Augmented Designs to Assess Principal Strata Causal Effects

Alessandra Mattei and Fabrizia Mealli

Department of Statistics "G. Parenti" – University of Florence

February 26, 2010

Motivation

- Many research questions involving causal inference are often concerned with understanding the causal pathways by which an exposure or treatment affects an outcome
  - Example: Drug treatment having side-effects (Pearl, 2001)
- Disentangling direct and indirect effects may be a difficult task, because the intermediate outcome is generally not under experimental control
- Traditional analyses are typically based on standard methods, resulting in estimates that generally lack causal interpretation
- Estimands are not clearly defined and the assumptions needed for a causal interpretation of the estimates are not always made explicit

Approaches to Casual Inference

- Decision theoretic approach (Dawid 2000, 2002);
- Causal graph or structural models framework (Pearl 2000);
- Potential outcomes framework (Neyman, 1923; Rubin 1974, 1977, 1978)

We focus on the potential outcomes framework, and use the concept of the principal stratification (Frangakis and Rubin, 2002) for addressing the topic of direct and indirect causal effects

Outline

- Theoretical background
- Augmented Randomized Design
- Partial Identification of Principal Strata Direct causal Effects
- An Illustrative Example
- Augmented Designs versus Standard Designs
- Concluding Remarks

Notation

- Units: $i = 1, 2, \ldots, n$
- Treatment variable: $Z_i \in \{C, T\}$
- Outcomes Variables
  - Intermediate variable: $S_i \in \{0, 1\}$
  - Primary Outcome: Y_i
- Potential outcomes for any given variable comprise the observable manifestation of this variable under each of the possible treatment assignments:

$$(S_i(C), S_i(T)) \quad \text{and} \quad (Y_i(C), Y_i(T)) \quad i = 1, \ldots, n$$

- $Y_i(z, s)$ = (a priori) counterfactual value for Y if, possibly contrary to fact, Z was set to z and S was set to s.
- Note that we assume the Stable Unit Treatment Value Assumption (SUTVA; Rubin, 1978, 1980, 1990).

A priori Direct and Indirect Causal Effects

(Robins and Greenland, 1992; Pearl, 2001)

- Average Controlled Direct Effect (*CDE*)

$$CDE(s) = E[Y_i(T, s) - Y_i(C, s)] \quad s = 0, 1$$

- Average Natural Direct Effect (*NDE*)

$$NDE(z) = E[Y_i(T, S_i(z)) - Y_i(C, S_i(z))] \qquad z = C, T$$

- Average Natural Indirect Effect (*NIE*)

$$NIE(z) = E[Y_i(z, S_i(T)) - Y_i(z, S_i(C))] \qquad z = C, T$$

- Total Effect = Natural Direct Effect + Natural Indirect Effect

$$E[Y_i(T) - Y_i(C)] = NIE(T) + NDE(C) = NDE(T) + NIE(C)$$

Principal Stratification and Principal Stratum Causal Effects
(Frangakis and Rubin, 2002)

The basic **principal stratification** *with respect to a posttreatment variable S is the partition of subjects into sets such that all subjects in the same set have the same vector $(S_i(C); S_i(T))$*

Principal Causal Effect (PCE):

$$PCE = E[Y_i(T) - Y_i(C) | S_i(C) = s_C, S_i(T) = s_T]$$

Principal Stratum Direct Effect ($PSDE$)

$$PSDE(s) = E\left[Y_i(T) - Y_i(C) | S_i(T) = S_i(C) = s\right]$$

- If $PSDE(s) = 0$, for each $s = 0, 1$, then all effect is indirect
- Total Effect = Weighted average of PCEs

$$E[Y_i(T) - Y_i(C)] = \sum_{(s_C, s_T)} PCE(s_C, s_T)\pi_{s_C, s_T}$$

π_{s_C, s_T} = proportion of subjects with $S_i(C) = s_C$ and $S_i(T) = s_T$

Augmented Design: The Encouragement Variable

- *PCE* analysis is challenging, due to the latent nature of principal strata
- Augmented designs: the treatment is randomized, and the mediating variable is not forced, but only randomly encouraged
- Encouragement variable: variable affecting the primary outcome, only through its effect on the intermediate outcome
- The encouragement variable is treated as an additional treatment
- Encouragement variable

$$W_i = \begin{cases} E & \text{if unit } i \text{ is encouraged;} \\ e, & \text{if unit } i \text{ is not encouraged.} \end{cases}$$

