Potential Outcomes Approach: A Brief Introduction

Booil Jo*
Dept of Psychiatry & Behavioral Sciences
Stanford University
booil@stanford.edu

Education 260, Statistics 209
May 16, 2006

Why Randomized Trials?

◦ Expensive and time consuming, recruitment is more difficult, ethical issues.

• To be able to make causal inferences (evidence).

• Gold Standard - typically compares an experimental treatment to a standard treatment or to a placebo group.

• Random assignment results in unbiased distribution of confounding variables (e.g., gender, age, education). Randomized groups are comparable at the baseline.

• Blinding results in unbiased distribution of placebo effects.

• The difference between groups in the outcome can be attributed to the treatment (i.e., treatment is the cause).

Intermediating Posttreatment Variables in Randomized Trials

• Posttreatment variables measured after the treatment assignment and before the assessment of final outcomes of interest.

• These variables can potentially mediate the effect of treatment assignment on the outcome.

• Mediators, proximal outcomes, surrogate markers...

◦ Option 1: ITT analysis ignoring these intermediate variables.

◦ Option 2: Take into account these intermediate variables - naive way or causal inference way.

Intermediating Posttreatment Variables in Randomized Trials

• Sometimes, we want to make inferences taking into account these intermittent posttreatment variables (vs. ITT analysis ignoring these variables).

  – Net treatment effect

  – As-treated effect

  – Mediation analysis (path analysis): direct & indirect effects

◦ Posttreatment variables are already **affected by treatment assignment** - not qualified as pretreatment covariates.

◦ Posttreatment variables are usually **not randomized** - no comparability across different values of these variables.

◦ Treatment effect estimates based on these analyses may not represent causal effects of treatments (Rosenbaum on net treatment effect, 1984; Holland on path analysis, 1988).

Potential Outcomes Approach (Rubin Causal Model)

• What it does: Based on intermediating posttreatment variables (S), formulate new variables that are **unaffected by treatment assignment** - just like pretreatment covariates.

• A new variable consists of categories of individuals based on observed intermediating posttreatment variables - stratification (just like stratifying individuals into male and female categories).

• Each category is defined based on potential values of S under all treatment conditions that are compared.

For example, if there are three treatment conditions ($Z = 0$ for the control, $Z = 1$ for treatment A, $Z = 2$ for treatment B), individual i will be classified into a category based on $(S_i(0), S_i(1), S_i(2))$. Here, $S_i(0)$ is the potential value of S if assigned to the control, $S_i(1)$ if assigned to treatment A, and $S_i(2)$ if assigned to treatment B condition.

Principal Stratification

• Each category defined based on potential values of S $(S_i(0), S_i(1), S_i(2))$ is unaffected by treatment - principal stratum.

• In each principal stratum, the outcome of interest (Y) can be compared across treatment conditions - principal effect. Any principal effect is a causal effect.

• The same principle applies to all possible categories (principal strata).

• Stratifying individuals into principal strata - principal stratification.

• Principal strata form a variable with multiple (or even continuous) categories. This variable is in general unobserved (latent, potential, missing) - just like a latent class variable. So let's call it C.

Exercise I. Angrist, Imbens, & Rubin Model

• Angrist, Imbens, & Rubin (1996, JASA).

• Interest in estimating treatment effects conditioning on treatment receipt status, particularly, treatment effects on individuals who would receive the treatment only if assigned to the treatment condition - complier average causal effect estimate (CACE).

• Analyses conditioning on treatment receipt status (which is a posttreatment variabe), in general, would not yield causal effect estimates.

• They used the instrumental variable (IV) approach. The difference from the conventional application of the IV approach is that they defined causal effects at individual level and that they clarified underlying assumptions of the IV approach.

• The more general principal stratification approach has its seeds in this approach.

Exercise I. Angrist, Imbens, & Rubin Model

• Treatment assignment (Z) has two levls ($Z_i = 1$ if individual i is assigned to the treatment, $Z_i = 0$ if assigned to the control condition).

