JOHNS HOPKINS
BLOOMBERG
SCHOOL of PUBLIC HEALTH

2012 Atlantic Causal Inference Conference (ACIC)

Johns Hopkins Bloomberg School of Public Health (JHSPH)
615 N Wolfe St.
Baltimore, MD

Sommer Hall

May 24-25, 2012

Organizing Committee:
Elizabeth Stuart (Chair; estuart@jhsph.edu)
Karen Bandeen-Roche
Constantine Frangakis
Nicholas Ialongo
Sherri Rose
Michael Rosenblum
Daniel Scharfstein

Sponsors:
The Center for Prevention and Early Intervention, JHSPH*
Department of Biostatistics, JHSPH

* Funded by the National Institutes of Mental Health and Drug Abuse

Schedule

Note: All scientific sessions will take place in Sommer Hall. All breaks and meals will be in Feinstone Hall. Breakfast and snacks on both days as well as lunch on Thursday will be provided. There is also an evening reception on Thursday.

Wednesday May 23, 7-10pm
Pre-conference evening informal gathering for junior researchers
Organizer: Michael Rosenblum
Red Star Bar and Grill (upstairs): 906 S. Wolfe St., Baltimore, MD,
http://www.yelp.com/biz/red-star-bar-and-grill-baltimore

Thursday May 24
8 – 9:30: Registration table open (Monument St. entrance to JHSPH)

8 – 9: Coffee and breakfast (Feinstone)

9-10:30 Symposium 1: Workshop on causal inference for high-dimensional data
Organizer and Chair: Dan Scharfstein, JHSPH
Speaker: Thomas Richardson, University of Washington

10:30-11: Coffee Break (Feinstone)

11-12:30 Symposium 2: Interference and spillover effects in causal inference
Organizer and Chair: Tyler VanderWeele, Harvard University
Speakers:
Michael Hudgens, University of North Carolina at Chapel Hill
Asymptotic distribution of causal effect estimators in the presence of interference
Guanglei Hong, University of Chicago
A Probabilistic Causal Model for Mediation with Interference
Peter Aronow, Yale University
Estimating Average Causal Effects Under General Inference
(Joint work with Cyrus Samii, New York University)
Discussant: Eric Tchetgen Tchetgen, Harvard University

12:30-2 Box Lunch (Feinstone)

2-3:30 Symposium 3: New Developments in Causal Inference for Longitudinal and Spatial Data
Organizer and Chair: Dylan Small, University of Pennsylvania
Speakers:
Kosuke Imai, Princeton University

On the use of linear fixed effects regression models for causal inference
http://imai.princeton.edu/research/FEmatch.html
(Joint work with In Song Kim, Princeton University)

James Dai, Fred Hutchinson Cancer Research Center
Partially hidden Markov model for time-varying principal stratification in HIV prevention trials

Cory Zigler, Harvard University
Estimating Causal Effects of Air Quality Regulations Using Principal Stratification for Spatially-Correlated Multivariate Intermediate Outcomes

Wei (Peter) Yang, University of Pennsylvania
Controlling the future methods for causal inference from longitudinal data

3:30-4 Coffee Break (Feinstone)

4-5:30 Symposium 4: Tom Ten Have Memorial Session
Winners of the 2011 Thomas R. Ten Have Award to Junior Researchers for exceptionally creative or skillful research on causal inference
Organizer and Chair: Elizabeth Stuart
Speakers:

Jose Zubizarreta, University of Pennsylvania
Using Mixed Integer Programming for Matching in Observational Studies

Liang Li, Cleveland Clinic
Matching weights and its application in propensity score analysis

Marc Ratkovic, Princeton University
Achieving Optimal Covariate Balance Under General Treatment Regimes

Luke Keele, Penn State University
Enhancing Geographic Discontinuities Through Matching
(Joint work with Rocio Titiunik, University of Michigan)

5:30-7 Reception and Poster Session (Feinstone; Organizer: Sherri Rose)

6:45 Presentation of 2012 Ten Have awards (Jay Kaufman and Sherri Rose)
Award Committee: Karen Bandeen-Roche (JHSPH), Jay Kaufman (McGill University), Susan Murphy (University of Michigan), Romain Neugebauer (Kaiser Permanente DOR), Dylan Small (University of Pennsylvania), Cory Zigler (Harvard University)

Friday May 25

8 – 9:30 Registration table open (Monument St. entrance to JHSPH)

8 – 9 Coffee and breakfast (Feinstone)

9-10:30 Symposium 5: Adaptive Designs for Causal Inference

Organizer and Chair: Sherri Rose, JHSPH

Co-organizer: Michael Rosenblum, JHSPH

Speakers:

Susan Murphy, University of Michigan

Piloting and Sizing Sequential Multiple Assignment Randomized Trials in Dynamic Treatment Regime Development

(Joint work with D. Almirall, S. Compton, M. Grunlicks-Stoessel, and N. Duan)

Mark van der Laan, University of California at Berkeley

TMLE of causal effects in adaptive group sequential randomized trials

(Joint work with Antoine Chambaz, Universite Paris Descartes).

Xiao-Hua Andrew Zhou, University of Washington

Optimal treatment Selection Using Biomarker Adjusted Treatment Effect (BATE) Curve

10:30-11:00 Coffee Break (Feinstone)

11-12:30 Symposium 6: Big-Data-Driven Medicine

Organizer and Chair: Dan Scharfstein, JHSPH

Speaker: David Madigan, Columbia University

Abstract: In our data-rich world, key medical decisions, ranging from a regulator's decision to curtail a drug to patient-specific treatment choices require optimal consideration of myriad inputs. Statistical/epidemiological methods that can harness real-world medical data in useful ways do exist, but much work remains to achieve the full potential of a truly data-driven medical environment. Key challenges ahead include scaleable causal inference and high-fidelity predictive models. I will describe some recent progress in the specific area of drug safety.

2012 ATLANTIC CAUSAL INFERENCE CONFERENCE POSTER SESSION

Poster abstracts on the following pages are ordered alphabetically by last/second name of the first author, or presenting author where this differed.

