lmtp: An R Package for Estimating the Causal Effects of Modified Treatment Policies and Beyond

Nicholas Williams   ntw2117@cumc.columbia.edu
Department of Epidemiology
Mailman School of Public Health
Columbia University
New York, NY 10032, USA

Iván Díaz   ild2005@med.cornell.edu
Division of Biostatistics and Epidemiology
Department of Population Health Sciences
Weill Cornell Medicine
New York, NY 10065, USA

Abstract

The majority of causal inference methods consider treatment effects based on counterfactual outcomes where exposure is deterministically assigned. When exposure is continuous, deterministic interventions may be irrelevant and impossible to bring about. As a solution, modified treatment policies offer a generalization that allows for the study of feasible interventions and offer a safeguard against positivity violations. The lmtp package implements the estimators of **?** for estimating causal effects based on non-parametric modified treatment policies in R. In addition to modified treatment policies, the package can be used to estimate effects of deterministic and dynamic interventions. The methods provided can also be applied to both point-treatment and longitudinal settings, and can account for time-varying exposure, covariates, and right censoring, thereby providing a very general tool for causal inference in longitudinal studies. Additionally, two of the provided estimators are based on flexible machine learning regression algorithms, and avoid bias due to parametric model misspecification while maintaining valid statistical inference.

Keywords: Causal inference, longitudinal data, non-parametric, modified treatment policies

1. Introduction

Most modern causal inference methods consider the effects of a treatment on a population mean outcome under interventions that set the treatment value deterministically. For example, the average treatment effect (ATE) considers the hypothetical difference in a population mean outcome if a dichotomous exposure was applied to all observations versus if it was applied to none. In the case of a continuous exposure, interventions that set the exposure to a static value deterministically are of little practical relevance. Furthermore, 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 value under consideration (**?**). This assumption is often violated when evaluating the effects of

deterministic interventions, and is usually exacerbated with longitudinal data as the number of time points grows.

Modified treatment policies (MTPs) are a class of stochastic treatment regimes that can be formulated to avoid the above problems (**??**). In a recent article (**?**), we generalized the theoretical framework for estimation of the effect of MTPs to the longitudinal setting, accounting for time-varying treatment, covariates, and right-censoring of the outcome. Briefly, MTPs are hypothetical interventions where the post-intervention value of treatment can depend on the actual observed treatment level and the unit's history. As such, MTPs are useful to assess the effect of continuous treatments. For example, **?** assess the effect of reducing surgery time by a predetermined amount (e.g., 5 minutes) for lung cancer patients, where the reduction is carried out only for those patients for whom the intervention is feasible. Furthermore, MTPs generalize many important effect estimands, such as the effect of a dynamic treatment rule in which the treatment level is assigned as a function of a unit's history. For example, dynamic treatment rules, a particular case of MTPs, may be used to estimate the effect of policies such as switching HIV treatment once the CD4 T-cell count goes below a predetermined threshold (**?**). MTPs also generalize many interesting causal effects such as the average treatment effects, the causal risk ratio, and causal odds ratio. In this article we describe how lmtp can be used for estimating the causal effects of MTPs, and present examples on the use of the software for several of the above cases.

The package lmtp implements four methods for estimating the effects of MTP. Two of these estimators, a targeted minimum-loss based estimator (**??**) and a sequentially doubly-robust estimator (SDR) (**??????**), are multiply-robust. TMLE and SDR are implemented using cross-fitting to allow for the use of flexible machine learning regression methodology (**?**). The package may be download from CRAN at `cran.r-project.org/package=lmtp`.

2. Notation and modified treatment policies

2.1 Data structure

In this article, we will use the notation of **?** with only slight modifications. Let i be the index of an observation from a data set with n total units and t be the index of time for a total number of time points τ. The observed data for observation i may be denoted as

$$Z_i = (W, L_1, A_1, L_2, A_2, ..., L_\tau, A_\tau, Y_{\tau+1}) \tag{1}$$

where W denotes baseline covariates, L_t denotes time-varying covariates, A_t denotes a vector of exposure and/or censoring variables and Y denotes an outcome measured at the end of study follow-up. We observe n i.i.d. copies of Z with distribution P. We use $A_t = a_t$ to denote a realization of a random variable. If right-censoring exists, A_t can be adapted so that $A_t = (A_{1,t}, A_{2,t})$ where $A_{1,t}$ equals one if an observation is still in the study at time t and zero otherwise, and $A_{2,t}$ denotes the exposure at time t. We use an overbar to indicate the history of a variable up until time t. We then use $H_t = (\bar{L}_t, \bar{A}_{t-1})$ to denote the history of all variables up until just before A_t.

