

Potential Outcomes Causal Model

John Stuart Mill (A System of Logic, Ratiocinative and Inductive)

↳ 5 methods for inferring causation. (Canons of Empirical Science)

1. Method of agreement → Factor necessary for effect
- * 2. Method of difference → Comparison of counterfactuals
3. Joint Method. → agreement + difference.
4. Method of concomitant variation (variation in one causes variation in the other)
5. Method of residues. (Through process of elimination)

Correlation \neq Causation

Regression alone is insufficient.

RANDOMIZATION

Aplawa - Neyman

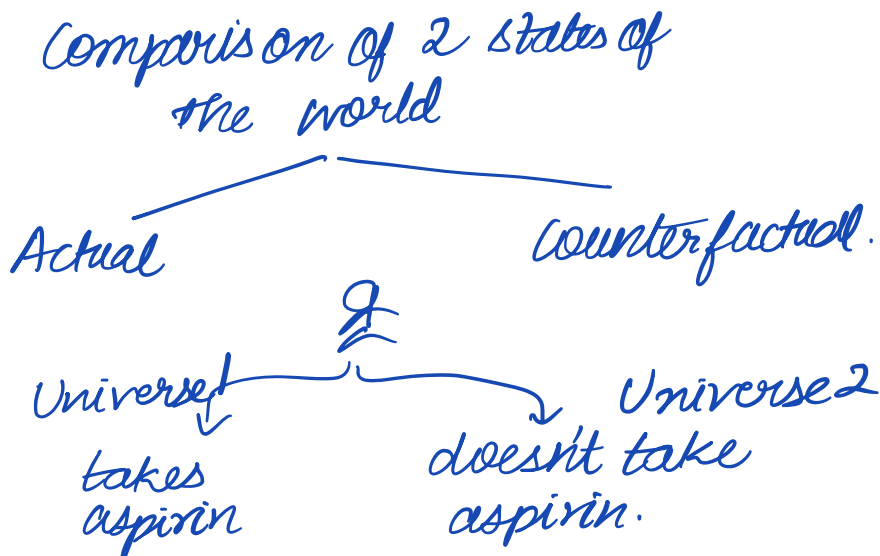
Potential outcomes notation.

(Social Sciences)

D. Rubin

Roland A. Fisher

Randomization for causal inference



Effect of aspirin = Difference in reported outcomes of Universe 1 & 2.

Counterfactual data \rightarrow missing data

"Potential outcomes exist ex-ante as a set of possibilities but once a decision is made, all but one outcome disappears"

$$D_i = \begin{cases} 1, & \text{unit } i \text{ receives treatment} \\ 0, & \text{unit } i \text{ doesn't receive treatment} \end{cases}$$

Potential outcomes \Rightarrow $Y_i^1 \rightarrow$ receives treatment
 $Y_i^0 \rightarrow$ doesn't
 $\xrightarrow{\text{control state}}$

Y_i } actual outcome.

$$Y_i = D_i Y_i^1 + (1 - D_i) Y_i^0$$

causal effect $\delta_i = Y_i^1 - Y_i^0$

Average Treatment Effect

$$\begin{aligned}ATE &= E[\delta_i] \\&= E[Y_i^1 - Y_i^0] \\&= E[Y_i^1] - E[Y_i^0]\end{aligned}$$

ATE is inherently unknowable. It cannot be calculated, only estimated.

\therefore Only 1 potential outcome can be known, the other is imagined and hence missing.

Average Treatment Effect for Treated Group

Population mean treatment effect for group of units that had been assigned to treatment.

ATT is unknowable. (but estimable)

$$\begin{aligned}ATT &= E[\delta_i | D_i = 1] \\&= E[Y_i^1 - Y_i^0 | D_i = 1] \\&= E[Y_i^1 | D_i = 1] - E[Y_i^0 | D_i = 1]\end{aligned}$$

Average treatment for untreated group.

$$\begin{aligned} ATU &= E[\delta_i | D_i = 0] \\ &= E[Y_i^1 - Y_i^0 | D_i = 0] \\ &= E[Y_i^1 | D_i = 0] - E[Y_i^0 | D_i = 0] \end{aligned}$$

In observational settings, given
heterogeneous treatment effects,

$$ATT \neq ATU \text{ (Probably)}$$

[Heterogeneous treatment effect
- Non random variability in the direction
or magnitude of treatment effect]

Concrete Example — 4.1.3. Scott Cunningham's
Book.

$$\begin{aligned} \text{Simple difference} &= ATE + \text{selection bias} \\ \text{in outcomes} &\quad + \text{heterogeneous} \\ &\quad \text{treatment effect} \\ &\quad \text{bias} \end{aligned}$$

If we assume no heterogeneity,
then $ATU = ATT$

$$\Rightarrow SDO = ATE + \text{selection bias}$$

Independence Assumption

$$(Y^1, Y^0) \perp D$$

treatment assignment has nothing
to do with gains of treatment

selection bias
&
heterogeneous
treatment effect

{ eliminated by
randomization of
treatment assignment.

If treatment is independent of potential
outcomes, $SDO = ATE$

But, this is not realistic. - People make choices
based on perceived
increase in gain.

\Rightarrow selection bias cannot
be eliminated.

SUTVA - Stable Unit Treatment Value Assumption.