Identification

The key step in causal inference

$$\underset{\substack{\text{(causal estimand}\\\text{causal parameter)}}}{\text{Causal Effect}} \xrightarrow{\text{Idenfication}} \text{Statistical parameter} \xrightarrow{\text{Estimation}} \text{Estimator / Estimate}$$

Method / Formula ↔ Estimator; Estimate ↔ Numerical result; Apply data

★ The identification proc. require certain assumption, known as identification assumption.

- Consistency
- Exchangeability

Def. The observable outcome Y is defined as

$$Y = \begin{cases} Y(1) & \text{if } A = 1 \\ Y(0) & A = 0 \end{cases}$$

In general, $Y = Y(a)$ if $A = a$

(Remark. paired definition

$$Y = Y(1)A + Y(0)(1-A)\ )$$

★ If two groups are exchangeable then any difference in outcomes between them can be attributed to the treatment, not to pre-existing difference.

$$P(Y(a) = 1 \mid A = 1) = P(Y(a) = 1 \mid A = 0) \quad \forall a = 0,1$$

or equivalently

$$\begin{cases} Y(1) \perp\!\!\!\perp A \\ Y(0) \perp\!\!\!\perp A \end{cases}$$

In general, $Y(a) \perp\!\!\!\perp A \quad \forall a$ (Counterfactual outcome and actual treatment are indep)

Subject : .... No. : 11 Date : ..../........./.........

★ Causal Inference as a fundamental missing data challenge.

missing $\Leftrightarrow$ potential outcome

Econ Program $\longrightarrow$ Income

$A = \begin{cases} 1 & \text{enter} \\ 0 & \text{not enter} \end{cases}$ $\qquad Y$

A	$Y(1)$	$Y(0)$
1	100	?
1	120	
1	220	
1	200	?
0	?	170
0		90
0		190
0	?	80

Representative ?

★ What does representative look like?

$\Rightarrow$ obs data and missing data take some values

A	$Y(1)$		$Y(0)$	
1	100		90	120
1	120		80	110
1	220		170	200
1	200		190	220
0	200	170	170	
0	120	90	90	
0	220	190	190	
0	100	70	80	

- For both potential outcomes, the dist^n of outcomes is the same regardless of the value of the treatment variable.

$$Y(a) \perp\!\!\!\perp A \quad \forall a = 0,1$$

- Lack of exchangeability

★ $Y(a) \perp\!\!\!\perp A \not\Rightarrow Y \perp\!\!\!\perp A$

$\begin{cases} Y(1) \perp\!\!\!\perp A \\ Y(0) \perp\!\!\!\perp A \end{cases}$ (exchangeability)

$E[Y(a) \mid A=1] = E[Y(a) \mid A=0]$
$\forall a = 0,1$ (mean exchangeability)

$(Y(1), Y(0)) \perp\!\!\!\perp A$
(full exchangeability)

Let $\phi(a) = \mathbb{E}[Y(a)] \quad a = 0, 1$

Causal Effect $= \phi(1) - \phi(0)$

$$\overset{\Delta}{=} \mathbb{E}(Y|A=1) - \mathbb{E}(Y|A=0)$$

identify

$$\begin{aligned} \phi(a) &= \mathbb{E}[Y(a)] \\ &= \mathbb{E}[Y(a) \mid A=a] \quad \left(\text{Exchangeability } Y(a) \perp\!\!\!\perp A \ \forall a\right) \\ &= \mathbb{E}(Y \mid A=a) \quad \left(\text{Consistency } Y = Y(a) \text{ if } A=a\right) \end{aligned}$$

Latin → confundere which mean mixing

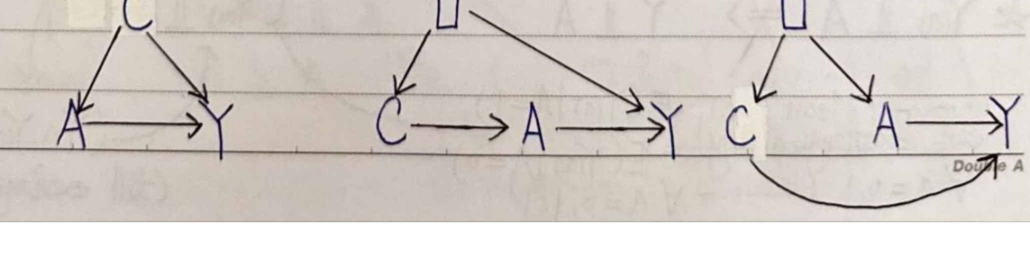
ξ Confounding

- Effect or association between exposure and outcome is distorted by the presence of another variable.

