9. (*To be discussed in the Jan. 16 meeting*. Oakes’s influential recent critique is similar in spirit to other critiques of statistical modeling for causal inference that have been offered over the years. The purpose of my presentation in this meeting will be outline a collection of ideas and program of research that responds to criticisms like his. It would be helpful, if not necessary, for participants to have read in advance the indicated sections of the paper. )

## 2 Foundations

### 2.1 The critique of causal inference with regression

There have been many critiques of causal inference using statistical models. Freedman’s are crisply presented and technically impeccable, and they well articulate some important and centrally recurring arguments:

- “Statistical Models and Shoe Leather” (Freedman, 1991);
- “Graphical Models for Causation, and the Identification Problem” (Freedman, 2004, particularly pp.267–9).

Supplemental reading: For a different kind of critique, an evaluation the performance of regression models for causal inference, see “Evaluating the Econometric Evaluations of Training Programs with Experimental Data” (LaLonde, 1986).

*To be discussed at the January 21 meeting.*

### 2.2 Alternative views of the role of statistics

- “Choice as an Alternative to Control In Observational Studies” (Rosenbaum, 1999, esp. § 1–3, pp.259–73; p.301);
- Supplemental: “The Tyranny of Continuous Models in a World of Discrete Data” (Holland, 1979).

Although not conceived or presented as a defense of causal inference with observational studies, Rosenbaum’s “Choice as an Alternative to Control In Observational Studies” can be read in that way: its picture of the role of statistics in science differs subtly but importantly from what critics like Freedman and Oakes have in mind, and supports statistical approaches that better withstand the critics’ attacks. Holland’s paper nicely articulates a view alternate to Freedman’s of the role of regression specifications, which may be of interest to those who are bothered by Freedman’s critique but hesitant to reduce their use of regression models.

*To be discussed at the January 28 meeting.*

## 2.3 The potential outcomes approach to causal inference

- “Statistics and Causal Inference” (Holland, 1986b);
- Cox’s and Rubin’s comments on “Statistics and Causal Inference,” and Holland’s rejoinder (Rubin, 1986; Cox, 1986; Holland, 1986a).

Rubin, Holland, Neyman and most others who’ve written about potential outcomes and Statistics add other bits of conceptual structure to the potential outcomes framework. Holland’s (1986b) paper is notable in part for presenting the essence of the approach, with a minimum of additional overhead. For reference, not as an assignment, here are some presentations of some of these different developments of the ideas:

- “Rubin Causal Model” (Imbens and Rubin, 2008), for Rubin’s approach;
- “On the application of probability theory to agricultural experiments. Essay on principles. Section 9” (Neyman, 1990), for Neyman’s approach, or
- *Statistics*, note 12 to chapter 27 (Freedman et al., 1998, pp.A-32–4), for Freedman and collaborators’ exegesis of Neyman’s view;
- “Causal Inference, Path Analysis, and Recursive Structural Equations Models” (Holland, 1988); and
- *Making Things Happen: A Theory of Causal Explanation* (Woodward, 2005) for a development oriented toward philosophical concerns about causation.

*To be discussed at the February 4 meeting.*

## 3 Causal inference without an assumed “response surface”

### 3.1 Fisher-type hypothesis tests in randomized experiments

- *Observational Studies* § 2.1–4 (Rosenbaum, 2002b, pp.19–40).
- Supplemental, on instrumental variables from a Fisherian point of view: “Identification of causal effects using instrumental variables: Comment” (Rosenbaum, 1996).
- Supplemental, on randomization-based and other modes of inference from a Bayesian point-of-view: “Formal Modes of Statistical Inference for Causal Effects” (Rubin, 1990).

*To be discussed at the February 11 meeting.*

### 3.2 Model-assisted, randomization-based hypothesis testing

- “Covariance adjustment in randomized experiments and observational studies” Rosenbaum (2002a).
- Supplemental: *Observational Studies* (Rosenbaum, 2002b), § 2.5–7.

*To be discussed at the February 18 meeting.*

### **3.3 Assignment to treatment by groups**

- “Randomization inference in a group-randomized trial of treatments for depression: covariate adjustment, noncompliance and quantile effects” (Small et al., 2007);
- “Attributing Effects to A Cluster Randomized Get-Out-The-Vote Campaign” (Hansen and Bowers, 2009).
- Supplemental: “On Design Considerations and Randomization-based Inference for Community Intervention Trials”(Gail et al., 1996).

