8월 21일 Simulation code 내용 정리

1. Moodie et al이 제시한 DR Estimator of ATT와 coding한 ATT Estimator가 동일함을 확인. --- ok!

1) R coding 때 참고한 DR Estimator for ATT (MercantantT and LT, 2014)

$$\hat{\tau}_{dr}^{ATT} = \sum_{i=1}^{N} \left[Y_i Z_i - \frac{Y_i (1-Z_i)\hat{e}_i + \hat{m}_0(X_i)(Z_i - \hat{e}_i)}{1-\hat{e}_i} \right] / N_1$$

N_1: Z=1인 개체들의 수

$\hat{e}_i$: Propensity Score

$\hat{m}_0(X_i)$: $E[\hat{Y|Z=0, X}]$

2) Moodie, et al 이 제시한 DR Estimator for ATT

$$E[\hat{Y^1|Z=1}] = \frac{\sum_{i=1}^{n} Z_i Y_i}{\sum_{i=1}^{n} Z_i}, \quad E[\hat{Y^0|Z=1}] = \frac{\sum_{i=1}^{n} (1-Z_i)\hat{w}(X_i)(Y_i - \hat{\mu}(0, X_i)) + Z_i \hat{\mu}(0, X_i)}{\sum_{i=1}^{n} Z_i}$$

/ 여기서, $w(X_i) = \frac{\hat{e}_i}{1-\hat{e}_i}$, $\hat{\mu}(0, X_i) = E[\hat{Y|Z=0, X_i}]$

$\hat{\mu}(0, X_i)$: $m_0(X_i)$와 동일

⇒ $m_0(X_i)$로 표기

Check

① $E[\hat{Y^1|Z=1}]$은 ok!

② $\sum_{i=1}^{n} (1-Z_i)\left(\frac{e_i}{1-e_i}\right)(Y_i - m_0(X_i)) + Z_i \times m_0(X_i) = \sum_{i=1}^{n} \frac{(1-Z_i)e_i(Y_i - m_0(X_i))}{1-e_i} + Z_i \times m_0(X_i) = \sum_{i=1}^{n} \frac{(1-Z_i)e_i(Y_i - m_0(X_i)) + (1-e_i)Z_i m_0(X_i)}{1-e_i}$

$= \sum_{i=1}^{n} \frac{Y_i(1-Z_i)e_i - m_0(X_i)e_i(1-Z_i) + (1-e_i)Z_i m_0(X_i)}{1-e_i}$

$= -m_0(X_i)e_i + \cancel{m_0(X_i)e_i} Z_i + Z_i m_0(X_i) - \cancel{m_0(X_i)e_i} Z_i$

$= m_0(X_i)(Z_i - e_i)$

$= \sum_{i=1}^{n} \frac{Y_i(1-Z_i)e_i + m_0(X_i)(Z_i - e_i)}{1-e_i}$ ⇒ 동일!

2. Variance estimator of DR Estimator of ATE & ATT 함수 생성 위한 공식 정리

: 참고한 부분이 강의노트와 Moodie et al 논문의 Appendix (참고 : M-estimation은 loss function을 최소화하는 추정량을 찾는 방법이다.)

: 이때, 두 부분에서 제시한 estimating function이 같은지 의문이 듦. --- 같지 않음!!!!!! 강의 노트 estimating equation 사용하자.

$$\sum_{i=1}^{N} \Psi_i(\theta) = \sum_{i=1}^{N} \begin{bmatrix} \nu_1 - \{Z_i Y_i - (Z_i - e_i)\mu_1(X_i; \alpha_1)\}/e_i \\ \nu_0 - \{(1 - Z_i)Y_i + (Z_i - e_i)\mu_0(X_i; \alpha_0)\}/(1 - e_i) \\ Z_i S_1(Y_i, X_i; \alpha_1) \\ (1 - Z_i)S_0(Y_i, X_i; \alpha_0) \\ S_\beta(X_i; \beta) \end{bmatrix} = 0$$

where S_1, S_0, S_β are score function of the outcome models and PS model, $\theta = (\nu_1, \nu_0, \alpha_0', \alpha_1', \beta')'$

$$U_i(\beta, \alpha) = \{\omega_i(\alpha)(y_i - \mu_i(\beta))\}$$

: 여기서 $w_i(\alpha)$는 i번째 개체의 treatment model(PS model), $\mu_i(\beta)$은 i번째 개체의 outcome model 의미

-> 만약 위 식을 estimating function으로 사용한다면, ATT일 때와 ATE일 때 차이가 없을 것 같음. --- 그건 맞으나, 강의 노트 기반에서는 달라질 것이다.

2-1. Moodie et al 공식 정리

Moodie et al 의 논문 이용해 Variance Estimator of DR Estimator formula 정리

$$U_T(\alpha,\beta) = \{ \underbrace{w_T(\beta)}_{\text{성향점수}} (y_T - \underbrace{\mu_T(\alpha)}_{\text{outcome model}}) \}$$

/ treatment model ; $\log\left(\frac{P(E=1|X)}{1-P(E=1|X)}\right) = \beta^T X$

outcome model ; $y_T = \alpha^T x_T + \varepsilon_T \Rightarrow E[y_T|x_T] = \alpha^T x_T$

