Introduction to Longitudinal Modified Treatment Policies

A solution for studying complex, continuous, and/or time-varying exposures

Kat Hoffman

2024-02-02

Overview

- Discussing a tutorial paper on Longtudinal Modified Treatment Policies (accepted with minor revisions to *Epidemiology*)
- Target audience: epidemiologists and applied statisticians
- Based on methodology proposed in *Díaz et al. (2021)*

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION
2023, VOL. 118, NO. 542, 846–857: Theory and Methods
https://doi.org/10.1080/01621459.2021.1955691

Taylor & Francis
Taylor & Francis Group

Check for updates

Nonparametric Causal Effects Based on Longitudinal Modified Treatment Policies

Iván Díaz[a], Nicholas Williams[a], Katherine L. Hoffman[a], and Edward J. Schenck[b]

[a]Division of Biostatistics, Department of Population Health Sciences, Weill Cornell Medicine, New York; [b]Division of Pulmonary & Critical Care Medicine, Department of Medicine, Weill Cornell Medicine, New York

ABSTRACT

Most causal inference methods consider counterfactual variables under interventions that set the exposure to a fixed value. With continuous or multi-valued treatments or exposures, such counterfactuals may be of little practical interest because no feasible intervention can be implemented that would bring them about. Longitudinal modified treatment policies (LMTPs) are a recently developed nonparametric alternative that yield effects of immediate practical relevance with an interpretation in terms of meaningful interventions such as reducing or increasing the exposure by a given amount. LMTPs also have the advantage that they can be designed to satisfy the positivity assumption required for causal inference. We present a novel sequential regression formula that identifies the LMTP causal effect, study properties of the LMTP statistical estimand such as the efficient influence function and the efficiency bound, and propose four different estimators. Two of our estimators are efficient, and one is sequentially doubly robust in the sense that it is consistent if, for each time point, either an outcome regression or a treatment mechanism is consistently estimated. We perform numerical studies of the estimators, and present the results of our motivating study on hypoxemia and mortality in intubated Intensive Care Unit (ICU) patients. Software implementing our methods is provided in the form of the open source R package `lmtp` freely available on GitHub (*https://github.com/nt-williams/lmtp*) and CRAN.

ARTICLE HISTORY
Received June 2020
Accepted July 2021

KEYWORDS
Continuous exposures; Longitudinal data; Modified treatment policies; Sequential double robustness; Targeted minimum loss-based estimation

Methodology motivation

- Many causal inference methods and tutorials focus on binary exposures at a single time point
  - Continuous/multi-level categorical exposures common in applied research, but methods, software, teaching materials are more limited
  - Many studies have time-varying exposures, but methods are even less common
- Positivity assumption is essential to causal inference
  - Violations are common in cases of categorical and continuous exposures
  - Exacerbated when there are multiple time points

One solution: Longitudinal Modified Treatment Policies (LMTPs)

- Diaz et al. (2021) proposed longitudinal interventions which depend on an individual's *natural value of treatment*
  - Natural value of treatment: the value treatment would take at time t if an intervention was discontinued right before time t
  - Provided identification result and doubly/sequentially robust estimation algoritms
- Methodology generalizes static, dynamic, and some stochastic interventions, so can accommodate:
  - binary, categorical, continuous, and multiple exposures
  - binary, continuous, time-to-event outcomes, competing risks, informative right-censoring, clustering
  - point-in-time and time-varying settings
- LMTPs help address violations of the positivity assumption, because we define an alternative interventions for which positivity holds by design

Motivation: Tutorial

1. Review static and dynamic interventions, and introduce (longitudinal) modified treatment policies
2. High-level theory:
   - Identification in point-in-time and time-varying settings
   - Estimation procedures
3. Application:
   - Provide examples of research questions that could be (or already have!) been addressed using LMTPs
   - Illustrate application of an LMTP to estimate the effect of intubation timing on mortality in COVID-19 patients, using a real-world longitudinal observational data set

