Principal Stratification

Zhichao Jiang

2023 summer short course

July 2023

Outline

- Potential outcomes and ACE
- Post-treatment variable
- Principal stratification
  - noncompliance
  - truncation by death
  - surrogate evaluation
- Identification and statistical inference
  - binary instrumental variable model
  - partial identification
  - principal ignorability
  - auxiliary independence
- Recent applications of principal stratification

Outline

- Potential outcomes and ACE

Post-treatment variable

Principal stratification

noncompliance
truncation by death
surrogate evaluation

Identification and statistical inference

binary instrumental variable model
partial identification
principal ignorability

Recent applications of principal stratification

Potential outcome and average causal effect

- Observed data: treatment Z_i, outcome Y_i
- Potential outcomes: $Y_i(1)$ and $Y_i(0)$
- Observed outcome: $Y_i(Z_i) \rightsquigarrow$ only one potential outcome is observed for each unit

Potential outcome and average causal effect

- Observed data: treatment Z_i, outcome Y_i
- Potential outcomes: $Y_i(1)$ and $Y_i(0)$
- Observed outcome: $Y_i(Z_i) \rightsquigarrow$ only one potential outcome is observed for each unit

- Individual causal effect: $Y_i(1) - Y_i(0) \rightsquigarrow$ difficult to estimate
- Average causal effect (ACE): $\mathbb{E}\{Y_i(1) - Y_i(0)\}$

Identification and inference

- Two methodological issues of causal inference:
    1. Identification: what can we learn if we have infinite amount of data? $\rightsquigarrow$ study design uncertainty
    2. Inference: what can we learn about identifiable quantities from a finite sample? $\rightsquigarrow$ statistical uncertainty

Identification and inference

- Two methodological issues of causal inference:
  1. Identification: what can we learn if we have infinite amount of data? $\rightsquigarrow$ study design uncertainty
  2. Inference: what can we learn about identifiable quantities from a finite sample? $\rightsquigarrow$ statistical uncertainty
- Both are important, but identification precedes inference

$$\text{Sample} \underbrace{\longrightarrow}_{\text{statistical inference}} \text{Population} \underbrace{\longrightarrow}_{\text{identification}} \text{Target parameters}$$

Identification and inference

- Two methodological issues of causal inference:
  1. Identification: what can we learn if we have infinite amount of data? $\rightsquigarrow$ study design uncertainty
  2. Inference: what can we learn about identifiable quantities from a finite sample? $\rightsquigarrow$ statistical uncertainty
- Both are important, but identification precedes inference

$$\text{Sample} \underbrace{\longrightarrow}_{\text{statistical inference}} \text{Population} \underbrace{\longrightarrow}_{\text{identification}} \text{Target parameters}$$

- In order to achieve identification, assumptions are unavoidable, but we need to figure out what assumptions are plausible in practice $\rightsquigarrow$ design trumps analysis

Identification for ACE

- Randomized experiment: $Z_i \perp\!\!\!\perp Y_i(z)$

$$\text{ACE} = \mathbb{E}(Y_i \mid Z_i = 1) - \mathbb{E}(Y_i \mid Z_i = 0)$$

Identification for ACE

- Randomized experiment: $Z_i \perp\!\!\!\perp Y_i(z)$

$$\text{ACE} = \mathbb{E}(Y_i \mid Z_i = 1) - \mathbb{E}(Y_i \mid Z_i = 0)$$

- Unconfounded observational study: $Z_i \perp\!\!\!\perp Y_i(z) \mid \mathbf{X}_i$

$$\begin{aligned}
\text{ACE} &= \mathbb{E}\{\mathbb{E}(Y_i \mid Z_i = 1, \mathbf{X}_i)\} - \mathbb{E}\{\mathbb{E}(Y_i \mid Z_i = 0, \mathbf{X}_i)\} \quad \text{(outcome reg.)} \\
&= \mathbb{E}\left\{\frac{Z_i Y_i}{\mathbb{P}(Z_i = 1 \mid \mathbf{X}_i)}\right\} - \mathbb{E}\left\{\frac{(1 - Z_i) Y_i}{1 - \mathbb{P}(Z_i = 1 \mid \mathbf{X}_i)}\right\} \quad \text{(IPW)} \\
&= \mathbb{E}\left[\frac{Z_i\{Y_i - \mu_1(X_i)\}}{\mathbb{P}(Z_i = 1 \mid \mathbf{X}_i)}\right\} - \mathbb{E}\left\{\frac{(1 - Z_i)\{Y_i - \mu_0(X_i)\}}{1 - \mathbb{P}(Z_i = 1 \mid \mathbf{X}_i)}\right] \\
&\quad + \mathbb{E}\{\mu_1(X_i) - \mu_0(X_i)\} \quad \text{(doubly robust)}
\end{aligned}$$

Identification for ACE

- Randomized experiment: $Z_i \perp\!\!\!\perp Y_i(z)$

$$\text{ACE} = \mathbb{E}(Y_i \mid Z_i = 1) - \mathbb{E}(Y_i \mid Z_i = 0)$$

- Unconfounded observational study: $Z_i \perp\!\!\!\perp Y_i(z) \mid \mathbf{X}_i$

$$\begin{aligned}
\text{ACE} &= \mathbb{E}\{\mathbb{E}(Y_i \mid Z_i = 1, \mathbf{X}_i)\} - \mathbb{E}\{\mathbb{E}(Y_i \mid Z_i = 0, \mathbf{X}_i)\} \quad \text{(outcome reg.)} \\
&= \mathbb{E}\left\{\frac{Z_i Y_i}{\mathbb{P}(Z_i = 1 \mid \mathbf{X}_i)}\right\} - \mathbb{E}\left\{\frac{(1 - Z_i) Y_i}{1 - \mathbb{P}(Z_i = 1 \mid \mathbf{X}_i)}\right\} \quad \text{(IPW)} \\
&= \mathbb{E}\left[\frac{Z_i\{Y_i - \mu_1(X_i)\}}{\mathbb{P}(Z_i = 1 \mid \mathbf{X}_i)}\right\} - \mathbb{E}\left\{\frac{(1 - Z_i)\{Y_i - \mu_0(X_i)\}}{1 - \mathbb{P}(Z_i = 1 \mid \mathbf{X}_i)}\right] \\
&\quad + \mathbb{E}\{\mu_1(X_i) - \mu_0(X_i)\} \quad \text{(doubly robust)}
\end{aligned}$$

- Latent confounding: instrumental variable, DID, synthetic control, proximal inference...

Outline

- Potential outcomes and ACE
- Post-treatment variable

Principal stratification

noncompliance
truncation by death
surrogate evaluation

Identification and statistical inference

binary instrumental variable model
partial identification
principal ignorability
auxiliary independence

Recent applications of principal stratification

Post-treatment variable

- Post-treatment variable (S): variables that are affected by the treatment (Z) and possibly affect the outcome (Y)

Post-treatment variable

- Post-treatment variable (S): variables that are affected by the treatment (Z) and possibly affect the outcome (Y)
- Studies with post-treatment variables include
  - clinical trials where the treatment receipt may be different from the treatment assignment

Post-treatment variable

- Post-treatment variable (S): variables that are affected by the treatment (Z) and possibly affect the outcome (Y)
- Studies with post-treatment variables include
  - clinical trials where the treatment receipt may be different from the treatment assignment
  - studies with long follow-up, where the dropout of the units is a post-treatment variable

Post-treatment variable

- Post-treatment variable (S): variables that are affected by the treatment (Z) and possibly affect the outcome (Y)
- Studies with post-treatment variables include
  - clinical trials where the treatment receipt may be different from the treatment assignment
  - studies with long follow-up, where the dropout of the units is a post-treatment variable
  - clinical trials, where the outcome may be truncated by death

Post-treatment variable

- Post-treatment variable (S): variables that are affected by the treatment (Z) and possibly affect the outcome (Y)
- Studies with post-treatment variables include
  - clinical trials where the treatment receipt may be different from the treatment assignment
  - studies with long follow-up, where the dropout of the units is a post-treatment variable
  - clinical trials, where the outcome may be truncated by death
  - studies with time-consuming or costly outcomes, where the surrogate (biomarker) is a post-treatment variable

Post-treatment variable

- Post-treatment variable (S): variables that are affected by the treatment (Z) and possibly affect the outcome (Y)
- Studies with post-treatment variables include
  - clinical trials where the treatment receipt may be different from the treatment assignment
  - studies with long follow-up, where the dropout of the units is a post-treatment variable
  - clinical trials, where the outcome may be truncated by death
  - studies with time-consuming or costly outcomes, where the surrogate (biomarker) is a post-treatment variable
- Adjusting for the post-treatment variable is necessary
  - post-treatment variable encodes characteristics of the unit as well as of the treatment
  - ACE is not well defined, or cannot answer the question of interest

Post-treatment variable adjustment

- Truncation by death in randomized trials
    - S_i: survival indicator

Post-treatment variable adjustment

- Truncation by death in randomized trials
  - S_i: survival indicator
  - Y_i is well defined only if $S_i = 1 \rightsquigarrow \mathbb{E}\{Y_i(1) - Y_i(0)\}$ not well defined

Post-treatment variable adjustment

- Truncation by death in randomized trials
  - S_i: survival indicator
  - Y_i is well defined only if $S_i = 1 \rightsquigarrow \mathbb{E}\{Y_i(1) - Y_i(0)\}$ not well defined
  - Naive method: $\mathbb{E}(Y_i \mid Z_i = 1, S_i = 1) - \mathbb{E}(Y_i \mid Z_i = 0, S_i = 1)$

Post-treatment variable adjustment

- Truncation by death in randomized trials
  - S_i: survival indicator
  - Y_i is well defined only if $S_i = 1 \rightsquigarrow \mathbb{E}\{Y_i(1) - Y_i(0)\}$ not well defined
  - Naive method: $\mathbb{E}(Y_i \mid Z_i = 1, S_i = 1) - \mathbb{E}(Y_i \mid Z_i = 0, S_i = 1)$

- Noncompliance in randomized trials
  - Z_i: treatment assignment; S_i: treatment receipt

Post-treatment variable adjustment

- Truncation by death in randomized trials
  - S_i: survival indicator
  - Y_i is well defined only if $S_i = 1 \rightsquigarrow \mathbb{E}\{Y_i(1) - Y_i(0)\}$ not well defined
  - Naive method: $\mathbb{E}(Y_i \mid Z_i = 1, S_i = 1) - \mathbb{E}(Y_i \mid Z_i = 0, S_i = 1)$

- Noncompliance in randomized trials
  - Z_i: treatment assignment; S_i: treatment receipt
  - intention-to-treat (ITT) effect: effect of $Z \neq$ treatment effect

Post-treatment variable adjustment

- Truncation by death in randomized trials
  - S_i: survival indicator
  - Y_i is well defined only if $S_i = 1 \rightsquigarrow \mathbb{E}\{Y_i(1) - Y_i(0)\}$ not well defined
  - Naive method: $\mathbb{E}(Y_i \mid Z_i = 1, S_i = 1) - \mathbb{E}(Y_i \mid Z_i = 0, S_i = 1)$

- Noncompliance in randomized trials
  - Z_i: treatment assignment; S_i: treatment receipt
  - intention-to-treat (ITT) effect: effect of $Z \neq$ treatment effect
  - as-treated analysis: $\mathbb{E}(Y_i \mid S_i = 1) - \mathbb{E}(Y_i \mid S_i = 0)$
  - per-protocol analysis: $\mathbb{E}(Y_i \mid Z_i = 1, S_i = 1) - \mathbb{E}(Y_i \mid Z_i = 0, S_i = 0)$

Post-treatment variable adjustment

- Truncation by death in randomized trials
  - S_i: survival indicator
  - Y_i is well defined only if $S_i = 1 \rightsquigarrow \mathbb{E}\{Y_i(1) - Y_i(0)\}$ not well defined
  - Naive method: $\mathbb{E}(Y_i \mid Z_i = 1, S_i = 1) - \mathbb{E}(Y_i \mid Z_i = 0, S_i = 1)$
- Noncompliance in randomized trials
  - Z_i: treatment assignment; S_i: treatment receipt
  - intention-to-treat (ITT) effect: effect of $Z \neq$ treatment effect
  - as-treated analysis: $\mathbb{E}(Y_i \mid S_i = 1) - \mathbb{E}(Y_i \mid S_i = 0)$
  - per-protocol analysis: $\mathbb{E}(Y_i \mid Z_i = 1, S_i = 1) - \mathbb{E}(Y_i \mid Z_i = 0, S_i = 0)$
- Surrogate
  - ACE provides no information about surrogate S

Post-treatment variable adjustment

- Truncation by death in randomized trials
  - S_i: survival indicator
  - Y_i is well defined only if $S_i = 1 \rightsquigarrow \mathbb{E}\{Y_i(1) - Y_i(0)\}$ not well defined
  - Naive method: $\mathbb{E}(Y_i \mid Z_i = 1, S_i = 1) - \mathbb{E}(Y_i \mid Z_i = 0, S_i = 1)$

- Noncompliance in randomized trials
  - Z_i: treatment assignment; S_i: treatment receipt
  - intention-to-treat (ITT) effect: effect of $Z \neq$ treatment effect
  - as-treated analysis: $\mathbb{E}(Y_i \mid S_i = 1) - \mathbb{E}(Y_i \mid S_i = 0)$
  - per-protocol analysis: $\mathbb{E}(Y_i \mid Z_i = 1, S_i = 1) - \mathbb{E}(Y_i \mid Z_i = 0, S_i = 0)$

- Surrogate
  - ACE provides no information about surrogate S
  - criterion for a good surrogate: $Z \perp\!\!\!\perp Y \mid S$?

Issue of adjusting for observed post-treatment variable

- Potential values: $Y_i(z)$ and $S_i(z)$
- Observed variables: $S_i = S_i(Z_i)$ and $Y_i = Y_i(Z_i)$
- Randomized experiments: $\{Y_i(1), Y_i(0), S_i(1), S_i(0)\} \perp\!\!\!\perp Z_i$

Issue of adjusting for observed post-treatment variable

- Potential values: $Y_i(z)$ and $S_i(z)$
- Observed variables: $S_i = S_i(Z_i)$ and $Y_i = Y_i(Z_i)$
- Randomized experiments: $\{Y_i(1), Y_i(0), S_i(1), S_i(0)\} \perp\!\!\!\perp Z_i$

- A causal effect is defined to be the comparison between potential outcomes for the same units
  - Comparison of $\{Y_i(1), i \in \text{set}_1\}$ and $\{Y_i(0), i \in \text{set}_0\}$
  - Two sets should be identical: $\text{set}_1 = \text{set}_0$
  - ACE: $\mathbb{E}(Y_i \mid Z_i = 1) - \mathbb{E}(Y_i \mid Z_i = 0) = \mathbb{E}\{Y_i(1)\} - \mathbb{E}\{Y_i(0)\}$

Issue of adjusting for observed post-treatment variable

- Comparison of $\mathbb{E}(Y_i \mid Z_i = 1, S_i = 1)$ and $\mathbb{E}(Y_i \mid Z_i = 0, S_i = 1)$

Issue of adjusting for observed post-treatment variable

- Comparison of $\mathbb{E}(Y_i \mid Z_i = 1, S_i = 1)$ and $\mathbb{E}(Y_i \mid Z_i = 0, S_i = 1)$
    - $\mathbb{E}(Y_i \mid Z_i = z, S_i = 1) = \mathbb{E}\{Y_i(z) \mid Z_i = z, S_i(z) = 1\} = \mathbb{E}\{Y_i(z) \mid S_i(z) = 1\}$

Issue of adjusting for observed post-treatment variable

- Comparison of $\mathbb{E}(Y_i \mid Z_i = 1, S_i = 1)$ and $\mathbb{E}(Y_i \mid Z_i = 0, S_i = 1)$
  - $\mathbb{E}(Y_i \mid Z_i = z, S_i = 1) = \mathbb{E}\{Y_i(z) \mid Z_i = z, S_i(z) = 1\} = \mathbb{E}\{Y_i(z) \mid S_i(z) = 1\}$
  - comparison of sets $\{i : S_i(1) = 1\}$ and $\{i : S_i(0) = 1\}$

Issue of adjusting for observed post-treatment variable

- Comparison of $\mathbb{E}(Y_i \mid Z_i = 1, S_i = 1)$ and $\mathbb{E}(Y_i \mid Z_i = 0, S_i = 1)$
  - $\mathbb{E}(Y_i \mid Z_i = z, S_i = 1) = \mathbb{E}\{Y_i(z) \mid Z_i = z, S_i(z) = 1\} = \mathbb{E}\{Y_i(z) \mid S_i(z) = 1\}$
  - comparison of sets $\{i : S_i(1) = 1\}$ and $\{i : S_i(0) = 1\}$

- Comparison of $\mathbb{E}(Y_i \mid Z_i = 1, S_i = 1)$ and $\mathbb{E}(Y_i \mid Z_i = 0, S_i = 0)$

