BIS 537 Final Project Report

Can Meng and Waveley Qiu

2023-12-14

1. Introduction

While the ideal setting to conduct causal inference on survival outcomes are randomized experiments, as is true for most areas of research, it is often the case that such an experimental setting is difficult or impossible to achieve. As such, when treatment and control groups are not exchangeable, it may be necessary to implement causal inference methods to address the absence of direct counterfactual proxies.

Time-to-event outcomes are of great interest to investigators across a variety of clinical settings, however present analytical difficulties even within the constrains of a randomized controlled trial. In particular, the non-observation of an event yet loss of a subject's data, or censoring, presents another missing data issue that the investigator has need to consider. This issue is compounded in the setting of a non-randomized or observational study, where exchangeability between treatment groups cannot be assumed of the collected data.

Several procedures addressing this dual-level missing data issue in the estimation of time-to-event causal effects have been explored previously. Some of these procedures extend the idea of inverse probability weighting to the setting of survival analysis in constructing a combined treatment and censoring weights (Hernan 2000, Cheng 2022) to approximate exchangeability and estimating the causal estimand through a weighted estimator. Other methods approach the problem by obtaining survival outcomes for each subject through a jackknife procedure (Andersen 2017) and estimate the causal estimand based on those pseudo-observed endpoints. The pseudo-observaions approach has been further improved upon with the employment of propensity score weighting to close the distance between the collected sample and the target population (Li 2021) as well as with the recent proposal of a doubly-robust estimation technique (Wang 2023) to allow for some flexibility with regards to model misspecification.

In this study, we are interested in revisiting the weighting methods and are proposing a new procedure in which model misspecification may be implicitly adjusted for. In doing so, we will first employ the two-arm weighting method established by Cheng et.al (2022), in which weights based on the probability of treatment and the probability of censoring are calculated and used together in estimating the causal estimand of interest. We will then compare its performance to the results obtained from a single-arm sequential weighting method that we have developed, where weights are constructed iteratively, and the final compositely-weighted dataset will then be used to estimate the causal estimand.

The remainder of this report is organized as follows: in section 2, we introduce our causal survival estimand; in section 3, we discuss our simulation design; in section 4, we illustrate our results from simulation study; and section 5 concludes with a discussion.

2. Methods

2.1 Causal Survival Estimand

While there are several causal estimands that could be of interest, we will be focusing our study on the restricted average causal effect (RACE) for ease of interpretation. In Mao et al. 2018, this was defined as follows:

$$\begin{aligned}\Delta_{RACE} &= \frac{E[\omega(e_i)\min(T_{1i}, t^*)]}{E[\omega(e_i)]} - \frac{E[\omega(e_i)\min(T_{0i}, t^*)]}{E[\omega(e_i)]} \\ &= \int_0^{t^*} S_1(t)dt - \int_0^{t^*} S_0(t)dt\end{aligned}$$

This estimand is interpreted as the average difference in survival time between the treatment and control groups, if both potential outcomes are observed, under the upper bound time restriction of t^*.

In this study, we are interested in the upper bound time restriction of 5 – thus, our estimand will take the following form:

$$\begin{aligned}\Delta_{RACE} &= \frac{E[\omega(e_i)\min(T_{1i}, 5)]}{E[\omega(e_i)]} - \frac{E[\omega(e_i)\min(T_{0i}, 5)]}{E[\omega(e_i)]} \\ &= \int_0^{5} S_1(t)dt - \int_0^{5} S_0(t)dt\end{aligned}$$