Thomas R. Ten Have Award

Junior researchers presenting a poster as first author at the 2012 Atlantic Causal Inference Conference were invited to be considered for the Thomas R. Ten Have Award. This award recognizes "exceptionally creative or skillful research on causal inference." Awardees will be announced at the end of the poster session and will be honored next year with an invited talk at the 2013 Atlantic Causal Inference Conference. Travel funds to attend the 2013 conference will also be offered if available.

The 2012 Atlantic Causal Inference Conference organizing committee would like to thank the Thomas R. Ten Have Award committee:

Karen Bandeen-Roche
Jay Kaufman
Susan Murphy
Romain Neugebauer
Dylan Small
Cory Zigler

1 Causal inference in case-noncase studies: a Rubin causal model approach

Nikola Andric and Donald B. Rubin
Department of Statistics, Harvard University
TEN HAVE AWARD CANDIDATE

Abstract

Case-noncase studies – commonly referred to as case-control studies – are a popular study design in biostatistics and epidemiology. These studies are commonly used to screen for factors that may be associated with a rare disease under study. We focus on the subset of case-noncase studies in which the treatment of interest has been identified. Current methodologies for analyzing case-noncase studies (e.g. Cochran-Mantel-Haenszel test with extensions to matched analyses, conditional logistic regression), although they can provide associative insights, are generally inappropriate for causal conclusions. In this poster we propose a causal inference approach for case-noncase studies that is consistent with the well-established Rubin Causal Model framework for prospective studies. We believe that our approach fills a conceptual gap between prospective and retrospective studies, and has the additional benefit of transparency of the assumptions being made. We use our framework to investigate the adequacy of retrospective matching in case-noncase studies. We propose a procedure that separates the design and analysis phases in these types of studies, allowing researchers to draw objective and causal conclusions by validly controlling for pretreatment variables.

2 Estimating average causal effects under general interference between units

Peter M. Aronow[1] and Cyrus Samii[2]
[1] Department of Political Science, Yale University
[2] Wilf Family Department of Politics, New York University
TEN HAVE AWARD CANDIDATE

Abstract

This paper presents randomization-based methods for estimating average causal effects under arbitrary interference of known form. Conservative estimators of the randomization variance of the average treatment effects estimators are presented, as is justification for confidence intervals based on a normal approximation. Examples relevant to research in environmental protection, networks experiments, "viral marketing," two-stage disease prophylaxis trials, and stepped-wedge designs are presented.

3 Why match in individually and cluster randomized trials?

Laura Balzer[1], Maya L. Petersen[1,2], and Mark J. van der Laan[1]
[1] Division of Biostatistics
[2] Division of Epidemiology
University of California, Berkeley School of Public Health
TEN HAVE AWARD CANDIDATE

Abstract

The decision to match individuals or clusters in randomized trials is motivated by both practical and statistical concerns. Matching protects against chance imbalances in baseline covariate distributions and is thereby thought to improve study credibility. Matching is also implemented to increase study power. Analogue to Rose and van der Laan (2009), this article investigates the asymptotic efficiency of pair-matching individuals or clusters relative to not matching in randomized trials. We focus on estimating the average treatment effect. We use the efficient influence curve to understand the information provided by each design for estimation of the target causal parameter. Our approach is estimator-independent, avoids all parametric modeling assumptions, and applies equally to individually randomized and cluster randomized studies. Our theoretical results indicate that the pair-matched design is asymptotically less efficient than its unmatched counterpart. Our simulations confirm these results asymptotically and in finite samples.

4 An application of three causal inference methods for estimating the effect of HBV on CD4 count dynamics among U.S. military active duty and beneficiaries starting HAART

Ionut Bebu, Kenneth J. Wilkins, Octavio Mesner, Brian Agan, and Grace Macalino
Infectious Disease Clinical Research Program
Uniformed Services University of the Health Sciences

Abstract

Given similar risk factors and the prevalence of co-infections, it is important to understand the impact of Hepatitis B infection (HBV) among HIV positive individuals. The U.S. Military HIV Natural History Study (NHS) cohort offers unique features including being a population screened for HIV with early entry to care, open access to care and medications, racial diversity, high level of education, lack of IV drug use, and stable socioeconomic status. We sought to evaluate and compare the causal effect of HBV on CD4 trajectories in our population using three standard approaches. For this analysis, an existing dataset for alcohol use and HAART outcomes was used, where participants with at least one alcohol questionnaire (administered in 2006) were included. We limited analysis to include only those who initiated HAART and we excluded those with missing data. The causal longitudinal effect of HBV at HAART initiation (HI) is evaluated using three standard approaches: g-estimation, propensity scores (PS) and inverse probability weighting (IPW). For g-estimation, a spline mixed model with knots at 4 years and 8 years from HI, which also adjusts for other baseline covariates is fitted, and the

causal effect is estimated as the difference in expected responses obtained for each subject as if they were assigned to both levels of the exposure. For PS and IPW, we first compute the propensity of HBV at HI using a logistic regression adjusting for baseline covariates. For PS analysis, we stratify the subjects in different risk groups defined by the propensity scores and separate models are fitted within each group. The IPW approach consists of a spline mixed model with weights equal to the inverse of the propensity scores of the observed exposure levels. Of the 1282 subjects contributing12884 person-visits 93.6% were male; 42.4% African American and 42.8% Caucasian; 80% were dated seroconverters. At HI, median age was 34 years (IQR 29, 40), median duration of HIV infections was 2.1 years (IQR 0.33, 6.21), median CD4 was 436 cells/mm3 (IQR 318, 591), and median VL was 4.5 log10 copies/mL (IQR 3.9, 5.0). Potential confounders considered at HI included age, race/ethnicity, gender, CD4, VL and time between diagnosis and HI. Using the g-estimation approach, the HBV causal effect at 4 years after HI is -86 CD4 cells/mm3, -73 at 8 years after HI, and -62 after 12 years after HI. The risk groups defined by the quartiles of the propensity scores were well balanced with respect to the covariates at baseline. The HBV rates in the four groups were 14.4%, 23.4%, 42.8% and 64.2%, respectively. The HBV causal effects were found to be similar across the risk groups. These findings were consistent with the ones obtained using IPW. Confirming previous association studies, we found that the HBV infection at HAART has a direct attributable effect on CD4 count trajectories in our population. The estimated causal effects obtained using the three methods are consistent. Future work will address time varying confounders and further explore the causal effect of chronic HBV infection.

