Identification of the Causal Effect (Homework 2)

Yancheng Li

October 20, 2022

We model the relationship between social mobility and number of deaths due to Covid-19 as a stochastic process:

$$
\begin{aligned}
&(A_t, Y_t),\ t = 1, 2, \cdots, T,\ \text{where} \\
&A_t:\ \text{social mobility on week t} \\
&Y_t:\ \text{the number of deaths due to Covid-19 on week t}
\end{aligned}
$$

Define

$$
\begin{aligned}
\overline{A}_t &= (A_1, \cdots, A_t),\ t = 1, 2, \cdots, T \\
\overline{Y}_t &= (Y_1, \cdots, Y_t),\ t = 1, 2, \cdots, T
\end{aligned}
$$

Denote $Y_t(\overline{a}_t)$ as a counterfactual quantity representing the value of Y_t if we set $\overline{A}_t$ equal to some value $\overline{a}_t = (a_1, \cdots, a_t)$:

$$
\begin{aligned}
(\overline{A}_T, \overline{Y}_T):&\ \text{observed data} \\
\left\{Y(\overline{a}_T):\ \overline{a}_T \in \mathbb{R}^T\right\}:&\ \text{unobserved counterfactual random variables}
\end{aligned}
$$

The causal estimand we want to estimate/identify is

$$E[Y_t(\overline{a}_t)]$$

To make this identification, we make following assumptions:

1. Conditional independence: there is no unmeasured confounding. At each time, the treatment is independent of the counterfactuals given the past measured variables.

2. Counterfactual consistency: If $\overline{A}_t = \overline{a}_t$, then $Y_t = Y(\overline{a}_t)$.

3. Positivity: the distribution of treatment has a positive density.

4. Markov compatibility: the dependence of mobility on the past satisfies a Markov condition.

Then the identification is

$$E[Y_t(\overline{a}_t)] = \int \cdots \int E[Y_t|\overline{A}_t = \overline{a}_t, \overline{Y}_{t-1} = \overline{y}_{t-1}] \prod_{s=1}^{t-1} P(y_s|\overline{y}_{s-1}, \overline{a}_s) dy_s$$

The proof is given by the next page. Note that the application of the g-formula requires two assumptions: conditional independence and consistency.

$$
\begin{aligned}
E[Y_t(\overline{a}_t)] &= \int E[Y_t(\overline{a}_t)|\overline{Y}_{t-1} = \overline{y}_{t-1}]P(\overline{Y}_{t-1}(\overline{a}_{t-1}) = \overline{y}_{t-1})d\overline{y}_{t-1} && \text{(Law of total probability)} \\
&= \int E[Y_t|\overline{A}_t = \overline{a}_t, \overline{Y}_{t-1} = \overline{y}_{t-1}]P(\overline{Y}_{t-1}(\overline{a}_{t-1}) = \overline{y}_{t-1})d\overline{y}_{t-1} && (g\text{-formula}) \\
&= \int E[Y_t|\overline{A}_t = \overline{a}_t, \overline{Y}_{t-1} = \overline{y}_{t-1}]P(\overline{Y}_{t-1} = \overline{y}_{t-1}|\overline{a}_{t-1})d\overline{y}_{t-1} && (g\text{-formula}) \\
&= \int \cdots \int E[Y_t|\overline{A}_t = \overline{a}_t, \overline{Y}_{t-1} = \overline{y}_{t-1}] \prod_{s=1}^{t-1} P(y_s|\overline{y}_{s-1}, \overline{a}_s)dy_s && \text{(Positivity, Markov compatibility)}
\end{aligned}
$$

Interpretion: the final result is the famous Robins' g methods. And we are estimating a longitudinal effect.