Topics in Causal Inference

STAT41530

Jingshu Wang

Lecture 7

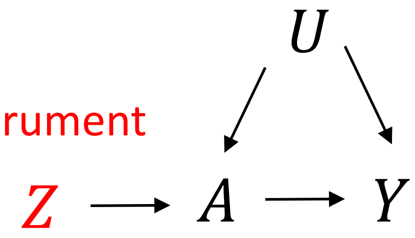
Topic:
Instrumental Variables (IV)

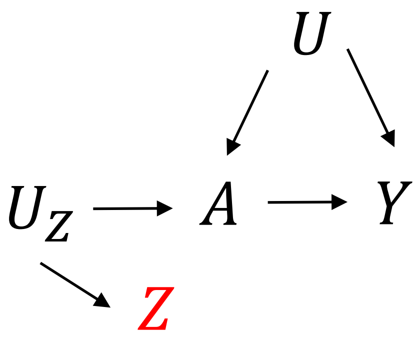
- Three conditions for valid IV
- Identifiability conditions using IV
  - Homogeneity
  - Monotonicity (principal stratification)
- Weak IV

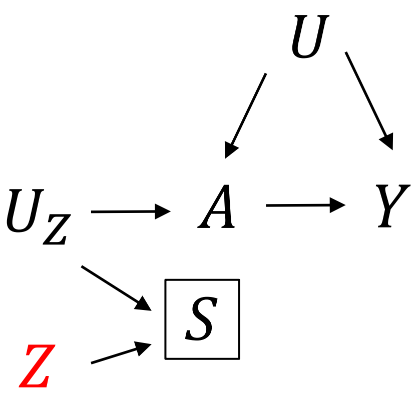
Instrumental variables (IV)

Identify causal effects without conditional exchangeability?

Z: causal instrument



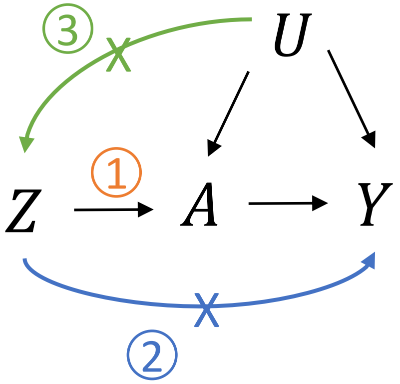


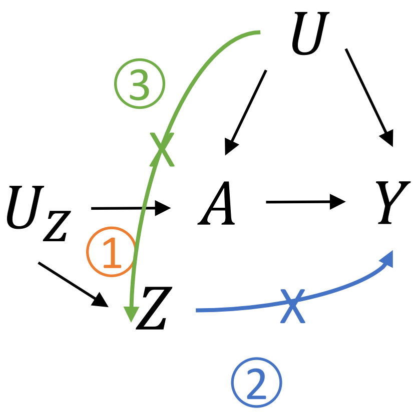


Z: proxy instrument

- Random assignment not available for treatment A, but for some other variable Z
- Non-compliance in randomized experiments

Three conditions for valid IV





① Relevance: $Z \not\perp A$, Z is associated with A

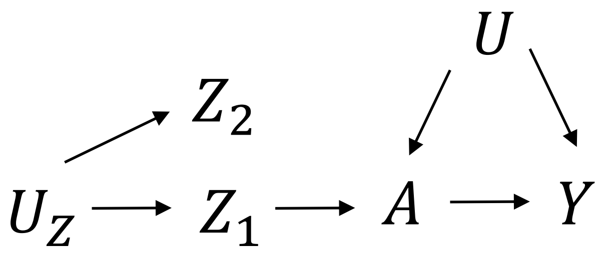
② Exclusion restriction: Z does not affect Y except through potential effects on A.

Structural equation $Y = f(A, Z, U, E_Y) = f(A, U, E_Y)$

③ Random assignment: Z and Y do not have any common causes, $Z \perp Y(a) \mid L$ for all a