5 A forest approach to defining a study population

Justin Bleich and Emil Pitkin
Department of Statistics, The Wharton School, University of Pennsylvania
TEN HAVE AWARD CANDIDATE

Abstract

Following the seminal work of Rosenbaum and Rubin (1983), matching based on the propensity score has become the predominant matching technique in observational studies. In order to reduce the asymptotic variance of the estimated average treatment effect (ATE), overlap between the propensity score distributions of the treatment and control groups is desirable. Tree-based approaches describe the overlapping populations in terms of their covariates, rather than propensity scores. We employ a novel method to define a study population that is as close to optimal as possible, which relies on a bootstrapping approach to search through a forest of prospective trees. Applications to real data are presented.

6 Surrogacy assessment using principal stratification when surrogate and outcome measures are multivariate normal

Anna S.C. Conlon, Jeremy M.G. Taylor, and Michael R. Elliot
Department of Biostatistics, University of Michigan
TEN HAVE AWARD CANDIDATE

Abstract

In clinical trials, a surrogate outcome variable (S) can be measured before the outcome of interest (T) and may provide early information regarding the treatment (Z) effect on T. Most previous methods for surrogate validation rely on models for the conditional distribution of T given Z and S. However, since S is a post-randomization variable, these methods do not result in a causal interpretation. Using the potential surrogacy framework introduced by Frangakis and Rubin (2002), we propose a Bayesian estimation strategy for surrogate validation when the joint distribution of potential surrogate and outcome measures is multivariate normal. We model the joint conditional distribution of the potential outcomes of T, given the potential outcomes of S and propose surrogacy validation measures from this model. By conditioning on principal strata of S, the resulting estimates are causal. As the model is not fully identifiable from the data, we propose some reasonable prior distributions and assumptions that can be placed on weakly identified parameters to aid in estimation. We explore the relationship between our surrogacy measures and the traditional surrogacy measures proposed by Prentice (1989). The method is applied to data from a macular degeneration study, previously analyzed by Buyse, et al. (2000) and data from an ovarian cancer study.

7 Assessment of the causal effect of policies based on stochastic interventions

Ivan Diaz and Mark J. van der Laan
Division of Biostatistics, University of California, Berkeley School of Public Health
TEN HAVE AWARD CANDIDATE

Abstract

Estimating the causal effect of an intervention on a population typically involves defining parameters in a nonparametric structural equation model (Pearl, 2000, Causality: Models, Reasoning, and Inference) in which the treatment or exposure is assigned in a deterministic way. We define a new causal parameter that measures the effect of an intervention that intends to alter distribution of an exposure in a stochastic manner. This parameter provides a powerful tool for assessing the impact of policies, since it takes into account that population interventions generally result in stochastically assigned exposures. The statistical parameter that identifies the causal parameter of interest is established, as well as its efficient influence function under the non parametric model. Inverse probability of treatment weighting (IPTW), augmented IPTW (A-IPTW), and targeted maximum likelihood estimators (TMLE) are developed. An application example assessing the causal effect of interventions on physical activity on overall cause mortality is presented.

8 Evaluation of treatments with heterogeneous effects using principal strata survival classes

Brian L. Egleston[1] and Mark K. Buyyounousk[2]
[1] Biostatistics and Bioinformatics Facility
[2] Department of Radiation Oncology
Fox Chase Cancer Center

Abstract

Combination radiotherapy and androgen deprivation therapy (ADT) is commonly used in the treatment of prostate cancer. While survival is typically long, there is growing concern that ADT may increase the risk of intercurrent death due to effects on the heart. We present a method for estimating the effect of treatment on five and ten year survival outcomes using a principal stratification approach. We use Cox proportional hazards regressions and estimators of the baseline survivor function to estimate individual survival probabilities at 5 and 10 years under different treatment assignments. We then use a number of conditional independence assumptions to estimate probabilities of being in survival classes that would benefit from treatment or be harmed by treatment. A sensitivity analysis approach is presented that depicts how deviations from our conditional independence assumptions affect our estimates. We present a data example of androgen deprivation therapy for the treatment of prostate cancer in which we find that more men benefit from therapy than are harmed, but the proportion who are harmed might be quite sizable.

9 How much compliance is enough? Estimating the complier average causal effect (CACE) for treatment efficacy with different definitions of compliance

Scott F. Grey
College of Public Health, Kent State University

Abstract

A recent solution to estimating treatment efficacy in studies with non-compliance has been the development of CACE estimates. Based on principal stratification, these models classify subjects who have received an adequate amount of the treatment as potential compliers and then compares them to control subjects who have an equal probability of being classified as compliers if they had been randomized to treatment. No studies have systematically examined how sensitive CACE estimates are to different definitions of compliance. This study hypothesizes that incorrect definitions of compliance can bias CACE estimates and seeks to determine under what circumstances bias can occur. The standard CACE method was compared to a partial compliance framework where there can be multiple principal strata of partial potential compliance and there is a true minimum partial potential compliance principal stratum where subjects would receive the minimum treatment exposure necessary to have a relevant outcome effect. In this framework, subjects can be incorrectly classified as non-compliers and compliers. Mathematical investigations and numeric analysis suggest that when subjects are incorrectly classified as compliers, CACE estimates are minimally affected. On the other hand,

when are incorrectly classified as non-compliers, CACE estimates can be grossly inflated. These results remain when CACE estimates were calculated using the exclusion restriction or a covariate, when exclusion restriction is true and when is false. Missing data, a common occurrence in research that is often related to noncompliance was found to somewhat attenuate the amount of bias seen in CACE estimates, but the extent of that attenuation appeared to be limited. These findings suggest that misclassifying true non-compliers as compliers will introduce a substantial, but not necessarily large amount of bias into CACE estimates, but that misclassifying true compliers as non-compliers will introduce a very large amount of bias into CACE estimates. This divergence below or above the true minimum partial potential compliance principal stratum may provide researchers with a method of identifying the true partial compliance principal strata using sensitivity analysis. This approach was tested using data from a large cluster randomized field trial, and appeared to be able to provide an estimate of the true partial compliance minimum, but the derived estimate did not obtain statistical significance, making it a questionable value.

