BIOSTAT 830
Statistical Methods for Causal Inference and Dynamic Treatment Regimes
Fall, 2015

Instructor	Lu Wang, Associate Professor of Biostatistics M4132, SPH II, University of Michigan Phone: 734-647-6935 Email: luwang@umich.edu
Hours and Locations	Class – Tuesday and Thursday, 12:30pm-14:00pm (SPH II M4332) No classes on Sep 15 and Oct 13. Office hour - Tuesday, 11:30-12:30 (SPH II M4132)
Prerequisites	Biostat 601, 602, 650, 651, and 653. Stat 610, 611.
Class website	https://ctools.umich.edu/portal/
Evaluation	Homework Assignments: 20% Class participation: 35% Leading Discussion and Presentation: 45%

Course Materials

Lecture notes and the corresponding readings (required)

Recommended books to read:

- Pearl, Judea (2000). Causality: Models, Reasoning, and Inference. Cambridge University Press.
- Spirtes P, Glymour C, Scheines R (1993). Causation, Prediction, and Search. Lecture Notes in Statistics 81. New York: Springer-Verlag.
- van der Laan MJ, Robins JM (2003). Unified Methods for Censored Longitudinal Data and Causality. New York: Springer Verlag.
- Morgan S and Winship C. (2014) Counterfactuals and Causal Inference: Methods and Principles for Social Research. Cambridge University Press.
- Chakraborty B and Moodie E. (2015) Statistical Methods for Dynamic Treatment Regimes: Reinforcement Learning, Causal Inference, and Personalized Medicine. Springer New York.
- Imbens G and Rubin R. (2015) Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction. Cambridge University Press.

Description of the Course

This course discusses statistical theory and methodology aimed at addressing causal inquiries from observational data and complex randomized designs. The second half of this course will mainly focus on statistical methods for evaluating dynamic treatment regimes, which are tailored individually for personalized treatment.

Two key modeling approaches are studied: the directed acyclic graph (DAG) models and models for counterfactual variables (structural models).

i) The theory of DAG is used

   a. to formally define and connect key causal concepts like confounding, exchangeability, overall effects, direct effects, intermediate variables, instrumental variables, non-compliance and selection bias; and

   b. to derive conditions for identifiability of causal contrasts. DAGs are also used to demonstrate that parameters of standard regression models do not have interpretation as causal contrasts for the effects of time-varying exposures in the presence of time dependent covariates that are simultaneously confounders and intermediate variables.

ii) Two broad classes of structural models are then introduced whose parameters have causal interpretation: marginal structural models and structural nested models. The statistical theory of estimation under these models is presented.

   This includes the derivation of

   a. the class of inverse weighted probability estimators for parameters of marginal structural models,

   b. locally semiparametric efficient, doubly-robust estimators for causal parameters, and

   c. g-estimators for parameters of structural nested mean models .

Course Objectives

At the end of the course the students will be able to:

1) Formulate causal contrasts of interest for addressing specific scientific inquiries.

2) Derive graphical models for investigating the conditions under which the causal contrasts of interest are identified from data collected under specific study designs.

3) Formulate adequate structural models for making inference about the causal contrasts of interest.

4) Derive efficient, double robust, estimators for the causal contrasts of interest under the postulated structural models.

5) Understand the statistical formulation and methods for evaluating dynamic treatment regimes.

Tentative Topics and Readings

Statistical DAGs, and Causal DAGs:

Required Readings:

- *Verma, T. and Pearl. J., "Causal Networks: Semantics and Expressiveness," UCLA Cognitive Systems Laboratory Technical Report 870032 (R-65), June 1987, in Proceedings, 4th Workshop on Uncertainty in Artificial Intelligence, Minneapolis, MN, Mountain View, CA, 352-359, August 1988. Also in R. Shachter, T.S. Levitt, and L.N. Kanal (Eds.), Uncertainty in AI 4, Elsevier Science Publishers, 69-76, 1990.*
- *D. Geiger, Verma, T.S. & Pearl, J., "Identifying Independence in Bayesian Networks," UCLA Cognitive Systems Laboratory, Technical Report CSD-890028 (R-116). In Networks, Vol. 20, No. 5, 507-534, 1990.*
- *J. Pearl, "Causal Diagrams for Empirical Research" Biometrika, 82(4), 669--710, December 1995.*
- *Pearl J. Robins JM. Probabilistic evaluation of sequential plans from causal models with hidden variables. In: Uncertainty in Artificial Intelligence.San Fancisco:Morgan Kaufmann, 1995:444-453.*

