

§ 4.4 IPTW estimation

如何估计 marginal structural model 中的参数?

4.1 类比标准回归模型中的参数估计.

(1) Estimation in Regression Model.

$$Y = X\beta + \varepsilon \quad \text{linear}$$

$$\text{最小二乘: } \min_{\beta} \|Y - X\beta\|^2$$

$$\Rightarrow X(Y - X\beta) = 0$$

$$\sum_{i=1}^n X_i(Y_i - X_i^T \hat{\beta}) = 0$$

(2) Estimation in generalized linear Model

$$E(Y_i | X_i) = \mu_i = g^{-1}(X_i^T \beta)$$

$$\text{ex: } \text{logit}(E(Y)) = X\beta$$

$$\Rightarrow \sum_{i=1}^n \frac{\partial \mu_i}{\partial \beta} V_i^{-1} (Y_i - \mu_i(\beta)) = 0$$

(3) Estimation in MSMs (Linear)

$$g(E(Y^a)) = \psi_0 + \psi_1 a \Rightarrow E(Y^a) = g^{-1}(\psi_0 + \psi_1 a)$$

① 跟 generalized linear Model 很像.

② 不等价于 regression Model. 在 regression Model 中, 我们以 "observed treatment A" 作为条件:

$E(Y|A) = g^{-1}(\psi_0 + \psi_1 A)$ 这里的 A 是固定的, 针对的是 subpopulation, 而对于 MSM, a 是变的, 我们可以设定它为任意值.

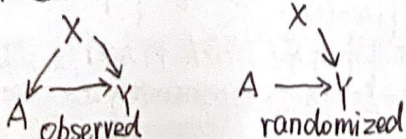
MSM: a-setting 可以是任意值.

regression model: A-conditioning

? 为什么 MSM 与 regression model 是不等同的?

因为有 Confounding 的存在!

如果是在 randomized trial 中, no confounding



影响, 那么可以拟合 regression model, 模型中的参数就表示 causal Effect.

! 这就给建模提供了一种思路: 既然在 randomized trial 中可以用 regression model, 我们可以尽量去构建 randomized trial, 然后做回归.

在前几课 (§ 4.2) 中, 我们知道可以使用 IPTW 构建 pseudo-population 来近似 randomized trial 的情形 Fig. [Estimation in MSMs]

• Pseudo-population is free from confounding

(在 ignorability 和 positivity 的假设下)

$$\sum_{i=1}^n \frac{\partial \mu_i}{\partial \beta} V_i^{-1} (Y_i - \mu_i(\beta)) = 0 \quad (\text{generalized model})$$

$$\sum_{i=1}^n \frac{\partial \mu_i}{\partial \beta} V_i^{-1} W_i (Y_i - \mu_i(\beta)) = 0 \quad (\text{MSMs})$$

(μ_i 是线性回归中 E(Y_i | X_i) 的条件均值)

$$W_i = \frac{1}{A_i I(A=1|X_i) + (1-A_i) I(A=0|X_i)}$$

A_i=1 示性 A_i=0

4.4.2 Steps in estimating parameters from MSM.

Step 1: Estimating propensity score.

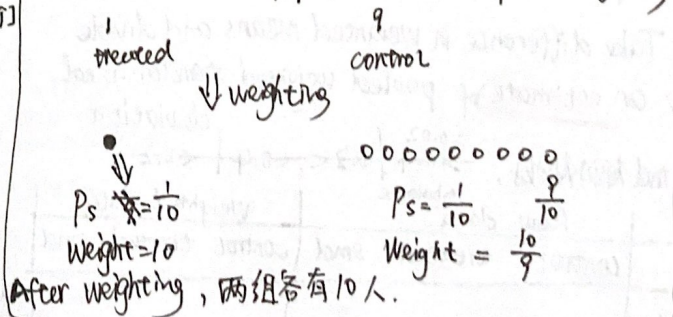
Step 2: Create weights $\frac{P(A=1|X_i)}{P(A=0|X_i)} \leftarrow P_s \text{ for treated}$

Step 3: Specify the MSM of interest. $P_s \text{ for control}$

(Linear or logit or with modification or ...)

Step 4: Use software to fit a weighted generalized linear model.

★ 人为地扩大了 population size, after weighting



Step 5: Use asymptotic variance estimator

因为 pseudo-population (DR bootstrapping)

> sample size