10 Perils and prospects of using aggregate area level socioeconomic information as a proxy for individual level socioeconomic confounders in instrumental variables regression survival classes

Jesse Y. Hsu[1], Scott A. Lorch[2], and Dylan S. Small[1]
[1] Department of Statistics, The Wharton School
[2] Department of Pediatrics, Perelman School of Medicine
University of Pennsylvania
TEN HAVE AWARD CANDIDATE

Abstract

A frequent concern in making statistical inference for causal effects of a policy or treatment based on observational studies is that there are unmeasured confounding variables. The instrumental variable method is an approach to estimating a causal relationship in the presence of unmeasured confounding variables. A valid instrumental variable needs to be independent of the unmeasured confounding variables. It is important to control for the confounding variable if it is correlated with the instrument. In health services research, socioeconomic status variables are often considered as confounding variables. In recent studies, distance to a specialty care center has been used as an instrument for the effect of specialty care vs. general care. Because the instrument may be correlated with socioeconomic status variables, it is important that socioeconomic status variables are controlled for in the instrumental variables regression. However, health data sets often lack individual socioeconomic information but contain area average socioeconomic information from the US Census, e.g., average income or education level in a county. We study the effects on the bias of the two stage least squares estimates in instrumental variables regression when using an area-level variable as a controlled confounding variable that may be correlated with the instrument. We present simulation results and an application to a study of perinatal care for premature infants.

11 Average causal effect estimation allowing covariate measurement error

Yi Huang[1], Karen Bandeen-Roche[2], Xiaoyu Dong[3], Andrew Raim[1], and Cunlin Wang[3]
[1] Department of Mathematics and Statistics, University of Maryland, Baltimore County
[2] Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health
[3] Center for Drug Evaluation and Research, U.S. Food and Drug Administration
TEN HAVE AWARD CANDIDATE

Abstract

The covariates are often measured with error in biomedical and policy studies, which is a violation of the strong ignorability assumption. The naive approach is to ignore the error and use the observed covariates in current propensity score framework for average causal effect (ACE) estimation. However, after extending the existing causal framework incorporating assumptions allowing errors-in-covariates, we showed that the naive approach typically produces biased ACE inference. In this talk, we developed a finite mixture model framework for ACE estimation with continuous outcomes, which captures the uncertainty in propensity score subclassification from unobserved measurement error using the joint likelihood. The proposed approach will estimate the propensity score subgroup membership and subgroup-specific treatment effect jointly. Simulations studies and one real application (using the recent data from Infant Feeding Practice Study II) are used to show its performance and implementation. In summary, the proposed method extended the current propensity score subclassification approach to accommodate the cases where covariates are measured with errors.

12 Covariate balancing propensity score

Kosuke Imai and Marc Ratkovic
Department of Politics, Princeton University

Abstract

The propensity score plays a central role in a variety of settings for causal inference. In particular, matching and weighting methods based on the estimated propensity score have become increasingly common in observational studies. Despite their popularity and theoretical appeal, the main practical difficulty of these methods is that the propensity score must be estimated. Researchers have found that slight misspecification of the propensity score model can result in substantial bias of estimated treatment effects. In this paper, we introduce covariate balancing propensity score (CBPS) estimation, which simultaneously optimizes the covariate balance and the prediction of treatment assignment. We exploit the dual characteristics of the propensity score as a covariate balancing score and the conditional probability of treatment assignment and estimate the CBPS within the generalized method of moments or empirical likelihood framework. We find that the CBPS dramatically improves the poor empirical performance of propensity score matching and weighting methods reported in the literature. We also show that the CBPS can be extended to a number of other important settings, including the estimation of generalized propensity score for non-binary treatments, causal inference in longitudinal settings, and the generalization of experimental and instrumental variable estimates to a target population.

13 An unidentifiability issue in g-estimation and suggestions to fix it

Yang Jiang and Dylan S. Small
Department of Statistics, The Wharton School, University of Pennsylvania
TEN HAVE AWARD CANDIDATE

Abstract

In causal inference for longitudinal data, g-estimation gives a consistent estimator of the treatment effect under the sequential randomization assumption. Often, g-estimation is used along with a parameterized potential outcome model and a logistic model for the propensity score. However, the mle may not exist for some extreme treatment effect we plug in the model and this leads to some unidentifiability issues. We investigate and compare several methods for fixing this problem. Then we extend the discussion to the case of multiple treatments and relaxed assumptions (future ignorability assumption instead of sequential randomization assumption).

14 A Mendelian randomization approach to assessing a causal relation between malaria and stunting in children

Hyunseung Kang and Dylan S. Small
Department of Statistics, The Wharton School, University of Pennsylvania
TEN HAVE AWARD CANDIDATE

Abstract

While a strong association between repeated malaria episodes and stunting in children has been observed, there is considerable difficulty in elucidating a causal relationship. Randomized experiments are impractical, presence of confounders including malnutrition, breastfeeding habits, and socioeconomic status introduce bias, and there is reason to believe reverse causality may be in play. In this study, we utilize Mendelian randomization, a type of instrumental variables method, where the sickle cell trait is used as the instrument. The trait is an ideal choice because, by Mendel's law of inheritance, the trait is randomly assigned to children at birth. In this study, we present a potential outcomes model and explore conditions for estimation of the causal parameters. We also apply our work on a data set containing 1070 children from the Ashanti Region in Ghana who were recruited and followed up with genetic, anthropometric, and demographic measurements until two years of age.

