Statistical Learning
Instrumental Variable

Outline

- Introduction and motivating examples:
  IV design as compromise between randomized trial vs observational study
- Three types of causal effects: CACE, ETT/ATT, ACE
- Evaluating IV assumptions

Introduction

- Studies in the social and health sciences commonly aim to determine the causal effect of a point exposure.
- Although the double-blind randomized study design remains the gold standard for unbiased evaluation of the effects of an exposure, observational studies are often conducted for practical or ethical reasons.
- The main challenge with drawing causal inferences from observational studies stems from their inability as a study design, to categorically rule out the possibility that differences in outcome measures between exposed and unexposed persons, may be due to systematic background differences in selecting exposure status, that also predict the outcome.

Introduction

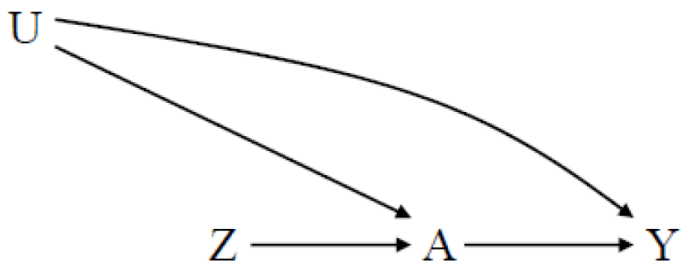
- Such confounding can compromise conclusions drawn from an observational study, but may likewise operate in a randomized trial with non-compliance, where individuals may choose for reasons unknown to the investigator, not to adhere to their treatment assignment.
- Confounding bias is a major concern for evaluating effects of exposures in observational studies and randomized experiments with non-compliance, and the development of methodology to adequately address this issue remains a priority for several disciplines, including statistics, management, econometrics, biostatistics, epidemiology, sociology, etc.

Introduction

- The instrumental variable (IV) approach refers to a particular set of methods that allow one to recover, under certain assumptions, a causal effect of an exposure in the presence of unmeasured confounding.
- The IV approach has a longstanding tradition in econometrics going back to the original work of Wright (1928) and Goldberger (1972) who initially developed the approach in the context of linear structural equation modeling.
- These ideas have more recently been formalized using potential outcomes or counterfactual variables, by Imbens & Angrist (1994), Robins (1994), Angrist, Imbens and Rubin (1996) and Heckman (1997).

Instrumental variables

- A key contribution of the counterfactual language to the IV approach is that it allows one to formally define the causal effect of interest and to clearly articulate assumptions needed to identify this effect.
- Fundamental to all IV methods is the key assumption that one has observed a pre-exposure unconfounded instrumental variable, known to satisfy the exclusion restriction, that the IV affects the outcome only through its effects on the exposure.
- e.g. Exposure A, Outcome Y, Instrumental variable Z, U unobserved confounders; the IV directed acyclic graph (IV DAG) is given by



- A variable satisfying these assumptions can be hard to find, but if a valid IV is identified, it may be used to reduce the evidentiary gap between observational and experimental study designs.

Example: MTO

- Causal question:
  Does moving from very high poverty public housing developments benefit the health of mothers or their children?
- A randomized experiment: Moving To Opportunity (MTO) study:
  - Z = families randomized to receive a voucher to alleviate the cost of moving from high poverty to low poverty neighborhood vs. no voucher.
  - A = Adhered to the random assignment
  - Outcome Y: Psychological distress of the kids measured after 10 yrs of follow-up
- Since A not randomized, cannot rule out U unobserved confounding of effects of moving from high poverty to low poverty neighborhood

Example: malaria vaccine

- Phase III trial of RTS,S/AS01 Malaria Vaccine in Infants in 7 African countries (NEJM,2012)
  Children of 6 to 12 weeks randomized to receive candidate vaccine vs meningococcal control vaccine.
- Each immunization program consisted of 3 doses one month apart.
- Z is randomized to RTS vaccine arm vs control vaccine, A indicates adherence, and Y is malaria status in 12 months of follow-up.

Intent-to-treat vs per-protocol vs as treated

- Both examples presented above, are randomized trials with Z representing randomization and A adherence.
- Intent-to-treat analysis: Ignore A and focus on $Z - Y$ relation, this is the preferred approach in most placebo controlled randomized because under the sharp null of no causal effect of the intervention, one is guaranteed not to report an effect beyond α level of say 0.05. Only test for an association between assignment to receive a voucher in MTO and Psych distress.
- per-protocol analysis compares only treated and untreated among persons who complied to their assigned treatment, i.e., the effect of Z on Y restricted to the sample with $\{A = Z\}$
- As treated analysis ignores the randomization and focuses on the effect of treatment taken on the outcome, i.e., the effect of A on Y.
- per-protocol and as treated analyses may be subject to confounding bias.

