

# <Simulation Scenario Code result check 기록>

- 12월 19일 Version

## <What TO DO>

1) 이전 과정에서 방법론 별 ATE, ATT 추정치의 차이가  $1e-06$ 보다 작은 경우가 꽤 되는 것으로 확인됨.  
-- 이는 이상한 상황이므로, N수가 작은 dataset을 생성해 추정치 차이 다시 확인해보기

2) 이전 확인 과정에서 DR ATT Sandwich robust Variance estimator의 Coverage는 1인 반면,  
SD Ratio가 10 이상으로 너무 큰 값이 나왔었다.  
-- Data의 Exposure ratio 늘려서 Coverage가 0.95가 되는 시점에서의 SD Ratio가 1인지 확인

## <Result>

1) 방법론 별 ATT, ATE 추정치의 차이가 없는 경우는 Exposure 변수 값이 1인 관측치의 수가 단 "1개"이다.

(n=15일 때)

```
> print(IPW_estimator_store)
      1      2      3      4 5      6      7      8 9      10
ATE 0.9134119 -1.1254240 1.561832 -0.05288077 NA -0.1597022 1.87224221 0.4170177 NA -0.4196455
ATT 0.6932026 -0.5423111 1.491913 -0.24350131 NA -0.4021742 0.09914152 0.5109750 NA 1.4338184
> print(DR_estimator_store)
      1      2      3      4 5      6 7      8 9      10
ATE 0.9154454 -0.6162573 1.483538 0.1499822 NA -0.2370464 -0.3226926 0.6497529 NA 0.6127509
ATT 0.7999995 -0.5671439 1.496958 -0.4598745 NA -0.4155262 -0.3226926 0.3576403 NA 1.4306761
```

cf) 추정치가 NA가 나온 이유는 해당 Data에서 처치변수 값이 1인 관측치가 없기 때문임을 확인함.

↳ 7번째 반복에서 얻은 Data set : 해당 Data에서 처치변수 값이 1인 관측치가 오직 1개임.

```
> dt7
      B      C      A      Y
1 0 0.20337299 0 -0.169194416
2 0 -1.38962491 0 -0.102240662
3 0 0.17849501 0 0.433405481
4 1 0.78854659 0 0.125613436
5 1 0.64613728 0 0.999004850
6 0 -1.87336361 0 -2.204225774
7 1 2.16671861 0 3.004329455
8 0 2.35635075 1 2.255076350
9 1 -0.07276940 0 1.109278249
10 0 0.52061180 0 2.086708844
11 0 -0.32830236 0 -0.715851449
12 1 -0.96274814 0 -0.669333200
13 1 0.03450127 0 0.985152913
14 0 0.22175280 0 0.473664182
15 0 0.44441221 0 0.003366045
```

(n=20일 때)

```
> print(IPW_estimator_store)
```

	1	2	3	4	5	6	7	8	9	10
ATE	0.7951137	0.1356370	1.155443	-0.1218353	NA	-0.0858072	1.550262	1.451296	-0.7843466	1.001104
ATT	0.6863842	0.6150319	1.377520	0.2344872	NA	-0.1613589	0.924076	1.533572	-0.8872043	1.911918

```
> print(DR_estimator_store)
```

	1	2	3	4	5	6	7	8	9	10
ATE	0.7647846	0.5408827	1.497323	0.56118316	NA	-0.07065337	-0.8310306	1.74110	-0.8920884	2.109960
ATT	0.7213782	0.5998683	1.268581	-0.02985104	NA	-0.16469170	-0.8310306	1.37307	-0.8942615	1.846404

cf) 추정치가 NA가 나온 이유는 해당 Data에서 처치변수 값이 1인 관측치가 없기 때문이라고 추측.

↳ 7번째 반복에서 얻은 Data set : 해당 Data에서 처치변수 값이 1인 관측치가 오직 1개임.

```
> dt7
```

	B	C	A	Y
1	0	0.20337299	0	-0.347322174
2	0	-1.38962491	0	-1.370686739
3	0	0.17849501	0	0.679496664
4	1	0.78854659	0	1.272000809
5	1	0.64613728	0	0.548656314
6	0	-1.87336361	0	-1.748642780
7	1	2.16671861	0	2.098231025
8	0	2.35635075	1	1.818397105
9	1	-0.07276940	0	0.245978642
10	0	0.52061180	0	0.225517334
11	0	-0.32830236	0	-0.009815418
12	1	-0.96274814	0	-0.816928645
13	1	0.03450127	0	0.879167743
14	0	0.22175280	0	0.926579538
15	0	0.44441221	0	0.536208930
16	0	0.39664716	0	0.515610351
17	1	1.37793175	0	3.238333732
18	1	0.65089065	0	-0.169759172
19	1	-1.09170496	0	-1.320055644
20	1	0.59005992	0	-0.287994691

→ n의 범위를 5 ~ 20까지 고려해보면서 IPW 방법론을 사용해 얻은 ATE, ATT 추정치 차이가  $1e-06$  이하인 경우는 찾을 수 없었다. 그러나, 추정치가 근사한 경우를 살펴보면, 모두 해당 Data에서 처치변수 값이 1인 관측치가 오직 1개임을 관찰함.