2.2 Modified treatment policies

We use the potential outcomes framework to define the causal effect of interest using our established data structure. We consider a hypothetical policy where $\bar{A}$ is set to a regime d defined as $A_t^d = d_t(A_t, H_t^d)$, where $H_t^d = (\bar{L}_t, \bar{A}_{t-1}^d)$, for a set of user-given regimes $d_t : t \in \{1, ..., \tau\}$. The defining characteristic that makes regime d_t a modified treatment policy is that it depends on the *natural value* of treatment $\bar{A}_t$, that is, the value that the treatment would have taken under no intervention. However, when the function d_t only depends on H_t, the LMTP reduces to the *dynamic treatment regimes* studied in the literature. Furthermore, when d_t is a constant that and does not depend on either A_t or H_t, then LMTPs reduce to the conventional static rules studied in the causal inference literature (e.g., **??**). Below we present examples of all these interventions.

First, consider a study of the effect of physical activity on mortality in the elderly. Assume that each patient is monitored at several time points, and that a measure of physical activity such as the metabolic equivalent of task (MET) (**?**) is measured together with a number of lifestyle, health status, and demographic variables. In this setup, a natural question to ask would be "what is the effect on mortality of an intervention that increases physical activity by δ units for patients whose socioeconomic and health status allows it?" Formally, consider a longitudinal study with loss-to-follow-up. Let $A_t = (A_{1,t}, A_{2,t})$ where $A_{1,t}$ equals one if an observation is still in the study at time t and zero otherwise, and $A_{2,t}$ denote a continuous exposure at time t that can be changed through some intervention. A modified treatment policy that increases $A_{2,t}$, whenever it is feasible to do so, can be defined as

$$d_t(a_t, h_t) = \begin{cases} (1, a_{2,t} + \delta_t) & \text{if } a_{2,t} + \delta_t \leq u_t(h_t) \\ (1, a_{2,t}) & \text{if } a_{2,t} + \delta_t > u_t(h_t) \end{cases} \tag{2}$$

where $u_t(h_t)$ defines the maximum level of physical activity allowed for a patient with characteristics h_t. Note that we also consider an intervention on $A_{1,t}$ because we are interested in a hypothetical world where there is no loss-to-follow-up. In this case the hypothetical exposure after intervention, A_t^d depends on the actually observed exposure, A_t. This is in contrast to a deterministic intervention where A_t^d would be set to some pre-specified value with probability one.

For dynamic treatment rules, consider a hypothetical longitudinal study where two different antiviral treatments are administered to HIV positive patients. Sometimes an antiviral drug works at first, until the virus develops resistance, at which point it is necessary to change the treatment regime. Assume we are interested in assessing a policy with two treatments encoded as $A_t \in \{0, 1\}$, and we want to assess the effect of a regime that would switch the antiviral treatment as soon as the CD4 T cell count drops bellow 300. Let $A_t = (A_{1,t}, A_{2,t})$ where $A_{1,t}$ equals one if an observation is still in the study at time t and zero otherwise, and $A_{2,t}$ denotes the treatment arm at time t. Let L_t denote the CD4 T cell count at time t. In this case, one may decide to assess the effect of the rule

$$d_t(h_t) = \begin{cases} (1, 1 - a_{2,t-1}) & \text{if } l_t < 300 \\ (1, a_{2,t-1}) & \text{if } l_t \geq 300. \end{cases} \tag{3}$$

In contrast to the previous rule (2), the dynamic treatment rule (3) does not depend on the natural value of treatment at time t, it only depends on the history. The software and methods presented here handle both cases seamlessly.