Example: Mendelian randomization

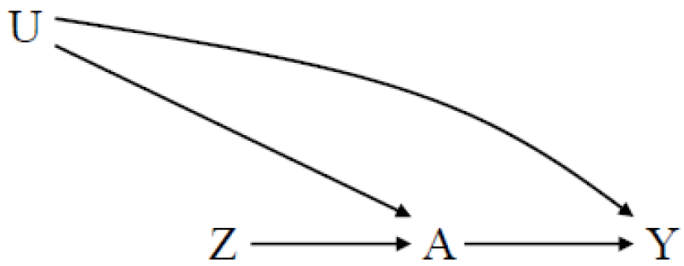
- Use randomized genotypes as instrumental variables (IVs) to estimate the causal health effects of phenotypes influenced by those genotypes
- e.g. $Z =$ FTO genetic variant, $A =$ Body Mass Index, $Y =$ Depressive Symptoms or other mental disorder, U environmental and behavioral factors unknown and/or unobserved.

Other IV examples

- Other Examples of IV:
  - Before/after policy change, e.g., labeling rules, pharmacy rules, especially if not implemented universally
  - Physician preference
  - Local area treatment patterns
  - Distance to service provider
  - Characteristic that makes patient ineligible for treatment

IV DAG

- The key assumption encoded in the IV DAG cannot be checked without an additional assumption or observing U



- That is because while the assumption implies

$$Y \perp Z \mid U, A$$

- We have that

$$Y \not\perp Z \mid A$$

- Why is that?

Linear models

- A variety of causal estimands have been shown to be identified in the IV approach, but typically require making an additional assumption, even with a valid IV.
- In the simplest context, one may assume models for Y and A are linear
- That is

$$E(Y \mid A, U, Z) = \beta_0 + \beta_A A + \beta_U U$$

where ε_Y is independent of (A, U)

- Note that by the exclusion restriction the IV Z does not appear in the model above
- However assuming that A is continuous, say $A = BMI$.

Linear models

- To gain insight over IV approach, we may compute the implied model for $E(Y \mid Z)$

$$\begin{aligned} E(Y \mid Z) &= E(E(Y \mid A, U, Z) \mid Z) \\ &= E\left(\beta_0 + \beta_A A + \beta_U U \mid Z\right) \\ &= \beta_0 + \beta_A E(A \mid Z) + \beta_U E(U \mid Z) \end{aligned}$$

- By IV assumption $E(U \mid Z) = E(U)$, and therefore

$$E(Y \mid Z) = \beta_0^* + \beta_A E(A \mid Z)$$

where $\beta_0^* = \beta_0 + \beta_U E(U)$.

Linear models

- Suppose further that $E(A \mid U, Z) = \alpha_0 + \alpha_Z Z + \alpha_u U$; then

$$\begin{aligned} E(A \mid Z) &= E(E(A \mid Z, U) \mid Z) \\ &= E\left(\alpha_0 + \alpha_Z Z + \alpha_u U \mid Z\right) \\ &= \alpha_0 + \alpha_Z Z + \alpha_u \underbrace{E(U \mid Z)}_{=E(U)} \\ &= \alpha_0^* + \alpha_Z Z \end{aligned}$$

  where $\alpha_0^* = \alpha_0 + \alpha_u E(U)$.

- We then have that

$$\begin{aligned} E(Y \mid Z) &= \beta_0^* + \beta_A E(A \mid Z) \\ &= \beta_0^* + \beta_A \left\{\alpha_0^* + \alpha_Z Z\right\} \\ &= \beta_0^{**} + \beta_A \alpha_Z Z \end{aligned}$$

  where $\beta_0^{**} = \beta_0^* + \beta_A \alpha_0^*$.

Linear models

- We conclude that

$$E(Y \mid Z) = \beta_0^{**} + \beta_Z^* Z \quad (1)$$

where

$$\beta_Z^* = \beta_A \alpha_Z \quad (2)$$

- Equation (1) is often referred to as the reduced form obtained by regression Y on the IV Z, ignoring the exposure A.

Linear models

- Equation (2) teaches us that the total effect β_Z^* of Z on Y is the product of the effect α_Z of Z on A and the effect β_A of A on Y
- Therefore

$$\begin{aligned}\beta_A &= \frac{\beta_Z^*}{\alpha_Z} \\ &= \frac{\text{Causal effect of } Z \text{ on } Y}{\text{Causal effect of } Z \text{ on } A}\end{aligned}$$

is recovered. This estimand is known as the instrumental variable estimand

Linear models

- The linear model

$$E(Y \mid A, U, Z) = \beta_0 + \beta_A A + \beta_U U$$

makes strong homogeneity assumptions concerning an unobserved confounder U, likewise for

$$E(A \mid Z, U) = \alpha_0 + \alpha_Z Z + \alpha_u U$$

which is most natural for A continuous.