Augmented Design: SUTVA and Potential Outcomes

Assumption 1. Stable Unit Treatment Value Assumption (SUTVA)

Let $\mathbf{Z} = (Z_1, \ldots, Z_n)'$ and $\mathbf{W} = (W_1, \ldots, W_n)'$.

$$Z_i = Z_i' \text{ and } W_i = W_i' \Longrightarrow \begin{cases} S_i(\mathbf{Z}, \mathbf{W}) = S_i(\mathbf{Z}', \mathbf{W}'); \\ Y_i(\mathbf{Z}, \mathbf{W}) = Y_i(\mathbf{Z}', \mathbf{W}') \end{cases}$$

Intermediate Potential Outcomes:

$$S_i(C, e), S_i(C, E), S_i(T, e), S_i(T, E)$$

Response Potential Outcomes

$$Y_i(C, e), Y_i(C, E), Y_i(T, e), Y_i(T, E)$$

Augmented Design: Principal Stratification

The basic principal stratification P_0 with respect to posttreatment variable S is the partition of units $i = 1, \ldots, n$ such that, all units within any set of P_0, all units have the same vector $(S_i(C,e), S_i(C,E), S_i(T,e), S_i(T,E))$.

A principal stratification P with respect to posttreatment variable S is a partition of the units whose sets are unions of sets in the basic principal stratification P_0.

Augmented Design: Principal Strata with Two Binary Treatments and a Binary Intermediate Variable

G_i	$S_i(C,e)$	$S_i(C,E)$	$S_i(T,e)$	$S_i(T,E)$	G_i	$S_i(C,e)$	$S_i(C,E)$	$S_i(T,e)$	$S_i(T,E)$
1	0	0	0	0	**9**	0	1	1	0
2	1	0	0	0	**10**	0	1	0	1
3	0	1	0	0	**11**	0	0	1	1
4	0	0	1	0	**12**	1	1	1	0
5	0	0	0	1	**13**	1	1	0	1
6	1	1	0	0	**14**	1	0	1	1
7	1	0	1	0	**15**	0	1	1	1
8	1	0	0	1	**16**	1	1	1	1

$\pi_g =$ Proportion of units belonging to principal stratum $g, g = 1, \ldots, 16$

Augmented Design: Principal Stratum Direct Effects

The average Principal Stratum Direct Effect of Z on Y at level s, $s \in \{0, 1\}$, denoted $PSDE(s; w, w')$, is defined as

$$PSDE(s; w, w') = E\left[Y_i(T, w') - Y_i(C, w) | S_i(T, w') = S_i(C, w) = s\right] \quad w, w' \in \{e, E\}.$$

Note that,

$$\begin{aligned} PSDE(s; w, w') &= E\left[Y_i(T, w') - Y_i(C, w) | S_i(T, w') = S_i(C, w) = s\right] \\ &= E\Big[E\left[Y_i(T, w') - Y_i(C, w) | S_i(T, w') = S_i(C, w) = s, S_i(T, w), S_i(C, w')\right]\Big] \\ &= E\left[PSDE(s; w, w') | S_i(T, w), S_i(C, w')\right] \equiv E\left[PSDE_{G_{s,w,w'}}(s; w, w')\right] \end{aligned}$$

$G_{s,w,w'}$: principal stratum comprising units with $S_i(T, w') = S_i(C, w) = s$