In the case of a single time point setting where the data structure is $Z = (W, A, Y)$, it follows trivially from the above definitions that the average treatment effect from a cross-sectional study, defined as $\mathsf{E}[Y(1) - Y(0)]$, can be estimated using MTPs by simply letting $\tau = 1$ and contrasting two MTPs $d(A) = 1$ and $d(A) = 0$. The **lmtp** package presented in this article allows the contrast of different MTPs using differences, ratios, and odds ratios. We provide examples of each of these contrasts in §4 below.

In what follows we focus on estimating the the causal effect of MTP d on outcome Y, using **lmtp**, through the causal parameter

$$\theta = \mathsf{E}\{Y(\bar{A}^d)\}, \tag{4}$$

where $Y(\bar{A}^d)$ is the counterfactual outcome in a world, where possibly contrary to fact, each entry of $\bar{A}$ was modified according to the MTP d. When Y is continuous, θ is the mean population value of Y under MTP d; when Y is dichotomous, θ is the population proportion of event Y under MTP d. Similarly, when Y is the indicator of an event by end of the study, θ is defined as the cumulative incidence of Y under MTP d.

2.3 Identification

The ability to estimate θ depends on the ability to identify an expression for θ as a function of the data generating distribution P using only the observed data Z and not counterfactual variables Y^d. A full review of these identification assumptions is outside the scope of this article but is presented in our technical paper (**?**). Briefly, the following standard assumptions must hold

Assumption 1 (Consistency) $\bar{A} = \bar{a} \implies Y = Y(\bar{a})$ *for all* $\bar{a} \in \operatorname{supp} \bar{A}$

Assumption 2 (Exchangeability) $A_t \perp\!\!\!\perp Y(\bar{a})|H_t$ *for all* $\bar{a} \in \operatorname{supp} \bar{A}$ *and* $t \in \{1, ..., \tau\}$

Assumption 3 (Positivity) *If* $(a_t, h_t) \in \operatorname{supp}\{A_t, H_t\}$ *then* $(d(a_t, h_t), h_t) \in \operatorname{supp}\{A_t, H_t\}$ *for* $t \in \{1, ..., \tau\}$

The consistency assumption states that the potential outcome for an observation under their observed exposure is the value of the outcome that we did actually observe. The exchangeability assumption is often also referred to as the no-unmeasured confounding assumption; it is satisfied if all common causes of the exposure and outcome are measured and adjusted for. Of particular importance to this article is the positivity assumption which states that the distribution of the exposure under the MTP is supported in the data. Concretely, in a study with a continuous exposure and loss-to-follow-up, the positivity assumption states that if an observation with covariate history h_t and exposure a_t who was not lost-to-follow-up at time t exists then there is also an observation with covariate history h_t who was not lost-to-follow-up at time t but whose exposure was observed as $d(a_t, h_t)$ that also exists. This assumption is often an issue when working with continuous exposures and/or multiple time points in the context of static or dynamic interventions.

One strength of MTPs is that if the support of the exposure variable is known, then they may be formulated to avoid violations of the positivity assumption.

Under these identification assumptions, the causal parameter θ is identified as follows. Set $\mathsf{m}_{\tau+1} = Y$. For $t = \tau, \ldots, 1$, recursively define

$$\mathsf{m}_t : (a_t, h_t) \mapsto \mathsf{E}\left[\mathsf{m}_{t+1}(A_{t+1}^d, H_{t+1}) \mid A_t = a_t, H_t = h_t\right]. \tag{5}$$

Then θ is identified as $\mathsf{E}[\mathsf{m}_1(A_1^d, L_1)]$. In addition to the function m_t, the description of the estimation methods below will benefit from the definition of the ratio between the density of the post-intervention exposure, A_t^d, and the density of the observed exposure, A_t. That is, define $\mathsf{g}_t^d(a_t \mid h_t)$ as the density of A_t^d conditional on $H_t = h_t$ evaluated at h_t. Likewise define $\mathsf{g}_t(a_t \mid h_t)$ as the conditional density of A_t. Our estimation procedure will make use of the density ratio defined as:

$$\mathsf{r}_t(a_t, h_t) = \frac{\mathsf{g}_t^d(a_t \mid h_t)}{\mathsf{g}_t(a_t \mid h_t)}.$$

This density ratio reduces to the well-known inverse probability weights used in estimation of the ATE in the single-time point setting where $Z = (W, A, Y)$ and A is binary.

