1 Definitions

$\hat{\tau}$ is the estimate of the average treatment effect on the treated, calculated by nearest-neighbor matching on the original sample.

$\hat{\tau}^b$ is the estimate of the average treatment effect on the treated, calculated by nearest-neighbor matching on the bootstrapped (wild) sample.

$\hat{\tau}_i$ and $\hat{\tau}_i^b$ are the difference between $Y_i(1)$ and $\widehat{Y_i(0)}$ on the original and bootstrapped samples, respectively.

ε_i is the residual for each treated unit i - that is, $\varepsilon_i = Y_i(1) - \hat{\tau} - \widehat{Y_i(0)}$.

ε_i^* is a random variable that takes on values ε_i and $-\varepsilon_i$ with equal probability.

2 Preamble

Throughout, I am conditioning on $\{Y, W, X\}$ - the entire original sample. I suppress the notation for this in the interests of legibility.

Without loss of generality, the sample is ordered so that the N_1 treated units are indexed $i \in \{1, 2, ..., N_1\}$ and the N_0 untreated units are indexed $i \in \{N_1+1, N_1+2, ..., N_1+N_0\}$.

3 Variance of $\hat{\tau}^b$

As in the original sample, $\hat{\tau}^b$ is constructed by averaging $\hat{\tau}_i^b$ over all treated units:

$$\hat{\tau}^b = \frac{1}{N_1}\sum_{i=1}^{N_1}\hat{\tau}_i^b \tag{1}$$

By the definition of $\hat{\tau}_i^b$, (1) can be rewritten as:

$$\begin{aligned}
\hat{\tau}^b &= \frac{1}{N_1}\sum_{i=1}^{N_1}\left(Y_i(1) + \varepsilon_i^* - \widehat{Y_i(0)}\right) && (2)\\
&= \frac{1}{N_1}\sum_{i=1}^{N_1}Y_i(1) + \frac{1}{N_1}\sum_{i=1}^{N_1}\varepsilon_i^* - \frac{1}{N_1}\sum_{i=1}^{N_1}\widehat{Y_i(0)}\\
&= \frac{1}{N_1}\sum_{i=1}^{N_1}Y_i(1) + \frac{1}{N_1}\sum_{i=1}^{N_1}\varepsilon_i^* - \frac{1}{N_1}\sum_{i=1}^{N_1}\widehat{Y_i(0)} && (3)
\end{aligned}$$

Thus:

$$\begin{aligned}
\mathbb{E}\left[\hat{\tau}^b\right] &= \mathbb{E}\left[\frac{1}{N_1}\sum_{i=1}^{N_1}Y_i(1) + \frac{1}{N_1}\sum_{i=1}^{N_1}\varepsilon_i^* - \frac{1}{N_1}\sum_{i=1}^{N_1}\widehat{Y_i(0)}\right]\\
&= \mathbb{E}\left[\frac{1}{N_1}\sum_{i=1}^{N_1}Y_i(1)\right] + \mathbb{E}\left[\frac{1}{N_1}\sum_{i=1}^{N_1}\varepsilon_i^*\right] - \mathbb{E}\left[\frac{1}{N_1}\sum_{i=1}^{N_1}\widehat{Y_i(0)}\right] && (4)
\end{aligned}$$

By the Law of Large Numbers and the fact that ε_i^* is mean-zero:

$$\begin{aligned}\mathbb{E}\left[\hat{\tau}^b\right] &= \mathbb{E}\left[Y_i(1)\right] + 0 - \mathbb{E}\left[\widehat{Y_i(0)}\right] \\ &= \mathbb{E}\left[Y_i(1)\right] - \mathbb{E}\left[\widehat{Y_i(0)}\right] = \hat{\tau} \end{aligned} \tag{5}$$

Thus:

$$\begin{aligned}\mathbb{V}\left[\hat{\tau}^b\right] &= \mathbb{E}\left[\left(\hat{\tau}^b - \mathbb{E}\left[\hat{\tau}^b\right]\right)^2\right] = \mathbb{E}\left[\left(\hat{\tau}^b - \hat{\tau}\right)^2\right] \\ &= \mathbb{E}\left[\left(\frac{1}{N_1}\sum_{i=1}^{N_1} Y_i(1) + \frac{1}{N_1}\sum_{i=1}^{N_1}\varepsilon_i^* - \frac{1}{N_1}\sum_{i=1}^{N_1}\widehat{Y_i(0)} - \hat{\tau}\right)^2\right] \end{aligned} \tag{6}$$

This next step I view as somewhat shaky - the argument is that $\hat{\tau}$ is numerically identical to $\frac{1}{N_1}\sum_{i=1}^{N_1} Y_i(1) - \frac{1}{N_1}\sum_{i=1}^{N_1}\widehat{Y_i(0)}$:

$$\mathbb{V}\left[\hat{\tau}^b\right] = \mathbb{E}\left[\left(\frac{1}{N_1}\sum_{i=1}^{N_1}\varepsilon_i^*\right)^2\right] \tag{7}$$

The problem here is that I think I can go from (7) (by noting that $\varepsilon_i^*\varepsilon_j^* = 0$ for $i \neq j$) to:

$$\mathbb{V}\left[\hat{\tau}^b\right] = \frac{1}{N_1}\mathbb{E}\left[\frac{1}{N_1}\sum_{i=1}^{N_1}(\varepsilon_i^*)^2\right] \tag{8}$$

But now an application of the LLN gives that the expectation converges to $\mathbb{E}\left[(\varepsilon_i^*)^2\right]$, but this is 0, so (8) is 0, which can't be right.