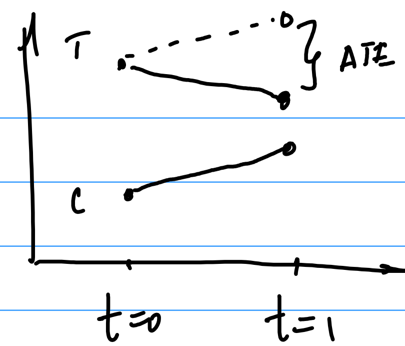
Diff - in - diff



$$y_i = \begin{cases} y_i(1), & D_i = 1 \\ y_i(0), & D_i = 0 \end{cases}$$

$$y_i = y_i(0) + D_i \underbrace{(y_i(1) - y_i(0))}_{\text{individual TE}}$$

$$E(y_i(1) - y_i(0)) - \text{ATE}$$

$$E(y_i \mid D_i = 1) - E(y_i \mid D_i = 0) =$$

$$E(y_i(1) \mid D_i = 1) - E(y_i(0) \mid D_i = 0) =$$

$$E(y_i(1) \mid D_i = 1) - E(y_i(0) \mid D_i = 1) +$$

$$+ E(y_i(0) \mid D_i = 1) - E(y_i(0) \mid D_i = 0) =$$

$$\underbrace{E(y_i(1) - y_i(0) \mid D_i = 1)}_{\text{ATET}} +$$

$$\underbrace{E(y_i(0) \mid D_i = 1) - E(y_i(0) \mid D_i = 0)}_{\text{selection bias}}$$

$$\widehat{\text{ATET}} = \bar{y}_1 - \bar{y}_0$$

$$\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 \cdot D_i$$

$$\hat{\beta}_1 = \bar{y}_0 \qquad \hat{\beta}_2 = \bar{y}_1 - \bar{y}_0$$

Ex. 1

$$\hat{\beta}_2 = \frac{\hat{Cov}(y, D)}{\hat{Var}(D)} = \frac{\overline{Dy} - \bar{D} \cdot \bar{y}}{\overline{D^2} - (\bar{D})^2}$$

$$\overline{Dy} - \bar{D} \cdot \bar{y} = \frac{0 \cdot \overset{\# n_0}{\sum_{D_i=0}} y_i + 1 \cdot \overset{\# n_1}{\sum_{D_i=1}} y_i}{n}$$

$$- \frac{n_1}{n} \cdot \frac{\sum_{D_i=0} y_i + \sum_{D_i=1} y_i}{n} =$$

$$\frac{(n_0 + n_1) \cdot \sum_{D_i=1} y_i - n_1 \cdot \sum_{D_i=0} y_i - n_1 \cdot \sum_{D_i=1} y_i}{n^2} =$$

$$\frac{n_0 \sum_{D_i=1} y_i - n_1 \cdot \sum_{D_i=0} y_i}{n^2}$$

$$\overline{D^2} - (\bar{D})^2 = \frac{n_1}{n} - \left(\frac{n_1}{n}\right)^2 = \frac{n_1(n_0 + n_1) - n_1^2}{n^2}$$

$$= \frac{n_0 n_1}{n^2}$$

$$\hat{\beta}_2 = \frac{n_0 \sum_{D_i=1} y_i - n_1 \cdot \sum_{D_i=0} y_i}{n^2} : \frac{n_0 n_1}{n^2} =$$

$$= \frac{\sum_{D_i=1} y_i}{n_1} - \frac{\sum_{D_i=0} y_i}{n_0} = \bar{y}_1 - \bar{y}_0$$

$$\hat{\beta}_1 = \bar{y} - \hat{\beta}_2 \cdot \bar{D} = \frac{\sum_{D_i=0} y_i}{n} + \frac{\sum_{D_i=1} y_i}{n} + (\bar{y}_1 - \bar{y}_0) \cdot \frac{n_1}{n} =$$

$$= \frac{1}{n} \cdot \left(1 + \frac{n_1}{n_0}\right) \cdot \sum_{D_i=0} y_i = \frac{\sum_{D_i=0} y_i}{n} = \bar{y}_0$$

Ex.2

$$\operatorname{Var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum (D_i - \bar{D})^2} = \frac{\sigma^2}{\sum (D_i - \alpha)^2} = \qquad \alpha = \frac{n_1}{n}$$

$$= \frac{\sigma^2}{n_1 (1-\alpha)^2 + n_0 \cdot \alpha^2} = \frac{\sigma^2}{n\alpha(1-\alpha)^2 + n(1-\alpha)\alpha^2} =$$

$$= \frac{\sigma^2}{n\alpha(1-\alpha)} \qquad \alpha = 1/2$$

DID

D.V. for treatment group

D.V. for time ($t=1$)

$$y_{it} = \beta_0 + \beta_1 \cdot X_i \cdot \beta_2 \cdot Z_i + \beta_3 \cdot X_i \cdot Z_i + \varepsilon_{it}$$

DID

$$\Delta y_i = \beta_0 + \delta \cdot X_i + \varepsilon_{it}$$

Matching

$y_{1,k}$ - treatment group $\qquad E(y_i(1) - y_i(0))$

$y_{0,k}$ - control group

1. Simple matching

$$\Delta^M = \sum_k w_k (\bar{y}_{1,k} - \bar{y}_{0,k})$$

2. Nearest Neighbour matching

$$A_i = \{ j \mid \min_j \|x_i - x_j\| \}$$

3. Propensity Score matching

$$P(D_i = 1 \mid x_i^{(1)}, \dots, x_i^{(K)})$$