Estimating PATT with noncompliance

April 18, 2015

1 Assumptions

Let Y_{ist} be the potential outcome for individual i in group s, where $s = 0$ for the population and $s = 1$ for the randomized control trial, and t be the treatment assigned. Let T_i denote the treatment assigned and D_i denote treatment received. Let W_i be individual i's observable covariates related to the sample selection mechanism for membership in the RCT. Let C_i be an indicator for individual i's compliance to treatment. For a generic value, we drop the subscript i.

We assume that in the RCT, treatment is assigned at random. Then for individuals with $C_i = 1$, we observe $D_i = T_i$. In the population, we suppose that treatment is made available to individuals according to some rule based on their covariates; treatment assignment is not completely random. Individuals with $T_i = 0$ do not receive treatment, while those with $T_i = 1$ may decide whether or not to accept treatment. We only observe D, not T, in the population. Among the population controls, we can't distinguish non-compliers (individuals with $T_i = 1$ and $D_i = 0$) from compliers (those with $T_i = 0$ and $D_i = 0$). Compliance is only observable for individuals assigned to treatment in the RCT.

We make the following assumptions:

- Consistency under parallel studies: for all i and for $t = 0, 1$,

$$Y_{i0t} = Y_{i1t} \tag{1}$$

- Strong ignorability of sample assignment for treated:

$$(Y_{01}, Y_{11}) \perp\!\!\!\perp S \mid (W, T = 1, C = 1), 0 < \mathbb{P}(S = 1 \mid W, T = 1, C = 1) < 1 \tag{2}$$

  This means that the potential outcomes for treatment are independent of sample assignment for individuals with the same covariates W and assignment to treatment.

- Strong ignorability of sample assignment for controls:

$$(Y_{00}, Y_{10}) \perp\!\!\!\perp S \mid (W, T = 1, C = 1), 0 < \mathbb{P}(S = 1 \mid W, T = 1, C = 1) < 1 \tag{3}$$

  This means that the potential outcomes for control are independent of sample assignment for individuals with the same covariates W and assignment to treatment.

- Stable unit treatment value assumption (SUTVA):

$$Y_{ist}^{L_i} = Y_{ist}^{L_j}, \forall i \neq j \tag{4}$$

  where L_j is the treatment and sample assignment vector for unit j. This means that the treatment assignment for all other individuals j does not affect the potential outcomes of individual i.

- Conditional independence of compliance and assignment:

$$C \perp\!\!\!\perp T = 1 \mid W, 0 < \mathbb{P}(C = 1 \mid W) < 1 \tag{5}$$

  This means that compliance is independent of sample and treatment assignment for all individuals with covariates W.

- Monotonicity:

$$T_i \geq D_i, \forall i \tag{6}$$

  This assumption implies that there are no defiers and that crossover is only possible from treatment to control.

- Exclusion restriction (ER): For non-compliers

$$Y_{11} = Y_{10} \tag{7}$$

  A treatment effect may only be non-zero for compliers.

2 Population Average Treatment Effect on the Treated

2.1 Parameter

The estimand of interest is

$$\tau_{\text{PATT}} = \mathbb{E}\left(Y_{01} - Y_{00} \mid S = 0, D = 1\right) \tag{8}$$

This is the average treatment effect on those in the population who receive treatment. It includes individuals who actually receive the treatment, but does not include those who are eligible for treatment and do not accept it (non-compliers).

Theorem 1. *Under assumptions* (1) - (7)*,*

$$\tau_{PATT} = \mathbb{E}_{01}\left[\mathbb{E}\left(Y_{11} \mid S = 1, D = 1, W\right)\right] - \mathbb{E}_{01}\left[\mathbb{E}\left(Y_{10} \mid S = 1, T = 0, C = 1, W\right)\right]$$

where $\mathbb{E}_{01}\left[\mathbb{E}(\cdot \mid \ldots, W)\right]$ *denotes the expectation with respect to the distribution of* W *in the target population.*

Proof. We separate the expectation into two terms and treat each individually.

$$\begin{aligned}
\mathbb{E}\left(Y_{01} \mid S = 0, D = 1\right) &= \mathbb{E}\left(Y_{11} \mid S = 0, D = 1\right) && \text{by (1)} \\
&= \mathbb{E}\left(Y_{11} \mid S = 0, T = 1, C = 1\right) && \text{by (6)} \\
&= \mathbb{E}_{01}\left[\mathbb{E}\left(Y_{11} \mid S = 1, T = 1, C = 1, W\right)\right] && \text{by (2)} \\
&= \mathbb{E}_{01}\left[\mathbb{E}\left(Y_{11} \mid S = 1, D = 1, W\right)\right]
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}\left(Y_{00} \mid S = 0, D = 1\right) &= \mathbb{E}\left(Y_{10} \mid S = 0, D = 1\right) && \text{by (1)} \\
&= \mathbb{E}\left(Y_{10} \mid S = 0, T = 1, C = 1\right) && \text{by (6)} \\
&= \mathbb{E}_{01}\left[\mathbb{E}\left(Y_{10} \mid S = 1, T = 1, C = 1, W\right)\right] && \text{by (3)} \\
&= \mathbb{E}_{01}\left[\mathbb{E}\left(Y_{10} \mid S = 1, T = 0, C = 1, W\right)\right]
\end{aligned}$$

The last line follows because of the randomization carried out in the RCT. This ensures $Y_{10} \perp\!\!\!\perp T \mid (W, S = 1)$. □

2.2 Estimation

Theorem 1 poses two challenges in practice. First, we must construct an estimate of the inner expectation over potential outcomes in the RCT. Here, we use the SuperLearner ensemble method to estimate the response curve for compliers, given their treatment assignment and covariates. We estimate the outer expectation by taking empirical means. Thus, the first term in the expression for τ_{PATT} is estimated by the weighted average of mean responses in the treatment group in the RCT. The second term is estimated by the weighted average of the mean control response for compliers assigned to control in the RCT. We compute these by evaluating the response curve at each point defined by a complier in the population.

The second challenge is that we cannot observe which individuals are included in the estimation of the second term. We cannot tell who is a complier or non-complier in the control group, as they receive $D = 0$ in either case. We must estimate the second term by predicting who in the control group would be a complier, had they been assigned to treatment. The exact procedure for classification isn't important, as long as predictions are made as accurate as possible.