

Consistency check about μ_0 and μ_1

: 7월 22일 Version

1. Consistency assumption check about μ_0 and μ_1

-- data generating 과정에서 unobservable variable "U" part는 제외

-- 두 가지 Potential outcome 생성 : 관측되지 않은 counterfactual을 E 값 조정해 생성
 $\hookrightarrow Y^1, Y^0$ 생성

1-1) μ_1

: true value는 0.9025169

: #of obs는 1,000,000개로 하여 $E[Y|E=1]$ 계산한 결과, true value와 같은 값 나옴. (Consistency assumption ok!)

1-2) μ_0

: true value는 0.210288

: #of obs를 1,000,000개로 하여 $E[Y|E=0]$ 계산한 결과, 값이 0.2022499로 true value와 약 0.008정도 차이 나는 점 확인 / 다른 추정량인 $\log(1.5)*P(B|E=1) + \log(2)*E[C|E=1]$ 계산한 결과 값이 0.2095174로 true value와 0.012정도 차이 나는 것을 확인

: #of obs를 10,000,000개, exposure ratio를 0.76으로 증가하여 μ_0 의 true value 다시 확인한 결과, 값이 0.2935713이었다. 그런데 다시 $E[Y|E=0]$ 계산한 결과, true value와 0.38정도 차이가 난다. (exposure ratio가 증가하니, $E[Y|E=0]$ 값이 감소함을 확인) / 반면, 다른 추정량인 $\log(1.5)*P(B|E=1) + \log(2)*E[C|E=1]$ 계산한 결과 값이 0.2924246로 true value와 0.011정도 차이 나는 것을 확인

---- Consistency assumption 만족한다고 할 수 있는 점인가 ----

2. Consistency check about μ_0 and μ_1 estimator

-- data generating, estimator 계산 과정에서 unobservable variable "U" part 제거 + exposure ratio를 0.76으로 증가시킴

-- #of obs를 5000, 20000, 80000, 320000개로 변화시켜가며 추정량의 bias와 variance 계산

-- 추정량 종류 3가지 고려

$$\begin{aligned} \text{Ver 04} \quad & \begin{cases} \hat{\mu}_0 = \log 1.5 \times P(B|E=1) + \log 2 \cdot E[C|E=1] \\ \hat{\mu}_1 = \hat{\mu}_0 + \log 2 \end{cases} \quad / \quad \text{Ver 05} \quad \begin{cases} \hat{\mu}_1 = \text{mean}(\text{result_model} \$ \text{fitted.values} \\ \quad \quad \quad [data \$ E == 1]) \\ \hat{\mu}_0 = \text{mean}(\text{result_model} \$ \text{fitted.values} \\ \quad \quad \quad [data \$ E == 0]) \end{cases} \\ \text{Ver 08} \quad & \begin{cases} \hat{\mu}_1 = \frac{\sum \hat{w}_i A_i Y_i}{\sum \hat{w}_i A_i} \\ \hat{\mu}_0 = \frac{\sum \hat{w}_i (1-A_i) Y_i}{\sum \hat{w}_i (1-A_i)} \end{cases} \quad (\hat{w}_i \text{ is weight of } |D|) \end{aligned}$$

[Result]

1) Ver04

	bias of μ_0 estimator ver1	variance of μ_0 estimator ver1	bias of μ_1 estimator ver1	variance of μ_1 estimator ver1
#of obs = 5000	0.0009414112	1.543259e-04	0.0009200590	1.543259e-04
#of obs = 20000	0.0003707005	8.050011e-05	0.0003493484	8.050011e-05
#of obs = 80000	-0.0002539948	8.028361e-06	-0.0002753470	8.028361e-06
#of obs = 320000	-0.0829123317	1.421640e-05	-0.0829336838	1.421640e-05

Bias의 절대값이 증가

Bias의 절대값이 증가

2) Ver 05

	bias of μ_0 estimator ver2	variance of μ_0 estimator ver2	bias of μ_1 estimator ver2	variance of μ_1 estimator ver2
#of obs = 5000	0.0020575828	1.383410e-03	-0.0001869002	3.892865e-04
#of obs = 20000	0.0060420632	5.567719e-04	0.0012009109	1.743383e-04
#of obs = 80000	0.0005457123	9.376294e-05	-0.0009471757	2.091511e-05
#of obs = 320000	-0.0828426877	1.615657e-05	-0.0828963942	4.129419e-05

Bias의 절대값이 증가
 증가

3) Ver 08

	bias of μ_0 estimator ver3	variance of μ_0 estimator ver3	bias of μ_1 estimator ver3	variance of μ_1 estimator ver3
#of obs = 5000	0.0020575828	1.383410e-03	-0.0001869002	3.892865e-04
#of obs = 20000	0.0060420632	5.567719e-04	0.0012009109	1.743383e-04
#of obs = 80000	0.0005457123	9.376294e-05	-0.0009471757	2.091511e-05
#of obs = 320000	-0.0828426877	1.615657e-05	-0.0828963942	4.129419e-05

