lmtp: An R Package for Estimating the Causal Effects of Feasible Interventions Based on Modified Treatment Policies

Nicholas Williams[*,a], Ivan Diaz[a]

[a] *Division of Biostatistics, Department of Population Health Sciences, Weill Cornell Medical College, 402 E. 67th Street, New York, NY, 10065*

Abstract

In many situations, it is impossible to estimate the causal effects of a policy intervention using randomization, resulting in a reliance on causal inference methods using observational data. The majority of these methods consider counterfactual outcomes where exposure is set deterministically. However, these treatment effects may not be particularly relevant and are often infeasible to bring about, especially in the case of continuous exposures. Modified treatment policies offer a non-parametric alternative to deterministic treatment effects that allow for the study of feasible interventions and offer a safegaurd against positivity violations. In this article, we introduce the lmtp `R` package. The lmtp package provides doubly-robust, non-parametric estimators of dynamic point-treatment and longitudinal modified treatment policies for binary, categorical, and continuous exposures with missing outcomes while leveraging ensemble machine learning for estimation.

Introduction

Most modern causal inference methods consider the effects of a treatment on population mean outcomes under interventions that set the treatment value deterministically. For example, the average treatment effect considers the hypothetical difference in population mean outcomes if a binary treatment was applied to all observations versus if it was applied to no observations. In the case of a categorical or continuous exposure, it is unlikely any policy intervention could bring this about. In addition, the estimation of causal effects requires the so called positivity assumption which states that all observations have a greater than zero chance of experiencing the exposure levels. This assumption is often violated when evaluating the effects of deterministic interventions and is usually exacerbated with longitudinal data.

In response to the aforementioned issues, alternative causal effects have been defined in the literature that can be formulated to avoid violations of the positivity assumption[INSERT CITATIONS HERE FOR THESE PAPERS]. Of immediate relevance to this article are *modified treatment policies*. Building off work by [INSERT DIAZ AND VAN DER LAAN] Modified treatment policies were first introduced by [INSERT ROTNISKY] and were extended to the

[*] Corresponding Author

Email address: `niw4001@med.cornell.edu` (Nicholas Williams)

Preprint submitted to Patterns *June 26, 2020*

longitudinal setting by [INSERT DIAZ, WILLIAMS, HOFFMAN]. Modified treatment policies are defined as interventions that can depend on the *natural* value of the exposure. [INSERT SECTION COMMENTING ON STATIC VS DYNAMIC INTERVENTIONS].

The `R` package **lmtp** (available for download from GitHub at https://github.com/nt-williams/lmtp) implements four estimators for estimating the effects of modified treatment policies in both the point-treatment and longitudinal studies. Two of these estimators, a targeted minimum-loss based estimator (TMLE) and a sequentially doubly-robust estimator (SDR), are doubly-robust. [INSERT EXPLANATION OF WHAT THIS MEANS]. The software allows for static or dynamic interventions, binary, categorical, and continuous exposures, and binary and continuous outcomes with right censoring. In addition, **lmtp** is capable of estimating deterministic causal effects such as the average treatment effect. Estimation procedures are carried out using the Super Learner algorithm by way of the `sl3` package [INSERT CITATION FOR SL3] and cross-fitting is used to maintain $n^{1/2}$ convergence [INSERT CITATIONS FOR WHY WE USE CROSS-FITTING]. In this article we describe how the **lmtp** package can be used for estimating static and dynamic deterministic and modified treatment policy effects for a variety of research questions. Code examples are provided.

Results

Required Data Structure

Overview of function parameters

Specifying policies of interest

Using the Super Learner

Outcomes of interest

Example 1: Static, Binary Trt., Continuous Outcome.

Example 2: MTP, Longitudinal, Continuous Trt., Binary Outcome.

Example 3: Dyanmic Deterministic, Longitudinal, Categorical Trt..

Example 4: MTP, Survival Outcome w/Censoring, Time-varying confounders.

Discussion

References