- However, under these strong homogeneity assumptions

$$\begin{aligned} \beta_A &= E\left(Y_1 - Y_0 \mid U\right) \\ &= E\left(Y_1 - Y_0\right) \end{aligned}$$

is the population average causal effect, where Y_a is the counterfactual outcome were the person exposed to a.

Beyond linearity

- The linearity assumptions made in the previous slides are often overly restrictive, fortunately, these assumptions are not necessary, and one may consider alternative assumptions, which may lead to identification of alternative causal effects.
- Imbens and Angrist (1994) and Angrist, Imbens and Rubin (1996) formulate the IV approach using counterfactuals to define the effect of treatment on individuals whose treatment status can be manipulated by the IV, also known as the complier average treatment effect.
- The key assumption they require is that of monotonicity assumption (defined later) about the effects of the IV on the exposure is sufficient for nonparametric identification of this particular causal effect.

Beyond linearity

- In a separate strand of work, Robins (1994) formulates the IV approach using counterfactuals to define the effect of treatment on the treated.
- He shows that this effect is identified assuming no heterogeneity with respect to the IV in a structural mean model, which he calls the "no current treatment value interaction" assumption.
- This approach has further been developed under similar identifying assumptions using additive, multiplicative and logistic models.

IV estimand for binary treatment: Wald estimand

- With A binary, the IV estimand reduces to

$$\frac{\text{effect of } Z \text{ on } Y}{\text{effect of } Z \text{ on } A} = \frac{E(Y \mid Z=1) - E(Y \mid Z=0)}{\Pr(A=1 \mid Z=1) - \Pr(A=1 \mid Z=0)}$$

- e.g. Z randomized treatment, A is adherence, then Wald estimand is

$$\frac{\text{effect of randomization on } Y = \text{ ITT effect}}{\text{effect of randomization on compliance}} = \frac{E(Y \mid Z=1) - E(Y \mid Z=0)}{\Pr(A=1 \mid Z=1) - \Pr(A=1 \mid Z=0)}$$

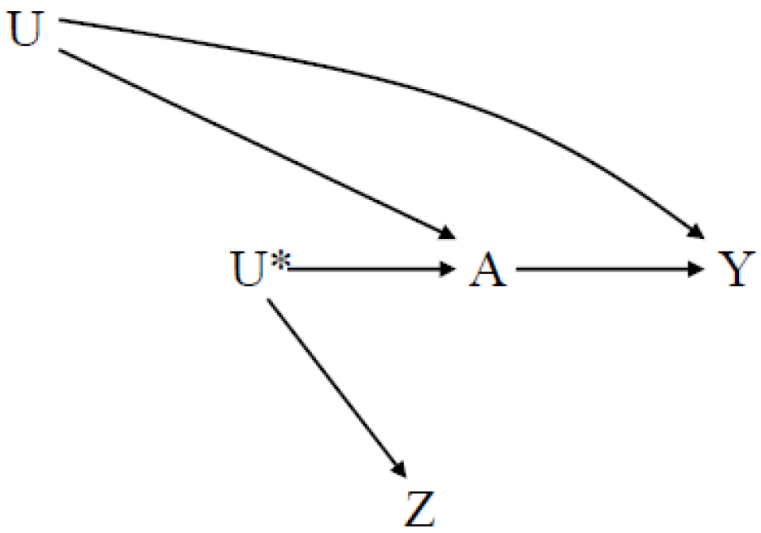
Wald estimand

- ITT effect in the numerator is inflated by a denominator that decreases with the degree of nonpcompliance.
- Equal to ITT if perfect compliance since $\Pr(A = 1 \mid Z = 1) = 1$ and $\Pr(A = 1 \mid Z = 0) = 0$
- What is its causal interpretation without assuming the previous linear model?
  - Assumptions encoded in the causal DAG not enough to identify a causal effect.
  - Interpretation of Wald estimand depends on what additional assumption is made.