Issue of adjusting for observed post-treatment variable

- Comparison of $\mathbb{E}(Y_i \mid Z_i = 1, S_i = 1)$ and $\mathbb{E}(Y_i \mid Z_i = 0, S_i = 1)$
  - $\mathbb{E}(Y_i \mid Z_i = z, S_i = 1) = \mathbb{E}\{Y_i(z) \mid Z_i = z, S_i(z) = 1\} = \mathbb{E}\{Y_i(z) \mid S_i(z) = 1\}$
  - comparison of sets $\{i : S_i(1) = 1\}$ and $\{i : S_i(0) = 1\}$
- Comparison of $\mathbb{E}(Y_i \mid Z_i = 1, S_i = 1)$ and $\mathbb{E}(Y_i \mid Z_i = 0, S_i = 0)$
  - $\mathbb{E}(Y_i \mid Z_i = 0, S_i = 0) = \mathbb{E}\{Y_i(0) \mid Z_i = 0, S_i(0) = 0\} = \mathbb{E}\{Y_i(0) \mid S_i(0) = 0\}$

Issue of adjusting for observed post-treatment variable

- Comparison of $\mathbb{E}(Y_i \mid Z_i = 1, S_i = 1)$ and $\mathbb{E}(Y_i \mid Z_i = 0, S_i = 1)$
  - $\mathbb{E}(Y_i \mid Z_i = z, S_i = 1) = \mathbb{E}\{Y_i(z) \mid Z_i = z, S_i(z) = 1\} = \mathbb{E}\{Y_i(z) \mid S_i(z) = 1\}$
  - comparison of sets $\{i : S_i(1) = 1\}$ and $\{i : S_i(0) = 1\}$

- Comparison of $\mathbb{E}(Y_i \mid Z_i = 1, S_i = 1)$ and $\mathbb{E}(Y_i \mid Z_i = 0, S_i = 0)$
  - $\mathbb{E}(Y_i \mid Z_i = 0, S_i = 0) = \mathbb{E}\{Y_i(0) \mid Z_i = 0, S_i(0) = 0\} = \mathbb{E}\{Y_i(0) \mid S_i(0) = 0\}$
  - comparison of sets $\{i : S_i(1) = 1\}$ and $\{i : S_i(0) = 0\}$

Issue of adjusting for observed post-treatment variable

- Comparison of $\mathbb{E}(Y_i \mid Z_i = 1, S_i = 1)$ and $\mathbb{E}(Y_i \mid Z_i = 0, S_i = 1)$
  - $\mathbb{E}(Y_i \mid Z_i = z, S_i = 1) = \mathbb{E}\{Y_i(z) \mid Z_i = z, S_i(z) = 1\} = \mathbb{E}\{Y_i(z) \mid S_i(z) = 1\}$
  - comparison of sets $\{i : S_i(1) = 1\}$ and $\{i : S_i(0) = 1\}$

- Comparison of $\mathbb{E}(Y_i \mid Z_i = 1, S_i = 1)$ and $\mathbb{E}(Y_i \mid Z_i = 0, S_i = 0)$
  - $\mathbb{E}(Y_i \mid Z_i = 0, S_i = 0) = \mathbb{E}\{Y_i(0) \mid Z_i = 0, S_i(0) = 0\} = \mathbb{E}\{Y_i(0) \mid S_i(0) = 0\}$
  - comparison of sets $\{i : S_i(1) = 1\}$ and $\{i : S_i(0) = 0\}$

- Adjusting for the observed post-treatment variable does NOT yield valid causal quantities

Example

- Treatment: Z_i; survival indicator: S_i; outcome: Y_i

Example

- Treatment: Z_i; survival indicator: S_i; outcome: Y_i
- Treatment has no effect on outcome: $Y_i(1) = Y_i(0)$ for all i

Example

- Treatment: Z_i; survival indicator: S_i; outcome: Y_i
- Treatment has no effect on outcome: $Y_i(1) = Y_i(0)$ for all i
- Individuals with $S_i(0) = 1$ are healthier than those with $S_i(1) = 1$

Example

- Treatment: Z_i; survival indicator: S_i; outcome: Y_i

- Treatment has no effect on outcome: $Y_i(1) = Y_i(0)$ for all i
- Individuals with $S_i(0) = 1$ are healthier than those with $S_i(1) = 1$
  - Larger $Y_i(z)$ for individuals with $S_i(0) = 1$

Example

- Treatment: Z_i; survival indicator: S_i; outcome: Y_i
- Treatment has no effect on outcome: $Y_i(1) = Y_i(0)$ for all i
- Individuals with $S_i(0) = 1$ are healthier than those with $S_i(1) = 1$
  - Larger $Y_i(z)$ for individuals with $S_i(0) = 1$
- Analysis using the survived individuals

$$
\begin{aligned}
& \mathbb{E}(Y_i \mid Z_i = 1, S_i = 1) - \mathbb{E}(Y_i \mid Z_i = 0, S_i = 1) \\
= \; & \mathbb{E}\{Y_i(1) \mid S_i(1) = 1\} - \mathbb{E}\{Y_i(0) \mid S_i(0) = 1\} \\
= \; & \mathbb{E}\{Y_i(0) \mid S_i(1) = 1\} - \mathbb{E}\{Y_i(0) \mid S_i(0) = 1\} \\
< \; & 0
\end{aligned}
$$

The hazard of hazard ratio (Hernán, 2010)

- Treatment Z_i and time-to-event outcome Y_i

The hazard of hazard ratio (Hernán, 2010)

- Treatment Z_i and time-to-event outcome Y_i
- Hazard: the event rate at time t conditional on survival until time t or later

$$\lim_{\Delta t \to 0} \frac{\mathbb{P}(t \leq Y < t + \Delta t \mid Y \geq t)}{\Delta t}$$

The hazard of hazard ratio (Hernán, 2010)

- Treatment Z_i and time-to-event outcome Y_i
- Hazard: the event rate at time t conditional on survival until time t or later

$$\lim_{\Delta t \to 0} \frac{\mathbb{P}(t \leq Y < t + \Delta t \mid Y \geq t)}{\Delta t}$$

- Survival analysis assumes models for hazard, e.g., Cox model, additive hazard model

The hazard of hazard ratio (Hernán, 2010)

- Treatment Z_i and time-to-event outcome Y_i
- Hazard: the event rate at time t conditional on survival until time t or later

$$\lim_{\Delta t \to 0} \frac{\mathbb{P}(t \leq Y < t + \Delta t \mid Y \geq t)}{\Delta t}$$

- Survival analysis assumes models for hazard, e.g., Cox model, additive hazard model
- Hazard ratio between the treatment and control compares

$$\lim_{\Delta t \to 0} \frac{\mathbb{P}(t \leq Y(1) < t + \Delta t \mid Y(1) \geq t)}{\Delta t}, \quad \lim_{\Delta t \to 0} \frac{\mathbb{P}(t \leq Y(0) < t + \Delta t \mid Y(0) \geq t)}{\Delta t}$$

which compares the populations $\{i : Y(1) \geq t\}$ and $\{i : Y(0) \geq t\}$

The hazard of hazard ratio (Hernán, 2010)

- Treatment Z_i and time-to-event outcome Y_i
- Hazard: the event rate at time t conditional on survival until time t or later

$$\lim_{\Delta t \to 0} \frac{\mathbb{P}(t \leq Y < t + \Delta t \mid Y \geq t)}{\Delta t}$$

- Survival analysis assumes models for hazard, e.g., Cox model, additive hazard model
- Hazard ratio between the treatment and control compares

$$\lim_{\Delta t \to 0} \frac{\mathbb{P}(t \leq Y(1) < t + \Delta t \mid Y(1) \geq t)}{\Delta t}, \quad \lim_{\Delta t \to 0} \frac{\mathbb{P}(t \leq Y(0) < t + \Delta t \mid Y(0) \geq t)}{\Delta t}$$

  which compares the populations $\{i : Y(1) \geq t\}$ and $\{i : Y(0) \geq t\}$
- Hazard ratio has a built-in selection bias

Outline

- Potential outcomes and ACE
- Post-treatment variable
- Principal stratification
    - noncompliance
    - truncation by death
    - surrogate evaluation
- Identification and statistical inference
    - binary instrumental variable model
    - partial identification
    - principal ignorability
    - auxiliary independence
- Recent applications of principal stratification

Principal stratification

- Principal stratification: stratification based on the joint potential values $\{S_i(1), S_i(0)\}$, denoted by U_i
  - U_i is an unobserved variable
  - U_i is unaffected by the treatment, similar to a covariate
  - binary $S_i \rightsquigarrow$ four valued U_i
  - principal effect: comparison of $\{Y_i(1) : U_i = u\}$ and $\{Y_i(0) : U_i = u\}$

Principal stratification

- Principal stratification: stratification based on the joint potential values $\{S_i(1), S_i(0)\}$, denoted by U_i
  - U_i is an unobserved variable
  - U_i is unaffected by the treatment, similar to a covariate
  - binary $S_i \rightsquigarrow$ four valued U_i
  - principal effect: comparison of $\{Y_i(1) : U_i = u\}$ and $\{Y_i(0) : U_i = u\}$

- Any principal effect is a causal effect
  - principal causal effect: $\text{ACE}_u = \mathbb{E}\{Y_i(1) - Y_i(0) \mid U_i = u\}$
  - scientific meanings differ in different applications

$$\text{ACE} = \sum_u \mathbb{E}\{Y_i(1) - Y_i(0) \mid U_i = u\}\mathbb{P}(U_i = u)$$

Outline

- Recap: potential outcomes and ACE
- Post-treatment variable
- Principal stratification
  - noncompliance
  - truncation by death
  - surrogate evaluation
- Identification and statistical inference
  - binary instrumental variable model
  - partial identification
  - principal ignorability
  - auxiliary independence
- Recent applications of principal stratification

Encouragement Design

- Often, for ethical and logistical reasons, we cannot force all experimental units to follow the randomized treatment assignment
  1. some in the treatment group refuse to take the treatment
  2. some in the control group manage to receive the treatment

Encouragement Design

- Often, for ethical and logistical reasons, we cannot force all experimental units to follow the randomized treatment assignment
  1. some in the treatment group refuse to take the treatment
  2. some in the control group manage to receive the treatment

- Encouragement design: randomize the encouragement to receive the treatment rather than the receipt of the treatment
  - job training program, insurance program

Encouragement Design

- Often, for ethical and logistical reasons, we cannot force all experimental units to follow the randomized treatment assignment
  1. some in the treatment group refuse to take the treatment
  2. some in the control group manage to receive the treatment

- Encouragement design: randomize the encouragement to receive the treatment rather than the receipt of the treatment
  - job training program, insurance program

- Treatment effect: effect of the actually received treatment

Noncompliance

- Treatment assignment Z_i; treatment receipt S_i

Noncompliance

- Treatment assignment Z_i; treatment receipt S_i
- Four principal strata represent compliance behavior:
  - compliers $(S_i(1), S_i(0)) = (1, 0)$
  - non-compliers $\begin{cases} \textit{always}-\textit{takers} & (S_i(1), S_i(0)) = (1, 1) \\ \textit{never}-\textit{takers} & (S_i(1), S_i(0)) = (0, 0) \\ \textit{defiers} & (S_i(1), S_i(0)) = (0, 1) \end{cases}$
- Observed strata and compliance behavior:

	$Z_i = 1$	$Z_i = 0$
$S_i = 1$	Complier/Always-taker	Defier/Always-taker
$S_i = 0$	Defier/Never-taker	Complier/Never-taker

Principal causal effects

- Four principal causal effects
  - complier average causal effect (CACE): $\mathbb{E}\{Y_i(1) - Y_i(0) \mid U_i = (1, 0)\}$
  - always takers: $\mathbb{E}\{Y_i(1) - Y_i(0) \mid U_i = (1, 1)\}$
  - never takers: $\mathbb{E}\{Y_i(1) - Y_i(0) \mid U_i = (0, 0)\}$
  - defiers: $\mathbb{E}\{Y_i(1) - Y_i(0) \mid U_i = (0, 1)\}$

Principal causal effects

- Four principal causal effects
  - complier average causal effect (CACE): $\mathbb{E}\{Y_i(1) - Y_i(0) \mid U_i = (1,0)\}$
  - always takers: $\mathbb{E}\{Y_i(1) - Y_i(0) \mid U_i = (1,1)\}$
  - never takers: $\mathbb{E}\{Y_i(1) - Y_i(0) \mid U_i = (0,0)\}$
  - defiers: $\mathbb{E}\{Y_i(1) - Y_i(0) \mid U_i = (0,1)\}$

- For compliers, $Z_i = S_i \rightsquigarrow$
  CACE $=$ treatment effect for compliers

Principal causal effects

- Four principal causal effects
  - complier average causal effect (CACE): $\mathbb{E}\{Y_i(1) - Y_i(0) \mid U_i = (1,0)\}$
  - always takers: $\mathbb{E}\{Y_i(1) - Y_i(0) \mid U_i = (1,1)\}$
  - never takers: $\mathbb{E}\{Y_i(1) - Y_i(0) \mid U_i = (0,0)\}$
  - defiers: $\mathbb{E}\{Y_i(1) - Y_i(0) \mid U_i = (0,1)\}$

- For compliers, $Z_i = S_i \rightsquigarrow$
  CACE $=$ treatment effect for compliers

- CACE $\neq$ overall treatment effect unless the treatment effect for non-compliers equals CACE

Special case: one-sided noncompliance

- Individuals in the control have no access to the treatment
- In some vaccine trials, vaccine antigens must be present to induce a specific immune response

Special case: one-sided noncompliance

- Individuals in the control have no access to the treatment
- In some vaccine trials, vaccine antigens must be present to induce a specific immune response

- One-sided noncompliance: $S_i(0) = \text{constant}$

Special case: one-sided noncompliance

- Individuals in the control have no access to the treatment
- In some vaccine trials, vaccine antigens must be present to induce a specific immune response

- One-sided noncompliance: $S_i(0) = \text{constant}$
  - principal stratum simplifies as $U_i = S_i(1)$
  - two principal strata: $S_i(1) = 1$ and $S_i(1) = 0$

Special case: one-sided noncompliance

- Individuals in the control have no access to the treatment
- In some vaccine trials, vaccine antigens must be present to induce a specific immune response

- One-sided noncompliance: $S_i(0) = \text{constant}$
  - principal stratum simplifies as $U_i = S_i(1)$
  - two principal strata: $S_i(1) = 1$ and $S_i(1) = 0$

- $\text{CACE} = \mathbb{E}\{Y_i(1) - Y_i(0) \mid S_i(1) = 1\}$

Outline

- Recap: potential outcomes and ACE
- Post-treatment variable
- Principal stratification
  - noncompliance
  - truncation by death
  - surrogate evaluation
- Identification and statistical inference
  - binary instrumental variable model
  - partial identification
  - principal ignorability
  - auxiliary independence
- Recent applications of principal stratification

Truncation by death

- Example 1: the effect of treatment (Z) on quality of life (Y) $\rightsquigarrow$ some patients may die before the outcome is observed
- Example 2: the effect of job training on hourly wage $\rightsquigarrow$ some subjects may be unemployed
- In example 1, S_i is the survival status and in example 2, S_i is the employment status

Truncation by death

- Example 1: the effect of treatment (Z) on quality of life (Y) $\rightsquigarrow$ some patients may die before the outcome is observed
- Example 2: the effect of job training on hourly wage $\rightsquigarrow$ some subjects may be unemployed
- In example 1, S_i is the survival status and in example 2, S_i is the employment status

- Traditional method treats the truncation by death problem as a standard missing data problem (censoring by death)
  - Heckman selection model: models for $\mathbb{E}(Y_i \mid Z_i, \mathbf{X}_i)$ and $\mathbb{E}(S_i \mid \mathbf{X}_i)$
  - assumes that the outcome is well-defined for all units

Principal stratification for truncation by death problem

- Four principal strata:
  - always survivors: $(S_i(1) = 1, S_i(0) = 1) \rightsquigarrow Y_i$ always well defined
  - $(S_i(1) = 1, S_i(0) = 0) \rightsquigarrow Y_i$ not well defined when $Z_i = 0$
  - $(S_i(1) = 0, S_i(0) = 1) \rightsquigarrow Y_i$ not well defined when $Z_i = 1$
  - $(S_i(1) = 0, S_i(0) = 0) \rightsquigarrow Y_i$ never well defined

Principal stratification for truncation by death problem

- Four principal strata:
  - always survivors: $(S_i(1) = 1, S_i(0) = 1) \rightsquigarrow Y_i$ always well defined
  - $(S_i(1) = 1, S_i(0) = 0) \rightsquigarrow Y_i$ not well defined when $Z_i = 0$
  - $(S_i(1) = 0, S_i(0) = 1) \rightsquigarrow Y_i$ not well defined when $Z_i = 1$
  - $(S_i(1) = 0, S_i(0) = 0) \rightsquigarrow Y_i$ never well defined

- Comparison of $Y_i(1)$ and $Y_i(0)$ is valid only for always survivors

Principal stratification for truncation by death problem

- Four principal strata:
  - always survivors: $(S_i(1) = 1, S_i(0) = 1) \rightsquigarrow Y_i$ always well defined
  - $(S_i(1) = 1, S_i(0) = 0) \rightsquigarrow Y_i$ not well defined when $Z_i = 0$
  - $(S_i(1) = 0, S_i(0) = 1) \rightsquigarrow Y_i$ not well defined when $Z_i = 1$
  - $(S_i(1) = 0, S_i(0) = 0) \rightsquigarrow Y_i$ never well defined