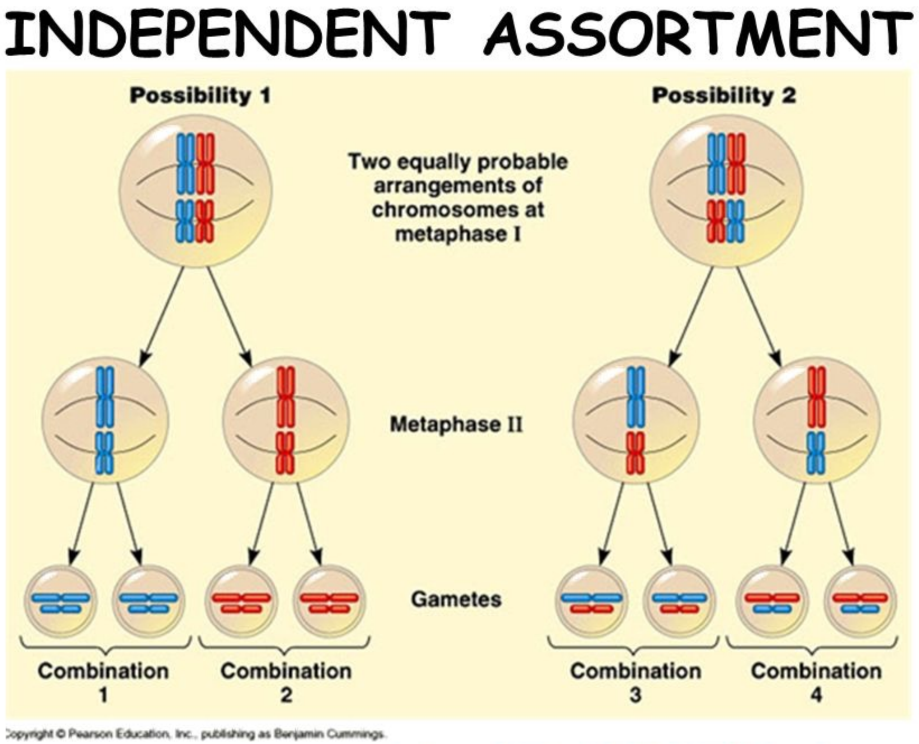
Examples of IV

- Genetic mutations (Mendelian Randomization)



- Access to treatment
  - Treatment assignment
  - Encouragement / Rewards
  - Cigarettes' price
  - Distance to the hospital

Mendel's law of Inheritance



http://fig.cox.miami.edu/~cmallery/150/mitosis/c13x9independent-assortment.jpg

Valid IV $\Longrightarrow$ Identification of causal effect of A?

NO!

- With valid IV, we can only identify the upper/lower bounds of the causal effect of A on Y

  (Robins 1989, Manski 1990)

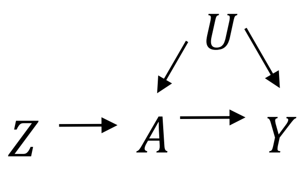
  If Z, A, Y are binary ,then the range of $\tau = E[Y(1)] - E[Y(0)]$ has width

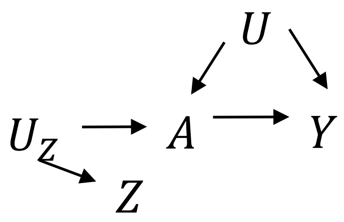
$$P(A = 1 \mid Z = 0) + P(A = 0 \mid Z = 1)$$

- We need either of the two further assumptions
  - Homogeneity
  - Monotonicity
- In this lecture, we focus on the case where $\boldsymbol{A}, \boldsymbol{Z}$ **are both binary**

  (Palmer 2011 reviewed more complicated scenarios)

The standard IV estimator





- The usual IV estimand

$$\tau_{IV} = \frac{E[Y \mid Z=1] - E[Y \mid Z=0]}{E[A \mid Z=1] - E[A \mid Z=0]} = \frac{\mathrm{Cov}(Y,Z)}{\mathrm{Cov}(A,Z)}$$

The IV estimand bypass the need to adjust for the confounder by inflating the intention-to-treat effect

- The standard IV estimator

$$\hat{\tau} = \frac{\hat{E}[Y \mid Z=1] - \hat{E}[Y \mid Z=0]}{\hat{E}[A \mid Z=1] - \hat{E}[A \mid Z=0]} = \frac{\frac{1}{n_1}\sum Y_i 1_{Z_i=1} - \frac{1}{n_0}\sum Y_i 1_{Z_i=0}}{\frac{1}{n_1}\sum A_i 1_{Z_i=1} - \frac{1}{n_0}\sum A_i 1_{Z_i=0}}$$

- Two-step least square
  - Step 1: $E[A \mid Z] = \alpha_0 + \alpha_1 Z$ regress A on Z, get $\hat{E}[A \mid Z]$
  - Step 2: $E[Y \mid Z] = \gamma_0 + \gamma_1 Z = \frac{\gamma_1}{\alpha_1}(\alpha_0 + \alpha_1 Z) + \gamma_0 - \frac{\alpha_0\gamma_1}{\alpha_1} = \tau_{IV} E[A \mid Z] + \gamma_0 - \frac{\alpha_0\gamma_1}{\alpha_1}$

    regress Y on $\hat{E}[A \mid Z]$

When does the IV estimand have causal interpretation?

Homogeneity assumption

- Robins mentioned 4 possible homogeneity conditions (section 16.3)
- The simplest (most stringent) condition: constant treatment effect of A

$$Y(a) = f(a, U, E_Y) = \tau a + \tilde{f}(U, E_Y)$$

- This means that for any individual: $Y_i(1) - Y_i(0) = \tau$

  Such homogeneity assumption is impossible for binary outcome (unless $\tau = 0$)

$$\text{As } Y = \tau A + \tilde{f}(U, E_Y), \quad Z \perp \tilde{f}(U, E_Y)$$

$$\text{Then } \mathbb{E}[Y \mid Z] = \mathbb{E}\left[\tau A + \tilde{f}(U, E_Y) \mid Z\right] = \tau \mathbb{E}[A \mid Z] + \mathbb{E}\left[\tilde{f}(U, E_Y)\right]$$

$$\tau_{IV} = \frac{\mathbb{E}[Y \mid Z=1] - \mathbb{E}[Y \mid Z=0]}{\mathbb{E}[A \mid Z=1] - \mathbb{E}[A \mid Z=0]} = \tau$$