- Confounding is also form a bias.

- An informal definition
  A confounder is any variable that can be used to adjust for confounding.

- In traditional approach.
  1) It is associated w/ treatment.
  2) 〃 outcome conditional on treatment.
  3) It does not lie on a causal pathway between treatment and outcome.



Case (a)

C has causal effect on A
C has direct causal effect on Y
C does not lie on causal pathway between A and Y

case (b)

C has causal effect on A
C shares cause U w/ Y
C does not lie on the causal pathway between A and Y

case (c)

C shares cause U w/ A
C has causal effect on Y
C does not lie on the causal pathway between A and Y

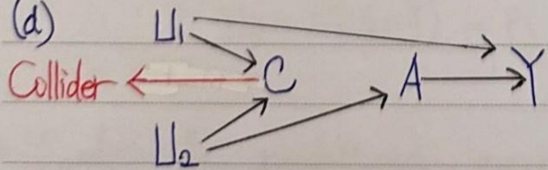
Thm.

Consider a set of nodes $C \subset \{v_1, \ldots, v_N\}$ s.t $C \cap A = \emptyset$
Assume that following conditions hold:

i) No element of C is a descendent of A
ii) C blocks all back door paths from A to Y

then $Y(a) \perp\!\!\!\perp A \mid C \quad \forall a$

(d)



$\Rightarrow$ No confounding

There are no common causes of A and Y.
The backdoor path between A and Y through C is blocked because C is a collider.

The C in (d) meets the rules for traditional approach
C shares common cause U_2 w/ A
" U_1 " Y
C does not lie on the causal pathway between A and Y

★ Adjust for C results in a biased est'or of the causal effect due to seletion bias.

How to correct it?
(1) Experimental Study { RCT ← Randomized Control Trials
Matching (必需要知道有興趣的干擾因子)

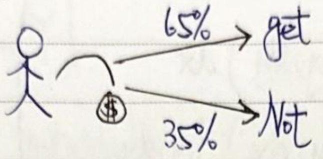
(2) Observational Study { Stratification { Restriction / 分層分析 ; 模型法
G-method { Standardization ; g-formula
IPW est'or and MSM
G-estimation and structual nested model (Semi-parametric)

ξ Randomization
- Randomized experiments, like other real study general data w/ missing value of counterfactual outcome.
- Randomization ensures that those missing values occured by chance.

★ A ideal randomized experiment
i) No loss to follow-up
ii) Full adherence to the assigned treatment
iii) A single version of treatment
iv) Double blind assignment (双盲実驗)

design 1.

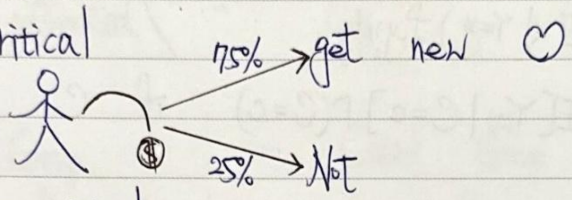
randomly selected 65% of individuals



design 2.

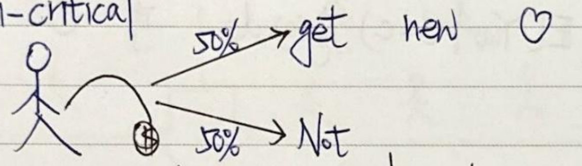
critical

C=1



non-critical

C=0



design 1 ⇒ marginally randomized experiments → single unconditional randomization prob. that is common to all individuals.