• Treatment receipt (S) has two levls ($S_i = 1$ if individual i receives the treatment, $S_i = 0$ otherwise).

• 4 compliance types are defined based on potential treatment receipt status S_i given treatment assignment Z. $S_i(1)$ if $Z = 1$, and $S_i(0)$ if $Z = 0$.

$$C_i = \begin{cases} c\,(\text{complier}) & \text{if } S_i(1) = 1, \text{ and } S_i(0) = 0 \\ n\,(\text{never-taker}) & \text{if } S_i(1) = 0, \text{ and } S_i(0) = 0 \\ d\,(\text{defier}) & \text{if } S_i(1) = 0, \text{ and } S_i(0) = 1 \\ a\,(\text{always-taker}) & \text{if } S_i(1) = 1, \text{ and } S_i(0) = 1. \end{cases}$$

Exercise II. When S has Three Levels

• Treatment assignment (Z) has two levls ($Z_i = 1$ if individual i is assigned to the treatment A, $Z_i = 2$ if assigned to the treatment B condition). E.g., experimental treatment vs. standard treatment.

• Let's assume that it is not possible to take both treatments.

• Treatment receipt (S) has three levls ($S_i = 0$ if individual i does not receive any treatment, $S_i = 1$ if individual i receives treatment A, $S_i = 2$ if individual i receives treatment B).

Exercise II. When S has Three Levels

• 9 compliance types are defined based on potential treatment receipt status S_i given treatment assignment Z. $S_i(1)$ if $Z = 1$, and $S_i(2)$ if $Z = 2$.

$$C_i = \begin{cases} \text{(never-taker)} & \text{if } S_i(1) = 0, \text{ and } S_i(2) = 0 \\ \text{(defier I)} & \text{if } S_i(1) = 0, \text{ and } S_i(2) = 1 \\ \text{(B complier)} & \text{if } S_i(1) = 0, \text{ and } S_i(2) = 2 \\ \text{(A complier)} & \text{if } S_i(1) = 1, \text{ and } S_i(2) = 0 \\ \text{(A always-taker)} & \text{if } S_i(1) = 1, \text{ and } S_i(2) = 1 \\ \text{(always-complier)} & \text{if } S_i(1) = 1, \text{ and } S_i(2) = 2 \\ \text{(defier II)} & \text{if } S_i(1) = 2, \text{ and } S_i(2) = 0 \\ \text{(defier III)} & \text{if } S_i(1) = 2, \text{ and } S_i(2) = 1 \\ \text{(B always-taker)} & \text{if } S_i(1) = 2, \text{ and } S_i(2) = 2. \end{cases}$$

• Principal startification makes it clear what kind of problem you have.

Exercise III. Mediation

• Treatment assignment (Z) has two levels ($Z_i = 1$ if individual i is assigned to the treatment, $Z_i = 0$ if assigned to the control condition). At the baseline all subjects are smokers.

• Treatment (smoking cessation program) targets smoking behavior, which will improver general health condition (Y).

• Smoking (S) has two levls ($S_i = 1$ if individual i quits smoking, $S_i = 0$ does not quit smoking).

• 4 smoking behavior improvement types are defined based on potential smoking status S_i given treatment assignment Z. $S_i(1)$ if $Z = 1$, and $S_i(0)$ if $Z = 0$.

$$C_i = \begin{cases} c\,(\text{forward-improver}) & \text{if } S_i(1) = 1, \text{ and } S_i(0) = 0 \\ n\,(\text{never-improver}) & \text{if } S_i(1) = 0, \text{ and } S_i(0) = 0 \\ d\,(\text{backward-improver}) & \text{if } S_i(1) = 0, \text{ and } S_i(0) = 1 \\ a\,(\text{always-improver}) & \text{if } S_i(1) = 1, \text{ and } S_i(0) = 1. \end{cases}$$

Identification of Causal Effects: AIR (1996) Model

• Principal stratification shows that there can be 4 causal effects defined for 4 principal strata (types of people).