Formal definition of IV

- Assumption IV.1. Z and exposure A are associated. $[Z \not\perp\!\!\!\perp A]$
  - Z has a causal effect on A
  - Z and A share common causes
- Assumption IV.2. Z affects the outcome Y only through A. $[Y_{z,a} = Y_a]$
  - No direct effect of Z on Y
  - This known as the exclusion restriction
- Assumption IV.3. Z does not share common causes with the outcome Y.
  - No confounding for the effect of Z on Y
  - If we write $[Y_{z,a} \perp\!\!\!\perp (Z, A) | U]$, then we need $[Z \perp\!\!\!\perp U]$

Non-causal IV



- The effect of Z on A is not causal, but still Z is a valid IV.
- Causal IV: $[Z \perp\!\!\!\perp (A_z, Y_{z,a})]$

Compliance types

- Suppose A and Z are both binary, let A_z denote the counterfactual A under IV value z.
- There are 4 types of individuals:
  - Never takers with $A_0 = 0, A_1 = 0$
  - Always takers with $A_0 = 1, A_1 = 1$
  - Defiers with $A_0 = 1, A_1 = 0$
  - Compliers with $A_0 = 0, A_1 = 1$

Monotonicity

- Assumption IV.4. There are no defiers, that is no individuals with $A_0 = 1, A_1 = 0$
- This assumption is sometime described as a monotonicity assumption

$$A_0 \leq A_1$$

- The assumption says that there is no individual in the population for who would be exposed, i.e., $A = 1$ under $Z = 0$, but would be unexposed, i.e., $A = 0$ under $Z = 1$.
- In the context of a randomized experiment, this would be a person who would so the exact opposite of what he/she is told to do, i.e., thus the name "defier".
- The assumption is a strong assumption and would need to be justified on a case by case basis

Identifying CACE

- We note that

$$\begin{aligned} & Y_{z=1} - Y_{z=0} \\ =& Y_{A_1,z=1} - Y_{A_0,z=0} \text{ (consistency)} \\ =& Y_{A_1} - Y_{A_0} \text{ (exclusion restriction)} \\ =& (Y_1 - Y_0) A_1 + Y_0 - (Y_1 - Y_0) A_0 - Y_0 \\ =& (Y_1 - Y_0)(A_1 - A_0) \end{aligned}$$

- Note that $(A_1 - A_0)$ takes values in $\{0, 1, -1\}$ however by monotonicity $A_1 \geq A_0$ therefore $(A_1 - A_0)$ takes values in $\{0, 1\}$. We conclude that

$$\begin{aligned} E(Y_{z=1} - Y_{z=0}) &= E[(Y_1 - Y_0)(A_1 - A_0)] \\ &= E[(Y_1 - Y_0) \mid A_1 > A_0] \Pr(A_1 > A_0) \end{aligned}$$

Identifying CACE

- We conclude that

$$E\left[\left(Y_1 - Y_0\right) \mid A_1 > A_0\right] = \frac{E\left(Y_{z=1} - Y_{z=0}\right)}{\Pr\left(A_1 > A_0\right)}$$

- The numerator is the reduced form
  $E\left(Y_{z=1} - Y_{z=0}\right) = E(Y \mid Z = 1) - E(Y \mid Z = 0)$
- The denominator is the proportion of compliers under monotonicity:

$$\begin{aligned}\Pr\left(A_1 > A_0\right) &= \Pr\left(A_1 = 1\right) - \Pr\left(A_1 = 1, A_0 = 1\right) \\ &= \Pr\left(A_1 = 1\right) - \Pr\left(A_0 = 1\right) \\ &= \Pr(A = 1 \mid Z = 1) - \Pr(A = 1 \mid Z = 0)\end{aligned}$$

therefore

$$\begin{aligned}CACE &= E\left[\left(Y_1 - Y_0\right) \mid A_1 > A_0\right] \\ &= \frac{E(Y \mid Z = 1) - E(Y \mid Z = 0)}{\Pr(A = 1 \mid Z = 1) - \Pr(A = 1 \mid Z = 0)}\end{aligned}$$

Complier causal effect

- Under Assumptions (IV.1)-(IV.4), i.e., the IV DAG, the causal effect that is identified by the Wald estimand is:

$$\frac{E(Y \mid Z=1) - E(Y \mid Z=0)}{\Pr(A=1 \mid Z=1) - \Pr(A=1 \mid Z=0)} = E\left(Y_1 - Y_0 \mid A_1 = 1, A_0 = 0\right)$$

- This is the average causal effect for compliers.
- In the context of a randomized trial, this is the causal effect for individuals who would adhere to their assignment.
- Sometimes also said to be the effect for individuals for whom treatment is manipulable.
- e.g. MTO study: Causal effect experienced by families who would moved if given a voucher, but not otherwise.