Augmented Design: Principal Stratum Direct Effects

$$
\begin{aligned}
PSDE(0; e, e) &= E\left[Y_i(T, e) - Y_i(C, e) | S_i(T, e) = S_i(C, e) = 0\right] \\
PSDE(0; e, E) &= E\left[Y_i(T, E) - Y_i(C, e) | S_i(T, E) = S_i(C, e) = 0\right] \\
PSDE(0; E, e) &= E\left[Y_i(T, e) - Y_i(C, E) | S_i(T, e) = S_i(C, E) = 0\right] \\
PSDE(0; E, E) &= E\left[Y_i(T, E) - Y_i(C, E) | S_i(T, E) = S_i(C, E) = 0\right]
\end{aligned}
$$

and

$$
\begin{aligned}
PSDE(1; e, e) &= E\left[Y_i(T, e) - Y_i(C, e) | S_i(T, e) = S_i(C, e) = 1\right] \\
PSDE(1; e, E) &= E\left[Y_i(T, E) - Y_i(C, e) | S_i(T, E) = S_i(C, e) = 1\right] \\
PSDE(1; E, e) &= E\left[Y_i(T, e) - Y_i(C, E) | S_i(T, e) = S_i(C, E) = 1\right] \\
PSDE(1; E, E) &= E\left[Y_i(T, E) - Y_i(C, E) | S_i(T, E) = S_i(C, E) = 1\right]
\end{aligned}
$$

Principal Strata and 'Basic' *PSDE*s

	Principal Stratum				$PSDE_{G_{s,w,w'}}(s;w,w')$			
1:	0	0	0	0	$PSDE_1(0;e,e)$	$PSDE_1(0;E,E)$	$PSDE_1(0;e,E)$	$PSDE_1(0;E,e)$
2:	1	0	0	0		$PSDE_2(0;E,E)$		$PSDE_2(0;E,e)$
3:	0	1	0	0	$PSDE_3(0;e,e)$		$PSDE_3(0;e,E)$	
4:	0	0	1	0		$PSDE_4(0;E,E)$	$PSDE_4(0;e,E)$	
5:	0	0	0	1	$PSDE_5(0;e,e)$			$PSDE_5(0;E,e)$
6:	1	1	0	0				
7:	1	0	1	0	$PSDE_7(1;e,e)$	$PSDE_7(0;E,E)$		
8:	1	0	0	1			$PSDE_8(1;e,E)$	$PSDE_8(0;E,e)$
9:	0	1	1	0			$PSDE_9(0;e,E)$	$PSDE_9(1;E,e)$
10:	0	1	0	1	$PSDE_{10}(0;e,e)$	$PSDE_{10}(1;E,E)$		
11:	0	0	1	1				
12:	1	1	1	0	$PSDE_{12}(1;e,e)$			$PSDE_{12}(1;E,e)$
13:	1	1	0	1		$PSDE_{13}(1;E,E)$	$PSDE_{13}(1;e,E)$	
14:	1	0	1	1	$PSDE_{14}(1;e,e)$		$PSDE_{14}(1;e,E)$	
15:	0	1	1	1		$PSDE_{15}(1;E,E)$		$PSDE_{15}(1;E,e)$
16:	1	1	1	1	$PSDE_{16}(1;e,e)$	$PSDE_{16}(1;E,E)$	$PSDE_{16}(1;e,E)$	$PSDE_{16}(1;E,e)$

Observed Data

For $i = 1, \ldots, n$

$$\begin{aligned} Z_i^{\text{obs}} &= \text{Observed treatment assignment} \\ W_i^{\text{obs}} &= \text{Observed encouragement assignment} \end{aligned}$$

and

$$\begin{aligned} S_i^{\text{obs}}\left(Z_i^{\text{obs}}, W_i^{\text{obs}}\right) &= \text{Actual intermediate outcome} \\ Y_i^{\text{obs}}\left(Z_i^{\text{obs}}, W_i^{\text{obs}}\right) &= \text{Actual primary outcome} \end{aligned}$$

$$\Downarrow$$

$$Z_i^{\text{obs}}, W_i^{\text{obs}}, S_i^{\text{obs}}, Y_i^{\text{obs}}$$

Group Classification based on Observed Data $(Z_i^{\text{obs}}, W_i^{\text{obs}}, S_i^{\text{obs}})$ and Associated Possible Latent Principal Strata

Observed Group			Latent Strata							
Z_i^{obs}	W_i^{obs}	S_i^{obs}	G_i							
C	e	0	1	3	4	5	9	10	11	15
C	e	1	2	6	7	8	12	13	14	16
C	E	0	1	2	4	5	7	8	11	14
C	E	1	3	6	9	10	12	13	15	16
T	e	0	1	2	3	5	6	8	10	13
T	e	1	4	7	9	11	12	14	15	16
T	E	0	1	2	3	4	6	7	9	12
T	E	1	5	8	10	11	13	14	15	16