" 2 conditionally " → several randomization probabilities that depend on the value of the variable C

★ A marginally randomized experiment is expected to result in exchangeability of treated and untreated

$$Y(a) \perp\!\!\!\perp A \quad \forall a=0,1$$

or

$$P(Y(a)=1 \mid A=1) = P(Y(a)=1 \mid A=0) \quad \forall a=0,1$$

★ Conditionally randomized experiments is simply the combination of separate marginally randomized experiments. It guarantees

$$Y(a) \perp\!\!\!\perp A \mid C \quad \forall a=0,1$$

Conditional exchangeability

§ Standardization

$$\left(\begin{aligned} \mathbb{E}(X) &= \int_x x f_X(x)dx \\ &= \int_x x \left(\int_y f_{XY}(x,y)\, dy \right) dx \\ &= \int_y \left(\int_x x f_{X|Y}(x|y) dx \right) f_Y(y) dy \\ &= \int_y \mathbb{E}(X|Y=y) f_Y(y) dy \end{aligned}\right)$$

$$\mathbb{E}[Y(a)] = \begin{cases} \sum_c \mathbb{E}[Y(a) \mid C=c]\, P(C=c) & \text{if } C \text{ discrete} \\ \int_c \mathbb{E}[Y(a) \mid C=c)\, f_C(c)\, dc & \text{if } C \text{ conti} \end{cases}$$

DAGs

$$\begin{tikzcd} & C \arrow[dl] \arrow[dr] & \\ A \arrow[rr] & & Y \end{tikzcd}$$

$$\begin{aligned} \phi(a) = \mathbb{E}[Y(a)] &= \sum_c \mathbb{E}[Y(a) \mid C=c]\, P(C=c) \\ &= \sum_c \mathbb{E}[Y(a) \mid A=a, C=c]\, P(C=c) \\ &= \sum_c \mathbb{E}(Y \mid A=a, C=c)\, P(C=c) \end{aligned}$$

conditional exchangeability by $Y(a) \perp\!\!\!\perp A \mid C$

by consistency

g-formula

the causal effect

$$\begin{aligned} &\mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] \\ &= \sum_c \left\{ \mathbb{E}(Y \mid A=1, C=c) - \mathbb{E}(Y \mid A=0, C=c) \right\} P(C=c) \end{aligned}$$

★ Inverse probability weighting (IPW)

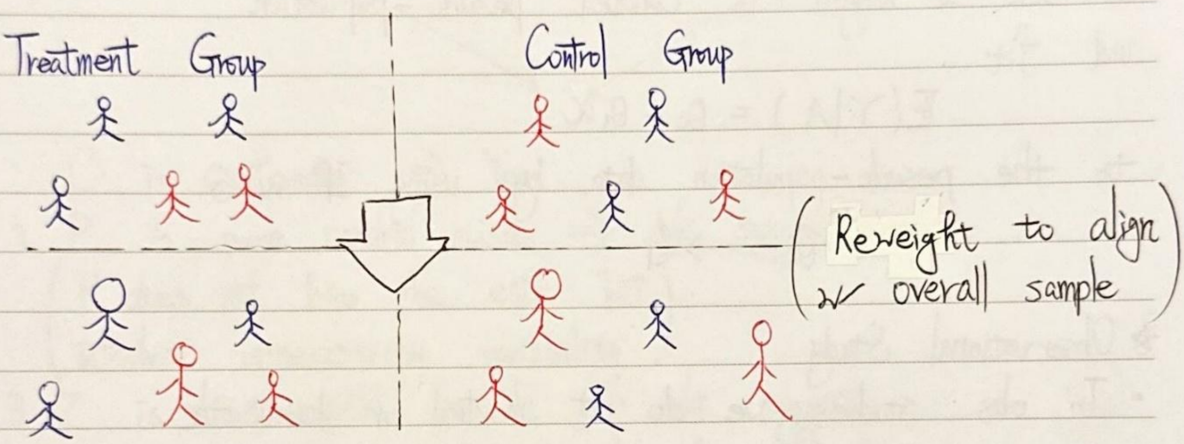
An individual's IP weight depends on + value of $\begin{cases} \text{Covariate } C \\ \text{treatment } A \end{cases}$

$$\begin{aligned} &\sum_c \mathbb{E}(Y \mid A=a, C=c)\, P(C=c) \\ &= \sum_c \left\{ \sum_y y \frac{P(Y=y, A=a, C=c)}{P(A=a \mid C=c)} \right\} \end{aligned}$$

$$= \sum_c \left\{ \sum_y \sum_{a^*} y\, I_{\{a^*=a\}} \frac{P(Y=y, A=a^*, C=c)}{P(A=a^* \mid C=c)} \right\}$$

$$= \mathbb{E}\left[\frac{I_{\{A=a\}}}{P(A|C)} Y \right]$$

- A treated individual w/ $C=c$ receives weight $\dfrac{1}{P(A=1 \mid C=c)}$
- An untreated " $C=c^*$ " $\dfrac{1}{P(A=1 \mid C=c^*)}$