2.2 Causal Estimators

In the separate weighting procedure, to incorporate both the inverse probability treatment weights (IPTW) and the inverse probability of censoring weights (IPCW), we employed the usage of the non-

parametric weighted Kaplan-Meier estimator defined by Cheng et al. (2022). This estimator is defined as follows:

$$\begin{aligned}\hat{\Delta}(t) &= \hat{S}(1) - \hat{S}(0) \\ &= \left(1 - \frac{\sum_{i=1}^{n} \omega_{IPTW_i} A_i \delta_i I(U_i \leq t) \omega_{IPCW_i}}{\sum_{i=1}^{n} \omega_{IPTW_i} A_i}\right) - \left(1 - \frac{\sum_{i=1}^{n} \omega_{IPTW_i} (1 - A_i) \delta_i I(U_i \leq t) \omega_{IPCW_i}}{\sum_{i=1}^{n} \omega_{IPTW_i} (1 - A_i)}\right)\end{aligned}$$

where T_i^S is the event time of interest, T_i^C is the censoring time, $U_i = min(T_i^S, T_i^C)$ is the observed time in study, δ_i is the censoring indicator, A_i is the treatment indicator, $I(U_i \leq t)$ is the indicator for the observed time being less than the upper limit of time t, and ω_{IPTW_i} and ω_{IPCW_i} are the inverse pIPTW and IPCW respectively. Then, our estimator would be defined as follows, at $t = 5$:

$$\hat{\Delta}(5) = \left(1 - \frac{\sum_{i=1}^{n} \omega_{IPTW_i} A_i \delta_i I(U_i \leq 5) \omega_{IPCW_i}}{\sum_{i=1}^{n} \omega_{IPTW_i} A_i}\right) - \left(1 - \frac{\sum_{i=1}^{n} \omega_{IPTW_i} (1 - A_i) \delta_i I(U_i \leq 5) \omega_{IPCW_i}}{\sum_{i=1}^{n} \omega_{IPTW_i} (1 - A_i)}\right)$$

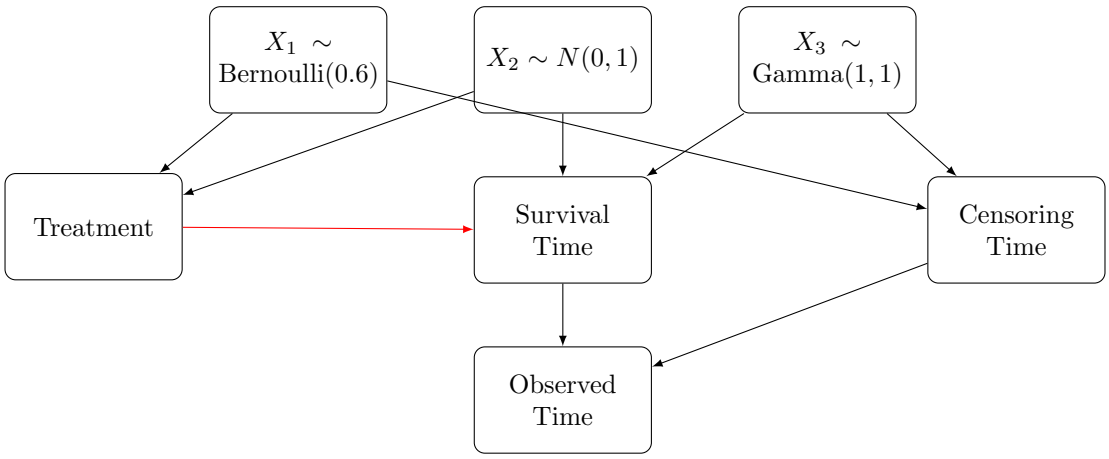
The estimators used in the sequentially weighted procedure will similarly be weighted Kaplan-Meier estimators – however, these will be distinct from the previously defined estimator in that there will be the need to adjust for one set of weights only. The estimator will thus have the following form, both generally and at $t = 5$:

$$\begin{aligned}\hat{\Delta}(t) &= \hat{S}(1) - \hat{S}(0) \\ &= \left(1 - \frac{\sum_{i=1}^{n} \omega_{IPTW \circ IPCW_i} A_i \delta_i I(U_i \leq t)}{\sum_{i=1}^{n} \omega_{IPTW \circ IPCW_i} A_i}\right) - \left(1 - \frac{\sum_{i=1}^{n} \omega_{IPTW \circ IPCW_i} (1 - A_i) \delta_i I(U_i \leq t)}{\sum_{i=1}^{n} \omega_{IPTW \circ IPCW_i} (1 - A_i)}\right) \\ \hat{\Delta}(5) &= \hat{S}(1) - \hat{S}(0) \\ &= \left(1 - \frac{\sum_{i=1}^{n} \omega_{IPTW \circ IPCW_i} A_i \delta_i I(U_i \leq 5)}{\sum_{i=1}^{n} \omega_{IPTW \circ IPCW_i} A_i}\right) - \left(1 - \frac{\sum_{i=1}^{n} \omega_{IPTW \circ IPCW_i} (1 - A_i) \delta_i I(U_i \leq 5)}{\sum_{i=1}^{n} \omega_{IPTW \circ IPCW_i} (1 - A_i)}\right)\end{aligned}$$

where $\omega_{IPTW \circ IPCW_i}$ is the composite treatment and censoring weight and all other notation is the same as defined previously.

3. Data Generation

We will generate data in accordance with the following relationships, and the causal effect of interest is highlighted in red:



Each component will be described in further detail in the following sections.

3.1 Covariates

As described in the previous diagram, we will be utilizing three covariates in this simulation study: X_1, X_2, X_3. These covariates have the following distributions: $X_1 \sim$ Bernoulli(0.6), $X_2 \sim N(0, 1)$, $X_3 \sim$ Gamma$(1, 1)$ and will explicitly be defined to not be time dependent.

In addition, these three covariates will all affect the critical measures of interest in this study (i.e., treatment, survival time, and censoring time). However, any given measure will only be affected by two of the three covariates and no two measures will be affected by the same combination of covariates.

3.2 True Propensity Score Model

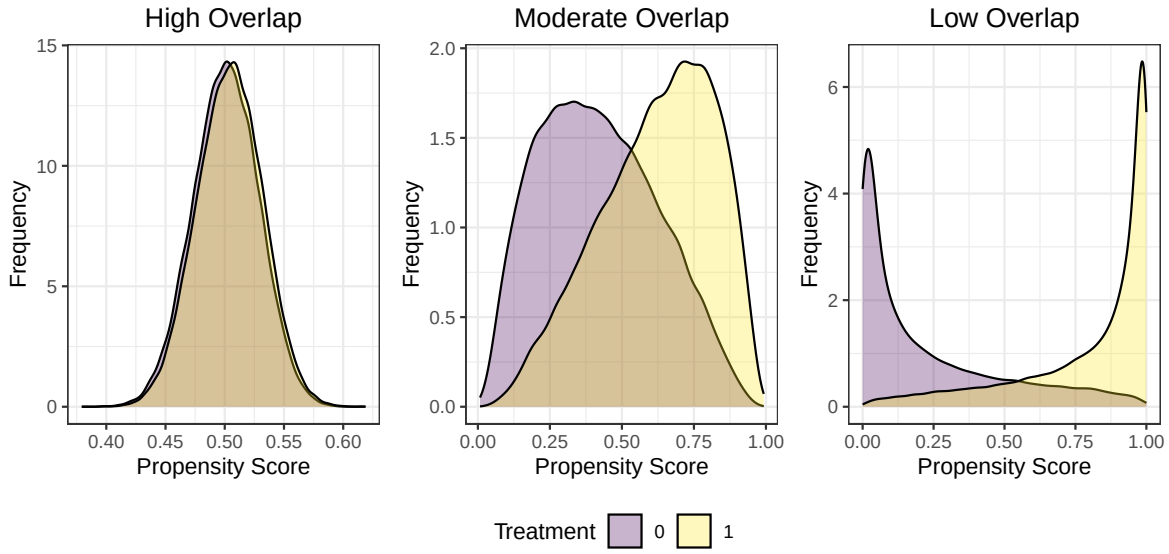
The true propensity score model will be defined using the following logistic regression, taking in covariates X_1 and X_2:

$$g^{-1}(E[Z = 1|\mathbf{X}_i]) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i},$$

$$\text{where } g(\mathbf{X_i}) = \frac{\exp(\boldsymbol{\beta}\mathbf{X}_i^T)}{1 + \exp(\boldsymbol{\beta}\mathbf{X}_i^T)} = E[Z = 1|\mathbf{X}_i] = e(\mathbf{X}_i)$$

In our simulation, β_1, β_2 will be held fixed at $\beta_1 = \beta_2 = 0.1$ for the setting of high overlap of overlap, $\beta_1 = \beta_2 = 1$ for the setting of medium overlap, and $\beta_1 = \beta_2 = 3$ for the setting of low overlap. For each of these settings, the β_0 that would produce an expected treatment proportion of 55% will be obtained through a numerical search over an interval.