• Let $C(t) = \{i \mid C_i = t\}$ for $t \in \{c, n, d, a\}$. The average causal effect of treatment assignment given C can be defined at the individual level as

$$ITT_t = \sum_{i \in C(t)} [Y_i(1, S_i(1)) - Y_i(0, S_i(0))]/N_t,$$

where $Y_i(1, S_i(1))$ denotes the potential outcome with S_i if $Z = 1$, and $Y_i(0, S_i(0))$ denotes the potential outcome with S_i if $Z = 0$. N_t is the number of individuals with $C(t)$.

Identification of Causal Effects: AIR (1996) Model

• Principal stratification shows there can be 4 causal effects defined for 4 principal strata (types of people).

• To do that, we need to identify 8 means (4 types in 2 conditions). However, we only observe 2 means (mean of the control, mean of the treatment condition individuals).

• Under monotonicity & exclusion restriction (in addition to randomization and SUTVA), the average causal effect of treatment assignment for compliers ($CACE$) can be identified.

$\Rightarrow$ Monotonicity: There are no defiers.

$\Rightarrow$ Exclusion restriction (ER): The effect of treatment is disallowed for never-takers and always-takers (allowed only for compliers). This assumption precludes the possibility of having direct effects of treatment assignment on outcomes without going through actual treatment receipt.

Identification of Causal Effects: AIR (1996) Model

• Monotonicity plays a critical role in identifying CACE.

• We can identify never-takers under $Z = 1$ because they are only type of people who would not receive the treatment $\Rightarrow \mu_{n1}$ identified.

• We can identify always-takers under $Z = 0$ because they are only type of people who would receive the treatment $\Rightarrow \mu_{a0}$ identified.

Note that, now we have

$$C_i = \begin{cases} c\,(\text{complier}) & \text{if } S_i(1) = 1, \text{ and } S_i(0) = 0 \\ n\,(\text{never-taker}) & \text{if } S_i(1) = 0, \text{ and } S_i(0) = 0 \\ a\,(\text{always-taker}) & \text{if } S_i(1) = 1, \text{ and } S_i(0) = 1. \end{cases}$$

Identification of Causal Effects: AIR (1996) Model

- Under monotonicity, there are no defiers - now only 6 means to identify.
- Assuming monotonicity, μ_{a0} and μ_{n1} are directly estimable from the data.
- Under ER, the effect of treatment is disallowed for never-takers and always-takers. In other words,

$$\begin{aligned} ITT_n &= \mu_{n1} - \mu_{n0} = 0, \\ ITT_a &= \mu_{a1} - \mu_{a0} = 0, \end{aligned}$$

where μ_{n1} denotes population mean potential outcome for never-takers if $Z = 1$, and μ_{n0} if $Z = 0$. μ_{a1} denotes population mean potential outcome for always-takers if $Z = 1$, and μ_{a0} if $Z = 0$.

- Based on monotonicity and ER, μ_{n1}, μ_{n0}, μ_{a0} and μ_{a1} are all identified.

Identification of Causal Effects: AIR (1996) Model

• Now, only two more means to identify (μ_{c1} and μ_{c0}) with two observable means (μ_1 and μ_0). Here, μ_{c1} denotes population mean potential outcome for compliers if $Z = 1$, and μ_{c0} if $Z = 0$. μ_1 denotes population mean potential outcome if $Z = 1$, and μ_0 if $Z = 0$.

• μ_{c1} is identifiable from $\mu_1 = \pi_c \, \mu_{c1} + \pi_n \, \mu_{n1} + \pi_a \, \mu_{a1}$.

• μ_{c0} is identifiable from $\mu_0 = \pi_c \, \mu_{c0} + \pi_n \, \mu_{n0} + \pi_a \, \mu_{a0}$.

Note that π_c, π_n, and π_a are population proportions of compliers, never-takers, and always-takers, which have corresponding sample proportions (directly estimable).

• CACE is defined as

$$\text{CACE} = ITT_c = \mu_{c1} - \mu_{c0}.$$

• The CACE estimate assuming monotonicity and ER can be biased if these assumptions are violated.

References

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association, 91*, 444-455.