15 Some pitfalls of using instrumental variables in observational health research: limited identification, weakness, and unmeasured confounding

Edward H. Kennedy[1] and James F. Burke[1,2]
[1] VA Center for Clinical Management Research, Ann Arbor VA Health Services Research & Development (HSR&D) Center of Excellence
[2] Robert Wood Johnson Foundation Clinical Scholars Program, University of Michigan
TEN HAVE AWARD CANDIDATE

Abstract

Instrumental variable (IV) methods are increasingly being used in observational health services and medical research. Distance- and utilization-based instruments have been employed to estimate effects on mortality of cardiac catheterization, stroke center admission, aprotinin during coronary bypass artery grafting, hospital complexity of services for trauma and prenatal care, hospital volume for mechanically ventilated patients, and more. IV methods are sometimes presented as a panacea for dealing with unmeasured confounding; however, in even simple settings they can require strong assumptions and yield limited inference due to identification issues. In this work we compare various IV approaches in the context of exploring the effect of magnetic resonance imaging (MRI) in preventing recurrent stroke, using a regional utilization-based instrument. Specifically, we compare (i) the popular two-stage least squares procedure, (ii) randomization inference (Rosenbaum 2002) after matching on the instrument propensity score (Tan 2006), (iii) an adapted randomization inference approach using a strengthened IV (Baiocchi et al. 2011), and (iv) methods for deriving bounds on the average treatment effect (Balke & Pearl 1997). We also consider structural mean models (Robins 1994, Hernan & Robins 2006, Vansteelandt et al. 2011). In these analyses, we pay particular attention to the interpretation of causal parameters (which can be complicated by strengthening, for example) and to problems with weak instruments. In addition, we describe the important role of sensitivity analysis, which we implement within the matching framework (Rosenbaum 2002). In settings for which it is a priori unclear whether the no unmeasured confounding assumption is more likely to be grossly violated for the treatment-outcome relationship or for the instrument-treatment and instrument-outcome relationships (as is the case in this application), we propose a direct comparison of corresponding sensitivity analyses. Such an approach can help to prioritize which among potentially conflicting conclusions is more trustworthy.

16 A Bayesian approach to the causal effect of multiple mediators

Chanmin Kim and Michael J. Daniels
Department of Statistics, University of Florida
TEN HAVE AWARD CANDIDATE

Abstract

We propose a Bayesian approach to estimate the natural direct and the joint indirect effect through multiple mediators in the setting of continuous mediators and a binary response. We can decompose the joint indirect effect into each individual indirect effect while preserving other effects, in particular the natural direct effect. Also, to increase efficiency of estimating marginal distributions of mediators, we incorporate baseline covariates. Several assumptions are introduced(with corresponding sensitivity parameters) to make unobservable effects identifiable from the observed data. We suggest strategies for eliciting sensitivity parameters and conduct simulations to assess violations of the assumptions.

17 The influence of village social cohesion on the gender gap in language skills: a panel study from India

Divya Nair and Nan Astone
Department of Population, Family and Reproductive Health,
Johns Hopkins Bloomberg School of Public Health
TEN HAVE AWARD CANDIDATE

Abstract

This paper examines the community determinants of cognitive development among a cohort of children in India (Andhra Pradesh) who were followed at 1, 5 and 8 years of age (n= 1900). Cognitive development is assessed via performance on the Peabody Picture Vocabulary Test that measures receptive vocabulary. The main social determinant of interest is village social cohesion. A challenge while examining the influences of social factors such as social cohesion on individual outcomes has been the failure to distinguish between social influences and other individual and contextual effects. Typically, the concern is that there is unobserved heterogeneity and that other factors determine both the nature of local associations across households and also influence child outcomes, and this makes estimators inconsistent. In response, we: (a) use three waves of data on children at eight years of age to model their current cognitive achievement in language skills, (b) we address endogeneity by using within-child fixed effect models and instrument parent-psychosocial measures with measures of community collective efficacy. We find that community social cohesion is a significant predictor of cognitive development after controlling for a host of relevant child and household characteristics (including, for example, child endowment, measures of socio-economic status and parent psychosocial responses). Community-type accounts for up to 26% of the variation in child cognition. We conclude by highlighting that in the rural Indian context, gender-differences in language skills are accentuated in communities with low social-cohesion.

18 Adaptive randomization

Julie Novak[1], Dylan S. Small[1], Benjamin French[2], and Scott Halpern[2,3]
[1] Department of Statistics, The Wharton School
[2] Department of Biostatistics and Epidemiology, Perelman School of Medicine
[3] Division of Pulmonary, Allergy and Critical Care, Perelman School of Medicine
University of Pennsylvania
TEN HAVE AWARD CANDIDATE

Abstract

Previous studies have shown that providing patients with financial incentives promotes healthier behavior. Behavioral and economic theory also suggest that the effectiveness of these incentives varies greatly depending on how and to whom we provide them. In particular, we focus on smoking cessation as our target healthy behavior, as smoking is the leading cause of preventable death in the United States. We develop methodology to optimize the structure of incentives by comparing efficacy, acceptance, and effectiveness of four financial incentives structures (versus usual care) with a randomized clinical trial. In order to accurately assess the differences between these financial structures, it is essential that the participants are randomly assigned to the incentives and the number of participants who accept each incentive to remain as similar as possible. The innovative method we propose for doing this is to update assignment probabilities into the five arms at a rate that increases accuracy and minimizes variability. The methodology is being developed for a nationwide smoking cessation study conducted by CVS.

19 Causal diagrams for interference and contagion

Elizabeth L. Ogburn and Tyler J. VanderWeele
Department of Epidemiology, Harvard School of Public Health
TEN HAVE AWARD CANDIDATE

Abstract

The term "interference" has been used to describe any setting in which one subject's exposure may affect another subject's outcome. We distinguish among three distinct causal mechanisms that give rise to interference. The first causal mechanism by which interference can operate is a direct causal effect of one individual's treatment on another individual's outcome. We call this direct interference. As an example, suppose that the outcome is obesity and the treatment is dietary counseling from a nutritionist. A treated individual can in turn "treat" his associates by imparting to them the information gained from the nutritionist; therefore if individual i receives treatment and individual j does not, individual j may be nevertheless be exposed to the treatment of individual i and his or her outcome will be affected accordingly. A second pathway by which one individual's treatment may affect another individual's outcome is via the first individual's outcome. For example, if the outcome is an infectious disease and the treatment is a prophylactic measure designed to prevent disease, then the treatment of individual i may affect the outcome of individual j by preventing individual i from contracting the

disease and thereby from passing it on. We call this type of interference contagion. It is differentiated from direct interference by the fact that it does not represent a direct causal pathway from the exposed individual to another individuals outcome, but rather a pathway mediated by the outcome of the exposed individual. The third pathway for interference is allocational interference. Treatment in this setting allocates individuals to groups; through interactions within a group individuals' characteristics may affect one another. An example that has been much discussed in the social science literature is the allocation of children to schools or to classrooms within schools. The performance and behavior of student i may affect the performance and behavior of student j in the same class, for example by distracting student j or by occupying the teachers attention with bad behavior, or by motivating student j with good study habits. In many settings more than one type of interference will be present simultaneously. The causal effects of interest differ according to which types of interference are present, as do the conditions under which causal effects are identifiable. We describe these differences, give criteria for the identification of important causal effects, and discuss applications to infectious diseases and social network data.

