Causal Inference in Statistics and the Quantitative Sciences

Erica E. M. Moodie (McGill University) and David A. Stephens (McGill University)

May 3–8, 2009

1 A Short Overview of the Field

Causal inference attempts to uncover the structure of the data and eliminate all non-causative explanations for an observed association. The goal of most, if not all, statistical inference is to uncover causal relationships. However it is not in general possible to conclude causality from a standard statistical inference procedure, it is merely possible to conclude that the observed association between two variables is not due to chance. Statistical inference procedures do not provide any information about which variable causes the other, or whether the apparent relationship between the two variables is due to another, confounding variable. An explicit introduction of the philosophy of and approaches to causation was first brought into the statistical sciences in 1986 by Paul Holland [1], although references to causal approaches exist in the literature up to 60 years prior (see, for example, [2, 3]). Since then, there has been an explosion of research into the area in a variety of disciplines including statistics (particularly biostatistics), computer science, and economics.

Causal inference in statistics is a broad area of research, under which many topics fall. In particular, the following themes were considered:

1. Inference and asymptotic theory

   Causal inference provides a natural testbed for classical asymptotic theory, in particular, semiparametric inference. Consistent estimation of causal parameters is guaranteed under certain standard regularity and sampling conditions, by the standard asymptotic theory of estimating equations. More interesting, however, is first, the issue of semiparametric efficiency - optimal asymptotic variances can be deduced by appealing to semiparametric arguments, but estimation is complicated by the presence of nuisance parameters in different components of the model - and secondly, the issue of "double robustness" [4], where consistent estimates of causal parameters follows even under misspecification of the mean model, provided a nuisance model such as the intervention or treatment model or a missingness mechanism model is correctly specified. In addition, non-standard asymptotic theory is required for certain non-regular problems, for instance, those that arise for non-differentiable estimating functions in the study of dynamic regimes [5].

2. Balancing scores and inverse weighting: advances in biostatistics

   The fundamental objective of causal inference is to balance the treatment groups so that the treated and untreated subjects are comparable with respect to confounding variables. There are two common approaches to achieving this balancing that are frequently employed in biostatistics. The first relies on modeling the probability of receiving treatment, so that comparisons between treatment groups may be made within strata of subjects who have similar profiles with respect to their likelihood of treatment exposure. Adjustment for the probability of receiving treatment is typically accomplished by weighted regression (Marginal Structural Models [6, 7]), adjustment in a regression model or matching (propensity scores [8]).

The second approach to causal comparisons is most relevant in the context of clinical trials. This approach aims to identify subjects who have complied with their randomly assigned treatment allocation and to compare response between treated and untreated subjects within strata of subjects who have similar profiles with respect to their likelihood of complying with the assigned treatment [9].

3. Instrumental variables and structural equation models: connecting statistics and econometrics

   As in the health sciences, economists are typically interested in causal relationships, such as determining whether a particular training program increases income. In an instrumental variables analysis, the key assumption of most causal methods – that all confounding variables have been recorded – is dropped; in its place, the analyst requires an instrument, i.e. a variable which predicts exposure but does not affect outcome via any other pathway [10]. Many important methods of causal inference including the instrumental variables approach to analysis and the Generalized Propensity Score – an extension of the traditional propensity score that facilitates the estimation of dose-response relationships – were developed by economists. These methods are particularly useful and are generally under-used by statisticians.

4. Adaptive treatment regimes

   Estimating the best sequence of treatment regime for a chronic illness such as hypertension or cancer presents many statistical challenges. In many such diseases, the potential for microbial resistance, toxic side-effects, and compliance with treatment over time can complicate the ability to decide when and how to recommend treatment changes. Typically, the individual tailoring of treatments has been done at the clinical level on an ad-hoc or experience-driven basis at the physician's discretion, and is not based on statistical evidence. The area of dynamic (or adaptive) treatment regimes, pioneered in the statistics literature by Dr. Susan Murphy [11] and Dr. James Robins [5], attempts to formalize the estimation of optimal decision rules for treatment over time, specific to time-varying patient characteristics. Sequential decision making problems such as the estimation of optimal adaptive treatment regimes have also been considered in the computer sciences, through methods in artificial intelligence, reinforcement learning, and control theory.