- Comparison of $Y_i(1)$ and $Y_i(0)$ is valid only for always survivors
- Survivor average causal effect (Rubin, 2006)

$$\text{SACE} = \mathbb{E}\{Y_i(1) - Y_i(0) \mid S_i(1) = 1, S_i(0) = 1\}$$

Principal stratification for truncation by death problem

- Four principal strata:
  - always survivors: $(S_i(1) = 1, S_i(0) = 1) \rightsquigarrow Y_i$ always well defined
  - $(S_i(1) = 1, S_i(0) = 0) \rightsquigarrow Y_i$ not well defined when $Z_i = 0$
  - $(S_i(1) = 0, S_i(0) = 1) \rightsquigarrow Y_i$ not well defined when $Z_i = 1$
  - $(S_i(1) = 0, S_i(0) = 0) \rightsquigarrow Y_i$ never well defined

- Comparison of $Y_i(1)$ and $Y_i(0)$ is valid only for always survivors
- Survivor average causal effect (Rubin, 2006)

$$\mathsf{SACE} = \mathbb{E}\{Y_i(1) - Y_i(0) \mid S_i(1) = 1, S_i(0) = 1\}$$

- Other principal causal effects are not well defined

Outline

- Potential outcomes and ACE
- Post-treatment variable
- Principal stratification
  - noncompliance
  - truncation by death
  - surrogate evaluation
- Identification and statistical inference
  - binary instrumental variable model
  - partial identification
  - principal ignorability
  - auxiliary independence
- Recent applications of principal stratification

Surrogate

- Outcomes may be time-consuming or costly to measure
  - Develop of measurement tool and training of people
  - Long term outcomes: long-term survival in clinical trials and long-run return to early life interventions

Surrogate

- Outcomes may be time-consuming or costly to measure
  - Develop of measurement tool and training of people
  - Long term outcomes: long-term survival in clinical trials and long-run return to early life interventions

- Surrogate: a proxy of outcome that is easy to measure
  - should be strongly associated with outcome? $\rightsquigarrow$ large $\text{cor}(S_i, Y_i)$
  - should be on the causal pathway from treatment to outcome? $Z_i \to S_i \to Y_i$

Surrogate

- Outcomes may be time-consuming or costly to measure
  - Develop of measurement tool and training of people
  - Long term outcomes: long-term survival in clinical trials and long-run return to early life interventions
- Surrogate: a proxy of outcome that is easy to measure
  - should be strongly associated with outcome? $\rightsquigarrow$ large $\mathrm{cor}(S_i, Y_i)$
  - should be on the causal pathway from treatment to outcome? $Z_i \to S_i \to Y_i$
- Examples of surrogates
  - CD4 cell count as surrogate for the survival status of HIV patients
  - short-term survival as surrogate for long-term survival
  - kindergarten test scores as surrogate for long-run return

Catastrophic consequence using an invalid surrogate

- Ventricular arrhythmia as surrogate for death related to cardiac complications
  - VA is associated with a fourfold increase in the risk for death
  - it was hypothesized that suppression of ventricular arrhythmias would reduce the rate of death

Catastrophic consequence using an invalid surrogate

- Ventricular arrhythmia as surrogate for death related to cardiac complications
  - VA is associated with a fourfold increase in the risk for death
  - it was hypothesized that suppression of ventricular arrhythmias would reduce the rate of death
- Drugs found to suppress arrhythmia were approved by the FDA
  - more than 200 000 persons per year took these drugs

Catastrophic consequence using an invalid surrogate

- Ventricular arrhythmia as surrogate for death related to cardiac complications
  - VA is associated with a fourfold increase in the risk for death
  - it was hypothesized that suppression of ventricular arrhythmias would reduce the rate of death
- Drugs found to suppress arrhythmia were approved by the FDA
  - more than 200 000 persons per year took these drugs
- Follow up trials showed that the drugs increased mortality
  - tens of thousands of patients died in America's worst drug disaster
  - the casualties were estimated to approach levels close to those of the war in Vietnam

Prentice's criteria of a statistical surrogate

- Treatment Z_i; surrogate S_i; outcome Y_i

Prentice's criteria of a statistical surrogate

- Treatment Z_i; surrogate S_i; outcome Y_i
- S_i is a statistical surrogate for a comparison of the effect of $Z_i = 1$ versus $Z_i = 0$ on Y_i, if $Z_i \perp\!\!\!\perp Y_i \mid S_i$ (Prentice 1989)

$$Z_i \longrightarrow S_i \longrightarrow Y_i \implies Z_i \perp\!\!\!\perp Y_i \mid S_i, \; S_i \not\!\perp\!\!\!\perp Y_i$$

Prentice's criteria of a statistical surrogate

- Treatment Z_i; surrogate S_i; outcome Y_i
- S_i is a statistical surrogate for a comparison of the effect of $Z_i = 1$ versus $Z_i = 0$ on Y_i, if $Z_i \perp\!\!\!\perp Y_i \mid S_i$ (Prentice 1989)

$$Z_i \longrightarrow S_i \longrightarrow Y_i \implies Z_i \perp\!\!\!\perp Y_i \mid S_i,\ S_i \not\perp\!\!\!\perp Y_i$$

$$\begin{aligned} & \mathbb{E}\{Y_i(1) - Y_i(0)\} \\ = \ & \mathbb{E}\{\mathbb{E}(Y_i \mid S_i, Z_i = 1) \mid Z_i = 1\} - \mathbb{E}\{\mathbb{E}(Y_i \mid S_i, Z_i = 0) \mid Z_i = 0\} \\ = \ & \mathbb{E}\{\mathbb{E}(Y_i \mid S_i) \mid Z_i = 1\} - \mathbb{E}\{\mathbb{E}(Y_i \mid S_i) \mid Z_i = 0\} \end{aligned}$$

Prentice's criteria of a statistical surrogate

- Treatment Z_i; surrogate S_i; outcome Y_i
- S_i is a statistical surrogate for a comparison of the effect of $Z_i = 1$ versus $Z_i = 0$ on Y_i, if $Z_i \perp\!\!\!\perp Y_i \mid S_i$ (Prentice 1989)

$$Z_i \longrightarrow S_i \longrightarrow Y_i \implies Z_i \perp\!\!\!\perp Y_i \mid S_i, S_i \not\perp\!\!\!\perp Y_i$$

$$\begin{aligned} & \mathbb{E}\{Y_i(1) - Y_i(0)\} \\ = \; & \mathbb{E}\{\mathbb{E}(Y_i \mid S_i, Z_i = 1) \mid Z_i = 1\} - \mathbb{E}\{\mathbb{E}(Y_i \mid S_i, Z_i = 0) \mid Z_i = 0\} \\ = \; & \mathbb{E}\{\mathbb{E}(Y_i \mid S_i) \mid Z_i = 1\} - \mathbb{E}\{\mathbb{E}(Y_i \mid S_i) \mid Z_i = 0\} \end{aligned}$$

  - $\mathbb{E}\{h(S_i(1)) - h(S_i(0))\} = 0$ implies $\mathbb{E}\{Y_i(1) - Y_i(0)\} = 0$

Prentice's criteria of a statistical surrogate

- Treatment Z_i; surrogate S_i; outcome Y_i
- S_i is a statistical surrogate for a comparison of the effect of $Z_i = 1$ versus $Z_i = 0$ on Y_i, if $Z_i \perp\!\!\!\perp Y_i \mid S_i$ (Prentice 1989)

$$Z_i \longrightarrow S_i \longrightarrow Y_i \implies Z_i \perp\!\!\!\perp Y_i \mid S_i, S_i \not\perp\!\!\!\perp Y_i$$

$$\begin{aligned} & \mathbb{E}\{Y_i(1) - Y_i(0)\} \\ = \; & \mathbb{E}\{\mathbb{E}(Y_i \mid S_i, Z_i = 1) \mid Z_i = 1\} - \mathbb{E}\{\mathbb{E}(Y_i \mid S_i, Z_i = 0) \mid Z_i = 0\} \\ = \; & \mathbb{E}\{\mathbb{E}(Y_i \mid S_i) \mid Z_i = 1\} - \mathbb{E}\{\mathbb{E}(Y_i \mid S_i) \mid Z_i = 0\} \end{aligned}$$

  - $\mathbb{E}\{h(S_i(1)) - h(S_i(0))\} = 0$ implies $\mathbb{E}\{Y_i(1) - Y_i(0)\} = 0$
  - there exist units with no causal effect of treatment on the surrogate but experience causal effects of treatment on outcome: $S_i(1) = S_i(0)$ but $Y_i(1) \neq Y_i(0)$

Properties of a good surrogate

- Intuitively, a good surrogate should satisfy:
  - if the treatment does not affect surrogate, it does not affect outcome
  - if the treatment affects surrogate, it affects outcome

- Formally, a good surrogate should satisfy:
  - causal necessity (Frangakis and Rubin 2002)

$$S_i(1) = S_i(0) \Longrightarrow Y_i(1) = Y_i(0)$$

  - causal sufficiency (Gilbert and Hudgens 2008)

$$S_i(1) \neq S_i(0) \Longrightarrow Y_i(1) \neq Y_i(0)$$

Principal stratification and principal causal effects

- A simple example: treatment comparison for AIDS patients, where CD4 is the candidate surrogate
- Four principal strata defined by $\{S_i(1), S_i(0)\}$

$$U_i \equiv \{S_i(1), S_i(0)\} \in \{HH, HL, LH, LL\}$$

Principal stratification and principal causal effects

- A simple example: treatment comparison for AIDS patients, where CD4 is the candidate surrogate
- Four principal strata defined by $\{S_i(1), S_i(0)\}$

$$U_i \equiv \{S_i(1), S_i(0)\} \in \{HH, HL, LH, LL\}$$

- Principal average causal effect

$$\mathrm{ACE}_u = \mathbb{E}\{Y_i(1) - Y_i(0) \mid U_i = u\}$$

Principal stratification and principal causal effects

- A simple example: treatment comparison for AIDS patients, where CD4 is the candidate surrogate
- Four principal strata defined by $\{S_i(1), S_i(0)\}$

$$U_i \equiv \{S_i(1), S_i(0)\} \in \{HH, HL, LH, LL\}$$

- Principal average causal effect

$$\mathsf{ACE}_u = \mathbb{E}\{Y_i(1) - Y_i(0) \mid U_i = u\}$$

- Criteria for a principal surrogate
    - causal necessity: $\mathsf{ACE}_{HH} = \mathsf{ACE}_{LL} = 0$
    - causal sufficiency: $\mathsf{ACE}_{HL} \neq 0$ and $\mathsf{ACE}_{LH} \neq 0$

Principal stratification and principal causal effects

- A simple example: treatment comparison for AIDS patients, where CD4 is the candidate surrogate
- Four principal strata defined by $\{S_i(1), S_i(0)\}$

$$U_i \equiv \{S_i(1), S_i(0)\} \in \{HH, HL, LH, LL\}$$

- Principal average causal effect

$$\mathrm{ACE}_u = \mathbb{E}\{Y_i(1) - Y_i(0) \mid U_i = u\}$$

- Criteria for a principal surrogate
  - causal necessity: $\mathrm{ACE}_{HH} = \mathrm{ACE}_{LL} = 0$
  - causal sufficiency: $\mathrm{ACE}_{HL} \neq 0$ and $\mathrm{ACE}_{LH} \neq 0$
- Lauritzen (2004) proposes the criteria of a strong surrogate, which is stronger than the criteria of a principal surrogate

Statistical surrogate vs. principal surrogate

- Statistical surrogate: $Z \perp\!\!\!\perp Y \mid S$
  principal surrogate: conditions on ACE_u

Statistical surrogate vs. principal surrogate

- Statistical surrogate: $Z \perp\!\!\!\perp Y \mid S$
  principal surrogate: conditions on ACE_u
    - If the post-treatment variable S_i is a principal surrogate, then it is not generally a statistical surrogate
    - If the post-treatment variable S_i is a statistical surrogate, then it is not generally a principal surrogate

Statistical surrogate vs. principal surrogate

- Statistical surrogate: $Z \perp\!\!\!\perp Y \mid S$
  principal surrogate: conditions on ACE_u
    - If the post-treatment variable S_i is a principal surrogate, then it is not generally a statistical surrogate
    - If the post-treatment variable S_i is a statistical surrogate, then it is not generally a principal surrogate

- Evaluation
    - statistical surrogate: direct estimation from the observed data
    - principal surrogate: requires the identification and estimation of principal causal effects

Can we avoid the disaster using these criteria?

$$\begin{tikzcd} & & U \arrow[dl] \arrow[dr] & \\ Z \arrow[r, "+"] & S \arrow[rr, "+"] & & Y \end{tikzcd}$$

- Surrogate "paradox" (Chen, Geng and Jia 2007, Ju and Geng 2010)

- For a principal or statistical surrogate, it is still possible that
  - Positive causal effect of Z on S: $\text{ACE}(Z \to S) > 0$
  - Positive causal effect of S on Y: $\text{ACE}(S \to Y) > 0$ and $\text{Cor}(S, Y) > 0$
  - Negative causal effet of Z on Y: $\text{ACE}(Z \to Y) < 0$

Can we avoid the disaster using these criteria?

$$\begin{tikzcd}
 & & U \arrow[dl] \arrow[dr] & \\
Z \arrow[r, "+"] & S \arrow[rr, "+"] & & Y
\end{tikzcd}$$

- Surrogate "paradox" (Chen, Geng and Jia 2007, Ju and Geng 2010)

- For a principal or statistical surrogate, it is still possible that
  - Positive causal effect of Z on S: $\text{ACE}(Z \rightarrow S) > 0$
  - Positive causal effect of S on Y: $\text{ACE}(S \rightarrow Y) > 0$ and $\text{Cor}(S, Y) > 0$
  - Negative causal effet of Z on Y: $\text{ACE}(Z \rightarrow Y) < 0$

- It is possible that drugs are beneficial for VA but increase mortality

Principal surrogates may suffer from surrogate paradox

- Proportions of principal strata, $\pi_u = \mathbb{P}(U_i = u)$:
  $(\pi_{HH}, \pi_{HL}, \pi_{LH}, \pi_{LL}) = (0.2, 0.4, 0.2, 0.2)$

$$
\begin{aligned}
& \mathsf{ACE}(Z \to S) = \mathbb{P}\{S_i(1) = 1\} - \mathbb{P}\{S_i(0) = 1\} \\
= \quad & (\pi_{HL} + \pi_{HH}) - (\pi_{HH} + \pi_{LH}) = \pi_{HL} - \pi_{LH} = 0.2 > 0
\end{aligned}
$$

Principal surrogates may suffer from surrogate paradox

- Proportions of principal strata, $\pi_u = \mathbb{P}(U_i = u)$:
  $(\pi_{HH}, \pi_{HL}, \pi_{LH}, \pi_{LL}) = (0.2, 0.4, 0.2, 0.2)$

$$\begin{aligned} \mathrm{ACE}(Z \to S) &= \mathbb{P}\{S_i(1) = 1\} - \mathbb{P}\{S_i(0) = 1\} \\ &= (\pi_{HL} + \pi_{HH}) - (\pi_{HH} + \pi_{LH}) = \pi_{HL} - \pi_{LH} = 0.2 > 0 \end{aligned}$$

- Principal causal effects

$$\mathrm{ACE}_{HH} = \mathrm{ACE}_{LL} = 0, \quad \mathrm{ACE}_{HL} = 0.1 > 0, \quad \mathrm{ACE}_{LH} = -0.3 < 0$$

Principal surrogates may suffer from surrogate paradox

- Proportions of principal strata, $\pi_u = \mathbb{P}(U_i = u)$:
  $(\pi_{HH}, \pi_{HL}, \pi_{LH}, \pi_{LL}) = (0.2, 0.4, 0.2, 0.2)$

$$
\begin{aligned}
& \mathsf{ACE}(Z \to S) = \mathbb{P}\{S_i(1) = 1\} - \mathbb{P}\{S_i(0) = 1\} \\
= \quad & (\pi_{HL} + \pi_{HH}) - (\pi_{HH} + \pi_{LH}) = \pi_{HL} - \pi_{LH} = 0.2 > 0
\end{aligned}
$$

- Principal causal effects

$$
\mathsf{ACE}_{HH} = \mathsf{ACE}_{LL} = 0, \quad \mathsf{ACE}_{HL} = 0.1 > 0, \quad \mathsf{ACE}_{LH} = -0.3 < 0
$$

- However, the ACE on the outcome is negative

$$
\mathsf{ACE}(Z \to Y) = \sum_u \pi_u \mathsf{ACE}_u = 0.4 \times 0.1 + 0.2 \times (-0.3) = -0.02 < 0
$$