Monotonicity can be refuted

- The assumption of monotonicity can be empirically refuted, because it implies

$$\Pr(A = 1 \mid Z = 1) \geq \Pr(A = 1 \mid Z = 0)$$

- So that the assumption would be known not to hold if the inequality is not empirically verified.
- But the monotonicity assumption might not be satisfied even though the above inequality holds

Complier group

- One limitation of this approach is that the compliers are not identified, since both counterfactuals are never observed for a person.
- Furthermore, in the context of a randomized trial, unclear whether compliers in the trial are representative of compliers in the population, moreover once the results from the trial are known, the proportion of compliers is likely to change.
- Can estimate the proportion of compliers under monotonicity

$$\Pr\left(A_1 > A_0\right) = \Pr(A = 1 \mid Z = 1) - \Pr(A = 1 \mid Z = 0)$$

New drug trial

- Assumption IV.4'. If as in placebo controlled randomized trial of a new drug not easily available,

$$A_0 = 0$$

so that monotonicity holds

$$A_0 \leq A_1$$

- Under assumptions (IV.1)-(IV.3) and (IV.4')

$$\begin{aligned}
&\frac{E(Y \mid Z=1) - E(Y \mid Z=0)}{\Pr(A=1 \mid Z=1) - \Pr(A=1 \mid Z=0)} \\
&= E\left(Y_1 - Y_0 \mid A_1 = 1, A_0 = 0\right) \\
&= E\left(Y_1 - Y_0 \mid A = 1, Z = 1\right)
\end{aligned}$$

The IV estimand also identifies the effect of treatment on the treated among those randomized to treatment.

Effect of treatment on the treated

- Alternative assumption to identify the effect of treatment on the treated (ETT) .
- Assumption IV.4*. No current treatment value interaction.

$$E\left(Y_1 - Y_0 \mid A = 1, Z = 1\right) = E\left(Y_1 - Y_0 \mid A = 1, Z = 0\right)$$

- Under assumptions IV.1-IV.3 and assumption IV.4* (IV DAG and no current treatment value interaction) then

$$\frac{E(Y \mid Z = 1) - E(Y \mid Z = 0)}{\Pr(A = 1 \mid Z = 1) - \Pr(A = 1 \mid Z = 0)} = E\left(Y_1 - Y_0 \mid A = 1\right)$$

ETT

- The effect of treatment on the treated has a long tradition in epidemiology
- Sometimes only relevant effect, that is the effect experience by individuals who are actually exposed.
- In the case of a bad exposure, this is the estimand a judicial court cares about.
- The advantage of the approach is that it does not require the monotonicity assumption, in fact, the relation between A and Z can be increasing for some individuals and decreasing for others
- But the approach requires ruling out the possibility of effect heterogeneity of the effect of Z in the treated.

Average causal effect (ACE)

- What sort of assumption allows one to interpret Wald/IV estimand as population ACE$= E\left(Y_1 - Y_0\right)$.
- From CACE to ACE: Under IV.1-IV.4 and the extra homogeneity assumption

$$E\left(Y_{a=1} - Y_{a=0} \mid A_1 = 1, A_0 = 0\right) = E\left(Y_{a=1} - Y_{a=0} \mid A_1 = 0, A_0 = 0\right)$$
$$= E\left(Y_{a=1} - Y_{a=0} \mid A_1 = 1, A_0 = 1\right)$$

Then

$$\begin{aligned} & E\left(Y_{a=1} - Y_{a=0} \mid A_{z=1} = 1, A_{z=0} = 0\right) \\ =& E\left(Y_{a=1} - Y_{a=0}\right) \\ =& \frac{E(Y \mid Z = 1) - E(Y \mid Z = 0)}{\Pr(A = 1 \mid Z = 1) - \Pr(A = 1 \mid Z = 0)} \end{aligned}$$

Average causal effect (ACE)

- From ETT to ACE: Under IV.1-IV.4* and the extra homogeneity assumption

$$E\left(Y_{a=1} - Y_{a=0} \mid A = 1\right) = E\left(Y_{a=1} - Y_{a=0} \mid A = 0\right)$$

Then

$$\begin{aligned} & E\left(Y_{a=1} - Y_{a=0} \mid A = 1\right) \\ =& E\left(Y_{a=1} - Y_{a=0}\right) \\ =& \frac{E(Y \mid Z = 1) - E(Y \mid Z = 0)}{\Pr(A = 1 \mid Z = 1) - \Pr(A = 1 \mid Z = 0)} \end{aligned}$$

Average causal effect (ACE)

- These homogeneity assumptions are somewhat hard to understand, i.e., what causal mechanisms could induce them? Recall that the IV DAG given above is assumed to hold which explicitly encodes assumptions about unmeasured confounder U.
- Consider assumption IV.4$^{\#}$ which requires only one of the following holds.
  - (a) $E\left(Y_{a=1} - Y_{a=0} \mid U\right)$ does not depend on U. No latent heterogeneity of ACE (no unmeasured confounder modifies $A \rightarrow Y$ effect.

    or
  - (b) $E(A \mid Z = 1, U) - E(A \mid Z = 0, U)$ does not depend on U. No unmeasured confounder predicts compliance type, i.e., no unmeasured confounder modifies $Z \rightarrow A$ association.