Structural Assumptions

Assumption 2. Randomization of the Treatment and the Encouragement

For all i,

$$\left(\begin{array}{c} S_i(C,e), S_i(C,E), S_i(T,e), S_i(T,E) \\ Y_i(C,e), Y_i(C,E), Y_i(T,e), Y_i(T,E) \end{array}\right) \perp\!\!\!\perp (Z_i, W_i)$$

Randomization implies that

$$\Big(Y_i(C,e), Y_i(C,E), Y_i(T,e), Y_i(T,E)\Big) \perp\!\!\!\perp (Z_i, W_i) \Big| \Big(S_i(C,e), S_i(C,E), S_i(T,e), S_i(T,E)\Big)$$

Structural Assumptions

Assumption 3. Conditional Stochastic Exclusion Restrictions w.r.t. the Encouragement

For $z \neq z' \in \{C, T\}$,

$$\begin{aligned} & Pr\left(Y_i(z,E) \mid S_i(z,E) = S_i(z,e), S_i(z',E), S_i(z',e)\right) \\ = \; & Pr\left(Y_i(z,e) \mid S_i(z,E) = S_i(z,e), S_i(z',E), S_i(z',e)\right) \end{aligned}$$

This Assumption implies that

$$\begin{aligned} PSDE_{G_{s,w,e}}(s; w, e) &= PSDE_{G_{s,w,e}}(s; w, E) \\ PSDE_{G_{s,e,w'}}(s; e, w') &= PSDE_{G_{s,E,w'}}(s; E, w') \end{aligned}$$

for $w, w' \in \{e, E\}$ such that

$$\begin{aligned} G_{s,e,w'}(s; e, w') &= \{i : S_i(C,e) = S_i(C,E) = s, S_i(T,\tilde{w}) = s_{T\tilde{w}}, S_i(T,w') = s\}, \; \tilde{w} \neq w' \\ G_{s,w,e}(s; w, e) &= \{i : S_i(C,w) = s, S_i(C,\bar{w}) = s_{C\bar{w}}, S_i(T,e) = S_i(T,E) = s\}, \; \bar{w} \neq w \end{aligned}$$

Structural Assumptions

Assumption 4. Nonzero Average Causal Effect of W on S

The average causal effect of W on S

$$E[\left(S_i\left(z, E\right) - S_i\left(z', e\right)\right] \qquad \text{for } z, z' = C, T$$

is not equal to zero

Structural Assumptions

Assumption 5. Monotonicity of S with respect to W

For all i

$$(i) \quad S_i(C,e) \leq S_i(C,E) \quad \text{and} \quad S_i(T,e) \leq S_i(T,E)$$

or

$$(ii) \quad S_i(C,e) \geq S_i(C,E) \quad \text{and} \quad S_i(T,e) \geq S_i(T,E)$$

- Without loss of generality, let $S_i(C,e) \leq S_i(C,E)$ and $S_i(T,e) \leq S_i(T,E)$ for all i.

Principal Strata (Left Panel) and Observed Groups with Associated Possible Latent Principal Strata (Right Panel) under Assumption 5.

G_i	$S_i(C,e)$	$S_i(C,E)$	$S_i(T,e)$	$S_i(T,E)$
1	0	0	0	0
3	0	1	0	0
5	0	0	0	1
6	1	1	0	0
10	0	1	0	1
11	0	0	1	1
13	1	1	0	1
15	0	1	1	1
16	1	1	1	1

Z_i^{obs}	W_i^{obs}	S_i^{obs}	Latent Strata (G_i)					
C	e	0	1	3	5	10	11	15
C	e	1	6	13	16			
C	E	0	1	5	11			
C	E	1	3	6	10	13	15	16
T	e	0	1	3	5	6	10	13
T	e	1	11	15	16			
T	E	0	1	3	6			
T	E	1	5	10	11	13	15	16

Structural Assumptions

Assumption 6. Monotonicity of S with respect to Z

$$S_i(C, e) \leq S_i(T, e) \qquad \text{and} \qquad S_i(C, E) \leq S_i(T, E).$$

Principal Strata (Left Panel) and Observed Groups with Associated Possible Latent Principal Strata (Right Panel) under Assumptions 5. and 6.