New criteria: when a surrogate satisfies causal necessity

- $\mathsf{ACE}(Z \to S) > 0$ should imply $\mathsf{ACE}(Z \to Y) > 0$

$$\mathsf{ACE}(Z \to Y) = \sum_u \mathsf{ACE}_u \cdot \mathbb{P}(U_i = u)$$

New criteria: when a surrogate satisfies causal necessity

- $\text{ACE}(Z \to S) > 0$ should imply $\text{ACE}(Z \to Y) > 0$

$$\text{ACE}(Z \to Y) = \sum_u \text{ACE}_u \cdot \mathbb{P}(U_i = u)$$

- If $\text{ACE}_{HL} + \text{ACE}_{LH} \geq 0$, then

$$\text{ACE}(Z \to Y) \geq \text{ACE}(Z \to S) \times \text{ACE}_{HL}$$

New criteria: when a surrogate satisfies causal necessity

- $\mathsf{ACE}(Z \rightarrow S) > 0$ should imply $\mathsf{ACE}(Z \rightarrow Y) > 0$

$$\mathsf{ACE}(Z \rightarrow Y) = \sum_{u} \mathsf{ACE}_u \cdot \mathbb{P}(U_i = u)$$

- If $\mathsf{ACE}_{HL} + \mathsf{ACE}_{LH} \geq 0$, then

$$\mathsf{ACE}(Z \rightarrow Y) \geq \mathsf{ACE}(Z \rightarrow S) \times \mathsf{ACE}_{HL}$$

  - $\mathsf{ACE}_{HL} > 0 \Longrightarrow$ lower bound of $\mathsf{ACE}(Z \rightarrow Y)$ and $\mathsf{ACE}(Z \rightarrow S)$ have the same sign

New criteria: when a surrogate satisfies causal necessity

- $\text{ACE}(Z \to S) > 0$ should imply $\text{ACE}(Z \to Y) > 0$

$$\text{ACE}(Z \to Y) = \sum_u \text{ACE}_u \cdot \mathbb{P}(U_i = u)$$

- If $\text{ACE}_{HL} + \text{ACE}_{LH} \geq 0$, then

$$\text{ACE}(Z \to Y) \geq \text{ACE}(Z \to S) \times \text{ACE}_{HL}$$

  - $\text{ACE}_{HL} > 0 \implies$ lower bound of $\text{ACE}(Z \to Y)$ and $\text{ACE}(Z \to S)$ have the same sign

- New criteria for avoiding surrogate paradox
  - causal necessity: $\text{ACE}_{HH} = \text{ACE}_{LL} = 0$
  - $\text{ACE}_{HL} > 0$ and $\text{ACE}_{HL} + \text{ACE}_{LH} \geq 0$

Surrogate evaluation in colon clinical trials (Jiang et al., 2016)

- Treatment Z; outcome Y: 5-year survival status
- Target: evaluate whether disease free survival with 3-year follow-up is a good surrogate (1 for not)

Surrogate evaluation in colon clinical trials (Jiang et al., 2016)

- Treatment Z; outcome Y: 5-year survival status
- Target: evaluate whether disease free survival with 3-year follow-up is a good surrogate (1 for not)

Surrogate evaluation in colon clinical trials (Jiang et al., 2016)

- Treatment Z; outcome Y: 5-year survival status
- Target: evaluate whether disease free survival with 3-year follow-up is a good surrogate (1 for not)

- Estimated principal causal effects
  - $\text{ACE}_{00} = -0.001$ (s.e.$= 0.042$) and $\text{ACE}_{11} = 0.015$ (s.e.$= 0.012$)
  - $\text{ACE}_{10} = 0.774$ (s.e.$= 0.037$) and $\text{ACE}_{01} = -0.750$ (s.e.$= 0.054$)

Surrogate evaluation in colon clinical trials (Jiang et al., 2016)

- Treatment Z; outcome Y: 5-year survival status
- Target: evaluate whether disease free survival with 3-year follow-up is a good surrogate (1 for not)

- Estimated principal causal effects
  - $\text{ACE}_{00} = -0.001$ (s.e.$= 0.042$) and $\text{ACE}_{11} = 0.015$ (s.e.$= 0.012$)
  - $\text{ACE}_{10} = 0.774$ (s.e.$= 0.037$) and $\text{ACE}_{01} = -0.750$ (s.e.$= 0.054$)

- DFS is a good surrogate

Mediation analysis

- Mediator S on the causal pathway from Z to Y

Mediation analysis

- Mediator S on the causal pathway from Z to Y
  - principal stratification does not require S on the causal pathway

Mediation analysis

- Mediator S on the causal pathway from Z to Y
  - principal stratification does not require S on the causal pathway
- Causal mechanism
  - Direct effect: Z affects Y not through S
  - indirect effect: Z affects Y through S

Mediation analysis

- Mediator S on the causal pathway from Z to Y
  - principal stratification does not require S on the causal pathway

- Causal mechanism
  - Direct effect: Z affects Y not through S
  - indirect effect: Z affects Y through S

- Potential values: $S(z)$, $Y(z,s)$, $Y(z,S(z'))$
  - $Y(z) = Y(z, S(z))$

Mediation analysis

- Mediator S on the causal pathway from Z to Y
  - principal stratification does not require S on the causal pathway

- Causal mechanism
  - Direct effect: Z affects Y not through S
  - indirect effect: Z affects Y through S

- Potential values: $S(z)$, $Y(z, s)$, $Y(z, S(z'))$
  - $Y(z) = Y(z, S(z))$

- Mediation analysis
  - Natural direct effect: $\mathbb{E}\{Y(1, S(1)) - Y(0, S(1))\}$
  - Natural indirect effect: $\mathbb{E}\{Y(0, S(1)) - Y(0, S(0))\}$

Mediation analysis and principal stratification

- Principal strata direct effect: $\mathbb{E}\{Y(1) - Y(0) \mid S(1) = S(0) = s\}$

Mediation analysis and principal stratification

- Principal strata direct effect: $\mathbb{E}\{Y(1) - Y(0) \mid S(1) = S(0) = s\}$
  - for units with $S(1) = S(0) = s$,
    $Y(1) - Y(0) = Y(1, S(1)) - Y(0, S(0)) = Y(1, S(1)) - Y(0, S(1))$

Mediation analysis and principal stratification

- Principal strata direct effect: $\mathbb{E}\{Y(1) - Y(0) \mid S(1) = S(0) = s\}$
  - for units with $S(1) = S(0) = s$,
    $Y(1) - Y(0) = Y(1, S(1)) - Y(0, S(0)) = Y(1, S(1)) - Y(0, S(1))$
  - Principal strata direct effect is the natural direct effect within principal strata

Mediation analysis and principal stratification

- Principal strata direct effect: $\mathbb{E}\{Y(1) - Y(0) \mid S(1) = S(0) = s\}$
  - for units with $S(1) = S(0) = s$,
    $Y(1) - Y(0) = Y(1, S(1)) - Y(0, S(0)) = Y(1, S(1)) - Y(0, S(1))$
  - Principal strata direct effect is the natural direct effect within principal strata

- Principal strata indirect effect: $\mathbb{E}\{Y(1) - Y(0) \mid S(1) = 1, S(0) = 0\}$?

$$
\begin{aligned}
& \mathbb{E}\{Y(1) - Y(0) \mid S(1) = 1, S(0) = 0\} \\
= \; & \mathbb{E}\{Y(1, S(1)) - Y(0, S(0)) \mid S(1) = 1, S(0) = 0\} \\
= \; & \mathbb{E}\{Y(1, 1) - Y(0, 0) \mid S(1) = 1, S(0) = 0\}
\end{aligned}
$$

Mediation analysis and principal stratification

- Principal strata direct effect: $\mathbb{E}\{Y(1) - Y(0) \mid S(1) = S(0) = s\}$
  - for units with $S(1) = S(0) = s$,
    $Y(1) - Y(0) = Y(1, S(1)) - Y(0, S(0)) = Y(1, S(1)) - Y(0, S(1))$
  - Principal strata direct effect is the natural direct effect within principal strata

- Principal strata indirect effect: $\mathbb{E}\{Y(1) - Y(0) \mid S(1) = 1, S(0) = 0\}$?

$$
\begin{aligned}
& \mathbb{E}\{Y(1) - Y(0) \mid S(1) = 1, S(0) = 0\} \\
= \; & \mathbb{E}\{Y(1, S(1)) - Y(0, S(0)) \mid S(1) = 1, S(0) = 0\} \\
= \; & \mathbb{E}\{Y(1, 1) - Y(0, 0) \mid S(1) = 1, S(0) = 0\}
\end{aligned}
$$

  - $\mathbb{E}\{Y(1) - Y(0) \mid S(1) = 1, S(0) = 0\}$ consists of both direct and indirect effects

Outline

- Potential outcomes and ACE
- Post-treatment variable
- Principal stratification
  - noncompliance
  - truncation by death
  - surrogate evaluation
- Identification and statistical inference
  - binary instrumental variable model
  - partial identification
  - principal ignorability
  - auxiliary independence
- Recent applications of principal stratification

Statistical inference and its difficulty

- For problems of non-compliance, truncation by death, and surrogate evaluation, we need to identify and estimate principal causal effects

Statistical inference and its difficulty

- For problems of non-compliance, truncation by death, and surrogate evaluation, we need to identify and estimate principal causal effects
  - we focus on the randomized experiment setting: $Z_i \perp\!\!\!\perp (Y_i(z), S_i(1), S_i(0))$

Statistical inference and its difficulty

- For problems of non-compliance, truncation by death, and surrogate evaluation, we need to identify and estimate principal causal effects
  - we focus on the randomized experiment setting: $Z_i \perp\!\!\!\perp (Y_i(z), S_i(1), S_i(0))$
  - key to identification: express ACE_u in terms of the observed distribution (Z, S, Y, X)

Statistical inference and its difficulty

- For problems of non-compliance, truncation by death, and surrogate evaluation, we need to identify and estimate principal causal effects
  - we focus on the randomized experiment setting: $Z_i \perp\!\!\!\perp (Y_i(z), S_i(1), S_i(0))$
  - key to identification: express ACE_u in terms of the observed distribution (Z, S, Y, X)

- Under randomization of Z

$$\text{ACE}_u = \mathbb{E}(Y_i = 1 \mid Z_i = 1, U_i) - \mathbb{E}(Y_i = 1 \mid Z_i = 0, U_i)$$

Statistical inference and its difficulty

- For problems of non-compliance, truncation by death, and surrogate evaluation, we need to identify and estimate principal causal effects
  - we focus on the randomized experiment setting: $Z_i \perp\!\!\!\perp (Y_i(z), S_i(1), S_i(0))$
  - key to identification: express ACE_u in terms of the observed distribution (Z, S, Y, X)

- Under randomization of Z

$$\text{ACE}_u = \mathbb{E}(Y_i = 1 \mid Z_i = 1, U_i) - \mathbb{E}(Y_i = 1 \mid Z_i = 0, U_i)$$

  - difficulty: $S_i(1)$ and $S_i(0)$ are not simultaneous observed $\rightsquigarrow$ U_i is latent

Outline

- Recap: potential outcomes and ACE
- Post-treatment variable
- Principal stratification
  - noncompliance
  - truncation by death
  - surrogate evaluation
- Identification and statistical inference
  - binary instrumental variable model
  - partial identification
  - principal ignorability
- Recent applications of principal stratification

Noncompliance

- Four principal strata (latent types):
  - compliers $(S_i(1), S_i(0)) = (1, 0)$,
  - non-compliers $\begin{cases} \textit{always} - \textit{takers} & (S_i(1), S_i(0)) = (1, 1), \\ \textit{never} - \textit{takers} & (S_i(1), S_i(0)) = (0, 0), \\ \textit{defiers} & (S_i(1), S_i(0)) = (0, 1) \end{cases}$
  - denote the compliance behavior (a, n, c, d) by $U_i \rightsquigarrow S_i$ is a function of Z_i and U_i

Noncompliance

- Four principal strata (latent types):
  - compliers $(S_i(1), S_i(0)) = (1, 0)$,
  - non-compliers $\begin{cases} \textit{always} - \textit{takers} & (S_i(1), S_i(0)) = (1, 1), \\ \textit{never} - \textit{takers} & (S_i(1), S_i(0)) = (0, 0), \\ \textit{defiers} & (S_i(1), S_i(0)) = (0, 1) \end{cases}$
  - denote the compliance behavior (a, n, c, d) by $U_i \rightsquigarrow S_i$ is a function of Z_i and U_i
- Observed strata and compliance behavior:

	$Z_i = 1$	$Z_i = 0$
$S_i = 1$	Complier/Always-taker	Defier/Always-taker
$S_i = 0$	Defier/Never-taker	Complier/Never-taker

Noncompliance

- Four principal strata (latent types):
  - compliers $(S_i(1), S_i(0)) = (1, 0)$,
  - non-compliers $\begin{cases} \textit{always} - \textit{takers} & (S_i(1), S_i(0)) = (1, 1), \\ \textit{never} - \textit{takers} & (S_i(1), S_i(0)) = (0, 0), \\ \textit{defiers} & (S_i(1), S_i(0)) = (0, 1) \end{cases}$
  - denote the compliance behavior (a, n, c, d) by $U_i \rightsquigarrow S_i$ is a function of Z_i and U_i
- Observed strata and compliance behavior:

	$Z_i = 1$	$Z_i = 0$
$S_i = 1$	Complier/Always-taker	Defier/Always-taker
$S_i = 0$	Defier/Never-taker	Complier/Never-taker

- $\text{CACE} = \text{ACE}_c = \mathbb{E}\{Y_i(1) - Y_i(0) \mid U_i = c\}$

Assumptions

- Randomization: $\{Y_i(z), S_i(1), S_i(0))\} \perp\!\!\!\perp Z_i \rightsquigarrow$ Identification of ACE$(Z \to S)$ and ACE$(Z \to Y)$

Assumptions

- Randomization: $\{Y_i(z), S_i(1), S_i(0))\} \perp\!\!\!\perp Z_i \rightsquigarrow$ Identification of ACE$(Z \rightarrow S)$ and ACE$(Z \rightarrow Y)$

- Monotonicity: $S_i(1) \geq S_i(0) \rightsquigarrow$ no defiers

Assumptions

- Randomization: $\{Y_i(z), S_i(1), S_i(0))\} \perp\!\!\!\perp Z_i \rightsquigarrow$ Identification of $\text{ACE}(Z \to S)$ and $\text{ACE}(Z \to Y)$
- Monotonicity: $S_i(1) \geq S_i(0) \rightsquigarrow$ no defiers
- Exclusion restriction: $\mathbb{E}\{Y_i(1) - Y_i(0) \mid U_i = u\} = 0$ for $u = a, n$

Assumptions

- Randomization: $\{Y_i(z), S_i(1), S_i(0))\} \perp\!\!\!\perp Z_i \rightsquigarrow$ Identification of $\text{ACE}(Z \to S)$ and $\text{ACE}(Z \to Y)$
- Monotonicity: $S_i(1) \geq S_i(0) \rightsquigarrow$ no defiers
- Exclusion restriction: $\mathbb{E}\{Y_i(1) - Y_i(0) \mid U_i = u\} = 0$ for $u = a, n$
- $\text{ACE}(Z \to S) = \mathbb{E}\{S_i(1) - S_i(0)\} > 0 \rightsquigarrow$ there exists compliers
    - $\mathbb{E}\{S_i(1) - S_i(0)\} = \mathbb{P}\{S_i(1) = 1\} - \mathbb{P}\{S_i(0) = 1\} = \mathbb{P}\{S_i(1) = 1, S_i(0) = 0\}$

Monotonicity: implications

- Observed strata and compliance behavior under monotonicity

	$Z_i = 1$	$Z_i = 0$
$S_i = 1$	Complier/Always-taker	Always-taker
$S_i = 0$	Never-taker	Complier/Never-taker

- $\mathbb{P}(\text{Never-taker}) = \mathbb{P}\{S_i(1) = 0\} = \mathbb{P}(S_i = 0 \mid Z_i = 1)$
  $\mathbb{P}(\text{always-taker}) = \mathbb{P}\{S_i(0) = 1\} = \mathbb{P}(S_i = 1 \mid Z_i = 0)$
  $\mathbb{P}(\text{complier}) = \mathbb{P}(S_i = 1 \mid Z_i = 1) - \mathbb{P}(S_i = 1 \mid Z_i = 0)$

Exclusion restriction: implications

- Exclusion restriction means outcome depends on the treatment assignment only through treatment receipt $\rightsquigarrow$ $\mathsf{ACE}_a = \mathsf{ACE}_n = 0$

Exclusion restriction: implications

- Exclusion restriction means outcome depends on the treatment assignment only through treatment receipt $\rightsquigarrow$ $\text{ACE}_a = \text{ACE}_n = 0$

- An alternative definition of exclusion restriction: $Y_i(z, s) = Y_i(s)$

Exclusion restriction: implications

- Exclusion restriction means outcome depends on the treatment assignment only through treatment receipt $\rightsquigarrow$ $\text{ACE}_a = \text{ACE}_n = 0$

- An alternative definition of exclusion restriction: $Y_i(z, s) = Y_i(s)$
  - always-taker and never-taker:
    $Y_i(1) = Y_i(1, S_i(1)) = Y_i(1, S_i(0)) = Y_i(0, S_i(0)) = Y_i(0)$
    $\text{ACE}_a = \text{ACE}_n = 0$
  - compliers: $\text{ACE}_c = \mathbb{E}\{Y_i(1) - Y_i(0) \mid U_i = c\}$, $Z_i = S_i$