3. Estimating modified treatment policy effects

3.1 Estimation methods

The **lmtp** package implements four estimation methods: a targeted minimum-loss based estimator (TMLE), a sequential doubly-robust estimator (SDR), an estimator based on the parametric G-formula, and an inverse probability weighted (IPW) estimator. We will only describe the use of TMLE, `lmtp_tmle`, and SDR, `lmtp_sdr`, as their use is strongly suggested over the others based on their advantageous theoretical properties which allow for machine learning regression while maintaining the ability to compute valid confidence intervals and p-values.

Targeted minimum-loss based estimation is a general framework for constructing asymptotically linear estimators leveraging machine learning, with an optimal bias-variance trade-off for the target causal parameter (**??**). In general, TMLE is constructed from a factorization of observed data likelihood into an outcome regression and an intervention mechanism. Using the outcome regression, an initial estimate of the target parameter is constructed and then *de-biased* by a fluctuation that depends on a function of the intervention mechanism. The sequential doubly-robust estimator is based on a unbiased transformation of the efficient influence function of the target estimand. For a thorough discussion of TMLE and SDR for static, dynamic, and modified treatment policies, we refer the reader to **????**.

TMLE and SDR require estimation of two nuisance parameters at each time point: an outcome mechanism and an intervention mechanism. Both TMLE and SDR are multiply-robust in that they allow certain configurations of nuisance parameters to be inconsistently estimated. Specifically, TMLE is considered $\tau + 1$-multiply robust in that it allows for inconsistent estimation of all the intervention mechanisms prior to any time point t, as long as all outcome mechanisms after time t are consistently estimated. SDR is 2^τ-robust in that at each time point, estimation of at most either the intervention mechanism or

outcome mechanism is allowed to be inconsistent. Both TMLE and SDR are efficient when all the treatment mechanism and outcome regression are consistently estimated at a given consistency rate, but the SDR has better protection against model misspecification (see **???**, for more details).

It is important to note that the SDR estimator can produce an estimate $\hat{\theta}$ outside of the bounds of the parameter space (e.g., probability estimates outside $[0, 1]$), while the TMLE guarantees that the estimate is within bounds of the parameter space. With this in mind and because for a single time-point TMLE and SDR are equally robust, we recommend use of TMLE for the case of a single time-point, while we recommend use of SDR for the longitudinal setting. All examples in this article will demonstrate use of both estimators.

3.2 Required data structure

Data is passed to **lmtp** estimators through the `data` argument. Data should be in wide format with one column per variable per time point under study (i.e., there should be one column for every variable in Z). These columns do not have to be in any specific order and the data set may contain variables that are not used in estimation. The names of treatment variables, censoring variables, baseline covariates, and time-varying covariates are specified using the `trt`, `cens`, `baseline`, and `time_vary` arguments respectively. The `trt`, `cens`, and `baseline` arguments accept character vectors and the `trt` and `cens` arguments should be ordered according to the time-ordering of the data generating mechanism. The `time_vary` argument accepts an unnamed list ordered according to the time-ordering of the model with each index containing the name of the time-varying covariates for the given time. The outcome variable is specified through the `outcome` argument.

Estimators are compatible with continuous, dichotomous and survival outcomes. In the case of a dichotomous or continuous outcome, only a single variable name should be passed to the `outcome` argument. For survival outcomes, a vector containing the names of the intermediate outcome and final outcome variables, ordered according to time, should be specified with the `outcome` argument. Dichotomous and survival outcomes should be coded using zero's and one's where one indicates the occurrence of an event and zero otherwise. If working with a survival outcome, once an observation experiences an outcome, all future outcome variables should also be coded with a one. The `outcome_type` argument should be set to `"continuous"` for continuous outcomes and `"binomial"` for dichotomous and survival outcomes. If the study is subject to loss-to-follow-up, the `cens` argument must be provided. Censoring indicators should be coded using zero's and one's where one indicates an observation is observed at the next time and zero indicates loss-to-follow-up. Once an observation's censoring status is switched to zero it cannot change back to one. Missing data before an observation is lost-to-follow-up is not allowed; a preprocessing step using multiple imputation is recommended for such variables.