G_i	$S_i(C,e)$	$S_i(C,E)$	$S_i(T,e)$	$S_i(T,E)$
1	0	0	0	0
5	0	0	0	1
10	0	1	0	1
11	0	0	1	1
15	0	1	1	1
16	1	1	1	1

Z_i^{obs}	W_i^{obs}	S_i^{obs}	Latent Strata (G_i)				
C	e	0	1	5	10	11	15
C	e	1			16		
C	E	0		1	5	11	
C	E	1		10	15	16	
T	e	0		1	5	10	
T	e	1		11	15	16	
T	E	0			1		
T	E	1	5	10	11	15	16

Large Sample Bounds for Principal Stratum Proportions

Under Assumption 1. through 6., we have

$$\begin{aligned}
\pi_1 &= 1 - P_{1|TE} \\
\pi_{16} &= P_{1|Ce} \\
\pi_{10} &= P_{1|TE} - P_{1|Te} - \pi_5 \\
\pi_{11} &= P_{1|TE} - P_{1|CE} - \pi_5 \\
\pi_{15} &= \pi_5 + \left(P_{1|CE} - P_{1|Ce}\right) - \left(P_{1|TE} - P_{1|Te}\right)
\end{aligned}$$

and

$$\max\left\{0; \left(P_{1|TE} - P_{1|Te}\right) - \left(P_{1|CE} - P_{1|Ce}\right)\right\} \leq \pi_5 \leq \min\left\{\left(P_{1|TE} - P_{1|CE}\right); \left(P_{1|TE} - P_{1|Te}\right)\right\}$$

where

$$P_{s|z,w} = Pr\left(S_i^{\mathrm{obs}} = s | Z_i^{\mathrm{obs}} = z, W_i^{\mathrm{obs}} = w\right) \quad z = C, T \text{ and } w = e, E$$

Large Sample Bounds for Principal Stratum Direct Effects

- Under Assumption 1. through 6., we have derived large sample bounds for each *PSDE*
- Large Sample Bounds for $PSDE(1; E, E)$

$$E\left[Y_i^{\text{obs}} | Z_i^{\text{obs}} = T, W_i^{\text{obs}} = E, S_i^{\text{obs}} = 1, Y_i^{\text{obs}} \leq y_{TE1}^{\pi_{10,15,16|TE1}}\right] - E\left[Y_i^{\text{obs}} | Z_i^{\text{obs}} = C, W_i^{\text{obs}} = E, S_i^{\text{obs}} = 1, \right]$$

$$\leq PSDE(1, E, E) \leq$$

$$E\left[Y_i^{\text{obs}} | Z_i^{\text{obs}} = T, W_i^{\text{obs}} = E, S_i^{\text{obs}} = 1, Y_i^{\text{obs}} \geq y_{TE1}^{1-\pi_{10,15,16|TE1}}\right] - E\left[Z_i^{\text{obs}} = C, W_i^{\text{obs}} = E, S_i^{\text{obs}} = 1, Y_i^{\text{obs}}\right]$$

where

$$\pi_{10,15,16|TE1} = Pr\left(G_i \in \{10, 15, 16\} | Z_i^{\text{obs}} = T, W_i^{\text{obs}} = E, S_i^{\text{obs}} = 1\right)$$

and y_{TE1}^{α} is the α^{th} quantile of the outcome distribution in the observed group with $Z_i^{\text{obs}} = T, W_i^{\text{obs}} = E$ and $S_i^{\text{obs}} = 1$,
$\alpha = \pi_{10,15,16|TE1}, 1 - \pi_{10,15,16|TE1}$

An Illustrative Example: Full Hypothetical Data under Assumptions 1. through 6.