Frangakis, C. E. & Rubin, D. B. (2002) Principal stratification in causal inference. *Biometrics, 58*, 21-29.

Compliance and Latent Compliance

• Noncompliance is common and difficult to control in randomized trials involving human participants.

• Depending on how noncompliance is handled in estimation approaches, different conclusions may be reached about the effect of the same treatment.

• Compliance status of individuals is not always observable (e.g., no treatment control, placebo control). Latent (missing) compliance status makes it difficult to estimate treatment effects taking into account compliance.

$\Rightarrow$ When the major interest is in estimating overall effectiveness of treatment, noncompliance is less problematic in the estimation of treatment effect (however, see Frangakis & Rubin, 1999).

$\Rightarrow$ When the major interest is in estimating efficacy of treatment, noncompliance is more problematic in the estimation of treatment effect.

Job Search Intervention Study (JOBS II)

(Vinokur, Price, & Schul, 1995)

• Designed to prevent poor mental health and to promote high quality reemployment. Conducted by Michigan Prevention Research Center. Study participants were randomly assigned to control/intervention conditions.

• Intervention condition consisted of 5 training sessions that included the application of decision-making processes, and job search skills. Control condition consisted of a booklet briefly describing job search methods and tips. **Therefore, compliance with the intervention treatments could not be observed among individuals assigned to the control condition.**

• A substantial proportion of individuals who were assigned to the intervention condition did not show up to the intervention (Compliance = 55%).

• The two major outcomes were employment and depression mainly caused by unemployment.

Johns Hopkins School Intervention Study
(Ialongo et al., 1999)

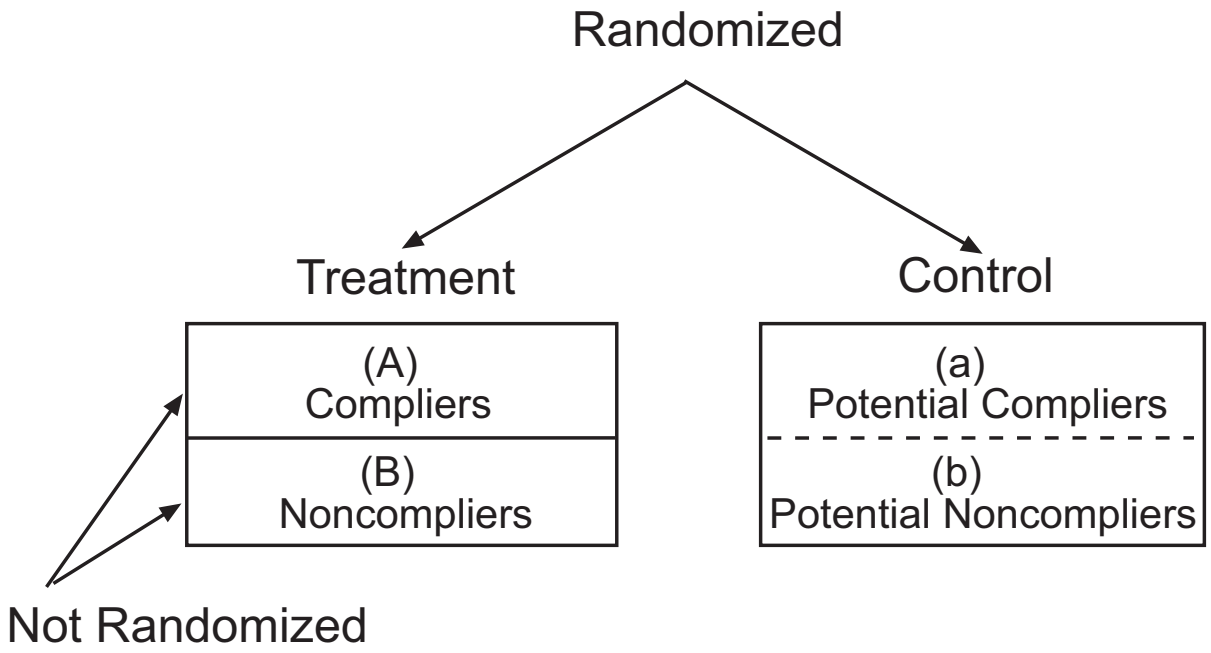
• Designed to improve academic achievement and to reduce early behavioral problems of school children. Teachers and first-grade children were randomly assigned to control/intervention conditions.