5. Bayesian causal inference

   There has been relatively little attention given to causal approaches such as marginal structural models in the Bayesian communities, although there are of course exceptions (e.g. [12]). Many of the methods of causal inference including regression on propensity scores, marginal structural models, and instrumental variables require two-step approaches in which a number of nuisance parameters much be estimated. A Bayesian approach would allow for cohesive propagation of the uncertainty in the models.

2 Presentation Highlights

The purpose of this inter-disciplinary workshop was threefold: to review recent advances in the causal inferences in statistic; to bring together researchers from related fields, in particular Economics, Computer Sciences, and Epidemiology, who work on causal inference methodology so that approaches and ideas may be shared; and finally, to increase the profile of causal inference amongst statisticians in Canada.

The workshop opened on the evening of May 3 with a social hour and posters presented by Bibhas Chakraborty (University of Michigan), Ashkan Ertefaie (McGill University), Sara Geneletti (Imperial College London), Jay Kaufman (McGill University), Benjamin Rich (McGill University), Susan Shortreed (McGill University), Elizabeth Stuart (Johns Hopkins University), and Yongling Xiao (McGill University).

The keynote speaker was Judea Pearl, a computer scientist and philosopher who formalized the topic of causal reasoning in his seminal book, "Causality: Models, Reasoning, and Inference" [13]. Dr. Pearl provided the workshop participants with a two-hour overview of causes and counterfacturals, introducing principles based on non-parametric structural equation models that are sufficient for solving many problems involving causal relationships.

Monday May 4, 2009

Pearl, Judea (University of California, Los Angeles) Discussed principles, based on non-parametric structural equation models enriched with ideas from logic and graph theory, that give rise to a formal calculus of counterfactuals and unify existing approaches to causation.

van der Laan, Mark (University of California, Berkeley) Presented an approach to causal effect estimation that uses cross-validation to select optimal combinations of many model fits, and subsequent targeted maximum likelihood estimation to target the fit towards the causal effect of interest. This approach takes away the need for specifying regression models, while still providing maximum likelihood based estimators and inference.

Geneletti, Sara (Imperial College London) Gave an overview of the decision theoretic framework for causal inference and discussed the pros and cons of this approach compared to one based on countrefactuals, arguing that DT provides a more concise, economical and justifiable approach to inference of treatment effects.

Neufeld, Eric (University of Saskatchewan) Presented visualizations offering an interesting pedagogical tool for explaining the ideas of causation, intervention, and confounding.

Tuesday May 5, 2009

Small, Dylan (University of Pennsylvania) Introduced the malaria attributable fraction (MAF), talked about difficulties in estimating this quantity, and presented a potential outcomes framework for defining and estimating the MAF, as well as a sensitivity analysis that assesses the sensitivity of inferences to departures from the assumption of random assignment of parasite densities.

VanderWeele, Tyler (University of Chicago) Demonstrated the use of marginal structural models, which can also be applied in the presence of time-dependent confounding, to test for sufficient cause interactions. He showed that lower bounds on the prevalence of such sufficient cause interactions could be determined.

Schaubel, Douglas (University of Michigan) Developed semiparametric methods to estimate the effect on restricted mean lifetime of a time-dependent treatment with application to data from a national organ transplant registry. The method involves weighting results from stratified proportional hazards models fitted using a generalization of case-cohort sampling. The evaluation of asymptotic and finite-sample properties of the proposed estimator was presented.

Henderson, Robin (University of Newcastle) Proposed a modelling and estimation strategy for optimal dynamic treatment regimes which incorporates the regret functions into a regression model for observed responses. This addresses problems of model building, checking and comparison that have had little or no attention so far.

Noorbaloochi, Siamak (Center for Chronic Disease Outcome Research) Showed how sufficiency and ancillarity concepts can be used to understand and construct methods to reduce bias due to imbalance in baseline predictors. Gave as an example the effective dimension reduction summaries provided by regression graphics as an alternative to propensity based analysis.

Wednesday May 6, 2009

Strumpf, Erin (McGill University) Provided an overview of two methods, instrumental variables and regression discontinuity, used by economists to identify causal effects of interest. Discussed how these methods have been used to address health-related questions.