The `k` argument controls a Markov assumption on the data generating mechanism. When `k = Inf`, the history H_t will be constructed using all previous time-point variables while setting `k` to any other value will restrict H_t to time-varying covariates from time $t - k - 1$ until $t-1$. Baseline confounders are always included in H_t. The `create_node_list` function may be used to inspect how variables will be used for estimation. It is specified with the same `trt`, `baseline`, `time_vary`, and `k` arguments as **lmtp** estimators and is used internally

to create a "node list" that encodes which variables should be used at each time point of estimation. For example, consider a study with the observed data structure

$$Z = (W_1, W_2, L_{1,1}, L_{1,2}, A_1, L_{2,1}, L_{2,2}, A_2, Y_3) \tag{6}$$

We can translate this data structure to R with

```
R> W <- c("W_1", "W_2")
R> A <- c("A_1", "A_2")
R> L <- list(c("L_11", "L_12"), c("L_21", "L_22"))
R> create_node_list(trt = A, baseline = W, time_vary = L, tau = 2)

$trt
$trt[[1]]
[1] "W_1"  "W_2"  "L_11" "L_12" "A_1"


$trt[[2]]
[1] "W_1"  "W_2"  "L_11" "L_12" "L_21" "L_22" "A_1"  "A_2"



$outcome
$outcome[[1]]
[1] "W_1"  "W_2"  "L_11" "L_12" "A_1"


$outcome[[2]]
[1] "W_1"  "W_2"  "L_11" "L_12" "A_1"  "L_21" "L_22" "A_2"
```

A list of lists is returned with the names of the variables in H_t to be used for estimation of the outcome regression and the treatment mechanism at every time t. Notice that variables A_1 and A_2 are included in the list of variables used for estimation of the treatment mechamism (`trt`). This is due to the fact that the nuisance parameter for the treatment mechanism is the density ratio r_t, which is a function of A_1 and A_2.

The density ratio is estimated based on a classification trick using an auxiliary variable Λ as a pseudo outcome and the treatment as a predictor. We now briefly describe how this density ratio estimation is done; the process is fully automated and hidden from the software user. Specifically, the TMLE and SDR estimation methods require estimation of the ratio of the densities of A_t^d and A_t, conditional on the history H_t, defined as r_t above. This is achieved through computing the odds in a classification problem in an augmented dataset with $2n$ observations where the outcome is the auxiliary variable Λ (defined below) and the predictors are the variables A_t and H_t. In the $2n$ augmented data set, the data structure at time t is redefined as

$$(H_{\lambda,i,t}, A_{\lambda,i,t}, \Lambda_{\lambda,i} : \lambda = 0, 1; i = 1, ..., n) \tag{7}$$

where $\Lambda_{\lambda,i} = \lambda_i$ indexes duplicate values. For all duplicated observations $\lambda \in \{0,1\}$ with the same i, $H_{\lambda,i,t}$ is the same. For $\lambda = 0$, $A_{\lambda,i,t}$ equals the observed exposure values $A_{i,t}$, whereas for $\lambda = 1$, $A_{\lambda,i,t}$ equals the exposure values under the MTP d, namely A_t^d. The

7

classification approach to density ratio estimation proceeds by estimating the conditional probability that $\Delta = 1$ in this dataset, and dividing it by the corresponding estimate of the conditional probability that $\Delta = 0$. Specifically, denoting p^λ the distribution of the data in the augmented dataset, we have:

$$\mathsf{r}_t(a_t, h_t) = \frac{\mathsf{p}^\lambda(a_t, h_t \mid \Lambda = 1)}{\mathsf{p}^\lambda(a_t, h_t \mid \Lambda = 0)} = \frac{\mathsf{p}^\lambda(\Lambda = 1 \mid A_t = a_t, H_t = h_t)}{\mathsf{p}^\lambda(\Lambda = 0 \mid A_t = a_t, H_t = h_t)}.$$

Further details on this algorithm may be found in our technical paper (**?**).