Additional Recommended Readings

- *Robins JM (1995). Comments on Judea Pearl's paper, "Causal diagrams for empirical research". Biometrika, 82:695-698.*
- *‘Greenland S, Pearl J, Robins JM (1999). Causal diagrams for epidemiologic research. Epidemiology, 10(1):37-48*
- *Robins JM (2003). Semantics of causal DAG models and the identification of direct and indirect effects. In Highly Structured Stochastic Systems, P. Green, N.L. Hjort, S. Richardson, Editors. NY: Oxford University Press, pp. 70-81.*
- *Hernán MA, Hernández-Díaz S, Robins JM (2004). A structural approach to selection bias. Epidemiology, 15:615-625.*

Causal Inference for Point Exposure/Treatment: G-formula and G-computation, Marginal Structural Model:

Required Readings:

- *Robins JM, Rotnitzky A, Zhao LP (1994). Estimation of regression coefficients when some regressors are not always observed. Journal of the American Statistical Association, 89:846-866. Reproduced courtesy of the American Statistical Association.*

- *Robins JM, Mark SD, Newey WK (1992). Estimating exposure effects by modelling the expectation of exposure conditional on confounders. Biometrics, 48:479-495.*
- *Robins JM, Hernán MA, Brumback B (2000). Marginal structural models and causal inference in epidemiology. Epidemiology, 11(5):550-560.*
- *Robins JM (1998). Marginal structural models. In: 1997 Proceedings of the Section on Bayesian Statistical Science, Alexandria, VA: American Statistical Association; pp. 1-10*
- *Robins JM (2000). Robust estimation in sequentially ignorable missing data and causal inference models. Proceedings of the American Statistical Association. Section on Bayesian Statistical Science 1999, pp. 6-10.*
- *Bang H, Robins J (2005). Doubly robust estimation in missing data and causal inference models. Biometrics, 61:692-972.*

Additional Recommended Readings

- *Hernán MA, Brumback B, Robins JM (2001). Marginal structural models to estimate the joint causal effect of nonrandomized treatments. Journal of the American Statistical Association -- Applications and Case Studies, 96(454):440-448.*
- *Robins JM, Rotnitzky A. (2001).Comment on the Bickel and Kwon article, "Inference for semiparametric models: Some questions and an answer" Statistica Sinica, 11(4):920-936. ["On Double Robustness."]*

<u>Causal Inference for Time-dependent Exposure/Treatment: Non-dynamic and dynamic regimes, non-ramdom and random regimes:</u>

Required Readings:

- *Robins JM (1997) Causal Inference from Complex Longitudinal Data. Latent Variable Modeling and Applications to Causality. Lecture Notes in Statistics (120), M. Berkane, Editor. NY: Springer Verlag, pp. 69-117.*
- *Pearl J. Robins JM. Probabilistic evaluation of sequential plans from causal models with hidden variables. In: Uncertainty in Artificial Intelligence.San Fancisco:Morgan Kaufmann, 1995:444-453*
- *Robins JM, Wasserman L (1997). Estimation of Effects of Sequential Treatments by Reparameterizing Directed Acyclic Graphs. Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence, Providence, RI, August 1-3, 1997. Geiger D., Shenoy P. (Eds.), Morgan Kaufmann, San Francisco, pp. 409-420.*
- *Neugebauer R. van der Laan (2006) G-computation estimation for causal inference with complex longitudinal data, Computational Statistics and Data Analysis ,51,1676-1697*

Additional Recommended Readings:

- *Robins JM (1986). A new approach to causal inference in mortality studies with sustained exposure periods - Application to control of the healthy worker survivor effect. Mathematical Modelling, 7:1393-1512.*
- *Robins JM (1987). A graphical approach to the identification and estimation of causal parameters in mortality studies with sustained*

exposure periods. Journal of Chronic Disease (40, Supplement), 2:139s-161s.