Identification

- ITT effect decomposition:

$$\begin{aligned} \text{ACE} &= \text{ACE}_c \times \Pr(\text{compliers}) + \text{ACE}_a \times \Pr(\text{always} - \text{takers}) \\ &\quad + \text{ACE}_n \times \Pr(\text{never} - \text{takers}) + \text{ACE}_d \times \Pr(\text{defiers}) \\ &= \text{ACE}_c \times \Pr(\text{compliers}) \end{aligned}$$

Identification

- ITT effect decomposition:

$$
\begin{aligned}
\text{ACE} &= \text{ACE}_c \times \Pr(\text{compliers}) + \text{ACE}_a \times \Pr(\text{always}-\text{takers}) \\
&\quad + \text{ACE}_n \times \Pr(\text{never}-\text{takers}) + \text{ACE}_d \times \Pr(\text{defiers}) \\
&= \text{ACE}_c \times \Pr(\text{compliers})
\end{aligned}
$$

- Identification:

$$
\begin{aligned}
\text{ACE}_c &= \frac{\text{ACE}(Z \to Y)}{\mathbb{P}(\text{compliers})} \\
&= \frac{\text{ACE}(Z \to Y)}{\text{ACE}(Z \to S)} \\
&= \frac{\mathbb{E}(Y_i \mid Z_i = 1) - \mathbb{E}(Y_i \mid Z_i = 0)}{\mathbb{E}(S_i \mid Z_i = 1) - \mathbb{E}(S_i \mid Z_i = 0)}
\end{aligned}
$$

Complier average causal effect

- CACE is identified by the ratio of two ITT effects
  - average treatment effect for compliers ($Z_i = S_i$)
  - always have the same sign as ACE($Z \to Y$)
  - more transferrable than ACE($Z \to Y$)

Complier average causal effect

- CACE is identified by the ratio of two ITT effects
  - average treatment effect for compliers ($Z_i = S_i$)
  - always have the same sign as ACE($Z \to Y$)
  - more transferrable than ACE($Z \to Y$)

- Encouragement design: randomly encourage people to take the treatment
  - CACE is the effect of people who would take the treatment only if encouraged

Complier average causal effect

- CACE is identified by the ratio of two ITT effects
  - average treatment effect for compliers ($Z_i = S_i$)
  - always have the same sign as ACE($Z \rightarrow Y$)
  - more transferrable than ACE($Z \rightarrow Y$)

- Encouragement design: randomly encourage people to take the treatment
  - CACE is the effect of people who would take the treatment only if encouraged
  - different encouragement yields different compliers

Estimation and inference

- Wald estimator: $\widehat{\text{CACE}} = \frac{\widehat{\mathbb{E}}(Y_i|Z_i=1)-\widehat{\mathbb{E}}(Y_i|Z_i=0)}{\widehat{\mathbb{E}}(S_i|Z_i=1)-\widehat{\mathbb{E}}(S_i|Z_i=0)}$

Estimation and inference

- Wald estimator: $\widehat{\text{CACE}} = \frac{\widehat{\mathbb{E}}(Y_i|Z_i=1)-\widehat{\mathbb{E}}(Y_i|Z_i=0)}{\widehat{\mathbb{E}}(S_i|Z_i=1)-\widehat{\mathbb{E}}(S_i|Z_i=0)}$

- Variance estimation and confidence interval: delta method

Estimation and inference

- Wald estimator: $\widehat{\text{CACE}} = \frac{\widehat{\mathbb{E}}(Y_i|Z_i=1)-\widehat{\mathbb{E}}(Y_i|Z_i=0)}{\widehat{\mathbb{E}}(S_i|Z_i=1)-\widehat{\mathbb{E}}(S_i|Z_i=0)}$

- Variance estimation and confidence interval: delta method

- Weak instrument: $\widehat{\text{CACE}}$ has poor properties when $\text{ACE}(Z \to S)$ is close to 0

Estimation and inference

- Wald estimator: $\widehat{\text{CACE}} = \frac{\widehat{\mathbb{E}}(Y_i|Z_i=1)-\widehat{\mathbb{E}}(Y_i|Z_i=0)}{\widehat{\mathbb{E}}(S_i|Z_i=1)-\widehat{\mathbb{E}}(S_i|Z_i=0)}$

- Variance estimation and confidence interval: delta method

- Weak instrument: $\widehat{\text{CACE}}$ has poor properties when $\text{ACE}(Z \to S)$ is close to 0
  - $\widehat{\text{CACE}}$ has finite sample bias and non-normal asymptotic distribution
  - confidence interval has poor coverage rate
  - when $\text{ACE}_D = 0$, $\widehat{\text{CACE}}$ has no first moment

Estimation and inference

- Wald estimator: $\widehat{\text{CACE}} = \frac{\widehat{\mathbb{E}}(Y_i|Z_i=1)-\widehat{\mathbb{E}}(Y_i|Z_i=0)}{\widehat{\mathbb{E}}(S_i|Z_i=1)-\widehat{\mathbb{E}}(S_i|Z_i=0)}$

- Variance estimation and confidence interval: delta method

- Weak instrument: $\widehat{\text{CACE}}$ has poor properties when $\text{ACE}(Z \to S)$ is close to 0
  - $\widehat{\text{CACE}}$ has finite sample bias and non-normal asymptotic distribution
  - confidence interval has poor coverage rate
  - when $\text{ACE}_D = 0$, $\widehat{\text{CACE}}$ has no first moment

- Observational studies: Abadies' Kappa

Evaluation of job training program

- A randomized field experiment investigating the efficacy of a job training intervention on unemployed workers
- Encouragement: Z_i; participation: S_i; job-search self-efficacy: Y_i

Evaluation of job training program

- A randomized field experiment investigating the efficacy of a job training intervention on unemployed workers
- Encouragement: Z_i; participation: S_i; job-search self-efficacy: Y_i
- Assumptions:
  - monotonicity: being encouraged would never discourage anyone from participating
  - exclusion restriction: being encouraged has no effect on other than through participation in the program

Evaluation of job training program

- A randomized field experiment investigating the efficacy of a job training intervention on unemployed workers
- Encouragement: Z_i; participation: S_i; job-search self-efficacy: Y_i
- Assumptions:
  - monotonicity: being encouraged would never discourage anyone from participating
  - exclusion restriction: being encouraged has no effect on other than through participation in the program
- CACE: effect of participation on job-search self-efficacy for people who would participate if and only if they are encouraged

Evaluation of job training program

- A randomized field experiment investigating the efficacy of a job training intervention on unemployed workers
- Encouragement: Z_i; participation: S_i; job-search self-efficacy: Y_i
- Assumptions:
  - monotonicity: being encouraged would never discourage anyone from participating
  - exclusion restriction: being encouraged has no effect on other than through participation in the program
- CACE: effect of participation on job-search self-efficacy for people who would participate if and only if they are encouraged
- ACE($Z \rightarrow Y$): est. $= 0.067$, s.e. $= 0.050$, 95% CI $= [-0.031,\ 0.166]$
- CACE: est. $= 0.109$, s.e. $= 0.081$, 95% CI $= [-0.050,\ 0.268]$

Effect of veteran status on earnings (Angrist, 1990, *AER*)

- There were five draft lotteries during the Vietnam War period. In each lottery, priority for induction was determined by a Random Sequence Number (RSN) from 1-365 that was assigned to birthdates in the cohort being drafted
- Men were called for induction by RSN up to a ceiling determined by the Defense Department, and only men with lottery numbers below the ceiling could have been drafted

Effect of veteran status on earnings (Angrist, 1990, *AER*)

- There were five draft lotteries during the Vietnam War period. In each lottery, priority for induction was determined by a Random Sequence Number (RSN) from 1-365 that was assigned to birthdates in the cohort being drafted
- Men were called for induction by RSN up to a ceiling determined by the Defense Department, and only men with lottery numbers below the ceiling could have been drafted
- Draft lottery RSNs were randomly assigned in a televised drawing held a few months before men reaching draft age were to be called
- Draft-eligibility ceilings were announced later in the year, once Defense Department manpower needs were known

Effect of veteran status on earnings (Angrist, 1990, *AER*)

- There were five draft lotteries during the Vietnam War period. In each lottery, priority for induction was determined by a Random Sequence Number (RSN) from 1-365 that was assigned to birthdates in the cohort being drafted
- Men were called for induction by RSN up to a ceiling determined by the Defense Department, and only men with lottery numbers below the ceiling could have been drafted
- Draft lottery RSNs were randomly assigned in a televised drawing held a few months before men reaching draft age were to be called
- Draft-eligibility ceilings were announced later in the year, once Defense Department manpower needs were known
- Subsequent selection from the draft-eligible pool was based on a number of criteria: physical examination and a mental aptitude test

Setup

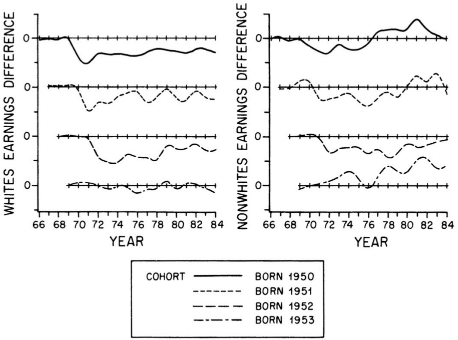
- Z_i: draft-eligibility; S_i: veteran status; outcome Y_i earnings in 1981-1984

Setup

- Z_i: draft-eligibility; S_i: veteran status; outcome Y_i earnings in 1981-1984
- Assumptions hold?

Setup

- Z_i: draft-eligibility; S_i: veteran status; outcome Y_i earnings in 1981-1984
- Assumptions hold?



Notes: The figure plots the difference in FICA taxable earnings by draft-eligibility status for the four cohorts born 1950–53. Each tick on the vertical axis represents $500 real (1978) dollars.

FIGURE 2. THE DIFFERENCE IN EARNINGS BY DRAFT-ELIGIBILITY STATUS

Wald estimates

Table 3—Wald Estimates

Cohort	Year	Draft-Eligibility Effects in Current $: FICA Earnings (1)	Adjusted FICA Earnings (2)	Total W-2 Earnings (3)	$\hat{p}^e - \hat{p}^n$ (4)	Service Effect in 1978 $ (5)
1950	1981	−435.8	−487.8	−589.6	0.159	−2,195.8
		(210.5)	(237.6)	(299.4)	(0.040)	(1,069.5)
	1982	−320.2	−396.1	−305.5		−1,678.3
		(235.8)	(281.7)	(345.4)		(1,193.6)
	1983	−349.5	−450.1	−512.9		−1,795.6
		(261.6)	(302.0)	(441.2)		(1,204.8)
	1984	−484.3	−638.7	−1,143.3		−2,517.7
		(286.8)	(336.5)	(492.2)		(1,326.5)
1951	1981	−358.3	−428.7	−71.6	0.136	−2,261.3
		(203.6)	(224.5)	(423.4)	(0.043)	(1,184.2)
	1982	−117.3	−278.5	−72.7		−1,386.6
		(229.1)	(264.1)	(372.1)		(1,312.1)
	1983	−314.0	−452.2	−896.5		-2,181.8
		(253.2)	(289.2)	(426.3)		(1,395.3)
	1984	−398.4	−573.3	−809.1		−2,647.9
		(279.2)	(331.1)	(380.9)		(1,529.2)
1952	1981	−342.8	−392.6	−440.5	0.105	−2,502.3
		(206.8)	(228.6)	(265.0)	(0.050)	(1,556.7)
	1982	−235.1	−255.2	−514.7		−1,626.5
		(232.3)	(264.5)	(296.5)		(1,685.8)
	1983	−437.7	−500.0	−915.7		−3,103.5
		(257.5)	(294.7)	(395.2)		(1,829.2)
	1984	−436.0	−560.0	−767.2		−3,323.8
		(281.9)	(330.1)	(376.0)		(1,959.3)

Physical activity and weight following car ownership in Beijing, China: quasi-experimental cross sectional study

Michael L Anderson,[1] Fangwen Lu,[2] Jun Yang[3]

- In January 2011, to deal with the problem of congestion, Beijing capped the number of new vehicles allowed at 240000 each year and introduced a vehicle permit (license plate) lottery
- After that date, only residents who entered and won the lottery were entitled to a license plate.
- The lottery was drawn monthly, and winners had to purchase a car within six months of winning. By mid-2012 the probability of winning fell below 2% a month

Effect of winning the lottery

- Z_i: winning the lottery; S_i: buying a car; Y_i: weekly transit rides, minute daily walking/bicycling, weight

Effect of winning the lottery

- Z_i: winning the lottery; S_i: buying a car; Y_i: weekly transit rides, minute daily walking/bicycling, weight
- Assumptions satisfied?

Effect of winning the lottery

- Z_i: winning the lottery; S_i: buying a car; Y_i: weekly transit rides, minute daily walking/bicycling, weight
- Assumptions satisfied?

Table 3 | Age stratified associations between winning the lottery and transit use, activity, and weight

	Time since winning (95% CI)		
Dependent variables	**0.1 years (minimum)**	**2.6 years (average)**	**5.1 years (maximum)**
Individuals aged ≥40			
Weekly transit rides	−2.18 (−4.13 to −0.24)	−2.1 (−3.35 to −0.85)	−2.02 (−5.16 to 1.12)
Minutes daily walking/bicycling	12.1 (−4.66 to 28.86)	−2.59 (−12.12 to 6.94)	−17.29 (−36.52 to 1.95)
Weight (kg)	1.29 (−5.07 to 7.65)	3.24 (−0.31 to 6.8)	5.2 (−2.59 to 12.99)
Individuals aged ≥50			
Weekly transit rides	−2.88 (−5.57 to −0.19)	−1.9 (−3.61 to −0.18)	−0.91 (−5.45 to 3.63)
Minutes daily walking/bicycling	27.4 (−0.28 to 55.08)	−1.19 (−13.76 to 11.38)	−29.78 (−54.08 to −5.49)
Weight (kg)	−1 (−8.4 to 6.4)	4.67 (0.04 to 9.31)	10.34 (0.49 to 20.19)

Effect of car ownership

- Ratio estimation for the effect of buying a car
- High compliance rate: 0.91 $(0.89, 0.94)$ $\rightsquigarrow$ effect of winning the lottery $\approx$ effect of buying a car

Effect of car ownership

- Ratio estimation for the effect of buying a car
- High compliance rate: 0.91 $(0.89, 0.94)$ $\rightsquigarrow$ effect of winning the lottery $\approx$ effect of buying a car

- Limitation
  - weight is self-reported

Effect of car ownership

- Ratio estimation for the effect of buying a car
- High compliance rate: 0.91 $(0.89, 0.94)$ $\rightsquigarrow$ effect of winning the lottery $\approx$ effect of buying a car

- Limitation
  - weight is self-reported
  - target population: people who want to and are able to buy a car

Outline

- Recap: potential outcomes and ACE
- Post-treatment variable
- Principal stratification
  - noncompliance
  - truncation by death
  - surrogate evaluation
- Identification and statistical inference
  - binary instrumental variable model
  - partial identification
  - principal ignorability
  - auxiliary independence
- Recent applications of principal stratification

Partial identification for truncation-by-death problem

- Treatment Z; survival status S_i; outcome Y_i

Partial identification for truncation-by-death problem

- Treatment Z; survival status S_i; outcome Y_i
- SACE: $\text{ACE}_{11} = \mathbb{E}\{Y_i(1) - Y_i(0) \mid S_i(1) = 1, S_i(0) = 1\}$

Partial identification for truncation-by-death problem

- Treatment Z; survival status S_i; outcome Y_i
- SACE: $\text{ACE}_{11} = \mathbb{E}\{Y_i(1) - Y_i(0) \mid S_i(1) = 1, S_i(0) = 1\}$
- Exclusion restriction implies SACE $= 0 \rightsquigarrow$ cannot be invoked

Partial identification for truncation-by-death problem

- Treatment Z; survival status S_i; outcome Y_i
- SACE: $\text{ACE}_{11} = \mathbb{E}\{Y_i(1) - Y_i(0) \mid S_i(1) = 1, S_i(0) = 1\}$
- Exclusion restriction implies $\text{SACE} = 0 \rightsquigarrow$ cannot be invoked
  - without exclusion restriction, SACE is not identifiable

Partial identification for truncation-by-death problem

- Treatment Z; survival status S_i; outcome Y_i
- SACE: $\mathrm{ACE}_{11} = \mathbb{E}\{Y_i(1) - Y_i(0) \mid S_i(1) = 1, S_i(0) = 1\}$
- Exclusion restriction implies SACE $= 0 \rightsquigarrow$ cannot be invoked
  - without exclusion restriction, SACE is not identifiable
  - alternative identification assumptions?

Partial identification for truncation-by-death problem

- Treatment Z; survival status S_i; outcome Y_i

- SACE: $\text{ACE}_{11} = \mathbb{E}\{Y_i(1) - Y_i(0) \mid S_i(1) = 1, S_i(0) = 1\}$

- Exclusion restriction implies SACE $= 0 \rightsquigarrow$ cannot be invoked
  - without exclusion restriction, SACE is not identifiable
  - alternative identification assumptions?