3.3 Machine Learnning

An attractive property of multiply-robust estimators is that they can incorporate flexible machine-learning algorithms for the estimation of nuisance parameters m_t and r_t while remaining $\sqrt{n}$-consistent. The super learner algorithm is an ensemble learner than incorporates a set of candidate models through a weighted convex-combination based on cross-validation (**?**). Asymptotically, this weighted combination of models, called the meta-learner, will outperform any single one of its components.

Our package uses the implementation of the super learner provided by the **sl3** package (**?**). Analysts must create **sl3** learner stacks which are then included in `lmtp_tmle` and `lmtp_sdr` calls with the `lrnrs_trt` and `lrnrs_outcome` arguments. The outcome variable type should guide users on selecting the appropriate candidate learners for use with the `lrnrs_outcome` argument. Regardless of whether an exposure is continuous, dichotomous, or categorical, the exposure mechanism is estimated using classification as discussed above, users should thus only include candidate learners capable of binary classification with the `lrnrs_trt` argument.

Candidate learners that rely on cross-validation for the tuning of hyper-parameters should support grouped data if used with `lrnrs_trt`. Because estimation of the treatment mechanism relies on the augmented $2n$ duplicated data set, duplicated observations must be put into the same fold during sample-splitting. This is done automatically by the package.

3.4 Additional arguments

Sample-splitting and cross-fitting is used with all methods to avoid certain technical conditions that may not hold for machine learning estimators (**??**). Specifically, the data are split in V volds and each nuisance parameter is estimated V times, excluding data in one fold each time. The number of folds V can be set with the `folds` argument; the minimum number of allowed folds is two. If data has a hierarchical structure, the `id` argument is used to indicate the name of a variable in the data set indicating unique groups. These identifiers will be used for generation of cross-validation folds and will be accounted for in standard error calculations. If a continuous outcome has known bounds, these bounds may be specified using the `bounds` argument with a length two numeric vector where the first index is the lower bound and the second index is the upper bound.

4. Examples

Example 1 *Longitudinal MTP with no loss-to-follow-up*

We have simulated data on $n = 5000$ observations over a 5-month period. Each observation has a continuous exposure (`A_1`, `A_2`, `A_3`, `A_4`) and covariate (`L_1`, `L_2`, `L_3`, `L_4`) recorded at months one through four and a dichotomous outcome (`Y`) at month five. We assume no loss-to-follow-up and no Markov assumption. This data set is installed with the package and is stored in the object `sim_t4`.

For this example, we are interested in the effect of a longitudinal MTP where at each month an observation's exposure decreases by one only if their observed exposure wouldn't be less than one if modified. Our data structure has no baseline confounders and we will use only GLMs for estimation so the only objects we must specify are the treatment variables, the time-varying covariates, the outcome variable, and the MTP shift function.

```
R> A <- c("A_1", "A_2", "A_3", "A_4")
R> L <- list(c("L_1"), c("L_2"), c("L_3"), c("L_4"))
R>
R> shift <- function(data, trt) {
R+   (data[[trt]] - 1) * (data[[trt]] - 1 >= 1) +
R+     data[[trt]] * (data[[trt]] - 1 < 1)
R+ }
R>
R> lmtp_tmle(sim_t4, A, "Y", time_vary = L, shift = f,
R+           intervention_type = "mtp", folds = 1)

LMTP Estimator: TMLE
   Trt. Policy: (f)

Population intervention effect
      Estimate: 0.2524
    Std. error: 0.0104
        95% CI: (0.2319, 0.2728)
```

Example 2 *Longitudinal MTP with right censoring*

For this example, we have a simulated dataset of $n = 1000$ observations. Data was simulated for three time points with a continuous time-varying exposure at times $t \in \{1, 2\}$ (`A1`, `A2`), a dichotomous time-varying covariate at times $t \in \{1, 2\}$ (`L1`, `L2`), and a dichotomous outcome (`Y`) at time $\tau + 1 = 3$. Loss-to-follow-up is present after time $t = 1$ so the data set contains censoring indicators (`C1`, `C2`). This data is installed with the package and is stored in the object `sim_cens`.