Average causal effect (ACE)

- Under assumptions IV.1-IV.4$^{\#}$ the population ACE is nonparametrically identified by the IV estimand.

$$E\left(Y_{a=1} - Y_{a=0}\right) = \frac{E(Y \mid Z=1) - E(Y \mid Z=0)}{\Pr(A=1 \mid Z=1) - \Pr(A=1 \mid Z=0)}$$

- The result under IV.4$^{\#}$ a) is not surprising because it literaly states that the average causal effect is the same for each value of U.
- The result under IV.4$^{\#}$ b) is somewhat surprising and unlike any of the other assumptions to identify ACE as it does not place any restriction on the level and nature of heterogeneity of the causal effect of A on Y. It only restricts the first stage relation such that the mechanism of $Z \to A$ is independent of that of $U \to A$.

Average causal effect (ACE)

- The result follows from the following expression we previously derived

$$Y_{z=1} - Y_{z=0} = (Y_1 - Y_0)(A_1 - A_0)$$

Then

$$\begin{aligned} E(Y_{z=1} - Y_{z=0} \mid U) =& E[(Y_1 - Y_0)(A_1 - A_0) \mid U] \\ =& E[(Y_1 - Y_0) \mid U] E[(A_1 - A_0) \mid U] \\ & (\text{no unmeasured confounding given } U) \end{aligned}$$

then under IV.4$^{\#}$ a)

$$E(Y_{z=1} - Y_{z=0} \mid U) = E[(Y_1 - Y_0)] E[(A_1 - A_0) \mid U]$$

and averaging over U gives the result. Likewise under IV.4$^{\#}$ b)

$$E(Y_{z=1} - Y_{z=0} \mid U) = E[(Y_1 - Y_0) \mid U] E[(A_1 - A_0)]$$

Averaging over U gives the result.

Illustration

Causal Inference: What If Book

- Exposure A: quitting smoking between baseline and 1982 (1: yes, 0: no)
- Continuous outcome Y: Weight gain: weight in 1982 minus baseline weight (in kg)
- Proposed instrument Z: Price of a pack of cigarettes in 1982 (adjusted for inflation to 2008 US \$) in the state in which the subject was born
- Dichotomized as "Price higher than \$1.50" (1: yes, 0: no)
- Y and Z available for 1476 subjects

Illustration

- Numerator of Wald Estimator

$$\begin{aligned}\widehat{E}[Y \mid Z=1] - \widehat{E}[Y \mid Z=0] \\ = 2.686 - 2.536 = 0.1503\end{aligned}$$

- Denominator of Wald Estimator

$$\begin{aligned}\widehat{Pr}[A=1 \mid Z=1] - \widehat{Pr}[A=1 \mid Z=0] \\ = 0.2578 - 0.1951 = 0.0627\end{aligned}$$

- The IV estimate

$$\frac{\widehat{E}(Y \mid Z=1) - \widehat{E}(Y \mid Z=0)}{\widehat{Pr}(A=1 \mid Z=1) - \widehat{Pr}(A=1 \mid Z=0)} = 0.1503/0.0627 = 2.40 \ kg$$

- How to interpret these findings under the IV DAG and monotonicity? under IV DAG and no current treatment value interaction? under IV DAG and no unmeasured confounder modifies either $A \rightarrow Y$ or $Z \rightarrow A$?

The role of covariates

Suppose that together with (A, Z, Y) we also observe a vector C of all common causes of (A, Y, Z). Then, the IV approach must account for such variables, in order to

- Account for confounding of the effects of Z on Y.
- Partially account for confounding of the effects of A on Y.
- Explain variation in the outcome Y to improve efficiency.

The role of covariates

- Under assumption IV.1-IV.3, we now have that the conditional IV estimand:

$$\frac{E(Y \mid Z=1,C) - E(Y \mid Z=0,C)}{\Pr(A=1 \mid Z=1,C) - \Pr(A=1 \mid Z=0,C)}$$

  identifies the conditional CACE under IV.4, ETT under IV.4* and ACE under IV.4$^{\#}$

- For estimation, in principle, one could proceed by computing

$$\frac{\widehat{E}(Y \mid Z=1,C) - \widehat{E}(Y \mid Z=0,C)}{\widehat{Pr}(A=1 \mid Z=1,C) - \widehat{Pr}(A=1 \mid Z=0,C)}$$

  which relies on estimating using linear regression and logistic regression $\widehat{E}(Y \mid Z=0,C)$ and $\widehat{Pr}(A=1 \mid Z=1,C)$.