Robins, James (Harvard School of Public Health) Showed that in certain special cases, a complete causal DAG can be discovered from data. This is an unusual result because the causal DAG is normally considered to be structural information necessarily external to the data. Under the assumption of faithfulness, conditional independence relationships among the observed variables impose constraints on possible DAGs for the data generating process—in the example considered, however, no such conditional relationships were present. Through manipulations of the joint density that correspond to specific operations on the associated (unknown) causal DAG, Robins showed that an exhaustive search could uncover conditional independence relationships that would have been present had one intervened on certain variables and that, in the special case considered, identify the causal DAG uniquely.

Goetghebeur, Els (Ghent University) Reviewed instrumental-variable-based methods of estimation for the causal odds ratio when outcomes are dichotomous. Comparisons made both formally and via simulation were presented.

Hirano, Keisuke (University of Arizona) Presented results on estimation and inference for partially identified models specified through moment inequalities that are of great interest in economics, but also closely related to problems in dynamic optimal treatment regimes.

Thursday May 7, 2009

Robins, James (Harvard School of Public Health) Presentation given by Dr. Robins in lieu of Andrea Rotnitzky. Proposed novel methods for using the data obtained from an observational database in one health care system to determine the optimal treatment regime for biologically similar subjects in a second health care system when the two health care systems differ in the frequency of, and reasons for, both laboratory tests and physician visits. Also proposed a novel method for estimating the optimal timing of expensive and/or painful diagnostic or prognostic tests.

Arjas, Elja (University of Helsinki and the National Institute for Health and Welfare) Presented a non-parametric Bayesian modeling/predictive inference approach to estimation of optimal dynamic treatment regimes. The proposed methods was illustrated using the Multicenter AIDS Cohort Study (MACS) data set.

Hernán, Miguel (Harvard School of Public Health) Presented an application of a *dynamic* marginal structural model. The application considered was determination of the optimal threshold in CD4 count for HAART initiation in HIV patients.

Joffe, Marshall (University of Pennsylvania) Outlined selective ignorability assumptions which can be used to derive valid causal inference in conjunction with structural nested models, illustrated on erythropoetin use and mortality among hemodialysis patients. Discussed the connection between selective ignorability assumptions and G-estimation with instrumental variables assumptions and estimation.

Gustafson, Paul (University of British Columbia) Considered the case of nonidentified models from a Bayesian perspective, with an emphasis on the example of instrumental variables analysis. Argued that in this context the posterior distribution on a parameter of interest may no longer concentrate to a single point as the sample size grows and it is therefore important to study the width of large-sample limiting posteriors, as well as their sensitivity to the choice of prior distribution.

Richardson, Thomas (University of Washington) Considered the problem of non-identifiability of the instrumental variable potential outcomes model in which the instrument, treatment and response are all binary using a Bayesian approach, also treated by Pearl in his book. After demonstrating sensitivity to the prior specification, went on to characterize the 15-dimensional parameter space for this problem in terms of 6 observed and 9 unobserved dimensions after re-parametrization.

Abadie, Alberto (Harvard University) Reviewed the topic of matching estimators and their distribution. Presented newly developed methods for the calculation of the asymptotic distribution of a broad class of matching estimators. In particular, the important case of the two-stage estimator obtained by matching on the propensity score were discussed.

Rosenblum, Michael (University of California, San Francisco) The targeted maximum likelihood estimator for the causal parameters of a marginal structural model was presented and advantageous properties of the estimator were discussed. A related method for diagnosing bias due to violation of experimental treatment assignment has been proposed. These methods were then applied to estimate the effect of medication adherence on viral suppression in a cohort of HIV positive, homeless individuals in San Francisco.

Friday May 8, 2009

Sekhon, Jasjeet (University of California, Berkeley) Concerning the problem of using matching to obtain good balance on observed covariates, proposed a non-parametric matching method based on an evolutionary search algorithm. Discussed advantages of this method over propensity score matching—in the example of Pulmonary Artery Catheterization considered, the proposed method is able to replicate the experimental results using observational data while propensity score matching does not. Also talked about difficulties and assumptions that apply to the use of matching in general.

McCandless, Lawrence (Simon Fraser University) Considered Bayesian techniques to adjust for unmeasured confounding. A novel methods for observational studies with binary covariates that models the confounding effects of measured and unmeasured confounders as exchangeable within a Bayesian framework was proposed, and its properties discussed.

