Causual Inference Note

Kirby CHEN

Academic Year 2024-2025

Contents

1 Fall 2019

1.1 Fri, Sept 6: Phenomenology of Microscopic Physics

- Newtonian mechanics (i.e., $\mathbf{F} = m\mathbf{a}$) is an excellent theory; it applies to the vast majority of human-scale (and even interplanetary-scale) physics.
- Apart from relativistic effects at very high velocities (special relativity) or in very strong gravitational fields (general relativity), Newtonian mechanics accurately describes a huge range of phenomena, but around the end of the Nineteenth Century people became aware of some physical effects for which there is no sensible Newtonian explanation.
- Examples include:
  - the **double slit experiment** (done with light by Thomas Young in 1801, and with electrons by Tonomura in 1986)
  - the photoelectric effect (analyzed by Einstein in 1905 — in fact his Nobel-winning work)
  - the “quantum Venn diagram” puzzle, involving the overlaps of three polarizing filters
  - the stability of the hydrogen atom (i.e., the fact that the electron doesn’t lose energy and spiral inward toward the proton).

Remark 1. *How now, brown cow?*

Definition 1. *The* Feynman kernel *is given by*

$$K(x_b, t_b; x_a, t_a) = \int_{x(t_a)=x_a}^{x(t_b)=x_b} e^{(i/\hbar)S[x(t)]} \, \mathscr{D}x(t).$$

2 PSM Model Setup

For an individual i, the outcome depends on whether they receive a certain treatment:

$$y_i = \begin{cases} y_{1i}, & \text{if } D_i = 1 \\ y_{0i}, & \text{if } D_i = 0 \end{cases} \tag{1}$$

- D_i indicates whether individual i receives the treatment, where 1 represents treated, and 0 represents untreated.
- y_{1i} represents the outcome for individual i if treated.
- y_{0i} represents the outcome for individual i if untreated.

Given the observable covariates x_i, the probability of an individual i receiving the treatment is defined as:

$$p(x_i) = \Pr(D_i = 1 \mid x = x_i) = E(D_i \mid x_i) \tag{2}$$

Based on Equations (1) and (2), the **Average Treatment Effect on the Treated (ATT)** is given by:

$$ATT = E[y_{1i} - y_{0i} \mid D_i = 1] \tag{3}$$

$$= E[E[y_{1i} - y_{0i} \mid D_i = 1, p(x_i)]] \tag{4}$$

$$= E[E[y_{1i} \mid D_i = 1, p(x_i)] - E[y_{0i} \mid D_i = 0, p(x_i)] \mid D_i = 1] \tag{5}$$

3 PSM Assumptions

3.1 Common Support Assumption

For any possible value of x_i, the propensity score must satisfy:

$$0 < p(x_i) < 1 \tag{6}$$

This assumption ensures that there is **overlap between the treated and control groups**, making it possible to find comparable units.

3.2 Balancing Assumption

$$D_i \perp (y_{1i}, y_{0i}) \mid p(x_i) \tag{7}$$

This assumption states that **conditional on the propensity score $p(x_i)$, treatment assignment is as good as random**. That is, for a given $p(x_i)$, there are no systematic differences between the treatment and control groups, meaning the treatment effect is entirely due to the treatment itself.