2SLS

- Instead of using the IV/Wald estimand, in practice, the most common approach to estimate the causal effect β_1 is to use two stage least square (2SLS).
- The approach entails consecutively fitting two linear regressions whereby:
  Stage 1: Fit a linear regression of A on Z and C and compute the predicted value

$$\widehat{A} = \widehat{E}(A \mid Z, C) = \widehat{\alpha}_0 + \widehat{\alpha}_1 Z + \widehat{\alpha}_2' C$$

  Stage 2: Fit a linear regression of Y on $\widehat{A}$ and C

$$E(Y \mid \widehat{A}, C) = \mu_0 + \mu_1 \widehat{A} + \mu_2' C$$

- The estimate $\widehat{\mu}_1$ converges to the causal effect of interest, e.g., $\beta_1 = E\left(Y_1 - Y_0 \mid A_1 = 1, A_0 = 0, C\right)$ under IV.1-IV.4

2SLS

- How does 2SLS work?
- Consider linear structural equation model for Y continuous

$$Y = \beta_0 + \beta_1 A + \beta_2' C + \epsilon$$

  with ϵ residual error
- Note that Z does not appear in this equation, how come?
- Note that because of unobserved confounding, ϵ is correlated with A even after correcting for C, and therefore does not have mean zero given A, C
- Therefore β_1 is not identified in this model

2SLS

- Suppose that $E(A \mid Z, C)$ is known and note however that,

$$\begin{aligned} Y &= \beta_0 + \beta_1 A + \beta_2' C + \epsilon \\ &= \beta_0 + \beta_1 E(A \mid Z, C) + \beta_2' C + \underbrace{\beta_1 \{A - E(A \mid Z, C)\} + \epsilon}_{=\varepsilon} \\ &= \beta_0 + \beta_1 E(A \mid Z, C) + \beta_2' C + \varepsilon \end{aligned}$$

  with new residual error ε

- This is a new linear regression model with covariates $E(A \mid Z, C)$ and C.
- Note that the effect of $E(A \mid Z, C)$ is β_1 the causal effect of interest.

2SLS

$$Y = \beta_0 + \beta_1 E(A \mid Z, C) + \beta_2' C + \varepsilon$$

- In this regression model, it can be verified that ε is uncorrelated with the new covariates $E(A \mid Z, C)$ and C and therefore the regression coefficients are all identified including the causal effect β_1.
- A condition for the above result is that $E(A \mid Z, C)$ depends on Z, i.e., Assumption IV.1

2SLS

$$Y = \beta_0 + \beta_1 E(A \mid Z, C) + \beta_2' C + \varepsilon$$

- Because $E(A \mid Z, C)$ is not generally known, it is estimated in the first stage of 2SLS. It is typically assumed that this regression is linear even if A is binary.
- This is because 2SLS is completely robust to first stage modeling error of $E(A \mid Z, C)$ provided one specifies a linear model for the latter estimated via OLS.
- In other words, it does not matter what linear model you use in the first stage, provided it includes Z.

Evaluating IV assumptions

- The strong assumptions needed to identify the causal effects of a phenotype on a disease via MR will often not hold exactly.
- These assumptions are not routinely systematically evaluated in MR applications, although such evaluation could add to the credibility of MR.
- Approach to falsify assumptions IV.2 and IV.3 (Glymour, Tchetgen Tchetgen, Robins, AJE, 2012):
  - Leverage prior causal assumption such as the known direction of confounding
  - Identify modifying subgroups
  - Instrumental inequality tests can detect extreme violations
  - Overidentification tests using several instruments
- Falsification tests may fail to reject a proposed instrument even when conditions IV.2 and IV.3 are violated.
- Nonetheless, implementing the test may increase confidence in validity of IV.