- Partial identification $\rightsquigarrow$ bounds on SACE

Partial identification for truncation-by-death problem

- Treatment Z; survival status S_i; outcome Y_i
- SACE: $\text{ACE}_{11} = \mathbb{E}\{Y_i(1) - Y_i(0) \mid S_i(1) = 1, S_i(0) = 1\}$
- Exclusion restriction implies $\text{SACE} = 0 \rightsquigarrow$ cannot be invoked
  - without exclusion restriction, SACE is not identifiable
  - alternative identification assumptions?
- Partial identification $\rightsquigarrow$ bounds on SACE
  - find all the possible values of SACE that are compatible with the observed data

Bounds for survivor average causal effect

- Randomization and monotonicity hold
- Proportions of principal strata:
  $\pi_{00} = \mathbb{P}(S_i = 0 \mid Z_i = 1)$, $\pi_{11} = \mathbb{P}(S_i = 1 \mid Z_i = 0)$, and
  $\pi_{10} = 1 - \pi_{00} - \pi_{11} = \mathbb{P}(S_i = 1 \mid Z_i = 1) - \mathbb{P}(S_i = 1 \mid Z_i = 0)$

Bounds for survivor average causal effect

- Randomization and monotonicity hold
- Proportions of principal strata:
  $\pi_{00} = \mathbb{P}(S_i = 0 \mid Z_i = 1)$, $\pi_{11} = \mathbb{P}(S_i = 1 \mid Z_i = 0)$, and
  $\pi_{10} = 1 - \pi_{00} - \pi_{11} = \mathbb{P}(S_i = 1 \mid Z_i = 1) - \mathbb{P}(S_i = 1 \mid Z_i = 0)$

- $S_i(1) = 1, S_i(0) = 1 \longleftrightarrow S_i(0) = 1$
  $\mathbb{E}\{Y_i(0) \mid S_i(1) = 1, S_i(0) = 1\} = \mathbb{E}\{Y_i(0) \mid S_i(0) = 1\} = \mathbb{E}(Y_i \mid Z_i = 0, S_i = 1)$

Bounds for survivor average causal effect

- Randomization and monotonicity hold
- Proportions of principal strata:
  $\pi_{00} = \mathbb{P}(S_i = 0 \mid Z_i = 1)$, $\pi_{11} = \mathbb{P}(S_i = 1 \mid Z_i = 0)$, and
  $\pi_{10} = 1 - \pi_{00} - \pi_{11} = \mathbb{P}(S_i = 1 \mid Z_i = 1) - \mathbb{P}(S_i = 1 \mid Z_i = 0)$

- $S_i(1) = 1, S_i(0) = 1 \longleftrightarrow S_i(0) = 1$
  $\mathbb{E}\{Y_i(0) \mid S_i(1) = 1, S_i(0) = 1\} = \mathbb{E}\{Y_i(0) \mid S_i(0) = 1\} = \mathbb{E}(Y_i \mid Z_i = 0, S_i = 1)$

- Key: bounds on $\mathbb{E}\{Y_i(1) \mid S_i(1) = 1, S_i(0) = 1\}$

Bounds on survivor average causal effect

- The observed stratum $(Z_i = 1, S_i = 1)$ is a mixture of two strata $U_i = 11$ and $U_i = 10$

$$
\begin{aligned}
& \mathbb{E}(Y_i = 1 \mid Z_i = 1, S_i = 1) = \mathbb{E}\{Y_i(1) = 1 \mid U = 11/10\} \\
= \quad & \mathbb{E}\{Y_i(1) \mid U = 11\}\mathbb{P}(U = 11 \mid U = 11/10) \\
& + \mathbb{E}\{Y_i(1) \mid U = 10\}\mathbb{P}(U = 10 \mid U = 11/10) \\
= \quad & \frac{\pi_{11}}{\pi_{11} + \pi_{10}} \mathbb{E}\{Y_i(1) \mid S_i(1) = 1, S_i(0) = 1\} \\
& + \frac{\pi_{10}}{\pi_{11} + \pi_{10}} \mathbb{E}\{Y_i(1) \mid S_i(1) = 1, S_i(0) = 0\}
\end{aligned}
$$

Bounds on survivor average causal effect

- The observed stratum $(Z_i = 1, S_i = 1)$ is a mixture of two strata $U_i = 11$ and $U_i = 10$

$$
\begin{aligned}
& \mathbb{E}(Y_i = 1 \mid Z_i = 1, S_i = 1) = \mathbb{E}\{Y_i(1) = 1 \mid U = 11/10\} \\
= \quad & \mathbb{E}\{Y_i(1) \mid U = 11\}\mathbb{P}(U = 11 \mid U = 11/10) \\
& + \mathbb{E}\{Y_i(1) \mid U = 10\}\mathbb{P}(U = 10 \mid U = 11/10) \\
= \quad & \frac{\pi_{11}}{\pi_{11} + \pi_{10}}\mathbb{E}\{Y_i(1) \mid S_i(1) = 1, S_i(0) = 1\} \\
& + \frac{\pi_{10}}{\pi_{11} + \pi_{10}}\mathbb{E}\{Y_i(1) \mid S_i(1) = 1, S_i(0) = 0\}
\end{aligned}
$$

- Y is bounded in $[l, u]$ $\rightsquigarrow$ bounds on $\mathbb{E}\{Y_i(1) \mid S_i(1) = 1, S_i(0) = 1\}$

$$
\begin{aligned}
\text{Upper} &= \frac{(\pi_{11} + \pi_{10})\mathbb{E}(Y_i = 1 \mid Z_i = 1, S_i = 1) - l \cdot \pi_{10}}{\pi_{11}} \\
\text{Lower} &= \frac{(\pi_{11} + \pi_{10})\mathbb{E}(Y_i = 1 \mid Z_i = 1, S_i = 1) - u \cdot \pi_{10}}{\pi_{11}}
\end{aligned}
$$

Bounds on survivor average causal effect

- The observed stratum $(Z_i = 1, S_i = 1)$ is a mixture of two strata $U_i = 11$ and $U_i = 10$

$$
\begin{aligned}
& \mathbb{E}(Y_i = 1 \mid Z_i = 1, S_i = 1) = \mathbb{E}\{Y_i(1) = 1 \mid U = 11/10\} \\
= \quad & \mathbb{E}\{Y_i(1) \mid U = 11\}\mathbb{P}(U = 11 \mid U = 11/10) \\
& + \mathbb{E}\{Y_i(1) \mid U = 10\}\mathbb{P}(U = 10 \mid U = 11/10) \\
= \quad & \frac{\pi_{11}}{\pi_{11} + \pi_{10}}\mathbb{E}\{Y_i(1) \mid S_i(1) = 1, S_i(0) = 1\} \\
& + \frac{\pi_{10}}{\pi_{11} + \pi_{10}}\mathbb{E}\{Y_i(1) \mid S_i(1) = 1, S_i(0) = 0\}
\end{aligned}
$$

- Y is bounded in $[l, u]$ $\rightsquigarrow$ bounds on $\mathbb{E}\{Y_i(1) \mid S_i(1) = 1, S_i(0) = 1\}$

$$
\begin{aligned}
\text{Upper} &= \frac{(\pi_{11} + \pi_{10})\mathbb{E}(Y_i = 1 \mid Z_i = 1, S_i = 1) - l \cdot \pi_{10}}{\pi_{11}} \\
\text{Lower} &= \frac{(\pi_{11} + \pi_{10})\mathbb{E}(Y_i = 1 \mid Z_i = 1, S_i = 1) - u \cdot \pi_{10}}{\pi_{11}}
\end{aligned}
$$

- Statistical inference is hard; even harder without monotonicity

Acute respiratory distress syndrome network study

$Z = 1$				$Z = 0$			
	$Y = 1$	$Y = 0$	total		$Y = 1$	$Y = 0$	total
$S = 1$	54	268	302	$S = 1$	59	218	277
$S = 0$	*	*	109	$S = 0$	*	*	152

- 861 patients with lung injury and acute respiratory distress syndrome were randomized to receive mechanical ventilation with either lower tidal volumes ($Z_i = 1$) or traditional tidal volumes ($Z_i = 0$)
- Outcome: breathe without assistance by day 28 (1 for not)

Acute respiratory distress syndrome network study

$Z=1$				$Z=0$			
	$Y=1$	$Y=0$	total		$Y=1$	$Y=0$	total
$S=1$	54	268	302	$S=1$	59	218	277
$S=0$	*	*	109	$S=0$	*	*	152

- Proportions of principal strata

$$\widehat{\pi}_{11} = \frac{277}{277+152} = 0.646, \quad \widehat{\pi}_{00} = \frac{109}{109+302} = 0.265, \quad \widehat{\pi}_{10} = 0.089$$

Acute respiratory distress syndrome network study

$Z=1$				$Z=0$			
	$Y=1$	$Y=0$	total		$Y=1$	$Y=0$	total
$S=1$	54	268	302	$S=1$	59	218	277
$S=0$	*	*	109	$S=0$	*	*	152

- Proportions of principal strata

$$\widehat{\pi}_{11} = \frac{277}{277+152} = 0.646, \quad \widehat{\pi}_{00} = \frac{109}{109+302} = 0.265, \quad \widehat{\pi}_{10} = 0.089$$

- Sample means of the outcome for survived patients

$$\widehat{\mathbb{E}}(Y_i \mid Z_i = 1, M_i = 1) = \frac{54}{302} = 0.179, \quad \widehat{\mathbb{E}}(Y_i \mid Z_i = 0, M_i = 1) = \frac{59}{277} = 0.213$$

Acute respiratory distress syndrome network study

$Z=1$				$Z=0$			
	$Y=1$	$Y=0$	total		$Y=1$	$Y=0$	total
$S=1$	54	268	302	$S=1$	59	218	277
$S=0$	*	*	109	$S=0$	*	*	152

- Proportions of principal strata

$$\widehat{\pi}_{11} = \frac{277}{277+152} = 0.646, \quad \widehat{\pi}_{00} = \frac{109}{109+302} = 0.265, \quad \widehat{\pi}_{10} = 0.089$$

- Sample means of the outcome for survived patients

$$\widehat{\mathbb{E}}(Y_i \mid Z_i = 1, M_i = 1) = \frac{54}{302} = 0.179, \quad \widehat{\mathbb{E}}(Y_i \mid Z_i = 0, M_i = 1) = \frac{59}{277} = 0.213$$

- Bounds on $\mathbb{E}\{Y_i(1) \mid S_i(1) = 1, S_i(0) = 1\}$: $[0.065, 0.203]$
- Bounds on SACE: $[-0.147, -0.010]$

Outline

- Potential outcomes and ACE
- Post-treatment variable
- Principal stratification
  - noncompliance
  - truncation by death
  - surrogate evaluation
- Identification and statistical inference
  - binary instrumental variable model
  - partial identification
  - principal ignorability
  - auxiliary independence
- Recent applications of principal stratification

Principal ignorability

- An alternative set of identification assumptions without assuming exclusion restriction

Principal ignorability

- An alternative set of identification assumptions without assuming exclusion restriction
  - truncation by death, surrogate evaluation, and etc.
  - similar idea as the treatment ignorability assumption for identifying ACE in observational studies

Principal ignorability

- An alternative set of identification assumptions without assuming exclusion restriction
  - truncation by death, surrogate evaluation, and etc.
  - similar idea as the treatment ignorability assumption for identifying ACE in observational studies

- Treatment ignorability ($Z_i \perp\!\!\!\perp Y_i(z) \mid X_i$) $\rightsquigarrow$ $\mathbb{E}\{Y_i(z) \mid Z_i = 1, \mathbf{X}_i\} = \mathbb{E}\{Y_i(z) \mid Z_i = 0, \mathbf{X}_i\}$

Principal ignorability

- An alternative set of identification assumptions without assuming exclusion restriction
  - truncation by death, surrogate evaluation, and etc.
  - similar idea as the treatment ignorability assumption for identifying ACE in observational studies

- Treatment ignorability $(Z_i \perp\!\!\!\perp Y_i(z) \mid X_i) \rightsquigarrow$ $\mathbb{E}\{Y_i(z) \mid Z_i = 1, \mathbf{X}_i\} = \mathbb{E}\{Y_i(z) \mid Z_i = 0, \mathbf{X}_i\}$

- Principal ignorability $Y_i(z) \perp\!\!\!\perp U_i \mid (Z_i = z, S_i = s, \mathbf{X}_i)$

Principal ignorability

- An alternative set of identification assumptions without assuming exclusion restriction
  - truncation by death, surrogate evaluation, and etc.
  - similar idea as the treatment ignorability assumption for identifying ACE in observational studies

- Treatment ignorability ($Z_i \perp\!\!\!\perp Y_i(z) \mid X_i$) $\rightsquigarrow$ $\mathbb{E}\{Y_i(z) \mid Z_i = 1, \mathbf{X}_i\} = \mathbb{E}\{Y_i(z) \mid Z_i = 0, \mathbf{X}_i\}$

- Principal ignorability $Y_i(z) \perp\!\!\!\perp U_i \mid (Z_i = z, S_i = s, \mathbf{X}_i)$
  - expected potential outcome is the same across different principal strata within each observed stratum
  - implies $\mathbb{E}\{Y_i(1) \mid U_i = 11, Z_i = 1, S_i = 1, \mathbf{X}_i\} = \mathbb{E}\{Y_i(1) \mid U_i = 10, Z_i = 1, S_i = 1, \mathbf{X}_i\}$

Identification assumptions

- Randomization: $Z_i \perp\!\!\!\perp \{Y_i(1), Y_i(0), S_i(1), S_i(0)\}$

Identification assumptions

- Randomization: $Z_i \perp\!\!\!\perp \{Y_i(1), Y_i(0), S_i(1), S_i(0)\}$

- Monotonicity: $S_i(1) \geq S_i(0)$
  - principal score
    $\mathbb{P}(U_i = 11 \mid \mathbf{X}_i) = \mathbb{P}(S_i = 1 \mid Z_i = 0, \mathbf{X}_i)$,
    $\mathbb{P}(U_i = 00 \mid \mathbf{X}_i) = \mathbb{P}(S_i = 0 \mid Z_i = 1, \mathbf{X}_i)$
    $\mathbb{P}(U_i = 10 \mid \mathbf{X}_i) = \mathbb{P}(S_i = 1 \mid Z_i = 1, \mathbf{X}_i) - \mathbb{P}(S_i = 1 \mid Z_i = 0, \mathbf{X}_i)$

Identification assumptions

- Randomization: $Z_i \perp\!\!\!\perp \{Y_i(1), Y_i(0), S_i(1), S_i(0)\}$
- Monotonicity: $S_i(1) \geq S_i(0)$
  - principal score
    $\mathbb{P}(U_i = 11 \mid \mathbf{X}_i) = \mathbb{P}(S_i = 1 \mid Z_i = 0, \mathbf{X}_i)$,
    $\mathbb{P}(U_i = 00 \mid \mathbf{X}_i) = \mathbb{P}(S_i = 0 \mid Z_i = 1, \mathbf{X}_i)$
    $\mathbb{P}(U_i = 10 \mid \mathbf{X}_i) = \mathbb{P}(S_i = 1 \mid Z_i = 1, \mathbf{X}_i) - \mathbb{P}(S_i = 1 \mid Z_i = 0, \mathbf{X}_i)$
- Principal ignorability: $Y_i(z) \perp\!\!\!\perp U_i \mid (Z_i = z, S_i = s, \mathbf{X}_i)$
  - $\mathbb{E}\{Y_i(1) \mid U_i = 11, Z_i = 1, S_i = 1, \mathbf{X}_i\} = \mathbb{E}\{Y_i(1) \mid Z_i = 1, S_i = 1, \mathbf{X}_i\}$
  - $\mathbb{E}\{Y_i(1) \mid U_i = 10, Z_i = 1, S_i = 1, \mathbf{X}_i\} = \mathbb{E}\{Y_i(1) \mid Z_i = 1, S_i = 1, \mathbf{X}_i\}$

Weighting method

Theorem

$$
\begin{aligned}
\mathrm{ACE}_{10} &= \mathbb{E}\{w_{1,10}(\mathbf{X})Y \mid Z=1, S=1\} - \mathbb{E}\{w_{0,10}(\mathbf{X})Y \mid Z=0, S=0\} \\
\mathrm{ACE}_{00} &= \mathbb{E}\{Y \mid Z=1, S=0\} - \mathbb{E}\{w_{0,00}(\mathbf{X})Y \mid Z=0, S=0\} \\
\mathrm{ACE}_{11} &= \mathbb{E}\{w_{1,11}(\mathbf{X})Y \mid Z=1, S=1\} - \mathbb{E}\{Y \mid Z=0, S=1\}
\end{aligned}
$$

$$e_u(\mathbf{X}) = \mathbb{P}(U_i = u \mid \mathbf{X}) \quad e_u = \mathbb{E}\{e_u(\mathbf{X})\}$$

$$w_{1,10} = \frac{e_{10}(\mathbf{X})}{e_{10}(\mathbf{X}) + e_{11}(\mathbf{X})} \Bigg/ \frac{e_{10}}{e_{10} + e_{11}} \quad w_{0,11} = \frac{e_{10}(\mathbf{X})}{e_{10}(\mathbf{X}) + e_{00}(\mathbf{X})} \Bigg/ \frac{e_{10}}{e_{10} + e_{00}}$$

$$w_{0,00} = \frac{e_{00}(\mathbf{X})}{e_{10}(\mathbf{X}) + e_{00}(\mathbf{X})} \Bigg/ \frac{e_{00}}{e_{10} + e_{00}} \quad w_{1,11} = \frac{e_{11}(\mathbf{X})}{e_{10}(\mathbf{X}) + e_{11}(\mathbf{X})} \Bigg/ \frac{e_{11}}{e_{10} + e_{11}}$$