$\Rightarrow w_T(\beta)$는 T^{th} 개체의 성향점수 의미 / $P(E=1|X) = \frac{e^{\beta^T x}}{1+e^{\beta^T x}}$ 이므로 $w_T(\beta) = \frac{e^{\beta^T x_T}}{1+e^{\beta^T x_T}}$

$$\Rightarrow U(\alpha,\beta) = \begin{bmatrix} w_1(\beta)(y_1-\mu_1(\alpha)) \\ w_2(\beta)(y_2-\mu_2(\alpha)) \\ w_3(\beta)(y_3-\mu_3(\alpha)) \\ \vdots \\ w_n(\beta)(y_n-\mu_n(\alpha)) \end{bmatrix} \Rightarrow \begin{bmatrix} w_1(\beta) & \cdots & & 0 \\ 0 & w_2(\beta) & \cdots & 0 \\ 0 & 0 & w_3(\beta) \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_n(\beta) \end{bmatrix} \begin{bmatrix} y_1-\mu_1(\alpha) \\ \vdots \\ y_n-\mu_n(\alpha) \end{bmatrix}$$

$\mu_1(\alpha) \to \alpha^T x_1$, $x_1 \to (1, E_T, B_T, C_T)$ 가 우리 situation.

; coding 할 때는 diag(ps) %*% (y-α^T x)

$$\Rightarrow U_{adj}(\beta,\alpha) = U(\beta,\alpha) - E\left[\frac{\partial}{\partial\alpha} U(\beta,\alpha)\right] \left(E[\underbrace{\dot{\ell}_\beta(\beta)}_{\text{Score function of treatment model 한번 더 미분}}]\right)^{-1} \underbrace{\ell'_\beta(\beta)}_{\text{Score ft}^n \text{ of trt model}}$$

[Treatment model]

: 원인변수 E에 대해 $E_i \sim Ber(p_i)$ 이고, $\mathrm{logit}(p) = \beta^T x$ 라 modeling

likelihood function $L(\beta;x) = \prod_{T=1}^{n} p_i^{E_i}(1-p_i)^{(1-E_T)} = \prod_{T=1}^{n} \left(\frac{1+e^{\beta^T x_T}}{1+e^{\beta^T x_T}}\right)^{E_T} \left(\frac{1}{1+e^{\beta^T x_T}}\right)^{1-E_T}$

$\log L(\beta;x) = \sum_{T=1}^{n} E_T(\beta^T x_i - \log(1+e^{\beta^T x_i})) + (1-E_T)\{\beta^T x_T - \log(1+e^{\beta^T x_T})\}$

$= \sum_{T=1}^{n} E_T\{\cancel{\beta^T x_T - \log(1+e^{\beta^T x_T})}\} + \beta^T x_T - \log(1+e^{\beta^T x_T}) - E_T\{\cancel{\beta^T x_T - \log(1+e^{\beta^T x_T})}\} = \sum_{T=1}^{n} \beta^T x_T - \log(1+e^{\beta^T x_T})$

Score function ; $S(\beta;x) = \frac{\partial}{\partial\beta} \log L(\beta;x) = \frac{\partial}{\partial\beta}\left[\sum_{T=1}^{n} \beta^T x_T - \log(1+e^{\beta^T x_T})\right] \to \frac{\partial}{\partial\beta}[\beta^T x - \log(1+e^{\beta^T x})]$

$= x - \frac{e^{\beta^T x}\cdot x}{1+e^{\beta^T x}} = x\left(1-\frac{e^{\beta^T x}}{1+e^{\beta^T x}}\right)$

$= x\left(\frac{1}{1+e^{\beta^T x}}\right)$

Score ft^{n}의 미분 ; $\frac{\partial}{\partial\beta} S(\beta;x) = \frac{\partial}{\partial\beta}\left[\frac{x}{1+e^{\beta^T x}}\right] = \frac{-x(1+e^{\beta^T x})\cdot e^{\beta^T x}\cdot x}{(1+e^{\beta^T x})^2}$

$= \frac{-x^T x \cdot e^{\beta^T x}}{(1+e^{\beta^T x})}$

$$\frac{\partial}{\partial\alpha} U(\beta,\alpha) = \begin{bmatrix} -w_1(\beta)x_1 \\ -w_2(\beta)x_2 \\ \vdots \\ -w_n(\beta)x_n \end{bmatrix} = \begin{bmatrix} w_1(\beta) & \cdots & 0 \\ 0 & w_2(\beta) \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_n(\beta) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \xrightarrow{\text{R code}} \text{diag(ps) \%*\% model.matrix(outcome-model)}$$

$\Rightarrow$ 이때 논문 Appendix 에 게시된 것은 $Var(\hat{\alpha})$ 이다

$$Var(\hat{\tau}_{dr}) = Var\left(\frac{\sum_{T=1}^{n} (E_T - (1-E_T)\underbrace{w(x_T)}_{Var(\beta)\text{ 필요}(?!)})(Y_T - \underbrace{\mu(0,x_T)}_{Var(\hat{\alpha})\text{ 필요}(?!)})}{\sum_{T=1}^{n} E_T}\right)$$

$\Rightarrow$ 이때 E_T, Y_T는 관측값으로 들어가며, $Var(E_T)$, $Var(Y_T)$는 없는 값...?

나의 전개) $\left(\frac{1}{\sum_{T=1}^{n} E_i}\right)^2 \Big[Var\left(\sum_{T=1}^{n}(E_T-(1-E_T)w(x_T))Y_T\right) + Var\left(\sum_{T=1}^{n}(E_T-(1-E_T)w(x_T))\mu(0,x_T)\right)$

$+ 2\cdot Cov\left[\left(\sum_{T=1}^{n}(E_T-(1-E_T)w(x_T))Y_T\right)\cdot\left(\sum_{T=1}^{n}(E_T-(1-E_T)w(x_T))\mu(0,x_T)\right)\right] \Big]$