Notation and setup

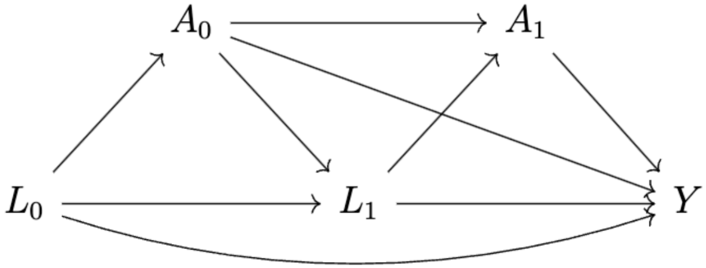
Notation

- $Z_1, ..., Z_n \overset{\text{iid}}{\sim} \mathsf{P}$
  - P represents a longitudinal process and may contain any number of time points, but for simplicity we will describe a distribution with only two time points, $t \in \{0, 1\}$
  - For each unit in the study, we observe a set of random variables $Z = (L_0, A_0, L_1, A_1, Y)$

Notation	Description	Structural Causal Equation
L_0	Baseline covariates	$L_0 \leftarrow f(U_{L_0})$
A_0	Treatment at first time point	$f(L_0) \leftarrow U_{L_0}$
L_1	Time-varying covariates	$f(L_0) \leftarrow U_{L_0}$
A_1	Treatment at second time point	$f(L_0) \leftarrow U_{L_0}$
Y	Outcome at defined study period end	$f(L_0) \leftarrow U_{L_0}$

Directed Acyclic Graph (DAG)

Simple DAG omitting unmeasured confounders:



Could have many more time points, high dimensional variables, competing events, censoring nodes etc.

Intervention notation

- H_t: history of data measured up to right before A_t
  - $H_0 = L_0$
  - $H_1 = (L_0, A_0, L_1)$
- Conceptualize treatment policies in terms of hypothetical interventions on nodes of the DAG
- Interventions: consider a user-given function $\mathsf{d}_0(a_0, h_0, \epsilon_0)$ which maps a treatment value a_0, a history h_0, and a possible randomizer ϵ_0 into a potential treatment value.

We refer to these longitudinal interventions, and the subsequent methods to identify and estimate effects under such interventions, as LMTPs.

Review of static and dynamic interventions

Static interventions

- All units receive the same treatment
  - For two time points, conceptualize a hypothetical world in which all units are treated at both time points $(\mathsf{d}_t = 1$ for $t \in \{0, 1\})$
  - Contrast to a hypothetical world in which no units are treated at either time point $(\mathsf{d}_t = 0\ t \in \{0, 1\})$
  - Gives rise to the well-known Average Treatment Effect (ATE)

Static intervention examples

- Hypothetical intervening on a population to:
  - enforce 30 minutes of moderate exercise for all individuals, every day
  - give all individuals an exact level of antibodies for a certain disease
  - setting a certain level of air quality each day, for all geographical areas of interest

Dynamic interventions

- Intervention depends only on a study unit's past covariates
  - Can include past treatment
- Often used in observational studies when study units need to meet an indication of interest for a treatment or policy to reasonably begin, e.g.
  - severity of illness indicator
  - socioeconomic threshold to begin a policy

Dynamic interventions

- One of the first uses of dynamic interventions was in the context of HIV, where investigators were interested in the effect of initiating antiretroviral therapy for a person with HIV if their CD4 count falls below a threshold, e.g. 200 cells/μl (Hernán et al. 2006)

$$\mathrm{d}_t(h_t) = \begin{cases} 1 & \text{if } l_t^* < 200 \text{ for all } s \geq t \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

where L_t^* is a variable in H_t that denotes CD4 T-cell count

Dynamic intervention examples

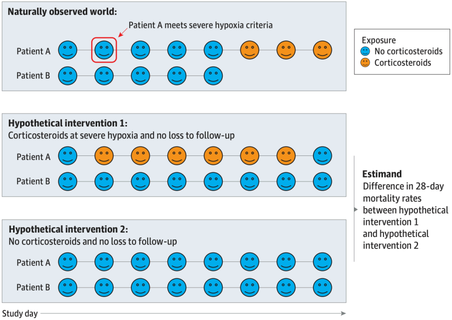
- Studying the effect of initiating a corticosteroids regimen for COVID-19 patients (Hoffman et al. 2022)
- Estimated mortality under a hypothetical policy where corticosteroids are administered for six days if and when a COVID-19 patient first meets a severity of illness criteria (i.e. low levels of blood oxygen)

$$\mathsf{d}_t(h_t) = \begin{cases} 1 & \text{if } l_s^* = 1 \text{ for any } s \in \{t-5, \ldots, t\} \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

where L_t^* is a variable in H_s that denotes the first instance of low levels of blood oxygen.