Estimation and inference

- Estimation steps
  1. estimate the principal score $\mathbb{P}(U_i \mid \mathbf{X}_i)$, e.g., logistic model
  2. plug the estimated principal score in the weighting formula

Estimation and inference

- Estimation steps
  1. estimate the principal score $\mathbb{P}(U_i \mid \mathbf{X}_i)$, e.g., logistic model
  2. plug the estimated principal score in the weighting formula

- Variance and confidence interval: bootstrap

Estimation and inference

- Estimation steps
  1. estimate the principal score $\mathbb{P}(U_i \mid \mathbf{X}_i)$, e.g., logistic model
  2. plug the estimated principal score in the weighting formula

- Variance and confidence interval: bootstrap

- A unification in observational studies (Jiang et al., 2022)
  - treatment ignorability, principal ignorability, monotonicity
  - triple robustness: propensity score, principal score, outcome model

Comparison of different identification strategies

- A1: $Z_i \perp\!\!\!\perp \{Y_i(1), Y_i(0), S_i(1), S_i(0)\}$;
  A2: $Z_i \perp\!\!\!\perp \{Y_i(1), Y_i(0), S_i(1), S_i(0)\} \mid \mathbf{X}_i$
- B1: monotonicity and exclusion restriction;
  B2: monotonicity and principal ignorability

Comparison of different identification strategies

- A1: $Z_i \perp\!\!\!\perp \{Y_i(1), Y_i(0), S_i(1), S_i(0)\}$;
  A2: $Z_i \perp\!\!\!\perp \{Y_i(1), Y_i(0), S_i(1), S_i(0)\} \mid \mathbf{X}_i$
- B1: monotonicity and exclusion restriction;
  B2: monotonicity and principal ignorability

- A1+ B1: non-compliance in randomized trials, model free

Comparison of different identification strategies

- A1: $Z_i \perp\!\!\!\perp \{Y_i(1), Y_i(0), S_i(1), S_i(0)\}$;
  A2: $Z_i \perp\!\!\!\perp \{Y_i(1), Y_i(0), S_i(1), S_i(0)\} \mid \mathbf{X}_i$
- B1: monotonicity and exclusion restriction;
  B2: monotonicity and principal ignorability

- A1+ B1: non-compliance in randomized trials, model free
- A2+ B1: non-compliance in observational studies, model for S_i or Y_i

Comparison of different identification strategies

- A1: $Z_i \perp\!\!\!\perp \{Y_i(1), Y_i(0), S_i(1), S_i(0)\}$;
  A2: $Z_i \perp\!\!\!\perp \{Y_i(1), Y_i(0), S_i(1), S_i(0)\} \mid \mathbf{X}_i$
- B1: monotonicity and exclusion restriction;
  B2: monotonicity and principal ignorability

- A1+ B1: non-compliance in randomized trials, model free
- A2+ B1: non-compliance in observational studies, model for S_i or Y_i
- A1+B2: randomized trials without exclusion restriction, model for S_i

Comparison of different identification strategies

- A1: $Z_i \perp\!\!\!\perp \{Y_i(1), Y_i(0), S_i(1), S_i(0)\}$;
  A2: $Z_i \perp\!\!\!\perp \{Y_i(1), Y_i(0), S_i(1), S_i(0)\} \mid \mathbf{X}_i$
- B1: monotonicity and exclusion restriction;
  B2: monotonicity and principal ignorability

- A1+ B1: non-compliance in randomized trials, model free
- A2+ B1: non-compliance in observational studies, model for S_i or Y_i
- A1+B2: randomized trials without exclusion restriction, model for S_i
- A2+B2: observational studies without exclusion restriction, model for two of Z_i, S_i, and Y_i

Comparison of different identification strategies

- A1: $Z_i \perp\!\!\!\perp \{Y_i(1), Y_i(0), S_i(1), S_i(0)\}$;
  A2: $Z_i \perp\!\!\!\perp \{Y_i(1), Y_i(0), S_i(1), S_i(0)\} \mid \mathbf{X}_i$
- B1: monotonicity and exclusion restriction;
  B2: monotonicity and principal ignorability

- A1+ B1: non-compliance in randomized trials, model free
- A2+ B1: non-compliance in observational studies, model for S_i or Y_i
- A1+B2: randomized trials without exclusion restriction, model for S_i
- A2+B2: observational studies without exclusion restriction, model for two of Z_i, S_i, and Y_i

- Other strategies: additional information (e.g., auxiliary independence), likelihood based inference, Bayesian analysis

Outline

- Potential outcomes and ACE
- Post-treatment variable
- Principal stratification
  - noncompliance
  - truncation by death
  - surrogate evaluation
- Identification and statistical inference
  - binary instrumental variable model
  - partial identification
  - principal ignorability
  - auxiliary independence
- Recent applications of principal stratification

Auxiliary variables

- Auxiliary independence: $Y_i(z) \perp\!\!\!\perp W_i \mid U_i$

Auxiliary variables

- Auxiliary independence: $Y_i(z) \perp\!\!\!\perp W_i \mid U_i$
- Augmented design in Follmann (2006)
  - HIV vaccine injection Z_i; immune response S_i; infection indicator Y_i
  - immune responses to rabies vaccine W_is: potential HIV infection status should not depend on a irrelevant vaccine

Auxiliary variables

- Auxiliary independence: $Y_i(z) \perp\!\!\!\perp W_i \mid U_i$
- Augmented design in Follmann (2006)
  - HIV vaccine injection Z_i; immune response S_i; infection indicator Y_i
  - immune responses to rabies vaccine W_is: potential HIV infection status should not depend on a irrelevant vaccine
- Identification with multiple trials in Jiang et al. (2016)
  - treatment Z_i; three-year cancer reoccurrence S_i; five-year survival Y_i
  - trial number W_i: potential survival status does not depend on the trial number given physical status

Identification assumptions

- Randomization: $Z_i \perp\!\!\!\perp \{Y_i(1), Y_i(0), S_i(1), S_i(0)\} \mid W_i$

Identification assumptions

- Randomization: $Z_i \perp\!\!\!\perp \{Y_i(1), Y_i(0), S_i(1), S_i(0)\} \mid W_i$

- Monotonicity: $S_i(1) \geq S_i(0)$

Identification assumptions

- Randomization: $Z_i \perp\!\!\!\perp \{Y_i(1), Y_i(0), S_i(1), S_i(0)\} \mid W_i$
- Monotonicity: $S_i(1) \geq S_i(0)$
- Auxiliary independence: $Y_i(z) \perp\!\!\!\perp W_i \mid U_i$
  - $\mathbb{E}\{Y_i(1) - Y_i(0) \mid U_i, W_i\} = \mathbb{E}\{Y_i(1) - Y_i(0) \mid W_i\}$

Identification with discrete S_i

- $S \in \{s_1, \ldots, s_K\}$ and $W \in \{w_1, \ldots, w_L\}$
- M_{s_0}: $K \times L$ matrix with (k, l)-th element $\mathbb{P}(S_1 = s_k \mid S_0 = s_0, W = w_l)$
- M_{s_1}: $K \times L$ matrix with (k, l)-th element $\mathbb{P}(S_0 = s_k \mid S_1 = s_1, W = w_l)$

Identification with discrete S_i

- $S \in \{s_1, \ldots, s_K\}$ and $W \in \{w_1, \ldots, w_L\}$
- M_{s_0}: $K \times L$ matrix with (k, l)-th element $\mathbb{P}(S_1 = s_k \mid S_0 = s_0, W = w_l)$
- M_{s_1}: $K \times L$ matrix with (k, l)-th element $\mathbb{P}(S_0 = s_k \mid S_1 = s_1, W = w_l)$

- $\mathbb{P}(Y_0 \mid S_1, S_0 = s_0)$ is identifiable if $\mathrm{rank}(M_{s_0}^\top M_{s_0}) = K$
- $\mathbb{P}(Y_1 \mid S_1 = s_1, S_0)$ is identifiable if $\mathrm{rank}(M_{s_1}^\top M_{s_1}) = K$

Special case

- Strong monotonicity holds $S_i(0) = 0$
- Binary S_i and W_i

Special case

- Strong monotonicity holds $S_i(0) = 0$
- Binary S_i and W_i

- Rank condition is equivalent to $S_i \not\perp\!\!\!\perp W_i \mid Z_i = 1$

Special case

- Strong monotonicity holds $S_i(0) = 0$
- Binary S_i and W_i

- Rank condition is equivalent to $S_i \not\perp\!\!\!\perp W_i \mid Z_i = 1$

- Solve $\mathbb{E}(Y_0 \mid S_1)$ from equations

$$
\begin{aligned}
\delta_1 &= \mathbb{E}(Y_0 \mid S_1 = 1)\theta_{11} + \mathbb{E}(Y_0 \mid S_1 = 0)\theta_{01} \\
\delta_0 &= \mathbb{E}(Y_0 \mid S_1 = 1)\theta_{10} + \mathbb{E}(Y_0 \mid S_1 = 0)\theta_{00}
\end{aligned}
$$

  - $\theta_{sw} = \mathbb{P}(S = s \mid Z = 1, W = w)$ and $\delta_w = \mathbb{E}(Y \mid Z = 0, W = w)$

Special case

- Strong monotonicity holds $S_i(0) = 0$
- Binary S_i and W_i

- Rank condition is equivalent to $S_i \not\perp\!\!\!\perp W_i \mid Z_i = 1$

- Solve $\mathbb{E}(Y_0 \mid S_1)$ from equations

$$
\begin{aligned}
\delta_1 &= \mathbb{E}(Y_0 \mid S_1 = 1)\theta_{11} + \mathbb{E}(Y_0 \mid S_1 = 0)\theta_{01} \\
\delta_0 &= \mathbb{E}(Y_0 \mid S_1 = 1)\theta_{10} + \mathbb{E}(Y_0 \mid S_1 = 0)\theta_{00}
\end{aligned}
$$

  - $\theta_{sw} = \mathbb{P}(S = s \mid Z = 1, W = w)$ and $\delta_w = \mathbb{E}(Y \mid Z = 0, W = w)$

- General case in Jiang and Ding (2021)

Outline

- Potential outcomes and ACE
- Post-treatment variable
- Principal stratification
  - noncompliance
  - truncation by death
  - surrogate evaluation
- Identification and statistical inference
  - binary instrumental variable model
  - partial identification
  - principal ignorability
  - auxiliary independence
- Recent applications of principal stratification

Principal stratification based on outcomes

- Principal stratification does not require S to be on the causal pathway from Z to Y

Principal stratification based on outcomes

- Principal stratification does not require S to be on the causal pathway from Z to Y
- Examples of principal stratification based on outcomes
  - proportion of units benefit from treatment: $\mathbb{P}\{Y(1) > Y(0)\}$
  - probability of necessity: $\mathbb{P}\{Y(0) = 0 \mid Z = 1, Y = 1\}$;
    probability of sufficiency: $\mathbb{P}\{Y(1) = 1 \mid Z = 0, Y = 0\}$

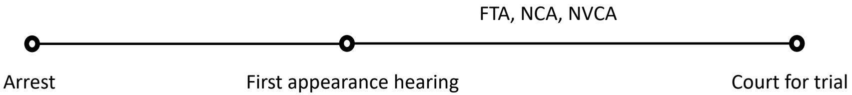
Principal stratification based on outcomes

- Principal stratification does not require S to be on the causal pathway from Z to Y

- Examples of principal stratification based on outcomes
  - proportion of units benefit from treatment: $\mathbb{P}\{Y(1) > Y(0)\}$
  - probability of necessity: $\mathbb{P}\{Y(0) = 0 \mid Z = 1, Y = 1\}$;
    probability of sufficiency: $\mathbb{P}\{Y(1) = 1 \mid Z = 0, Y = 0\}$

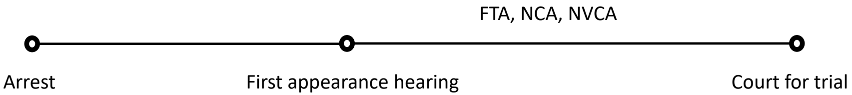
- Evaluation of algorithm-assisted human decision making
  - consequential decisions made by judges, doctors, etc.
  - pre-trial risk assessment instrument

First appearance hearings



- Judges decide pre-trial release conditions
  - bail and monitoring
  - many cases in one day

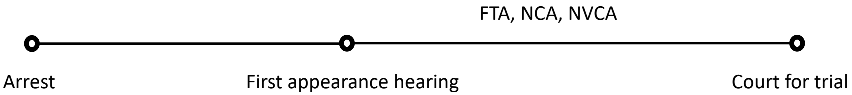
First appearance hearings



- Judges decide pre-trial release conditions
  - bail and monitoring
  - many cases in one day

- Presumption of innocence: judges balance between

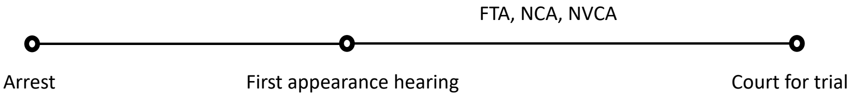
First appearance hearings



- Judges decide pre-trial release conditions
  - bail and monitoring
  - many cases in one day

- Presumption of innocence: judges balance between
  - cost of pre-trial detention
  - risk of arrestee

First appearance hearings



- Judges decide pre-trial release conditions
  - bail and monitoring
  - many cases in one day
- Presumption of innocence: judges balance between
  - cost of pre-trial detention
  - risk of arrestee
- Judges are required to consider three negative outcomes
  1. arrestee may fail to appear in trial court (FTA)
  2. arrestee may engage in new criminal activity (NCA)
  3. arrestee may engage in new violent criminal activity (NVCA)

Pretrial Public Safety Assessment (PSA)

- PSA as an algorithmic recommendation

- PSA scores
  1. calculated based on nine factors
  2. two 6-point scores for FTA and NCA
  3. one binary score for NVCA

Failure to Appear: Points

PSA FACTOR	RESPONSE	POINTS
Pending charge at the time of the arrest	No	0
	Yes	1
Prior conviction (misdemeanor or felony)	No	0
	Yes	1
Prior failure to appear in the past 2 years	No	0
	Yes, just 1	2
	Yes, 2 or more	4
Prior failure to appear older than 2 years	No	0
	Yes	1

Pretrial Public Safety Assessment (PSA)

- PSA as an algorithmic recommendation

- PSA scores
  1. calculated based on nine factors
  2. two 6-point scores for FTA and NCA
  3. one binary score for NVCA

- Decision Making Framework (DMF)
  - combines scores for bail recommendation:
    cash bail or signature bond

Failure to Appear: Points

PSA FACTOR	RESPONSE	POINTS
Pending charge at the time of the arrest	No	0
	Yes	1
Prior conviction (misdemeanor or felony)	No	0
	Yes	1
Prior failure to appear in the past 2 years	No	0
	Yes, just 1	2
	Yes, 2 or more	4
Prior failure to appear older than 2 years	No	0
	Yes	1

DANE COUNTY CLERK OF COURTS
Public Safety Assessment – Report

215 S Hamilton St #1000
Madison, WI 53703
Phone: (608) 266-4311

Name: [redacted]
Spillman Name Number: [redacted]
DOB: [redacted]
Gender: Male
Arrest Date: 03/25/2017
PSA Completion Date: 03/27/2017

New Violent Criminal Activity Flag

No

New Criminal Activity Scale

1	2	3	4	5	6

Failure to Appear Scale

1	2	3	4	5	6

Charge(s):

961.41(1)(D)(1) MFC DELIVER HEROIN <3 GMS F 3

Risk Factors:	**Responses:**
1. Age at Current Arrest	23 or Older
2. Current Violent Offense	No
a. Current Violent Offense & 20 Years Old or Younger	No
3. Pending Charge at the Time of the Offense	No
4. Prior Misdemeanor Conviction	Yes
5. Prior Felony Conviction	Yes
a. Prior Conviction	Yes
6. Prior Violent Conviction	2
7. Prior Failure to Appear Pretrial in Past 2 Years	0
8. Prior Failure to Appear Pretrial Older than 2 Years	Yes
9. Prior Sentence to Incarceration	Yes

Recommendations:

Release Recommendation - Signature bond

Conditions - Report to and comply with pretrial supervision

A field experiment for evaluating the PSA

- A field experiment in Dane county, Wisconsin
  - PSA is generated for each case using a computer system
  - randomly make PSA reports available to judges
  - mid-2017 – 2019 (randomization), 2-year follow-up

A field experiment for evaluating the PSA

- A field experiment in Dane county, Wisconsin
  - PSA is generated for each case using a computer system
  - randomly make PSA reports available to judges
  - mid-2017 – 2019 (randomization), 2-year follow-up

- Trichotomized ordinal decisions of bail amount
  1. signature bond
  2. $\leq$ \$1,000 cash bond (small)
  3. $>$ \$1,000 cash bond (large)