- *Robins JM, Greenland S, Hu F-C (1999). Estimation of the causal effect of a time-varying exposure on the marginal mean of a repeated binary outcome. Journal of the American Statistical Association - Applications and Case Studies, 94:687-700. Reproduced courtesy of the American Statistical Association.*
- *Gill RD, Robins JM (2001). Causal inference for complex longitudinal data: the continuous case. Annals of Statistics.*
- *Yu. Z, van der Laan. (2006) Construction of counterfactuals and the G-computation formula . To appear in Mathematical Methods of Statistics*
- *Hernán MA, Brumback B, Robins JM (2002). Estimating the causal effect of zidovudine on CD4 count with a marginal structural model for repeated measures. Statistics in Medicine, 21:1689-1709.*
- *Yu Z. van der Laan (2006) Double Robust Estimation in Longitudinal Marginal Structural Models, Journal of Statistical Planning and Inference 136, 3,1061-89.*
- *Brumback BA, Hernán MA, Haneuse SJPA, Robins JM (2004). Sensitivity analyses for unmeasured confounding assuming a marginal structural model for repeated measures. Statistics in Medicine, 23:749-767.*

Optimal dynamic treatment regimes:

- *Murphy S, van der Laan M, Robins JM and CPPRG.(2001). Marginal mean models for dynamic regimes. Journal of the American Statistical Association 96(456):1410-1423*
- *Murphy, Susan A*. *Optimal Dynamic Treatment Regimes*."*Journal of the Royal Statistical Society Series B-Statistical Methodology (with* discussion), *65 (2): 331-366. 2003.*
- *Robins JM (2004). Optimal structural nested models for optimal sequential decisions. In DY Lin and P Heagerty (Eds.), Proceedings of the Second Seattle Symposium on Biostatistics , New York. Springer.*
- *Erica E. M. Moodie, Thomas S. Richardson, David A. Stephens. Demystifying Optimal Dynamic Treatment Regimes Biometrics (OnlineEarly Articles).*
- *More is coming…..*

Other Potential Topics:

1. Structural Nested Models

- *Robins JM (1999). Marginal Structural Models versus Structural Nested Models as Tools for Causal Inference. Statistical Models in Epidemiology: The Environment and Clinical Trials. Halloran ME, Berry D, Eds, IMA Volume 116, NY: Springer-Verlag, pp. 95-134.*
- *Robins JM, Rotnitzky A.(2004). Estimation of treatment effects in randomised trials with non-compliance and a dichotomous outcome using structural mean models.Biometrika 91: 763-783.*

- *Robins JM (1994). Correcting for non-compliance in randomized trials using structural nested mean models. Communications in Statistics, 23:2379-2412.*
- *Zhiqiang Tan, "nonparametric likelihood and further development of inverse weighting and G-estimation for marginal and nested structural models" (September 2006). Johns Hopkins University, Dept. of Biostatistics Working Papers. Working Paper 117.*http://www.bepress.com/jhubiostat/paper117*.*

2. Causal Inference in Survival Analysis

- *Robins JM, Tsiatis AA (1992). Semiparametric estimation of an accelerated failure time model with time-dependent covariates. Biometrika, 79:311-319.*
- *Robins JM (1992). Estimation of the time-dependent accelerated failure time model in the presence of confounding factors. Biometrika, 79:321-34.*
- *Lok, J.J. , A.W. van der Vaart, R.D. Gill, A.W. van der Vaart, J.M. Robins. Estimating the causal effect of a time-varying treatment on time-to-event using structural nested failure time models. (2004) Statistica Neerlandica 58 (3), 271--295.*
- *Mark SD, Robins JM (1993). Estimating the causal effect of smoking cessation in the presence of confounding factors using a rank preserving structural failure time model. Statistics in Medicine, 12:1605-1628.*
- *Robins JM (1997). Structural nested failure time models. In: Survival Analysis, P.K. Andersen and N. Keiding, Section Editors. The Encyclopedia of Biostatistics, P. Armitage and T. Colton, Editors. Chichester, UK: John Wiley & Sons, pp. 4372-4389.6)*

- *Hernán MA, Brumback B, Robins JM (2000). Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. Epidemiology, 11(5):561-570.*