						Expected Values			
G_i	$S_i(C,e)$	$S_i(C,E)$	$S_i(T,e)$	$S_i(T,E)$	π_i	$Y_i(C,e)$	$Y_i(C,E)$	$Y_i(T,e)$	$Y_i(T,E)$
1	0	0	0	0	0.16	0.1	0.1	0.2	0.2
5	0	0	0	1	0.16	0.1	0.1	0.3	0.5
10	0	1	0	1	0.16	0.2	0.3	0.5	0.7
11	0	0	1	1	0.20	0.2	0.2	0.7	0.7
15	0	1	1	1	0.16	0.2	0.3	0.8	0.8
16	1	1	1	1	0.16	0.3	0.3	0.9	0.9

$$E[S_i(T,E)] = 0.84 \qquad E[S_i(C,E)] = 0.48$$
$$E[S_i(T,e)] = 0.52 \qquad E[S_i(C,e)] = 0.16$$

and

$$E[Y_i(T,E) - Y_i(C,E)] = 0.420 \quad \text{and} \quad E[Y_i(T,e) - Y_i(C,e)] = 0.388$$

An Illustrative Example: Principal Stratum Direct Effects

$PSDE_{G_{s,w,w'}}(s;w,w')$			$G_i = 1$	$G_i = 5$	$G_i = 10$	$G_i = 15$	$G_i = 16$	$PSDE(s;w,w')$
s	w	w'						Mean
0	e	e	0.1	0.2	0.3			0.20
0	E	E	0.1					0.10
0	e	E	0.1					0.10
0	E	e	0.1	0.2				0.15
1	e	e					0.6	0.60
1	e	E					0.6	0.60
1	E	e				0.5	0.6	0.55
1	E	E			0.4	0.5	0.6	0.50

An Illustrative Example: Summary Statistics of Hypothetical Observed Data

Z_i^{obs}	W_i^{obs}	S_i^{obs}	Observed Proportions	MEAN Rescue Medication Usage (S_i^{obs})	MEAN Disease Status (Y_i^{obs})
C	e		0.25	0.16	0.184
C	E		0.25	0.48	0.216
T	e		0.25	0.52	0.572
T	E		0.25	0.84	0.636
C	e	0	0.21	0	0.162
C	e	1	0.04	1	0.300
C	E	0	0.13	0	0.138
C	E	1	0.12	1	0.300
T	e	0	0.12	0	0.333
T	e	1	0.13	1	0.792
T	E	0	0.04	0	0.200
T	E	1	0.21	1	0.719

An Illustrative Example: Estimated Bounds

Principal Strata Proportions	Lower Bound	Upper Bound
π_1	0.16	
π_5	0.00	0.32
π_{10}	0.00	0.32
π_{11}	0.04	0.36
π_{15}	0.00	0.32
π_{16}	0.16	

Estimand $PSDE(s; w, w')$	Lower B.	Upper B.	Proportions of units with $S_i(T, w') = S_i(C, w) = s$ Lower B.	Upper B.
$PSDE(0; e, e)$	0.052	0.333	0.16	0.80
$PSDE(0; E, E)$=$PSDE(0; e, E)$	−0.239	0.200	0.16	
$PSDE(0; E, e)$	0.185	0.951	0.16	0.48
$PSDE(1; e, e)$=$PSDE(1; e, E)$	0.025	0.700	0.16	
$PSDE(1; E, e)$	0.475	1.000	0.16	0.48
$PSDE(1; E, E)$	0.208	0.700	0.16	0.80

An Illustrative Example: Standard Randomized Design

Full Hypothetical Data

				Expected Values		
$\widetilde{G}_i$	$\widetilde{S}_i(C)$	$\widetilde{S}_i(T)$	$\widetilde{\pi}_i$	$\widetilde{Y}_i(C)$	$\widetilde{Y}_i(T)$	$PSDE(s)$
1	0	0	0.48	0.13	0.33	0.2
2	0	1	0.36	0.20	0.74	
4	1	1	0.16	0.30	0.90	0.6

Observed Data

		Observed	Mean	
Z_i^{obs}	S_i^{obs}	Proportions	Rescue Medication Usage (S_i^{obs})	Disease Status(Y_i^{obs})
C		0.50	0.16	0.184
T		0.50	0.52	0.604
C	0	0.42	0	0.162
C	1	0.08	1	0.300
T	0	0.24	0	0.333
T	1	0.26	1	0.792