Glynn, Adam (Harvard University) Demonstrated that the observation of a post-treatment variable can improve our knowledge about total causal effects when treatment assignment is non-ignorable and the assumptions necessary for the front-door technique do not hold. Presented a Bayesian model that provides a framework for sensitivity analysis when the treatment is unobserved, but a post-treatment proxy is.

3 Outcome of the Meeting

The workshop successfully brought together researchers from Statistics, School of Government, Economics, Computer Science, and Epidemiology with a common interest in quantitative methods for causal inference. Participants came from Canada, the United States, England, Belgium, and Finland and represented a great range of career stages. The seminars presented were of an exceptional quality, and participants took advantage of the unscheduled time to exchange ideas informally.

A special issue of the *International Journal of Biostatistics* is forthcoming which will be devoted to publishing research material presented at the workshop or stemming from discussions which took place at the workshop.

References

[1] P. W. Holland, Statistics and Causal Inference *Journal of the American Statistical Association* **81** (1986), 945–960

[2] J. Neyman, On the application of probability theory to agricultural experiments. (Translation published in 1990.) *Statistical Science*, **5** (1923) 472–480.

[3] D. B. Rubin, Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies *Journal of Educational Psychology* **66** (1974), 688–701.

[4] H. Bang and J. M. Robins, Doubly Robust Estimation in Missing Data and Causal Inference Models, *Biometrics* **61** (2005), 962–973.

[5] J. M. Robins, Optimal structural nested models for optimal sequential decisions. In *Proceedings of the Second Seattle Symposium on Biostatistics*, D. Y. Lin and P. J. Heagerty (editors), Springer, New York, 2004.

[6] J. M. Robins, Association, Causation, and Marginal Structural Models, *Synthese* **121** (1999), 151-179.

[7] J. M. Robins, M. A. Hernán, and B. Brumback, Marginal Structural Models and Causal Inference in Epidemiology, *Epidemiology* **11** (2000), 550-560.

[8] P. R. Rosenbaum and D. B. Rubin, The Central Role of the Propensity Score in Observational Studies for Causal Effects, *Biometrika* **70** (1983), 41-55.

[9] M. M. Joffe, T. R. Ten Have, and C. Brensinger, The Compliance Score as a Regressor in Randomized Trials, *Biostatistics* **4** (2003), 327-340.

[10] J. D. Angrist, G. W. Imbens, and D. B. Rubin, Identification of Causal Effects Using Instrumental Variables *Journal of the American Statistical Association* **91** (1996), 444–472.

[11] S. A. Murphy, Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society, Series B* **65** (2003), 331–366.

[12] D. Rubin, Bayesian Inference for Causal Effects: The Role of Randomization *Annals of Statistics* **6** (1978), 34–58.

[13] J. Pearl *Causality: Models, Reasoning, and Inference*, Cambridge University Press, Cambridge, UK, 2000.

Participant Feedback

Michael Rosenblum:

> I wanted to express my many thanks for your having organized such a wonderful conference at Banff! It was amazing–I learned so much and met such amazing researchers in causal inference. The goal of connecting statisticians, computer scientists, and economists was certainly achieved.
>
> In particular, the presentations of Judea Pearl, Jamie Robins, and Alberto Abadie were outstanding. The time built in to allow for discussions was great as well.

Thomas Richardson:

> The causal conference was very good. A good mix of people.

Benjamin Rich:

> I wanted to tell you what a great time I had at the workshop. It was an amazing week that I won't forget. I feel very lucky to have participated and I wanted to thank you again for giving me that opportunity. I think the poster was well received, and I learned a great deal from the many excellent talks I heard throughout the week.

Marshall Joffe:

> Thank you very much again for including me in the program at Banff last week. This was one of the best meetings I have ever attended: there were many interesting talks, and there was ample opportunity to converse with and meet many interesting and bright people working in the field, and to learn about many new developments. Additionally, the setting was unbeatable.

Dylan Small:

> I had a great time at the causal conference. The talks were great and it was great having chance to spend time with friends and meet new friends.

Tyler VanderWeele

> I wanted to write to thank you again for organizing the Banff workshop. I really enjoyed my time there. You did a very nice job of bringing together people with different backgrounds. The sessions Monday through Wednesday were quite interesting and, although I then had to leave after that, I imagine Thursday and Friday went just as well. Thank you again for all of your work in putting the conference together. Banff is lovely!