- A milder homogeneity condition for identifying $\tau = E[Y(1)] - E[Y(0)]$:

  The average causal effect is the same at different level of U

$$E[Y(1) \mid U] - E[Y(0) \mid U] \equiv \tau$$

Homogeneity assumption example

- Non-constant treatment effect

  Structural equation (let $E_Y = (E_0, E_1), E(E_Y) = 0$)

  $$Y = \beta_0 + \tau A + U + AE_1 + (1 - A)E_0$$

  corresponding potential outcomes:

  $$Y(0) = \beta_0 + U + E_0, \qquad Y(1) = \beta_0 + \tau + U + E_1$$

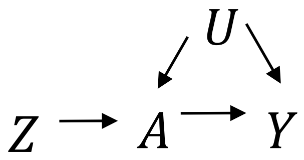
  treatment effect: $Y(1) - Y(0) = \tau + E_1 - E_0$

- Identification of average causal effect: $\tau_{IV} = \tau$ as $E[Y \mid Z] = \tau E[A \mid Z] + E[U]$

- Notice that the structural equation $Y = \beta_0 + \tau A + U + E_Y$ is different from the structural equation $Y = \beta_0 + \tau A + U + AE_1 + (1 - A)E_0$
  - If $E_0 \sim E_1$, then the two structural equations are not distinguishable from the observed data
  - The noise E_Y can have one or multiple dimensions

Principal stratification

- Only works for binary causal instrument



- Define potential outcomes of Z: $A(0), A(1), Y(Z=0), Y(Z=1)$
- Partition individuals into 4 groups (principal stratification)

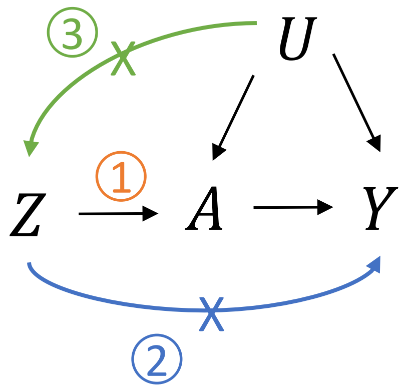
$$\begin{cases} A(1)=1, & A(0)=1 \\ A(1)=1, & A(0)=0 \\ A(1)=0, & A(0)=1 \\ A(1)=0, & A(0)=0 \end{cases}$$

Groups (G)	
Always takers	(a)
Complier	(c)
Never taker	(n)
Defier	(d)

- Define intention to treatment $ITT_{Y,g} = E[Y(Z=1) - Y(Z=0)|G=g]$ and the proportion of each group $\pi_g = P(G=g)$
- $ITT_{Y,a} = E[Y(1) - Y(1)|G=a] = 0 = E[Y(0) - Y(0)|G=n] = ITT_{Y,0}$
- Then the intention to treatment effect is

$$E[Y(Z=1)] - E[Y(Z=0)] = \pi_c ITT_{Y,c} + \pi_d ITT_{Y,d}$$

Monotonicity assumption



- Relevance $\Longrightarrow\ \pi_c \neq 0$
- Exclusion restriction $\Longrightarrow ITT_{Y,a} = ITT_{Y,n} = 0$
- Random assignment $\Longrightarrow E[Y(Z=1)] - E[Y(Z=0)] = E[Y|Z=1] - E[Y|Z=0]$
- Monotonicity (no defiers): $\pi_d = 0$

- Average causal effect of A on Y for the compliers (CACE: complier average causal effect):

$$E[Y(1) - Y(0)|G=c] = E[Y(Z=1) - Y(Z=0)|G=c] = ITT_{Y,c}$$

- As $\pi_d = 0$,

$$ITT_{Y,c} = \frac{E[Y(Z=1)] - E[Y(Z=0)]}{\pi_c} = \frac{E[Y \mid Z=1] - E[Y \mid Z=0]}{E[A \mid Z=1] - E[A \mid Z=0]} = \tau_{IV}$$

$$\mathbb{E}[A \mid Z=1] - \mathbb{E}[A \mid Z=0] = \mathbb{E}[A(1)] - \mathbb{E}[A(0)]$$
$$= \pi_a \cdot 0 + \pi_c \cdot 1 + \pi_n \cdot 0 + \pi_d \cdot (-1) = \pi_c$$

Weak instrument

- π_c is small ($\text{Cov}(A, Z)$ is small): sensitive to violations of valid IV
- $\hat{E}[A \mid Z]$ is inaccurate as an estimator of $E[A \mid Z]$

  two much uncertainty in the first-step $\Longrightarrow$ weak IV bias in the the two-step least square

- Rule of thumb: an instrument is considered weak if the F statistics for the $A - Z$ association is **less than 10**