Dynamic intervention examples

Figure 1. Illustrated Example of 2 Patients Under the 2 Hypothetical Treatment Regimens of the Target Trial Emulation



Patient A reached severe hypoxia criteria at study day 2 and was followed the entire study duration. Patient B never reached severe hypoxia criteria and was lost to follow-up after 5 study days. Under the dynamic corticosteroids regimen (intervention 1), patient A received 6 days of corticosteroids, and under intervention 2, they received no corticosteroids. Patient B did not receive corticosteroids under either treatment regimen; however, in both hypothetical worlds, they were observed for the entire study duration.

Dynamic intervention examples

- Pollution example: instead of studying a hypothetical intervention in which all counties in the US have an air quality of a certain measure, we design a dynamic intervention in which rural counties have an air quality of 20 and urban counties have an air quality of 40

PICTURES

Modified Treatment Policy

- Intervention function $t(a_t, h_t, \epsilon_t)$ non-trivially depends on the natural value of treatment a_t, and perhaps on h_t and/or ϵ_t
- Prior examples of interventions which depend on the natural value of treatment
  - Threshold intervention (CITE)
  - smoking cessation
  - we will focus on shift interventions

Identification

Identification assumptions

Positivity

- If it is possible to find an observation with history h_t with an exposure of a_t, then it is also possible to find an observation with history h_t with an exposure (a_t, h_t, ϵ_t)

Strong sequential randomization

- All common causes of the intervention variable A_t and $(U_{L,t+1}, U_{A,t+1})$ are measured and recorded in H_t.
  - Generally satisfied if H_t contains all common causes of A_t and $(L_{t+1}, A_{t+1}, \ldots, L_\tau, A_\tau, Y)$, where τ is the last time point in the study

Weak sequential randomizationDíaz et al. (2021)

- All common causes of the intervention variable A_t and $(U_{L,t+1})$ are measured and recorded in H_t (Taubman et al. 2009)

Identification formula

1. Start with the conditional expectation of the outcome Y given $A_1 = a_1$ and $H_1 = h_1$. Let this function be denoted $Q_1(a_1, h_1)$.
2. Evaluate the above conditional expectation of Y if A_1 were changed to A_1^{d}, which results in a pseudo outcome $\tilde{Y}_1 = Q_1(A_1^{\mathsf{d}}, H_1)$.
3. Let the true expectation of $\tilde{Y}_1$ conditional on $A_0 = a_0$ and $H_0 = h_0$ be denoted $Q_0(a_0, h_0)$.
4. Evaluate the above expectation of $\tilde{Y}_1$ if A_0 were changed to A_0^{d}, which results in $\tilde{Y}_0 = Q_0(A_0^{\mathsf{d}}, H_0)$.
5. Under the identifying assumptions, we have $\mathsf{E}[Y(\bar{A}_\tau^{\mathsf{d}})] = \mathsf{E}[\tilde{Y}_0]$.

Estimation

test

Once a causal estimand is defined and identified, the researcher's task is to estimate the statistical quantity, e.g. $\mathsf{E}[\tilde{Y}_0]$

Parametric estimation

- Simplest option: fit parametric outcome regressions for each step of g-formula identification result
  - "Plug-in" estimator called **parametric g-formula** or **g-computation**
- Alternatively, use estimator which relies on the exposure mechanism
  - Inverse Probability Weighting (IPW) estimator
  - IPW estimation involves reweighting the observed outcome by a quantity which represents the likelihood the intervention was received, conditional on covariates

G-computation

1. Fit a generalized linear model (GLM) for Y conditional on $A_1 = a_1$ and $H_1 = h_1$. Call this $\hat{Q}_1(a_1, h_1)$.

