

Consistency check for IPTW ATT variance - 6/17 ver

1) True Variance of IPTW ATT estimator ; 관측치가 총 10000개인 Data 생성

/ Data generating 때 사용했던 true relationship 이용!

→ Data 개수 더 늘리면

후에 분산 계산 때 error 발생

⇒ "E" 변수 생성할 때 사용한 " $\log p(z)$ " 변수 저장 / ps model 적합 때 사용하는 model matrix 라 대응하는

X 행렬 생성 ⇒ 이를 이용해 ATT 추정 위한 weight 계산! ⇒ 그 후, 생성한 함수 이용해 Σ^{true} 구함

$$\Sigma_{true} = 3.362494$$

2) Consistency check of function (⇒ 이전과 동일한 방식)

2-1) Version 1 ; 총 관측치가 100개인 Data에 대해 1000번 반복해 계산

2-2) Version 2 ; 총 관측치가 1000개인 Data에 대해 100번 replication

2-3) Version 3 ; 총 관측치가 10000개인 Data에 대해 100번 replication

2-4) Version 4 ; 총 관측치가 20000개인 Data에 대해 100번 replication

⇒ 추정치들의 Bias,

Standard Error 계산

[Result]

	bias of estimator	variance of estimator
#of obs = 100	10.061871	76.2246104
#of obs = 1000	9.667859	4.3535434
#of obs = 10000	9.842432	0.4669178
#of obs = 20000	9.863591	0.2909521

증거는
안됐다!

< Meeting 결과 정리 >

∴ true μ_1 , true μ_0 다시 계산해보자!

↑ B.C (Confounder) 에 dependent 한 부분
우리의 "Y" 변수 generating model ; $Y^a = \mu(B.C) + E \cdot \log 2 + \epsilon^a$ (이때, $E[\epsilon^a] = 0$, $\epsilon^a \perp X.A$ 가정)

$$\textcircled{1} \mu_1 = E[Y^1 | E=1] = E[\mu(B.C) + \log 2 \cdot 1 | E=1]$$

$$\textcircled{2} \mu_0 = E[Y^0 | E=1] = E[\mu(B.C) | E=1]$$

↪ 추정 필요!

$$\Rightarrow E[\mu(B.C) | E=1] = E[\log 1.2 B + \log 1.5 C | E=1] \quad / \quad \text{따라서 } P(B=1 | E=1) \text{과 } E[C | E=1]$$

추정이 필요하다 \Rightarrow crossprod 연산 없으므로 Data 수 크게 해서 Monte Carlo Approximation 해보자!

🍭 그리고 다시 Consistency test 진행! \Rightarrow PS Model 의 차이 비교도 보자!