The distribution of propensity scores according to each of these settings can be visualized with the following plots, demonstrating that the levels of overlap between treatment groups have indeed been achieved:



3.3 True Survival Time Model

The true survival time model will be defined using the following Cox-Weibull mode, taking in covariates X_1 and X_3:

$$\begin{aligned} h(t|\mathbf{X}_i) &= h_0(t)\exp(L_i), \\ \text{where } h_0(t) &= \lambda\nu t^{\nu-1}, \text{ and } L_i = a_0 Z_i + a_1 X_{1i} + a_3 X_{3i} \\ S(T_i^S = t) &= \exp(-\lambda t^v \exp(L_i)) \\ &= 1 - F(t) \end{aligned}$$

The survival time for subject i will then be drawn from:

$$T_i^S = \left(\frac{-\log(u_i^S)}{\lambda\exp(L_i)}\right)^{1/\nu},$$
$$\text{where } u_i^S \sim \text{Unif}(0,1)$$

From this model, the treatment effect will be determined by α_0. Two settings of α_0 were selected to correspond with a low and high treatment effect respectively.

To obtain the true treatment effects from each of these settings, we started with the conditional distributions of survival time that would be obtained from the defined Cox-Weibull model:

$$\begin{aligned} T(1)\Big|(X_1, X_3) &\sim h_{Z=1}(t|X_1, X_3) = \lambda v t^{v-1} \exp(\alpha_0 + \alpha_1 X_1 + \alpha_3 X_3) \\ &= \lambda^* \nu t^{\nu-1}, \text{ where } \lambda^* = \lambda \exp(\alpha_0 + \alpha_1 X_1 + \alpha_3 X_3) \\ T(0)\Big|(X_1, X_3) &\sim h_{Z=0}(t|X_1, X_3) = \lambda v t^{v-1} \exp(\alpha_1 X_1 + \alpha_3 X_3) \\ &= \lambda^* \nu t^{\nu-1}, \text{ where } \lambda^* = \lambda \exp(\alpha_1 X_1 + \alpha_3 X_3) \end{aligned}$$

Our target estimand is the average causal effect, or $E[T(1) - T(0)]$, which is unconditional on (X_1, X_3). By the law of total expectations and iterated expectations, we see the following:

$$\begin{aligned} E[T_i] &= E\Big[E\Big[T_i|X_1, X_3)\Big]\Big] \\ &= \int_{X_1, X_3} E\Big[T_i|X_1 = x_1, X_3 = x_3]P(X_1 = x_1, X_3 = x_3)d\mu(x_1, x_3) \\ &= \int_{X_1, X_3} E\Big[T_i|X_1 = x_1, X_3 = x_3]P(X_1 = x_1)P(X_3 = x_3)d\mu(x_1, x_3) \end{aligned}$$