   $\texttt{Q1\_hat} \leftarrow \texttt{glm}(\text{outcome} = Y, \text{predictors} = \{H_1, A_1\})$

2. Modify the data set used in step (1) so that the values in the column for A_1 are changed to A_1^{d}. Obtain the predictions for the model $\hat{Q}_1$ using this modified data set. These are pseudo-outcomes $\tilde{Y}_1$.

   $\texttt{pseudo\_Y1} \leftarrow \texttt{predict}(\text{fit} = \texttt{Q1\_hat}, \text{new data} = \{H_1, A_1 = A_1^{\mathsf{d}}\})$

3. Fit a generalized linear model (GLM) for $\tilde{Y}_1$ conditional on $A_0 = a_0$ and $H_0 = h_0$. Call this $\hat{Q}_0(a_0, h_0)$.

   $\texttt{Q0\_hat} \leftarrow \texttt{glm}(\text{outcome} = \texttt{pseudo\_Y1}, \text{predictors} = \{H_0, A_0\})$

4. Modify the data set used in step (3) so that the values in the column for A_0 are changed to A_0^{d}. Obtain the predictions for the model $\hat{Q}_0$ using this modified data set. These are pseudo-outcomes $\tilde{Y}_0$.

   $\texttt{pseudo\_Y0} \leftarrow \texttt{predict}(\text{fit} = \texttt{Q1\_hat}, \text{new data} = \{H_0, A_0 = A_0^{\mathsf{d}}\})$

5. Average $\tilde{Y}_0$, i.e. compute $\hat{\mathsf{E}}[\tilde{Y}_0]$.

Temp

->

Application examples

Illustrative application

Software/algos

Technicalities

Or **bold**, *italic*, or URL text.

Mathematics

Inline mode: $c^2 = a^2 + b^2$

Display mode:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Columns

We could also split content between two columns:

Left column

Right column

https://quarto.org/docs/presentations/revealjs/#multiple-columns

Code Highlighting

For continuous highlighting, use `from-to` (`6-8`).

For discontinuous highlighting, use `line1, line2, ...` (`1, 4`).

To highlight lines in a progressive manner, use `range1|range2` (`|1,4|6-8`).

```python
1  import numpy as np
2  import matplotlib.pyplot as plt
3  
4  r = np.arange(0, 2, 0.01)
5  theta = 2 * np.pi * r
6  fig, ax = plt.subplots(subplot_kw={'projection': 'polar'})
7  ax.plot(theta, r)
8  ax.set_rticks([0.5, 1, 1.5, 2])
9  ax.grid(True)
10 plt.show()
```

https://quarto.org/docs/presentations/revealjs/#code-blocks

Enable more revealjs features

The theme is built ontop of Quarto's `revealjs` implementation. So, any features of available are also usable within the theme. For example, if we wanted to incorporate the chalkboard feature. We would use:

```yaml
format:
  washington-revealjs:
    chalkboard: true
```

Summary

UW-themed presentation slide deck

The Quarto University of Washington Revealjs theme is an extension of Reveal.js and offers all of its features in the context of being brand friendly at UW.

Install the theme without this template:

```
quarto install extension kathoffman/quarto-uw
```

Install the theme with the template being present:

```
quarto use template kathoffman/quarto-uw
```

You can learn more about using RevealJS with Quarto at: https://quarto.org/docs/presentations/revealjs/

Díaz, Iván, Nicholas Williams, Katherine L. Hoffman, and Edward J. Schenck. 2021. "Nonparametric Causal Effects Based on Longitudinal Modified Treatment Policies." *Journal of the American Statistical Association*, September, 1–16. https://doi.org/10.1080/01621459.2021.1955691.

Hernán, Miguel A, Emilie Lanoy, Dominique Costagliola, and James M Robins. 2006. "Comparison of Dynamic Treatment Regimes via Inverse Probability Weighting." *Basic & Clinical Pharmacology & Toxicology* 98 (3): 237–42.