

〈Simulation Scenario 구현 결과〉

- 12월 7일 Version

〈What TO DO〉

: 현재까지는 Exposure ratio $p = 0.12$ / Scenario 1에 대해서만 파악

1) ATT, ATE 별 Simulation result 생성할 때 추정치가 동일한지 확인

2) Performance 지표 하나 더 추가

-- 각 분산 추정량(Naive variance estimator, Sandwich variance estimator)마다 $\frac{\widehat{var}(\hat{\beta})}{var(\hat{\beta})}$ 지표 계산

-- 이때, $var(\hat{\beta})$ 은 n 수를 크게, REPL=1000으로 해서 각 데이터 셋마다 분산 추정량을 구한 다음 총 1000개의 분산 추정량들의 평균값 계산 (Monte Carlo Approximation)

-- $\widehat{var}(\hat{\beta})$ 은 n=1000, REPL=1000인 Simulation 상황에서 얻은 각 분산 추정량들 1000개의 “평균값”

〈Result〉

1) n=1000, REPL=1000인 Simulation 상황에서 얻은 ATT, ATE 추정치 값들이 모두 동일하지 않음을 확인

-- 의문점 이전에는 추정치들이 동일한 경우가 있었는데 seed를 바꾸고 나서 결과가 달라짐.
seed 설정의 차이인지?

```
> table(ATE_1_result$Outcome_reg_est==ATT_1_result$Outcome_reg_est)
FALSE
1000
> table(ATE_1_result$IPW_est==ATT_1_result$IPW_est)
FALSE
1000
> table(ATE_1_result$DR_est==ATT_1_result$DR_est)
FALSE
1000
```

2) $var(\hat{\beta})$ 을 구하는데 시간 소모가 너무 크다

-- n = 100000, REPL = 1000으로 했을 때 아예 결과를 보지 못함.

-- n = 10000, REPL = 1000도 마찬가지.

-- REPL 수를 줄여야 하나?