Suppose we are interested in the additive effect of an MTP where exposure is increased by 0.5 at every time point for all observations. Instead of using a linear model, we will estimate the outcome regression and treatment mechanism using a super learner composed of a generalized linear model, a random forest (**?**), and multivariate adaptive regression splines (**?**).

```
R> A <- c("A1", "A2")
R> C <- c("C1", "C2")
R> L <- list(c("L1"), c("L2"))
R> Y <- "Y"
R> shift <- function(data, trt) {
R+   data[[trt]] + 0.5
R+ }
R> lrnrs <- c("SL.glm", "SL.ranger", "SL.earth")
R> lmtp_sdr(sim_cens, A, Y, time_vary = L, cens = C, shift = shift,
R+          learners_trt = lrnrs, learners_outcome = lrnrs)

LMTP Estimator: SDR
   Trt. Policy: (shift)

Population intervention effect
      Estimate: 0.899
    Std. error: 0.0096
        95% CI: (0.8802, 0.9178)
```

If loss-to-follow-up exists, we can estimate the population mean outcome under the observed exposures adjusting for informative loss-to-follow-up by specifying `shift = NULL`.

```
R> lmtp_sdr(sim_cens, A, Y, time_vary = L, cens = C, shift = NULL,
R+          learners_trt = lrnrs, learners_outcome = lrnrs)

FILL IN
```

Example 3 *Survival analysis and deterministic effects*

The **lmtp** package may also be used to estimate deterministic causal effects, such as the causal relative risk. Suppose we have time-to-event data on $n = 2000$ observations with a time-invariant dichotomous exposure followed for a period of seven days. We wish to estimate the causal relative risk of experiencing the event by day seven. This data is installed with the package and is stored in the object `sim_point_surv`.

```
R> A <- "trt"
R> W <- c("W1", "W2")
R> C <- paste0("C.", 0:5)
R> Y <- paste0("Y.", 1:6)
R> a_1 <- lmtp_tmle(sim_point_surv, A, Y, W, cens = C,
R+                  learners_trt = lrnrs, learners_outcome = lrnrs,
R+                  shift = static_binary_on)
R> a_0 <- lmtp_tmle(sim_point_surv, A, Y, W, cens = C,
R+                  learners_trt = lrnrs, learners_outcome = lrnrs,
R+                  shift = static_binary_off)
R> lmtp_contrast(a_1, ref = a_0, type = "rr")
```

```
  LMTP Contrast: relative risk
Null hypothesis: theta == 1

  theta shift   ref std.error conf.low conf.high p.value
1  1.22  0.81 0.664   0.0339     1.14       1.3  <0.001
```

Example 4 *Dynamic treatment regimes*

Dynamic treatment regimes are treatment rules where treatment is applied based on a fixed rule that depends on covariate history. The **lmtp** package is capable of estimating the effects of deterministic dynamic treatment rules and modified treatment policies that depend on covariate history.

For this example we will use the same data from example 1. However, we will extend the longitudinal MTP used in that example so that a shift in the exposure also depends on covariate history and time. Specifically, if $t = 1$ exposure decreases by one if an observation's realized exposure wouldn't be less than one if modified. If $t > 1$ and the value of the time-varying covariate at that time-point equals one then exposure should decrease by one if an observation's realized exposure wouldn't be less than one if modified; otherwise exposure shouldn't be modified.

```
R> x <- c("A_1", "A_2", "A_3", "A_4")
R> tv <- list(c("L_1"), c("L_2"), c("L_3"), c("L_4"))
R> y <- "Y"
R> shift <- function(data, trt) {
R+   (data[[trt]] - 1) * (data[[trt]] - 1 >= 1) +
R+     data[[trt]] * (data[[trt]] - 1 < 1)
R+ }
R> dynamic_shift <- function(data, trt) {
R+   if (trt == "A_1") {
R+     shift(data, trt)
R+   } else {
R+     ifelse(data[[sub("A", "L", trt)]] == 1,
R+            shift(data, trt),
R+            data[[trt]])
R+   }
R+ }
R> lmtp_sdr(sim_t4, x, y, time_vary = tv, shift = dynamic_shift, folds = 3,
R+          learners_outcome = lrnrs, learners_trt = lrnrs)

LMTP Estimator: SDR
   Trt. Policy: (dynamic_shift)

Population intervention effect
     Estimate: 0.3177
    Std. error: 0.0099
        95% CI: (0.2983, 0.3371)
```

5. Summary