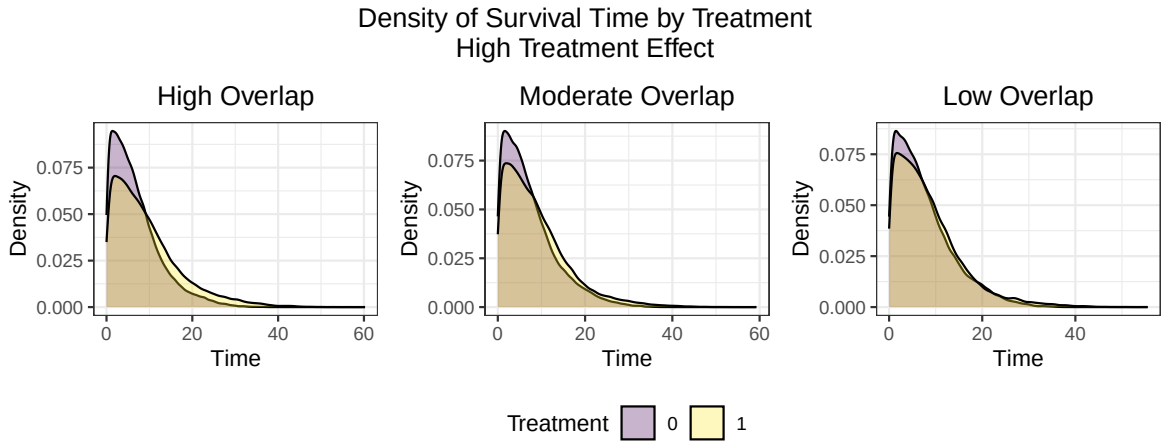
Let $X_{11}, ..., X_{1m} \sim P(X_1)$ and $X_{31}, ..., X_{3m} \sim P(X_3)$. Then, for a sufficiently large m, we note:

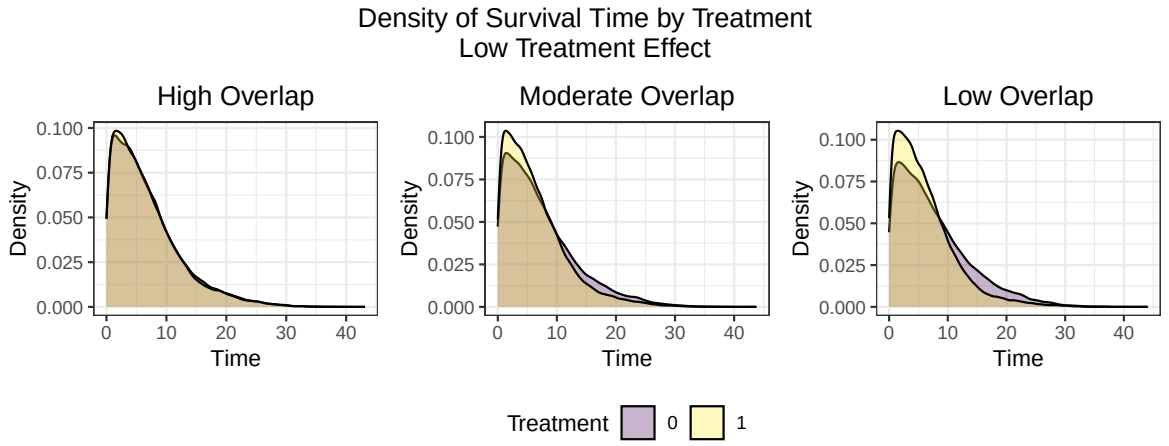
$$\begin{aligned} E[T_i] &= \int_{X_1, X_3} E\Big[T_i|X_1 = x_1, X_3 = x_3]P(X_1 = x_1)P(X_3 = x_3)d\mu(x_1, x_3) \\ &\approx \sum_{j=1}^{m} E[T_i|X_1 = x_{1j}, X_3 = x_{3j}] \end{aligned}$$

The true value estimates of $E[T(1)]$ and $E[T(0)]$ will be calculated by taking two samples of size $m = 1000000$ from the distributions of X_1 and X_3 to simulate the distributional behavior of each of these random variables, and sum the computed conditional expectations of $T(1)|(X_{1j}, X_{3j})$ and $T(0)|(X_{1k}, X_{3k})$ for each sample j and k, where $j = k = m$.

The true RACE values that we computed from this procedure were 0.127 for the "high" treatment effect setting and 0.014 for the "low" treatment effect setting.