• In Family-School Partnership Intervention condition, parents were asked to implement 66 take-home activities related to literacy and math over a six-month period. Nothing was offered to control condition. **Therefore, compliance with the intervention treatments could not be observed among individuals assigned to the control condition**.

• A large variation in completed activities (ranges 0 to 66), and over-reporting of compliance is also expected. The intervention may not show any desirable effects unless parents report a quite high level of compliance.

• Categorizing individuals into low and high compliers will provide a more meaningful intervention effect estimate than categorizing them into never-takers and compliers (97% reported they completed at least one activity).

Options in Dealing with Noncompliance



- Intent-to-Treat (ITT) Analysis: (A+B) vs. (a+b)
- As-Treated Analysis: (A) vs. (B+a+b)
- Per-Protocol Analysis: (A) vs. (a+b)
- CACE (Complier Average Causal Effect): (A) vs. (a)

Common Setting

- Randomized trials, where successful placebo control is unlikely.
- 2 conditions: intervention ($Z = 1$) and control ($Z = 0$)
- 2 compliance types (C_i)

  1) complier (c) - receives the intervention treatment if assigned, and does not if not assigned. $\pi_c =$ compliance rate.

  2) noncomplier (n) - does not receive the intervention treatment even if assigned to receive it. $1 - \pi_c = \pi_n =$ noncompliance rate.
- 2 observed average outcomes in $Z = 1$: μ_{c1} and μ_{n1}.
- 2 unobserved average outcomes in $Z = 0$: μ_{c0} and μ_{n0}.
- 1 observed outcome in $Z = 0$: μ_0 $(= \pi_c \, \mu_{c0} + \pi_n \, \mu_{n0})$.
- The estimators of interest are

$$ITT = \mu_1 - \mu_0 = \pi_c \, (\mu_{c1} - \mu_{c0}) + (1 - \pi_c) \, (\mu_{n1} - \mu_{n0}).$$

$$CACE = ITT_c = \mu_{c1} - \mu_{c0}.$$

JHU PIRC Study: N=284 (listwise deletion)

- Completed at least 45 activities - compliers.
- Outcome: change score (baseline - followup) of anti-social behavior .

$\hat{\mu}_0$	$\hat{\mu}_{c1}$	$\hat{\mu}_{n1}$	μ_1	$\hat{\pi}_c$
-0.319 (1.383)	-0.177 (1.214)	0.248 (1.271)	0.045 (1.259)	0.479

- The ITT estimate is

$$\widehat{ITT} = \hat{\mu}_1 - \hat{\mu}_0 = 0.045 - (-0.319) = 0.364.$$

- From $\mu_0 = \pi_c\,\mu_{c0} + \pi_n\,\mu_{n0}$ and ER ($\hat{\mu}_{n0} = \hat{\mu}_{n1}$), $\hat{\mu}_{c0} = \frac{\hat{\mu}_0 - \hat{\mu}_{n1}\,(1-\hat{\pi}_c)}{\hat{\pi}_c}$.
- From $\mu_1 = \pi_c\,\mu_{c1} + \pi_n\,\mu_{n1}$, we get $\hat{\mu}_{c1} = \frac{\hat{\mu}_1 - \hat{\mu}_{n1}\,(1-\hat{\pi}_c)}{\hat{\pi}_c}$.

- The $CACE$ estimate is

$$\widehat{CACE} = \hat{\mu}_{c1} - \hat{\mu}_{c0} = \frac{\hat{\mu}_1 - \hat{\mu}_0}{\hat{\pi}_c} = 0.364/0.457 = 0.760.$$

Plausibility of ER

• The ER assumption can be unrealistic in practice.