*To be discussed at the February 25 meeting.*

## **4 Post-stratifying observational studies to mimic block-randomized experiments**

### **4.1 Prelude: Matching for robustness of model-based adjustment**

- “Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies” (Rubin, 1979);
- “Matching to Remove Bias in Observational Studies” (Rubin, 1973).

*To be discussed at the March 4 meeting.*

### **4.2 “Hidden experiments” within observational studies**

- “The Central Role of the Propensity Score in Observational Studies for Causal Effects” (Rosenbaum and Rubin, 1983)
- “The Prognostic Analogue of the Propensity Score” (Hansen, 2008b)

*To be discussed at the March 18 meeting.*

### **4.3 Subclassification on the propensity score**

- “Estimating Causal Effects from Large Data Sets Using Propensity Scores” (Rubin, 1997).
- “Reducing Bias in Observational Studies using Subclassification on the Propensity Score” (Rosenbaum and Rubin, 1984).

*To be discussed at the March 25 meeting.*

### **4.4 Matched sampling**

- “Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score” (Rosenbaum and Rubin, 1985);
- “Practical Implications of Modes of Statistical Inference for Causal Effects and the Critical Role of the Assignment Mechanism” § 1–3 (Rubin, 1991, pp.1213-21).

*To be discussed at the April 1 meeting.*

## 4.5 Matching with multiple controls and full matching

- “Substantial Gains in Bias Reduction from Matching with a Variable Number of Controls” (Ming and Rosenbaum, 2000);
- “OPTMATCH: Flexible, Optimal Matching for Observational Studies” (Hansen, 2007).

Supplemental:

- “Full matching in an observational study of coaching for the SAT” (Hansen, 2004);
- “Comparison of Multivariate Matching Methods: Structures, Distances, and Algorithms” (Gu and Rosenbaum, 1993).

*To be discussed at the April 8 meeting.*

## 4.6 Balance tests (and associated controversies)

- “Misunderstandings among Experimentalists and Observationalists: Balance Test Fallacies in Causal Inference,” particularly § 7 (Imai et al., 2008);
- “The essential role of balance tests in propensity-matched observational studies: Comments on “A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003” by Peter Austin, *Statistics in Medicine*” (Hansen, 2008a).
- Supplemental: “Covariate balance in simple, stratified and clustered comparative studies,” § 1–4,5,6 (Hansen and Bowers, 2008, pp.219–29,231–3).

*To be discussed at the April 15 meeting.*

## 5 Closing lectures

Readings to be announced.

## References

- Cox, D. R. (1986), “Comment,” *Journal of the American Statistical Association*, 81, 963.
- Freedman, D. A. (1991), “Statistical Models and Shoe Leather,” *Sociological Methodology*, 21, 291–313.
- (2004), “Graphical Models for Causation, and the Identification Problem,” *Evaluation Review*, 28, 267.
- Freedman, D. A., Pisani, R., and Purves, R. (1998), *Statistics*, W.W. Norton and Co.
- Gail, M. H., Mark, S. D., Carroll, R. J., Green, S. B., and Pee, D. (1996), “On Design Considerations and Randomization-based Inference for Community Intervention Trials,” *Statistics in Medicine*, 15, 1069–1092.
- Gu, X. and Rosenbaum, P. R. (1993), “Comparison of Multivariate Matching Methods: Structures, Distances, and Algorithms,” *Journal of Computational and Graphical Statistics*, 2, 405–420.
- Hansen, B. B. (2004), “Full matching in an observational study of coaching for the SAT,” *Journal of the American Statistical Association*, 99, 609–618.
- (2007), “OPTMATCH: Flexible, Optimal Matching for Observational Studies,” *R News*, 7, 18–24.