20 Identifying causally meaningful interpretations of coefficients for Black-White race: an inductive analysis of directed acyclic graphs

Whitney R. Robinson[1], Katherine J. Hoggatt[2], and Jay Kaufman[3]
[1] Department of Epidemiology, UNC Gillings School of Global Public Health
[2] University of California, Los Angeles School of Public Health
[3] Department of Epidemiology, Biostatistics, and Occupational Health, McGill University
TEN HAVE AWARD CANDIDATE

Abstract

BACKGROUND: In the potential outcomes framework, regression coefficients for race cannot be interpreted as causal effects. OBJECTIVE: Using graphical analysis of a toy example, we attempt to identify causally meaningful estimands for race coefficients. METHODS: We use an inductive approach. We present a toy example examining the relationship between dichotomous Black-White race and rates of skin cancer mortality. Two additional variables, genetics and quality health care, are potential covariates. We posit three DAGs (directed acyclic graphs) that describe potential causal relationships among the four variables. We analyze each graph to identify what covariate adjustment strategy is warranted. Finally, given a particular causal structure, adjustment strategy, and empirical result, we evaluate whether the adjusted coefficient for race corresponds to a causally meaningful estimand. We assume no measurement error, no unmeasured confounding, no interactions between race and the covariates, and collapsibility of the rate difference. RESULTS: In our toy example, the unadjusted coefficient for Black race is -3.0/100,000 person-years. When genetics and health care are graphed as confounders of the race-mortality relationship, the estimated total effect of race is null. When genetics is graphed as a confounder and health care as a mediator, the estimated total effect is positive. When genetics and health care are graphed as mediators, the total effect is -3.0 /100,000 person-years; the indirect effect via genes is negative; the indirect effect via health care is positive; and the direct effect of race is null. Relying on the previous DAG, in which race was an exogenous node, we propose an alternative estimand for race coefficients: the

counterfactual inequality. Interpreting the coefficients this way, unadjusted coefficient for race would estimate the racial inequality in mortality given the observed distribution of covariates. The coefficient adjusted for either covariate would estimate the racial inequality given that the distribution of that covariate is set equal across the race groups. Finally, the coefficient adjusted for both covariates would estimate the racial inequality given that the distributions of both covariates were set equal across the race groups. CONCLUSION: Graphical analysis demonstrated that there were multiple potential interpretations of coefficients for race. Two estimands the total effect of race given that all covariates are confounders and the direct effect given that all covariates are mediators – were null and theoretically demonstrate the fallacy of race effects. Counterfactual inequalities were novel estimands for race coefficients. Counterfactual inequalities are population-level estimands that are intuitive and causally meaningful.

21 'High'-er education: substance use among students at community and 4-year colleges

Janet Rosenbaum
Maryland Population Research Center, University of Maryland, College Park
TEN HAVE AWARD CANDIDATE

Abstract

Community college has been the primary means of post-secondary educational access for disadvantaged students, but most post-secondary educational research focuses on traditional four year colleges. This study tested whether substance use is more common among students at 4 year or community colleges, whether the pattern persists after matching on pre-college factors that predict college attendance, and whether substance use predicts graduation likelihood in either group, using the National Longitudinal Study of Adolescent Health. Pre-college variables were measured in 1995, when respondents were ages 12–18, substance use was measured in 2001, and failure to attain any post-secondary degree was measured in 2008. Exact and nearest-neighbor Mahalanobis matching within propensity score calipers used demographic factors (gender, age, race/ethnicity), deviance (marijuana use, friends smoking, out- of-school suspension history, ever pregnant), and socioeconomic status (household income, test score, grade average, grades missing, school attachment, college expectancies), all measured at baseline. Matching balanced on all 15 factors. After matching, most forms of illegal drug use were more common among community college students than four year college students, but alcohol use was more common among four year college students than among community college students. Students at four year colleges who used a variety of substances were more likely to attain no post-secondary degree than students at two-year colleges. Students at 4 year colleges who used moderate amounts of alcohol were more likely to earn a post-secondary degree than students who used no alcohol or who used large amounts of alcohol, which is consistent with Tinto's theory that social integration predicts college completion. Students at two-year college were more likely to use substances, but substance use was less likely to interfere with their attaining a post-secondary degree. High school students with substance use problems may be more likely to earn a post-secondary degree if they begin at a community college, and transfer to a 4 year college only after earning a post-secondary degree.

22 Comments on the Neyman-Fisher controversy

Arman Sabbaghi and Donald B. Rubin
Department of Statistics, Harvard University
TEN HAVE AWARD CANDIDATE

Abstract

The Neyman-Fisher controversy considered here originated with the 1935 publication of Jerzy Neyman's "Statistical Problems in Agricultural Experimentation" in the Journal of the Royal Statistical Society. Neyman asserted in this article that the standard ANOVA F-test for randomized complete block designs is valid, whereas the analogous test for Latin squares is invalid in the sense of detecting differentiation among the treatments, when none existed on average, more often than desired (i.e. having a higher Type I error than advertised). Fundamental algebraic mistakes were made in this work, and Neyman's expressions for the expected mean residual sum of squares, for both randomized complete block and Latin square designs, are incorrect. Furthermore, Neyman's claim that the Type I error (when testing the null hypothesis of zero average treatment effects) is higher than desired if the expected mean treatment sum of squares is larger than the expected mean residual sum of squares, is incorrect. Our simple examples illustrating these mistakes show that one cannot, in general (i.e., without further assumptions on the potential outcomes), determine the Type I error of the standard ANOVA F-test based only on expected mean sums of squares.

