Supplementary Materials for "Estimands and their Estimators for Clinical Trials Impacted by the COVID-19 Pandemic: A Report from the NISS Ingram Olkin Forum Series on Unplanned Clinical Trial Disruptions"

A AIPW estimator for the neuroscience trial

For the neuroscience trial, the algorithm to obtain a double robust estimator is as follows:

1. At each time point t ($t \in \{1, \ldots, 8\}$), we estimate the conditional probability
$$P(C_t^* = 0 | A, \bar{\mathbf{C}}_{t-1}, \bar{\mathbf{X}}_{t-1}, \bar{Y}_{t-1})$$
by fitting a logistic regression among the patients with $\mathbf{C}_{t-1}^* = \mathbf{0}$.

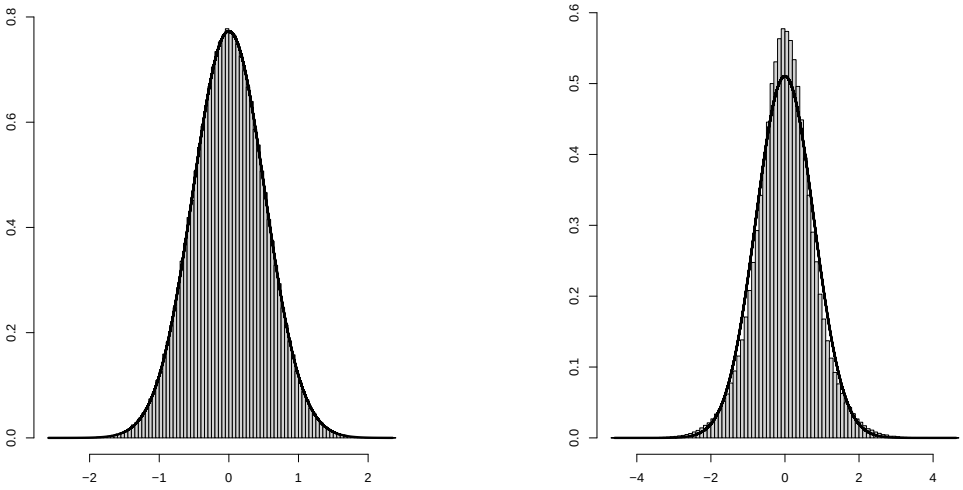
2. Fit a generalized linear model with canonical link (e.g., linear regression for continuous endpoint) for the outcome Y_8 among the treated ($A = 1$) complete cases given $\bar{\mathbf{X}}_7$ and $\bar{Y}_7$, using weights
$$\prod_{t=1}^{8} \frac{1}{P(C_t^* = 0 | A, \bar{\mathbf{C}}_{t-1}, \bar{\mathbf{X}}_{t-1}, \bar{Y}_{t-1})}.$$
Let $\hat{Y}_{i8}(\bar{\mathbf{X}}_{i7}, \bar{Y}_{i7})$ denote the fitted value for patient i (in the treatment arm) for whom no missing data is observed up to at least time 7 ($\mathbf{C}_7^* = \mathbf{0}$).

3. Recursively, for $t^* = 7, \ldots, 2, 1$: fit a generalized linear model with canonical link for $\hat{Y}_{it^*}(\bar{\mathbf{X}}_{i,t^*-1}, \bar{Y}_{i,t^*-1})$ among the patients with no missing data up to at least time t^* given $\bar{\mathbf{X}}_{t^*-1}$ and $\bar{Y}_{t^*-1}$ using weights
$$\prod_{t=1}^{t^*} \frac{1}{P(C_t^* = 0 | A, \bar{\mathbf{C}}_{t-1}, \bar{\mathbf{X}}_{t-1}, \bar{Y}_{t-1})}.$$
Let $\hat{Y}_{it^*}(\bar{\mathbf{X}}_{i,t^*-1}, \bar{Y}_{i,t^*-1})$ denote the fitted value for patient i (in the treatment arm) for whom no missing data is observed up to at least time $t^* - 1$ ($\mathbf{C}_{t^*-1}^* = \mathbf{0}$).

4. Take the sample average of the fitted values $\hat{Y}_8(X_0, Y_0)$ over **all** patients (treated and untreated).

B A Monte-Carlo Study on Combining Unbiased and Possibly Biased Estimators.

To illustrate the relative benefits of $\hat{\theta}^0(0)$ and $\hat{\theta}^0(\hat{\delta})$ versus $\hat{\theta}$, consider an illustrative Monte-Carlo study with the statistical model generated two samples: (1) the pre-pandemic sample, which is a sample with 100 paired standard normal random variables (X_1 representing the primary endpoint and Y_1 representing a surrogate endpoint) with correlation $cor(X_1, Y_1) = 0.9$ and (2) the pandemic sample with 1000 standard normal random variables (X_2 - surrogate endpoint assessed during the pandemic). The objective is to estimate the mean of Y ($EY = \theta$), which is equal to zero in this example.



(a) $\sqrt{100}\hat{\theta}^0(0)$ (b) $\sqrt{100}\hat{\theta}^0(\hat{\delta})$

Figure 1: Histogram and a normal approximation of the distribution; 500, 000 Monte-Carlo simulations.

The asymptotic distribution of $\hat{\theta}$ (mean of Y_1) is approximately normal, so that $\sqrt{100} \cdot \hat{\theta} \sim N(0, 1)$ leads to the width of 95% for $\sqrt{100} \cdot \hat{\theta}$ equal to $3.92 (= 2 \cdot 1.96)$. The asymptotic distribution of $\sqrt{100} \cdot \hat{\theta}^0(0)$ is also approximately normal with zero mean and variance $= 0.266358$, see Figure 1a. The distance between 2.5% and 97.5% level quantiles of the distribution of $\sqrt{100} \cdot \hat{\theta}^0(0)$ is equal to 2.03227. Wald's confidence interval ("mean estimate" ± 1.96 "standard deviation of the estimate") has an almost identical length ($= 2.023107$).

The asymptotic distribution of $\sqrt{100} \cdot \hat{\theta}^0(\hat{\delta})$ is not normal anymore and is shown in Figure 1b. The normal approximation allows us to visually evaluate the departure from normality. The absence of asymptotic normality, however, is not really a problem. Since

the asymptotic distribution is known it can be used for estimation, hypothesis testing, and for calculating confidence intervals. For example, the distance between the 2.5% and 97.5% level quantiles of the distribution of $\sqrt{100} \cdot \hat{\theta}^0(\hat{\delta})$ is equal to 3.20191. Wald's confidence interval has a shorter length (= 3.064861) associated with a less than 95% coverage.

This Monte-Carlo study demonstrates that if a data analyst is confident that pandemic data on a surrogate endpoint is unbiased, then it should be incorporated using minimum variance estimation ($\hat{\theta}^0(0)$). If, however, it can be biased, $\hat{\theta}^0(\hat{\delta})$ is a more appropriate method.