☆ Goal of IP weight is to create a pesudo-population
the IP weight

$$w^A = \frac{1}{P(A|C)}$$

are referred to as nonstablized weights.
and

$$SW^A = \frac{P(A)}{P(A|C)}$$

narrower CI

are referred to as stablized weights.

Σ Marginal Structural Model (MSM)

The model attempts to solve

$$\mathbb{E}[Y(a)] = \mu(a; \xi)$$

(1) Linear model

$$E[Y(a)] = \beta_0 + \beta_1 a$$

(2) Log-linear model

$$\log\{E[Y(a)]\} = \beta_0 + \beta_1 a \iff E[Y(a)] = e^{\beta_0 + \beta_1 a}$$

(3) Logistic model

$$\text{logit}\{E[Y(a)]\} = \beta_0 + \beta_1 a \iff E[Y(a)] = \frac{e^{\beta_0 + \beta_1 a}}{1 - e^{\beta_0 + \beta_1 a}}$$

We use IP weight to conduct pesudo-population and fit

$$E(Y|A) = \theta_0 + \theta_1 a$$

to the pesudo-population data by using IP WLS

$$\hat{\theta}_1 \longrightarrow \beta_1$$

§ Observational Study

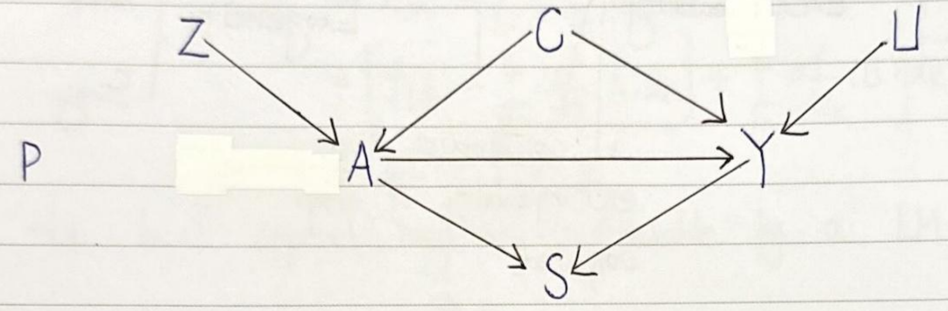
- In obs. studies, we do not control or know the assignment mechanism.
- Presence of measured and unmeasured confounders : unblanced between group.
- Some structural (often untestable) assumption must be made

★ Informally, an obs. study can be conceptualized as a conditionally randomized experiments if

1. Well-defined treatment $\iff$ Consistency
   The treatments being compared are clearly defined and match the ver. in data.
2. No unmeasured confounding $\iff$ Exchangeability
   Treatment assign depends only on observed covariates C
3. Positivity
   Every treatment has non-zero proby for all value of C

• These conditions are often heroic which explain why causal inferences from obs. study are viewed w/ suspicion.

★ What covariates should be adjust for obs. study?



1. C is confounder ⇒ Yes
2. P is pure random noise ⇒ Not suggest
   (P does not bias the esté but introduce unnecessary variability)
3. Z is instrumental variable ⇒ Not suggest
   (Z does not bias the esté although it increases variability. However, w/ unmeasured confounding, include it amplifies bias)
4. U affect Y only ⇒ Yes
   (Since U is predictive to Y, including them often improve precision)
5. S is collider ⇒ No

⇒ We should adjust for at least C to remove bias and more ideally, further adjust for U to reduce variance.

§ Causality in Econometric:

Choice vs chance

Endogenous Decision → Choice; Exogenous Variation → chance