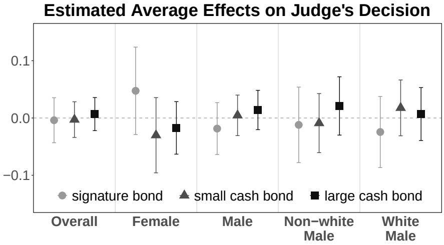
A field experiment for evaluating the PSA

- A field experiment in Dane county, Wisconsin
  - PSA is generated for each case using a computer system
  - randomly make PSA reports available to judges
  - mid-2017 – 2019 (randomization), 2-year follow-up

- Trichotomized ordinal decisions of bail amount
  1. signature bond
  2. $\leq$ \$1,000 cash bond (small)
  3. $>$ \$1,000 cash bond (large)

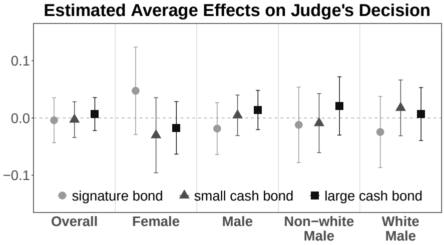
- Z_i: PSA provision indicator
  S_i: judge's decision (0 for signature bond, 1 for small cash, and 2 for large cash)
  Y_i: FTA, NCA, or NVCA

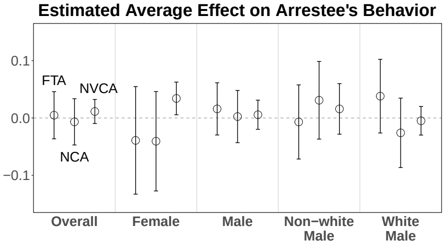
Intention-to-treat analysis of PSA provision



- ACE($Z \to S$) and ACE($Z \to Y$) $\rightsquigarrow$ overall effects of PSA provision

Intention-to-treat analysis of PSA provision





- ACE$(Z \rightarrow S)$ and ACE$(Z \rightarrow Y)$ $\rightsquigarrow$ overall effects of PSA provision
  - insignificant effects on judges' decisions
  - possible effect on NVCA for females

Issue of ITT analysis

- ITT effect on the negative outcomes
    - answers whether PSA provision helps prevent FTA, NCA, and NVCA

Issue of ITT analysis

- ITT effect on the negative outcomes
  - answers whether PSA provision helps prevent FTA, NCA, and NVCA
  - not enough for evaluating PSA

Issue of ITT analysis

- ITT effect on the negative outcomes
  - answers whether PSA provision helps prevent FTA, NCA, and NVCA
  - not enough for evaluating PSA

- ITT effect on the judge's decision
  - answers whether PSA provision makes decision harsher or more lenient

Issue of ITT analysis

- ITT effect on the negative outcomes
  - answers whether PSA provision helps prevent FTA, NCA, and NVCA
  - not enough for evaluating PSA

- ITT effect on the judge's decision
  - answers whether PSA provision makes decision harsher or more lenient
  - does not answer whether PSA provision helps make better decisions

Issue of ITT analysis

- ITT effect on the negative outcomes
  - answers whether PSA provision helps prevent FTA, NCA, and NVCA
  - not enough for evaluating PSA
- ITT effect on the judge's decision
  - answers whether PSA provision makes decision harsher or more lenient
  - does not answer whether PSA provision helps make better decisions
- Good decisions: detain risky arrestees and release safe arrestees

The meaning of a "good" decision varies

- A toy example

	if released	if detained
Arrestee A	NCA	no NCA
Arrestee B	no NCA	no NCA
Arrestee C	NCA	NCA

	without PSA	with PSA
Judge 1	release all	detain A,C
Judge 2	release all	detain B,C

The meaning of a "good" decision varies

- A toy example

	if released	if detained
Arrestee A	NCA	no NCA
Arrestee B	no NCA	no NCA
Arrestee C	NCA	NCA

	without PSA	with PSA
Judge 1	release all	detain A,C
Judge 2	release all	detain B,C

  - ITT effects of PSA are the same

The meaning of a "good" decision varies

- A toy example

	if released	if detained
Arrestee A	NCA	no NCA
Arrestee B	no NCA	no NCA
Arrestee C	NCA	NCA

	without PSA	with PSA
Judge 1	release all	detain A,C
Judge 2	release all	detain B,C

  - ITT effects of PSA are the same
  - Risk level: C $>$ A $>$ B
    - detaining A prevents an NCA
    - detaining B is unnecessary
    - detaining C does not help

The meaning of a "good" decision varies

- A toy example

	if released	if detained
Arrestee A	NCA	no NCA
Arrestee B	no NCA	no NCA
Arrestee C	NCA	NCA

	without PSA	with PSA
Judge 1	release all	detain A,C
Judge 2	release all	detain B,C

  - ITT effects of PSA are the same
  - Risk level: C $>$ A $>$ B
    - detaining A prevents an NCA
    - detaining B is unnecessary
    - detaining C does not help

- Risk levels depend on $\left\{ \begin{array}{l} \text{(potential) outcome if released} \\ \text{(potential) outcome if detained} \end{array} \right.$

Setup

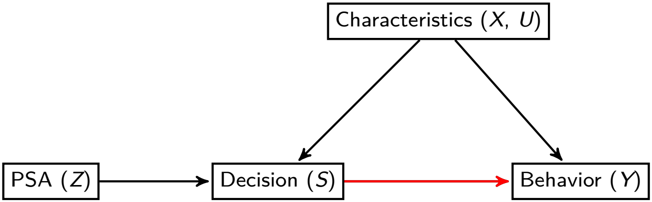
- Notation
  - Z_i: PSA provision indicator
  - S_i: 1 for detention, 0 for release
  - Y_i: binary outcome (e.g., NCA)
  - X_i: observed covariates
  - U_i: unobserved covariates

Setup

- Notation
  - Z_i: PSA provision indicator
  - S_i: 1 for detention, 0 for release
  - Y_i: binary outcome (e.g., NCA)
  - X_i: observed covariates
  - U_i: unobserved covariates

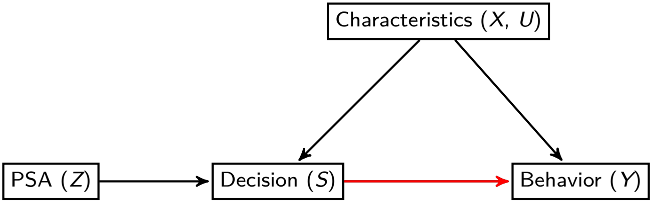
- Potential outcomes
  - $D_i(z)$: potential value of the decision when $Z_i = z$
  - $Y_i(z, s)$: potential outcome when $Z_i = z$ and $S_i = s$
  - No interference across cases: first arrests only

Assumptions



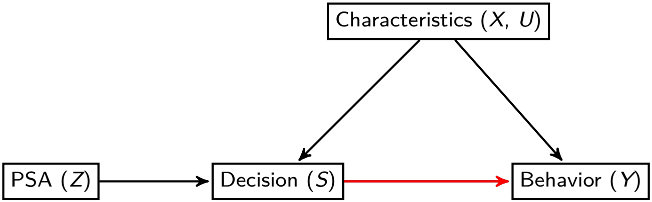
- Randomized treatment assignment: $\{S_i(z), Y_i(z, s), X_i, U_i\} \perp\!\!\!\perp Z_i$

Assumptions



- Randomized treatment assignment: $\{S_i(z), Y_i(z,s), X_i, U_i\} \perp\!\!\!\perp Z_i$
- Exclusion restriction: $Y_i(z,s) \; = \; Y_i(s)$

Assumptions



- Randomized treatment assignment: $\{S_i(z), Y_i(z, s), X_i, U_i\} \perp\!\!\!\perp Z_i$
- Exclusion restriction: $Y_i(z, s) \;=\; Y_i(s)$
- Monotonicity: $Y_i(0) \;\geq\; Y_i(1)$

Defining risk levels based on principal stratification

- Principal stratification
  - $(Y_i(1), Y_i(0)) = (0, 1)$: preventable cases
  - $(Y_i(1), Y_i(0)) = (1, 1)$: risky cases
  - $(Y_i(1), Y_i(0)) = (0, 0)$: safe cases
  - $\cancel{(Y_i(1), Y_i(0)) = (1, 0)}$: eliminated by monotonicity

Defining risk levels based on principal stratification

- Principal stratification
  - $(Y_i(1), Y_i(0)) = (0, 1)$: preventable cases
  - $(Y_i(1), Y_i(0)) = (1, 1)$: risky cases
  - $(Y_i(1), Y_i(0)) = (0, 0)$: safe cases
  - $\cancel{(Y_i(1), Y_i(0)) = (1, 0)}$: eliminated by monotonicity

- Average principal causal effects of PSA on judges' decisions:

$$
\begin{aligned}
\text{APCEp} &= \mathbb{E}\{S_i(1) - S_i(0) \mid Y_i(1) = 0, Y_i(0) = 1\}, \\
\text{APCEr} &= \mathbb{E}\{S_i(1) - S_i(0) \mid Y_i(1) = 1, Y_i(0) = 1\}, \\
\text{APCEs} &= \mathbb{E}\{S_i(1) - S_i(0) \mid Y_i(1) = 0, Y_i(0) = 0\}.
\end{aligned}
$$

Defining risk levels based on principal stratification

- Principal stratification
  - $(Y_i(1), Y_i(0)) = (0, 1)$: preventable cases
  - $(Y_i(1), Y_i(0)) = (1, 1)$: risky cases
  - $(Y_i(1), Y_i(0)) = (0, 0)$: safe cases
  - $\cancel{(Y_i(1), Y_i(0)) = (1, 0)}$: eliminated by monotonicity

- Average principal causal effects of PSA on judges' decisions:

$$
\begin{aligned}
\text{APCEp} &= \mathbb{E}\{S_i(1) - S_i(0) \mid Y_i(1) = 0, Y_i(0) = 1\}, \\
\text{APCEr} &= \mathbb{E}\{S_i(1) - S_i(0) \mid Y_i(1) = 1, Y_i(0) = 1\}, \\
\text{APCEs} &= \mathbb{E}\{S_i(1) - S_i(0) \mid Y_i(1) = 0, Y_i(0) = 0\}.
\end{aligned}
$$

- If PSA is helpful, we should have APCEp > 0 and APCEs < 0.
- The desirable sign of APCEr depends on various factors.

Partial identification

$$
\begin{aligned}
\text{APCEp} &= \frac{\mathbb{P}(Y_i = 1 \mid Z_i = 0) - \mathbb{P}(Y_i = 1 \mid Z_i = 1)}{\mathbb{P}\{Y_i(0) = 1\} - \mathbb{P}\{Y_i(1) = 1\}} \\
\text{APCEr} &= \frac{\mathbb{P}(S_i = 1, Y_i = 1 \mid Z_i = 1) - \mathbb{P}(S_i = 1, Y_i = 1 \mid Z_i = 0)}{\mathbb{P}\{Y_i(1) = 1\}} \\
\text{APCEs} &= \frac{\mathbb{P}(S_i = 0, Y_i = 0 \mid Z_i = 0) - \mathbb{P}(S_i = 0, Y_i = 0 \mid Z_i = 1)}{1 - \mathbb{P}\{Y_i(0) = 1\}}
\end{aligned}
$$

- The signs of principal causal effects are identifiable under the assumptions of randomization, exclusion restriction, and monotonicity.

Extension to Ordinal Decision

- Judges decisions are typically ordinal (e.g., bail amount)
  - $S_i = 0, 1, \ldots, k$: a bail of increasing amount
  - Monotonicity: $Y_i(s_1) \;\geq\; Y_i(s_2)$ for $s_1 \leq s_2$

Extension to Ordinal Decision

- Judges decisions are typically ordinal (e.g., bail amount)
  - $S_i = 0, 1, \ldots, k$: a bail of increasing amount
  - Monotonicity: $Y_i(s_1) \;\geq\; Y_i(s_2)$ for $s_1 \leq s_2$
- Principal strata based on an ordinal measure of risk

$$R_i \;=\; \begin{cases} \min\{s : Y_i(s) = 0\} & \text{if } Y_i(k) = 0 \\ k + 1 & \text{if } Y_i(k) = 1 \end{cases}$$

  - Least amount of bail that keeps an arrestee from committing NCA

Extension to Ordinal Decision

- Judges decisions are typically ordinal (e.g., bail amount)
  - $S_i = 0, 1, \ldots, k$: a bail of increasing amount
  - Monotonicity: $Y_i(s_1) \;\geq\; Y_i(s_2)$ for $s_1 \leq s_2$

- Principal strata based on an ordinal measure of risk

$$R_i \;=\; \begin{cases} \min\{s : Y_i(s) = 0\} & \text{if } Y_i(k) = 0 \\ k+1 & \text{if } Y_i(k) = 1 \end{cases}$$

  - Least amount of bail that keeps an arrestee from committing NCA
  - Example with $k = 2$

principal strata	$(Y_i(0), Y_i(1), Y_i(2))$	R_i
risky cases	$(1, 1, 1)$	3
preventable cases	$(1, 1, 0)$	2
easily preventable cases	$(1, 0, 0)$	1
safe cases	$(0, 0, 0)$	0

APCE for ordinal decision

- For people with $R_i = r$
  - judges make decision $S_i \geq r \rightsquigarrow$ not commit a crime
  - judges make decision $S_i < r \rightsquigarrow$ commit a crime

- Causal quantities of interest : reduction in the proportion of NCA attributable to PSA provision

$$\mathsf{ACEP}(r) \;=\; \Pr\{S_i(1) \geq r \mid R_i = r\} - \Pr\{S_i(0) \geq r \mid R_i = r\}$$

- Partial identification without unconfoundedness

Point identification

- Unconfoundedness: $Y_i(s) \perp\!\!\!\perp S_i \mid \mathbf{X}_i, Z_i = z$
- Violation of unconfoundedness
  - unobserved covariates between decision and outcome
  - sensitivity analysis

Point identification

- Unconfoundedness: $Y_i(s) \perp\!\!\!\perp S_i \mid \mathbf{X}_i, Z_i = z$
- Violation of unconfoundedness
  - unobserved covariates between decision and outcome
  - sensitivity analysis
- Principal score

$$e_r(\mathbf{x}) = \Pr(R_i = r \mid \mathbf{X}_i = \mathbf{x})$$

Point identification

- Unconfoundedness: $Y_i(s) \perp\!\!\!\perp S_i \mid \mathbf{X}_i, Z_i = z$
- Violation of unconfoundedness
  - unobserved covariates between decision and outcome
  - sensitivity analysis
- Principal score

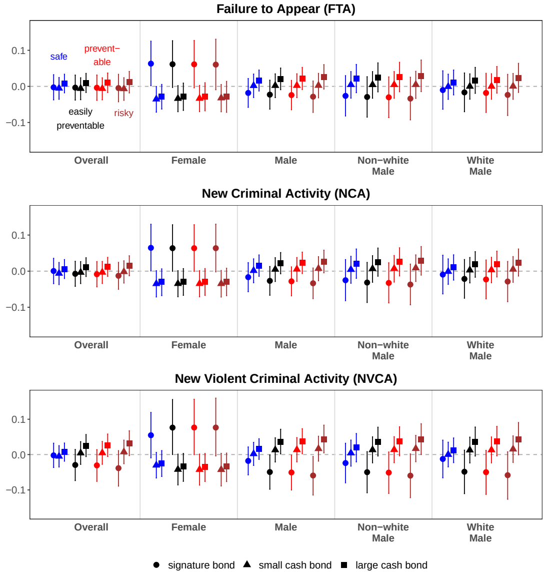
$$e_r(\mathbf{x}) \quad = \quad \Pr(R_i = r \mid \mathbf{X}_i = \mathbf{x})$$

- Identification formula

$$\mathsf{ACEP}(r) \;=\; \mathbb{E}\left[\underbrace{\frac{e_r(\mathbf{x})}{\mathbb{E}\{e_r(\mathbf{X}_i)\}}}_{\text{weight}} \mathbf{1}(S_i \geq r) \mid Z_i = 1\right] - \mathbb{E}\left[\underbrace{\frac{e_r(\mathbf{x})}{\mathbb{E}\{e_r(\mathbf{X}_i)\}}}_{\text{weight}} \mathbf{1}(S_i \geq r) \mid Z_i = 0\right]$$

Estimated average principal causal effects

$$\Pr\{S_i(1) = s \mid R_i = r\} - \Pr\{S_i(0) = s \mid R_i = r\}$$



Topics not covered

- Principal stratification in observational studies
  - identification assumptions
  - outcome regression, inverse probability weighting, multiply robust estimation
- Other identification strategies
  - parametric modeling
  - Bayesian analysis
  - using additional information, e.g., secondary outcome
- More complex settings
  - interference
  - data complications: missing data, selection bias, measurement error

Summary

- Post-treatment variable S_i: affected by the treatment
  - naive adjustment for the observed variable S_i does not yield a valid causal quantity
  - principal stratification defined by $S_i(1)$ and $S_i(0)$
  - application to non-compliance, truncation-by-death, and surrogate evaluation problems

- Various identification strategies
  - monotonicity and exclusion restriction
  - monotonicity and principal ignorability

- Extension
  - identification without monotonicity
  - discrete and continous S