Semiparametric double-robust estimation for continuous treatment effects: an application for road traffic accident analysis

Daniel J. Graham
Department of Civil Engineering,
Imperial College London, London, UK
Email: d.j.graham@imperial.ac.uk

1 Introduction

Propensity score based estimators are frequently used in inference problems for which it is possible to define a set of potential outcomes that correspond to different treatment regimes. Such regimes could, for instance, be different type of transport intervention or could relate to some exposure variable that varies in intensity across units in the sample. In this paper a semiparametric double robust estimator for continuous treatments is derived by combining generalised propensity score (GPS) estimates with an outcome regression model. The method is applied in a case study of the effects of area deprivation on Child Pedestrian Casualties (CPCs). The approach has general applicability for the assessment of continuous or multivalued interventions on transport related outcomes

2 Previous research

The typical set up for propensity score models under the potential outcomes framework is one in which the data available for estimation take the form of a random vector, $z_i = (y_i, d_i, x_i)$, $i = 1, ..., n$, where for the i-th unit of observation y_i denotes a response, d_i the treatment (or exposure) received, and x_i a vector of pre-treatment covariates. In the absence of experimental data we cannot assume that the treatment (or exposure) is assigned randomly, and consequently, simple comparisons of mean responses across

different treatment groups will not in general reveal a 'causal' effect due to potential for confounding. If, however, the vector of covariates x_i can be used to ensure conditional independence of potential outcomes and treatment assignment, then consistent 'causal' estimates of treatment effects can be obtained in a variety of ways.

The conditional independence, or unconfoundedness, assumption is key and in the case of binary treatments amounts to:

$$(Y_i(0), Y_i(1)) \perp\!\!\!\perp D_i | X_i, \tag{1}$$

where $Y(1)$ and $Y(0)$ indicate potential outcomes under treated or control status respectively.

The assumption of 'unconfoundedness' can be restated using a scalar known as the propensity score, which measures the conditional probability of assignment to the treatment given the covariates,

$$e(x) = Pr(D_i = 1 | X_i = x).$$

If unconfoundedness given X_i holds, and if the propensity score effectively balances the distribution of the observed covariates within strata of of the sample that have the same propensity score such that $X_i \perp\!\!\!\perp D_i | e(X_i)$, then $(Y_i(0), Y_i(1)) \perp\!\!\!\perp D_i | e(X_i)$.

The development of a number of useful nonparametric estimators for binary treatments, based for instance on matching, stratification and weighting, has relied on the propensity score as a minimal sufficient reduction of the potentially high dimensional covariate vector X_i. More recently, propensity score methods have allowed the potential outcomes framework to be extended to multi-valued and continuous treatments, in which a treatment $D = d$ can take values in k categories $\mathcal{D} \equiv (d_0, d_1, ..., d_k)$ or in some bounded interval in $\mathbb{R}$ [3, 2]. The relevant question here is what is the mean response to a given dose.

3 Methodological contribution

Double robust (DR) estimation combines an outcome regression (OR) model with inverse propensity score (PS) weighting to derive an average treatment effects (ATEs) estimator which is consistent and asymptomatically normal under correct specification of the OR or the PS model. DR estimators been studied extensively in the context of binary treatments (e.g. [6, 8, 7, 10, 5, 1, 4]).

In this paper we extend the binary DR approach of [9] to derive a semiparametric approximation for continuous dose-response functions. Analytical results and simulations show that our proposed estimator is DR for ATEs and can provide a good approximation to linear or nonlinear dose-response functions. A probability representation of the Generalized propensity score (GPS), introduced by [3] and [2], is used to ensure overlap in the support of the covariates and to achieve balancing such that heterogeneity in the sample is rendered nearly orthogonal to the treatment.

The continuous DR estimator is derived as follows. First, an OR model is augmented with a set of inverse PS covariates to provide separate bias correction estimating equations for defined strata of the treatment. The mean predicted values from the augmented OR model are then used to obtain consistent DR point estimates of ATEs at various doses. These in turn feature as the response variable in a penalized spline regression to recover an approximation to the potentially nonlinear dose-response for all $d \subseteq \mathbb{R}$ over the range of interest. Bootstrapping is used for inference.

4 Results

The paper demonstrates that the proposed continuous DR estimator is robust to problems of confounding or functional form misspecification. The DR estimator is applied in an analysis of the effect of exposure to area deprivation on the incidence of child pedestrian casualties (CPCs). The DR model is estimated on a sample trimmed to ensure overlap results are compared to those from an OR model obtained using data for the full sample. The results uncover a positive dose-response curve between increasing deprivation and the incidence of CPCs.

References

[1] Bang, H. and J. M. Robins (2005). Doubly robust estimation in missing data and causal inference models. *Biometrics* 61, 962972.

[2] Hirano, K. and G. W. Imbens (2004). The propensity score with continuous treatments. In A. Gelman and X. Meng (Eds.), *Applied Bayesian modeling and causal inference from incomplete data perspectives*, pp. 7384. New York: Wiley.

[3] Imbens, G. W. (2000). The role of the propensity score in estimating dose-response functions. *Biometrika* 87(3), 706710.

[4] Kang, J. D. Y. and J. L. Schafer (2007). Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data. *Statistical Science* 22(4), 523539.

[5] Lunceford, J. K. and M. Davidian (2004). Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. *Statistics in Medicine* 23, 29372960.

[6] Robins, J. M. (2000). Robust estimation in sequentially ignorable missing data and causal inference models. In *Proceedings of the American Statistical Association, Section on Bayesian Statistical Science*, pp. 610. Alexandria, VA: American Statistical Association.

[7] Robins, J. M. and A. Rotnitzky (2001). Comment on Inference for semiparametric models: some questions and an answer. *Statistical sinica* 11, 920936.

[8] Robins, J. M., A. Rotnitzky, and M. J. Van Der Laan (2000). Comment on the Murphy and Van der Vaart article On profile likelihood. *Journal of the American Statistical Association* 95, 431435.

[9] Scharfstein, D. O., A. Rotnitzky, and J. M. Robins (1999). Adjusting for nonignorable dropout using semiparametric nonresponse models. *Journal of the American Statistical Association* 94(448), 10961120. (with rejoinder 1135-1146).

[10] Van Der Laan, M. and J. M. Robins (2003). Unified methods for censored longitudinal data and causality. Berlin: Springer.