Estimated Bounds

Estimand	Estimate	Estimand	Lower Bound	Upper Bound
$\widetilde{\pi}_1$	0.48	$PSDE(0)$	0.051	0.333
$\widetilde{\pi}_2$	0.36	$PSDE(1)$	0.019	0.700
$\widetilde{\pi}_4$	0.16			

Augmented Designs versus Standard Designs
Estimands

Average (overall) Direct Effects in Standard Randomized Designs

$$ADE(s) = E\left[Y_i\left(T, S_i(T) = s\right)\right] - E\left[Y_i\left(C, S_i(C) = s\right)\right] \qquad s = 0, 1$$

Versus

Average (overall) Direct Effects in Augmented Randomized Designs

For $w = e, E$; $s = 0, 1$,

$$ADE(s; w) = E\left[Y_i(T, w, S_i(T, w) = s)\right] - E\left[Y_i(C, w, S_i(C, w) = s)\right]$$

Large Sample Bounds for $ADE(s)$

$$\left(E\left[Y_i^{\text{obs}}|Z_i^{\text{obs}}=T,S_i^{\text{obs}}=s\right]Pr\left(S_i^{\text{obs}}=s|Z_i^{\text{obs}}=T\right)+L_{Ts}Pr\left(S_i^{\text{obs}}=1-s|Z_i^{\text{obs}}=T\right)\right)$$
$$-\left(E\left[Y_i^{\text{obs}}|Z_i^{\text{obs}}=C,S_i^{\text{obs}}=s\right]Pr\left(S_i^{\text{obs}}=s|Z_i^{\text{obs}}=C\right)+U_{Cs}Pr\left(S_i^{\text{obs}}=1-s|Z_i^{\text{obs}}=C\right)\right)$$
$$\leq ADE(s) \leq$$
$$\left(E\left[Y_i^{\text{obs}}|Z_i^{\text{obs}}=T,S_i^{\text{obs}}=s\right]Pr\left(S_i^{\text{obs}}=s|Z_i^{\text{obs}}=T\right)+U_{Ts}Pr\left(S_i^{\text{obs}}=1-s|Z_i^{\text{obs}}=T\right)\right)$$
$$-\left(E\left[Y_i^{\text{obs}}|Z_i^{\text{obs}}=C,S_i^{\text{obs}}=s\right]Pr\left(S_i^{\text{obs}}=s|Z_i^{\text{obs}}=C\right)+L_{Cs}Pr\left(S_i^{\text{obs}}=1-s|Z_i^{\text{obs}}=C\right)\right)$$
$$\Downarrow$$

$$Bound\ width(s) =$$
$$(U_{Cs}-L_{Cs})\,Pr\left(S_i^{\text{obs}}=1-s|Z_i^{\text{obs}}=C\right)+(U_{Ts}-L_{Ts})\,Pr\left(S_i^{\text{obs}}=1-s|Z_i^{\text{obs}}=T\right)$$