- (2008a), “The essential role of balance tests in propensity-matched observational studies: Comments on “A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003” by Peter Austin, *Statistics in Medicine*,” *Statistics in Medicine*, 27, 2050–2054.
- (2008b), “The Prognostic Analogue of the Propensity Score,” *Biometrika*, 95, 481–488, doi: 10.1093/biomet/asn004.
- Hansen, B. B. and Bowers, J. (2008), “Covariate balance in simple, stratified and clustered comparative studies,” *Statistical Science*, 23, 219–236.
- (2009), “Attributing Effects to A Cluster Randomized Get-Out-The-Vote Campaign,” *Journal of the American Statistical Association*, to appear.
- Holland, P. (1988), “Causal Inference, Path Analysis, and Recursive Structural Equations Models,” *Sociological Methodology*, 18, 449–484.
- Holland, P. W. (1979), “The Tyranny of Continuous Models in a World of Discrete Data,” *IHS-Journal*, 3, 29–42.
- (1986a), “Statistics and Causal Inference: Rejoinder,” *Journal of the American Statistical Association*, 81, 968–970.
- (1986b), “Statistics and Causal Inference (with discussion),” *Journal of the American Statistical Association*, 81, 945–970.
- Imai, K., King, G., and Stuart, E. (2008), “Misunderstandings among Experimentalists and Observationalists: Balance Test Fallacies in Causal Inference,” *Journal of the Royal Statistical Society*, 171, 481–502.
- Imbens, G. and Rubin, D. (2008), “Rubin Causal Model,” in *The New Palgrave Dictionary of Economics*, eds. Durlauf, S. N. and Blume, L. E., Macmillan Publishers Ltd.
- LaLonde, R. J. (1986), “Evaluating the Econometric Evaluations of Training Programs With Experimental Data,” *American Economic Review*, 76, 604–620.
- Ming, K. and Rosenbaum, P. R. (2000), “Substantial Gains in Bias Reduction from Matching with a Variable Number of Controls,” *Biometrics*, 56, 118–124.
- Neyman, J. (1990), “On the application of probability theory to agricultural experiments. Essay on principles. Section 9,” *Statistical Science*, 5, 463–480, transl. by D.M. Dabrowska and T.P. Speed from 1923 Polish original.
- Oakes, J. M. (2004), “The (mis)estimation of neighborhood effects: causal inference for a practicable social epidemiology,” *Social Science & Medicine*, 58, 1929–1952.
- Rosenbaum, P. R. (1996), “Identification of causal effects using instrumental variables: Comment,” *Journal of the American Statistical Association*, 91, 465–468.
- (1999), “Choice as an Alternative to Control in Observational Studies (with discussion),” *Statistical Science*, 14, 259–304.
- (2002a), “Covariance adjustment in randomized experiments and observational studies,” *Statistical Science*, 17, 286–327.
- (2002b), *Observational Studies*, Springer-Verlag, 2nd ed.
- Rosenbaum, P. R. and Rubin, D. B. (1983), “The Central Role of the Propensity Score in Observational Studies for Causal Effects,” *Biometrika*, 70, 41–55.
- (1984), “Reducing Bias in Observational Studies using Subclassification on the Propensity Score,” *Journal of the American Statistical Association*, 79, 516–524.

- (1985), “Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score,” *American Statistician*, 39, 33–38.
- Rubin, D. B. (1973), “Matching to Remove Bias in Observational Studies (Corr: V30 P728),” *Biometrics*, 29, 159–183.
- (1979), “Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies,” *Journal of the American Statistical Association*, 74, 318–328.
- (1986), “Comments on ‘Statistics and Causal Inference’,” *Journal of the American Statistical Association*, 81, 961–962.
- (1990), “Formal Modes of Statistical Inference for Causal Effects,” *Journal of Statistical Planning and Inference*, 25, 279–292.
- (1991), “Practical Implications of Modes of Statistical Inference for Causal Effects and the Critical Role of the Assignment Mechanism,” *Biometrics*, 47, 1213–1234.
- (1997), “Estimating Causal Effects from Large Data Sets Using Propensity Scores,” *Annals of Internal Medicine*, 127, 757–763.
- Small, D., Ten Have, T. R., and Rosenbaum, P. (2007), “Randomization inference in a group-randomized trial of treatments for depression: covariate adjustment, noncompliance and quantile effects,” *Journal of the American Statistical Association*, to appear.
- Woodward, J. (2005), *Making Things Happen: A Theory of Causal Explanation*, Oxford University Press, USA.