⇒ Being assigned to the treatment condition may have a positive (feeling lucky, taken care of, optimistic?) or a negative (demoralized?) psychological effect on never-takers. This effect should be considered especially when trials are not placebo controlled. E.g., JOBS II

⇒ Always-takers may receive a treatment only partially when assigned to the control condition (difficulty of access, lack of supervision).

⇒ When treatment condition includes several doses or sessions, one may want to estimate differential treatment effects for low and high compliers. Potential low-compliers will not receive a treatment at all if assigned to the control condition. E.g., JHU Study

• ER is more likely to hold in trials where double blinding or blinding is successfully implemented.

• In other situations, it is often difficult to gauge scientific plausibility of ER (guess? previous studies? expert opinion?).

Co-presence of noncompliance and nonresponse

• Randomized trials involving human participants often suffer from both intervention noncompliance (no-show) and nonresponse (missing outcome information) at followup assessments.

• Individuals' compliance and outcome information is available only under the condition they were assigned to, and $unavailable$ under other conditions - these values can be considered as latent, missing, or potential values.

• However, based on random assignemnt, the causal inference is possible at the average level - Intention to treat analsyis as gold standard.

• ITT analysis is a reasonable method for estimating overall effectiveness of the intervention programs in the presence of noncompliance (since we are not particularly interested in differential intervention effects for individuals with different compliance types).

• However, robustness of ITT analysis can be threatened in the co-presence of noncompliance and nonresponse (missing outcomes).

Common Setting

- Randomized trials, where successful placebo control is unlikely.
- 2 conditions: intervention ($Z = 1$) and control ($Z = 0$)
- 2 compliance types (C_i)

  1) complier (c) - receives the intervention treatment if assigned, and does not if not assigned. $\pi_c =$ compliance rate.

  2) noncomplier (n) - does not receive the intervention treatment even if assigned to receive it. $1 - \pi_c = \pi_n =$ noncompliance rate.
- 2 observed average outcomes in $Z = 1$: μ_{c1} and μ_{n1}.
- 2 observed average responses in $Z = 1$: π^R_{c1} and π^R_{n1}.
- 2 unobserved average outcomes in $Z = 0$: μ_{c0} and μ_{n0}.
- 2 unobserved average responses in $Z = 0$: π^R_{c0} and π^R_{n0}.
- 1 observed outcome and 1 observed response rate in $Z = 0$: μ_0 and π^R_0.

Common Setting

- The target estimators of ITT and $CACE$ are defined as

$$ITT = \pi_c\,(\mu_{c1} - \mu_{c0}) + (1 - \pi_c)\,(\mu_{n1} - \mu_{n0}) \neq \mu_1^{obs} - \mu_0^{obs}, \tag{1}$$

$$CACE = \mu_{c1} - \mu_{c0}. \tag{2}$$

- Observed average outcome in the control condition can be written as

$$\mu_0^{obs} = \frac{\pi_{c0}^R}{\pi_0^R}\,\pi_c\,\mu_{c0} + \frac{\pi_{n0}^R}{\pi_0^R}\,(1 - \pi_c)\,\mu_{n0}. \tag{3}$$

- The average response rate in the control condition can be written as

$$\pi_0^R = \pi_{c0}^R\,\pi_c \,+\, \pi_{n0}^R\,(1 - \pi_c). \tag{4}$$

- The μ_{c0} can be derived from (3) as

$$\mu_{c0} = \frac{\mu_0^{obs}\,\pi_0^R - \mu_{n0}\,\pi_{n0}^R\,(1 - \pi_c)}{\pi_0^R - \pi_{n0}^R\,(1 - \pi_c)}, \tag{5}$$

where μ_{n0} and π_{n0}^R do not have corresponding sample statistics.

Some possible identifying assumptions for μ_{n0}

• Outcome exclusion restriction (OER): For never-takers, the distributions of the potential outcomes are independent of treatment assignment status (AIR, 1996) In the current setting, this implies that $\mu_{n1} = \mu_{n0}$. That is, never-taker average causal effect (NACE) = 0.