The distribution of survival time according to each of these settings and by levels of covariate overlap between treatment groups can be visualized with the following plots:





3.4 True Censoring Model

The censoring model will be defined using the following exponential model, taking in covariates X_2 and X_3, and is generated independently of T_i^C, T_i^S, and Z_i:

$$T_i^C \sim \text{Exponential}(\lambda \exp(K_i)),$$
$$\text{where } K_i = \gamma_0 + \gamma_2 X_{2i} + \gamma_3 X_{3i}$$

Algorithmically, the censoring time for subject i will be drawn from:

$$T_i^C = \frac{-\log(u^C)}{\lambda \exp(K_i)},$$
$$\text{where } u^C \sim \text{Unif}(0, 1)$$

For subject i, the observed time, T^{obs} is the minimum of T_i^C and T_i^S, or $T^{obs} = \min(T_i^C, T_i^S)$. In addition,

the censoring indicator for subject i, given the survival time T_i and censoring time T^C will be assigned as follows:

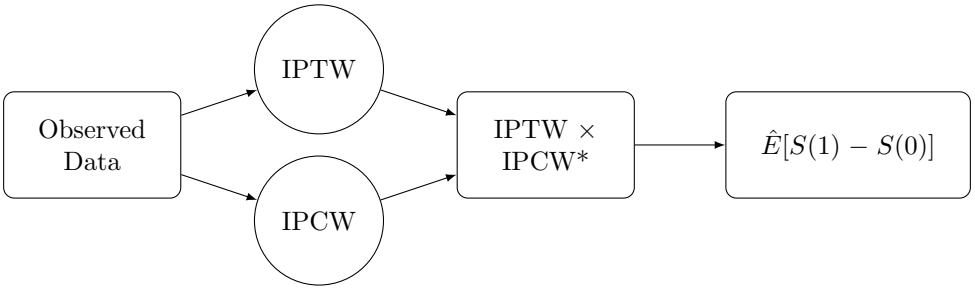
$$C_i(T_i^S, T_i^C) = \begin{cases} 1, & \text{where } T_i^S > T_i^C \\ 0. & \text{where } T_i^S \leq T_i^C \end{cases}$$

It is critical to note that although T_i^C is by definition independent of T_i^S and Z_i, the censoring indication, C_i is not.

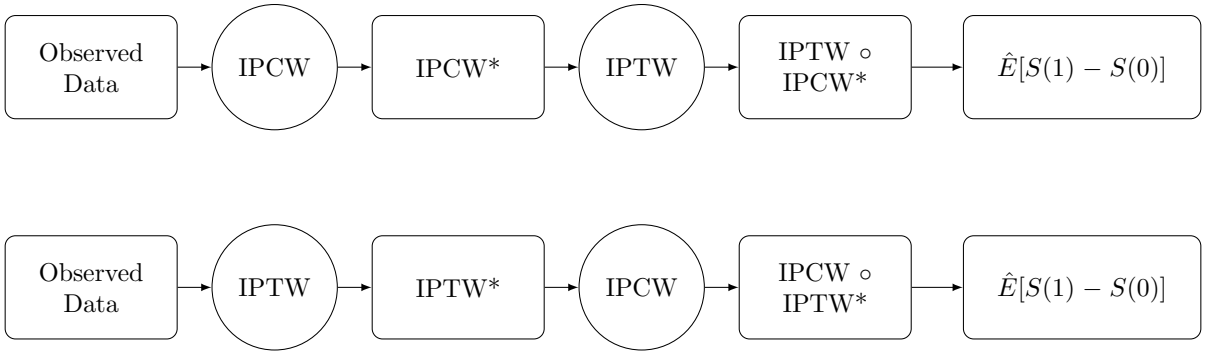
In a similar fashion to how the model intercept was selected in the true propensity score model in order to fix a treatment proportion in expectation, after fixing $\gamma_2 = 2$ and $\gamma_3 = 4$, a numerical search procedure was also run to select a γ_0 so that the proportion of censored subjects could be varied across two levels, corresponding to low (25%) and high (50%) levels of censoring. These selected intercepts varied according to the magnitude of treatment effect, as well as the defined censoring level, consequentially resulting in four defined γ_0 values corresponding to the four treatment effect and censoring level categories.

4. Analysis

4.1 Separate Weighting



4.2 Sequential Weighting



5. Results

6. Discussion