23 Analyzing regression discontinuity designs as randomized experiments

Adam Sales and Ben Hansen
Department of Statistics, University of Michigan
TEN HAVE AWARD CANDIDATE

Abstract

Regression discontinuity (RD) designs are pseudo-experimental scenarios wherein treatment assignment is a function of a continuous variable t: treatment is assigned when t is greater than (or less than) some known cutoff point. Conventionally, social scientists analyze RD designs by regressing the outcome of interest on t and treatment assignment. This approach has two distinct drawbacks: the first is that the resulting estimate is of a local average treatment effect – the limit of the treatment effect as t approaches the cutoff – whereas scientists are typically interested in average treatment effectsthe average treatment for a well-specified subgroup of the population. Secondly, doubt regarding the functional form of the relationship between t and the outcome has led researchers to limit their analyses to subjects whose t values fall within a small bandwidth of the cutoff; however, the best choice for the size of this bandwidth is an open question. This study introduces a new approach to analyzing RD designs by pre-processing the outcome variable, and then attempting to verify or identify a region in which they resemble data from a randomized experiment. This approach addresses the two important drawbacks of conventional RD analysis, by interpreting the resulting estimates as average treatment effects within a specified window of the cutoff (rather than at a point) and by suggesting a natural way to use covariate balance testing to validate or choose a bandwidth.

24 Marginal structural modeling of a survival outcome with targeted maximum likelihood estimation

Mireille Schnitzer[1], Mark J. van der Laan[2], Erica Moodie[1], and Robert Platt[1]
[1] Department of Epidemiology, Biostatistics, and Occupational Health, McGill University
[2] Division of Biostatistics, University of California, Berkeley School of Public Health
TEN HAVE AWARD CANDIDATE

Abstract

Targeted maximum likelihood estimation (TMLE) methods in causal inference (van der Laan & Rubin, 2006) have been developed to estimate exposure-specific mean outcomes in the general longitudinal context with simple extensions to the survival setting (van der Laan, 2010; van der Laan & Gruber, 2011). We present two different methods to model survival using Marginal Structural Models with TMLE: 1) modeling the log-odds of survival conditional on exposure and time, and 2) modeling the hazard. A form for the asymptotic variance is produced using semiparametric influence function theory. This methodology is presented in the context of an application in mortality for a cohort of patients co-infected with HIV and Hepatitis C.

25 Using generalized boosted models for propensity score estimation for multinomial treatments: a substance abuse treatment application

Daniel McCaffrey[1], Megan Schuler[2,*], Beth Ann Griffin[1], Rajeev Ramchand[1], and Daniel Almirall[3]
[1] RAND Corporation
[2] Department of Mental Health, Johns Hopkins Bloomberg School of Public Health
[3] Survey Research Center, Institute for Social Research, University of Michigan
[*] *Presenting author*

Abstract

Use of propensity score weighting when examining more than two treatment conditions has received limited attention in spite of theoretical advancements. Moreover, most applications involving more than two treatment groups depend on parametric estimation of the propensity score model when it has been shown that machine learning methods outperform (in terms mean squared error of resulting effects) the use of simple logistic regression models in the binary treatment case. This presentation will discuss the use of weighting to compare multiple treatment conditions, describe a method for estimating the multinomial treatment propensity scores that relies on generalized boosted models (GBM), and introduce a useful diagnostic criterion for assessing balance that does not rely on traditional p-values. The estimation technique and diagnostic criterion proposed will be illustrated using dataset which studies the relative effectiveness of four adolescent substance abuse treatment modalities on 12-month substance use outcomes.

26 Nested Markov models of directed acyclic mixed graphs

Thomas S. Richardson[1], Ilya Shpitser[2,*], and James M. Robins[2]
[1] Department of Statistics, University of Washington
[2] Department of Epidemiology, Harvard School of Public Health
[*] *Presenting author*

Abstract

Many datasets are plagued by unobserved confounders: hidden but relevant variables. The presence of such hidden variables obscures many conditional independence constraints on the observed margin, and greatly complicates data analysis. In this poster I introduce a new type of equality constraint which generalizes conditional independence, and which is a "natural" equality constraint for data generated from the marginal distribution of a DAG graphical model. I also introduce a new kind of graphical model, called the nested Markov model, which captures these constraints via a simple graphical criterion.

27 Assessing the effect of organ transplantation on the distribution of residual lifetime

David M. Vock[1], Anastasios A. Tsiatis[1], Marie Davidian[1], Eric B. Laber[1], Wayne M. Tsuang[2], C. Ashley Finlen-Copeland[2], and Scott M. Palmer[2]
[1] Department of Statistics, North Carolina State University
[2] Lung Transplant Program, Department of Medicine, Duke University Medical Center
TEN HAVE AWARD CANDIDATE

Abstract

Because the number of patients waiting for organ transplants exceeds the number of organs available, a better understanding of how transplantation affects the distribution of residual lifetime is needed to improve organ allocation. However, there has been little work to assess the survival benefit of transplantation from a causal perspective. Previous methods developed to estimate the causal effects of treatment in the presence of time-varying confounders have assumed that treatment assignment was independent across patients, which is not true for organ transplantation. We develop a version of G-estimation that accounts for the fact that treatment assignment is not independent across individuals to estimate the parameters of a structural nested failure time model. In addition, G-estimation for failure time models requires the use of artificial censoring, a technique where some subjects observed to fail are censored. Because artificial censoring reduces the information available and leads to non-smooth estimating equations, prior research has noted that finding the solutions to estimating equations can be difficult. We suggest some computational strategies to mitigate the problems typically encountered with artificial censoring. We derive the asymptotic properties of our estimator and confirm through simulation studies that our method leads to valid inference on the effect of transplantation on the distribution of residual lifetime. We demonstrate our method on the survival benefit of lung transplantation using data from the United Network for Organ Sharing (UNOS).