Large Sample Bounds for $ADE(s; w)$

$$\Big(E\left[Y_i^{\text{obs}}|Z_i^{\text{obs}}=T, W_i^{\text{obs}}=w, S_i^{\text{obs}}=s\right]Pr\left(S_i^{\text{obs}}=s|Z_i^{\text{obs}}=T, W_i^{\text{obs}}=w\right)$$
$$+L_{Tws}Pr\left(S_i^{\text{obs}}=1-s|Z_i^{\text{obs}}=T, W_i^{\text{obs}}=w\right)\Big)$$
$$-\Big(E\left[Y_i^{\text{obs}}|Z_i^{\text{obs}}=C, W_i^{\text{obs}}=w, S_i^{\text{obs}}=s\right]Pr\left(S_i^{\text{obs}}=s|Z_i^{\text{obs}}=C, W_i^{\text{obs}}=w\right)$$
$$+U_{Cws}\cdot Pr\left(S_i^{\text{obs}}=1-s|Z_i^{\text{obs}}=C, W_i^{\text{obs}}=w\right)\Big)$$
$$\leq ADE(s; w) \leq$$
$$\Big(E\left[Y_i^{\text{obs}}|Z_i^{\text{obs}}=T, W_i^{\text{obs}}=w, S_i^{\text{obs}}=s\right]Pr\left(S_i^{\text{obs}}=s|Z_i^{\text{obs}}=T, W_i^{\text{obs}}=w\right)$$
$$+U_{Tws}Pr\left(S_i^{\text{obs}}=1-s|Z_i^{\text{obs}}=T, W_i^{\text{obs}}=w\right)\Big)$$
$$-\Big(E\left[Y_i^{\text{obs}}|Z_i^{\text{obs}}=C, W_i^{\text{obs}}=w, S_i^{\text{obs}}=s\right]Pr\left(S_i^{\text{obs}}=s|Z_i^{\text{obs}}=C, W_i^{\text{obs}}=w\right)$$
$$+L_{Cws}\cdot Pr\left(S_i^{\text{obs}}=1-s|Z_i^{\text{obs}}=C, W_i^{\text{obs}}=w\right)\Big)$$
$$\Downarrow$$
$$\begin{aligned} Bound\ width(s; w) &= (U_{Tws}-L_{Tws})\,Pr\left(S_i^{\text{obs}}=1-s|Z_i^{\text{obs}}=T, W_i^{\text{obs}}=w\right) \\ &+ (U_{Cws}-L_{Cws})\,Pr\left(S_i^{\text{obs}}=1-s|Z_i^{\text{obs}}=C, W_i^{\text{obs}}=w\right) \end{aligned}$$

Augmented Designs versus Standard Designs

Suppose that for fixed values of $W = w^*$ and $S = s^*$,

$$U_{zw^*s^*} = U_{zs^*} \quad \text{and} \quad L_{zw^*s^*} = L_{zs^*} \quad z = C, T.$$

Then

$$Pr\left(S_i^{\text{obs}} = 1 - s^* | Z_i^{\text{obs}} = z, W_i^{\text{obs}} = w^*\right) \leq Pr\left(S_i^{\text{obs}} = 1 - s^* | Z_i^{\text{obs}} = z\right), \; z = C, T$$

$$\Downarrow$$

$$width(s^*; w^*) \leq width(s^*)$$

- The benefits of our augmented design versus a standard randomized design depend on the role of the encouragement

Concluding Remarks

Discussion

- Our augmented randomized design may help to identify and estimate direct and indirect effects
- The presence of an encouragement variable allows us to derive large sample bounds for the causal estimands of interest, which are narrower than those we would derive in the absence of the encouragement variable

Directions for future research

- Investigating additional assumptions to achieve point-identification of *PSDE*s
- Developing (semi-)parametric (Bayesian) estimation strategies
- Using our augmented randomized design as template for mediation analysis in observational studies

References

Dawid, A. P. (2000). Causal inference without countefactuals (with discussion). *JASA* **95**, 407 – 448.
Dawid, A. P. (2002). Influence diagrams for causal modelling and inference. *International Statistical Review* **2**, 161 – 189.
Frangakis, C. and Rubin, D. B. (2002). Principal stratification in causal inference. *Biometrics* **58**, 21 – 29.
Neyman, J. (1923). On the application of probability theory to agricultural experiments: essay on principles, Section 9. Translated in *Statistical Science* **5**, 465 – 480, 1990.
Pearl, J. (2000). *Causality*. Cambridge: Cambridge University Press.
Pearl, J. (2001). Direct and indirect effects. *In proceedings of 7th Conference on Uncertainty in artificial intelligence* (eds. J. S. Breese & D. Koller), 411 – 420. Morgan Kaufman, San Francisco, CA.
Rubin, D. B. (1978). Bayesian inference for causal effects. *Annals of Statistics* **6**, 34 – 58.
Rubin, D. B. (1980). Comment on 'Randomization Analysis of Experimental Data: The Fisher Randomization Test' by D. Basu. *JASA* **75**, 59 1 – 593.
Rubin, D. B. (1990). 'Comment: Neyman (1923) and Causal Inference in Experiments and Observational Studies". *Statistical Science* **5**, 472 – 480.