• Baseline restriction (BR): The distributions of the potential outcomes are unaffected by unobserved compliance status conditional on treatment assignment and observed treatment receipt status. In the current setting, this implies that $\mu_{c0} = \mu_{n0}$.

• Average effect restriction (AER): The distributions of the potential causal effects are independent of the compliance status. In the current setting, this implies that $CACE\,(\mu_{c1} - \mu_{c0}) = NACE\,(\mu_{n1} - \mu_{n0})$.

Some possible identifying assumptions for π_{n0}^R

◦ Missing at random (MAR): The probability of outcome being recorded is unassociated with the outcome conditional on treatment assignment and observed treatment receipt status. In the current setting, this implies that $\pi_{n0}^R = \pi_{c0}^R$.

◦ Response exclusion restriction (RER): For never-takers, response behavior is unaffected by treatment assignment status (Frangakis & Rubin, 1999). In the current setting, this implies that $\pi_{n0}^R = \pi_{n1}^R$.

◦ Stable complier response (SCR): For compliers, response behavior is unaffected by treatment assignment status. In the current setting, this implies that $\pi_{c0}^R = \pi_{c1}^R$.

JHU PIRC Study: N=440

• Outcome: change score (baseline - followup) of anti-social behavior .

$\hat{\mu}_0^{obs}$	$\hat{\mu}_{c1}$	$\hat{\mu}_{n1}$	$\hat{\pi}_0^R$	$\hat{\pi}_{c1}^R$	$\hat{\pi}_{n1}^R$	$\hat{\pi}_c$
-0.319	-0.177	0.248	0.781	0.911	0.833	0.457

• From (2) and (5), based on OER and MAR, a point estimator of $CACE$ is established as

$$CACE^{MAR.OER} = \mu_{1,1} - \left\{ \frac{\mu_0^{obs} - \mu_{0,1}(1-\pi_c)}{\pi_c} \right\}.$$

• From (1) and (5), based on OER and MAR, a point estimator of ITT is established as

$$ITT^{MAR.OER} = \pi_c \left[\mu_{1,1} - \left\{ \frac{\mu_0^{obs} - \mu_{0,1}(1-\pi_c)}{\pi_c} \right\} \right].$$

JHU PIRC Study

• Based on OER and RER, a point estimator of $CACE$ is established as

$$CACE^{RER.OER} \;=\; \mu_{1,1} \;-\; \left\{ \frac{\mu_0^{obs}\,\pi_0^R - \mu_{0,1}\,\pi_{0,1}^R\,(1-\pi_c)}{\pi_0^R - \pi_{0,1}^R\,(1-\pi_c)} \right\},$$

where ER and RER are together called the "compound exclusion restriction" (Frangakis & Rubin, 1999)

• Based on OER and RER, a point estimator of ITT is established as

$$ITT^{RER.OER} \;=\; \pi_c \left[\mu_{1,1} \;-\; \left\{ \frac{\mu_0^{obs}\,\pi_0^R - \mu_{0,1}\,\pi_{0,1}^R\,(1-\pi_c)}{\pi_0^R - \pi_{0,1}^R\,(1-\pi_c)} \right\} \right].$$

• Based on sample statistics, $\widehat{CACE}^{MAR.OER} = 0.816$ and $\widehat{CACE}^{RER.OER} = 0.923$, and $\widehat{CACE}_{obs}^{OER} = 0.760$.

• Based on sample statistics, $\widehat{ITT}^{MAR.OER} = 0.373$ and $\widehat{ITT}^{RER.OER} = 0.422$, and $\widehat{ITT}_{obs}^{OER} = 0.364$.

References

Frangakis, C. E. & Rubin, D. B. (1999). Addressing complications of intention-to-treat analysis in the presence of all-or-none treatment-noncompliance and subsequent missing outcomes. *Biometrika, 86*, 365-379.

Homework

• Establish 3 point estimators of $CACE$ assuming RER/AER, MAR/BR, and SCR/AER.

• Estimate $CACE$ based on these point estimators and sample statistics given on p.29.