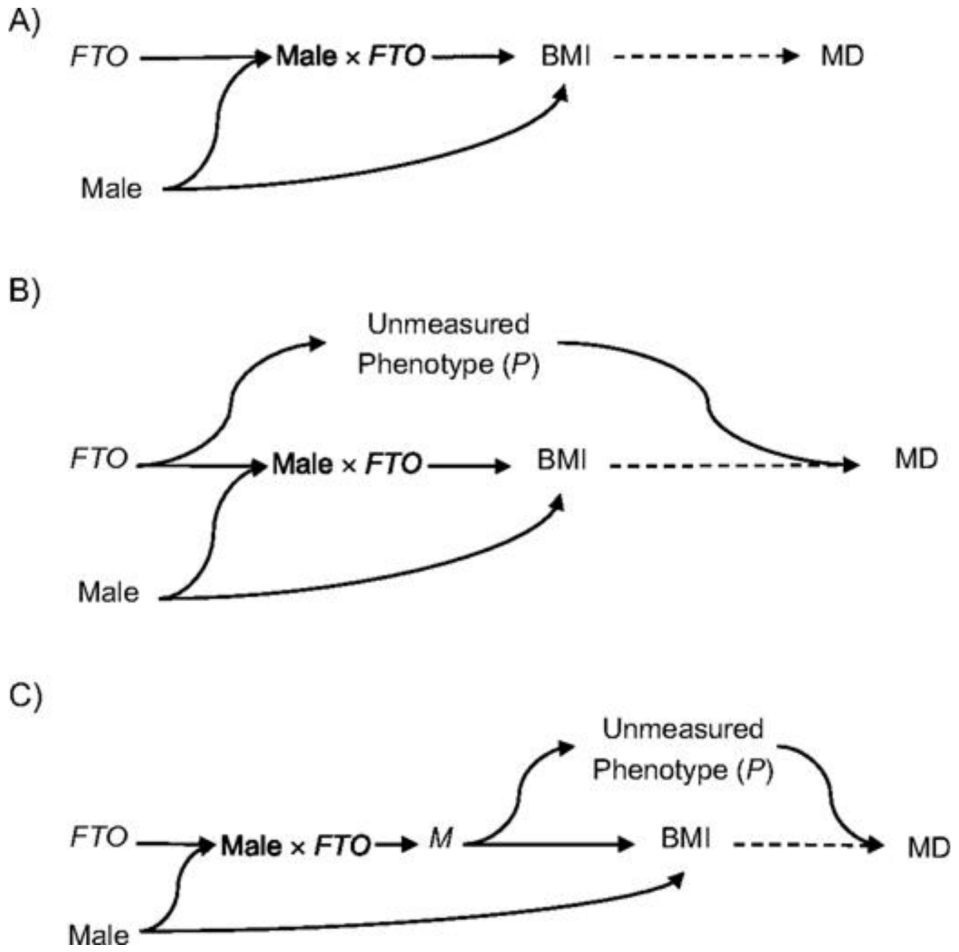
Leverage prior assumptions about confounding

- Often, field is only interested in knowing if a conventional effect estimate is biased up.
- Other direction of bias not of interest or not considered plausible.
- Compare IV effect estimate to conventional effect estimate. If conventional estimate is positively confounded, IV $<$ conventional.
- Relies on assumption regarding the direction of confounding, if you don't know the direction, the test is not informative.
- Not guaranteed to be consistent in nonlinear causal structures.

Modifying factors

- Identify factors that modify the IV-exposure relationship.
- Compare the IV effect estimate across groups in which the population association between Z and A is either silenced or reversed.
- This test could identify a biased instrument if the biasing pathway is active in both subgroups.

Modifying factors: example



IV inequality test

- Motivation: consider binary Y, A, Z, consider $E(Y_1) = \Pr(Y_1 = 1)$. It can be shown that the IV assumptions imply for each z

$$\Pr(Y = 1 \mid A = 1, z) \Pr(A = 1 \mid z) \leq \Pr(Y_1 = 1)$$
$$\Pr(Y_1 = 1) \leq \Pr(Y = 1 \mid A = 1, z) \Pr(A = 1 \mid z) + \Pr(A = 0 \mid z)$$

  therefore

$$L_1 = \max_z \Pr(Y = 1 \mid A = 1, z) \Pr(A = 1 \mid z) \leq \Pr(Y_1 = 1)$$
$$\leq \min_z \Pr(Y = 1 \mid A = 1, z) \Pr(A = 1 \mid z) + \Pr(A = 0 \mid z) = U_1$$

- IV inequality essentially test whether $U_1 < L_1$ in which case the IV assumptions IV.1-IV.3 cannot hold.

IV inequality test

- One can likewise test whether $U_0 < L_0$ for $\Pr(Y_0 = 1)$.
- The above inequalities only use the following implications of IV.2 and IV.3: $Y_a \perp Z, a = 0$. Can get a more powerful test by using $(Y_0, Y_1) \perp Z$: these are known as Bonet's inqualities. See Glymour et al (2012).
- These tests only apply to categorical phenotypes.
- They are not particularly powerful and only detect extreme forms of violations, but the more IVs are available the more sensitive the test becomes.

Overidentification tests

- Use multiple IVs to conduct overidentification tests.
- Other genes, or even polymorphisms of the same gene might provide additional instruments.
- Cannot detect violations of the IV assumptions if all instruments have identical biasing pathways.
- May also reject even when all IVs are valid if the model is incorrectly assumed to be linear or the phenotype is composite. The tests generally have low statistical power.