28 Causal inference in longitudinal comparative effectiveness studies with repeated measures of a continuous intermediate variable

Chen-Pin Wang[1], Booil Jo[2], and C. Hendricks Brown[3]
[1] Department of Epidemiology & Biostatistics, School of Medicine,
University of Texas Health Science Center, San Antonio
[2] Department of Psychiatry and Behavioral Sciences, Stanford School of Medicine
[3] Department of Epidemiology and Public Health, University of Miami Miller School of Medicine

Abstract

Lin, Ten Have, and Elliot (2008, 2009) proposed a principal stratification approach for a longitudinal randomized study to assess the treatment effect of a continuous outcome adjusting for repeated measures of a binary intermediate variable. Extending this model, we propose a principal stratification approach to assess causal effects in non-randomized longitudinal comparative effectiveness studies with a binary endpoint outcome and repeated measures of a continuous intermediate variable. Our motivation for this work comes from a comparison of the effect of two glucose-lowering medications on a clinical cohort of patients with type 2 diabetes. Here we consider a causal inference problem assessing how well the two medications work relative to one another on two binary endpoint outcomes, cardiovascular disease related hospitalization and all-cause mortality. Clinically, these glucose-lowering medications can have differential effects on the intermediate outcome, glucose level over time. Ultimately we want to examine whether the medication effect on the endpoint outcomes vary by glucose trajectories in response to the medications. The proposed method involves a 3-step model estimation procedure. Step 1 identifies principal strata associated with the intermediate variable using hybrid growth mixture modeling analyses (Jo, Wang, and Ialongo 2009). Step 2 obtains the stratum membership using the pseudoclass technique (Bandeen-Roche, Miglioretti, Zeger et al. 1997; Wang, Brown, Bandeen-Roche 2005), and derives the stratum-specific propensity scores for treatment assignment. Step 3 obtains the stratum-specific treatment effect on the endpoint outcome weighted by inverse propensity probabilities derived from Step 2.

29 Causal inference for the epidemiology of injury-associated infections: an approach for causal estimation within nested case-control studies of trauma-wounded individuals with outcome-dependent follow-up

Kenneth J. Wilkins and the Trauma Infectious Disease Outcomes Study (TIDOS)
and Case-Control Osteomyelitis Data Working Groups
Infectious Disease Clinical Research Program
Uniformed Services University of the Health Sciences

Abstract

We present an approach to causal inference within nested case-control designs, applied here using well-defined cohorts that involve distinct stages of outcome-dependent follow-up. A Department of Defense / Veterans Affairs (DoD/VA) Trauma Infectious Disease Outcomes Study (TIDOS) cohort motivates this approach. All deployed military personnel who sustain trauma injuries are followed from point of injury through the medical evacuation chain to DoD hospitals, and assessed for incident infections. Military point-of-care trauma registries, medical records and clinically-sampled specimens determine variables that are suspected to play a role as confounders for the point exposures of interest: blast injury, and antibiotic/pathogen-susceptibility concordance upon hospital admission. Bone infections (osteomyelitis) that develops subsequent to extremity wound infections, however, are often only ascertained after initial U.S. hospitalization. Estimating disease-exposure associations may be prone to bias, therefore, if trauma-wounded are only followed up for developing osteomyelitis within this limited period. Notably, extended surveillance is possible as almost half of eligible trauma-wounded within DoD consent to post-discharge followup; this prospective cohort is re-approached upon VA admission to consent to further follow-up. Investigators suspect that each individuals chances for such extended follow-up may depend in some way on their propensity for osteomyelitis. Disease-exposure association estimates are thus sensitive to unverifiable assumptions about this dependence, further complicating estimation of causal association measures. A companion retrospective case-control study of osteomyelitis affords incidence estimates for late-developing osteomyelitis via medical records, and such auxiliary data may guide investigators in selecting assumptions reasonable to adopt for the prospective cohort. An adaptation of targeted learning methods for two-stage designs (Rose & van der Laan 2011) allows one to examine how estimated associations vary under distinct (yet unverifiable) assumptions about the dependence between osteomyelitis and extended follow-up after initial hospitalization, illustrated here using preliminary (albeit incomplete) data from each study.

30 Prognostic score adjustments in multi-stage clinical trials of psychiatric medications

Frank Yoon
Mathematica Policy Research
TEN HAVE AWARD CANDIDATE

Abstract

Adjustments based on the propensity score achieve balance on covariates that are associated with treatment decisions. As a complement or alternative, adjustments on the prognostic score achieve balance on covariates related to outcomes; specifically, prognostic balance asserts that patients' covariates are not systematically associated with trial outcomes. Two major psychiatric studies of medication treatment for schizophrenia and major depression present an opportunity to explore the use of prognostic score adjustments: the Clinical Antipsychotic Trials of Intervention Effectiveness for Schizophrenia (CATIE) and the Sequential Trial of Alternatives to Relieve Depression (STAR*D). In these trials, patients and their providers were able to elicit choices in their multi-stage treatment strategies. For example, if a first medication was ineffective or presented harmful side effects, the patient could subsequently choose from alternative medication strategies. Because these choices are not random, adjustments must be made in order to infer the effect of alternative strategies. The objective of trial designs that incorporate patient and provider choice is to emulate real world settings in which medical treatment for severe mental illness does not usually involve a static intervention, but rather sequential decisions based on patient experience, including response to medications. Prognostic score methods hold potential to aid investigators in the analysis of complex studies that involve patient-centered medical decision making.

31 A comparison of approaches for assessing causal mediation

Donna L. Coffman[1], David P. MacKinnon[2], Yeying Zhu[3,*], and Debashis Ghosh[3]
[1] The Methodology Center, Pennsylvania State University
[2] Department of Psychology, Arizona State University
[3] Department of Statistics, Pennsylvania State University
[*] *Presenting author*

Abstract

Mediation occurs as part of a hypothesized causal chain of events: An intervention or treatment, T, has an effect on the mediator, M, which then affects an outcome variable, Y. Mediation is often assessed using a regression-based approach that relies on the assumption that there are no unmeasured confounders that influence both M and Y. This assumption holds if individuals are randomly assigned to levels of M but generally random assignment to M is not possible. Recently, three different approaches, all of which fall under the potential outcomes framework for causal inference, have been proposed for drawing more valid causal inference in mediation analyses. These approaches define the mediation effects as either principal strata effects (e.g., Rubin, 2004; Jo, 2008), natural effects (e.g., Pearl, 2001; Imai et al., 2010), or controlled effects (e.g., Robins & Greenland, 1992; VanderWeele, 2009). In this study, we illustrate that

each of these definitions answer different scientific questions and that each makes different assumptions about the manipulability of the mediator, the existence of direct effects (i.e., the effect of T on Y that is not due to M), iatrogenic effects of T on M, the existence of post-T confounders that have been influenced by T, and the existence of T*M interactions. We also describe how the traditional regression-based approach relates to the three approaches based on the potential outcomes framework. We assess the sensitivity of each of the three approaches to violations of the assumptions using a simulation study that systematically varies different aspects of these assumptions. The data generation for the simulation study is very general so as to not favor one approach over another.