An Introduction to the Theory of Causal (Total) Effects

Rolf Steyer

Friedrich Schiller Universität Jena
Institut für Psychology
Lehrstuhl für Methodenlehre und Evaluationsforschung

Mplus Users' Meeting, Utrecht

July 21, 2014

Introduction

Talking about causality in empirical statistical research we talk about:

- Regressions (i.e., conditional expectations) such as $E(Y \mid X)$, $E(Y \mid X, Z)$, or their representations by a path diagram, graph, etc.
- Omitted variables, other causal factors, etc.

We cannot understand and reasonably discuss causality unless the two kinds of variables mentioned above have a common mathematical representation.

This means: X, Y, Z, the regressions $E(Y \mid X)$ and $E(Y \mid X, Z)$, as well as the omitted variables are random variables on the same probability space. This space and its time structure, as well as the relations between the variables in our regressions on one side and the omitted variables on the other side need a mathematical representation.

Some Examples

Table 1: Joe and Ann With Self-Selection

Outcomes ω				Observables			Regressions					
Unit	Treatment	Success	$P(\{\omega\})$	Person variable U	Treatment variable X	Outcome variable Y	$E(Y\mid X,U) = P(Y{=}1\mid X,U)$	$E(Y\mid X) = P(Y{=}1\mid X)$	$E(X\mid U) = P(X{=}1\mid U)$	$E^{X=0}(Y\mid U) = \tau_0$	$E^{X=1}(Y\mid U) = \tau_1$	$E^{\overline{U}}(Y\mid X)$ Adjusted Regression
(*Joe*,	*no*,	−)	.144	*Joe*	0	0	.70	.60	.04	.70	.80	.45
(*Joe*,	*no*,	+)	.336	*Joe*	0	1	.70	.60	.04	.70	.80	.45
(*Joe*,	*yes*,	−)	.004	*Joe*	1	0	.80	.42	.04	.70	.80	.60
(*Joe*,	*yes*,	+)	.016	*Joe*	1	1	.80	.42	.04	.70	.80	.60
(*Ann*,	*no*,	−)	.096	*Ann*	0	0	.20	.60	.76	.20	.40	.45
(*Ann*,	*no*,	+)	.024	*Ann*	0	1	.20	.60	.76	.20	.40	.45
(*Ann*,	*yes*,	−)	.228	*Ann*	1	0	.40	.42	.76	.20	.40	.60
(*Ann*,	*yes*,	+)	.152	*Ann*	1	1	.40	.42	.76	.20	.40	.60

Table 2: Joe and Ann With Self-Selection – Compressed

U	$P(U{=}u)$	$E^{X=0}(Y \mid U) = \tau_0$	$E^{X=1}(Y \mid U) = \tau_1$	$P(X{=}1 \mid U)$
Joe	1/2	.70	.80	.04
Ann	1/2	.20	.40	.76

Table 3: Joe and Ann With Random Assignment

Outcomes ω		Random variables			Regressions					Adjusted Regression
Unit, Treatment, Success	$P(\{\omega\})$	Person variable U	Treatment variable X	Outcome variable Y	$E(Y \mid X, U) = P(Y{=}1 \mid X, U)$	$E(Y \mid X) = P(Y{=}1 \mid X)$	$E(X \mid U) = P(X{=}1 \mid U)$	$E^{X=0}(Y \mid U) = \tau_0$	$E^{X=1}(Y \mid U) = \tau_1$	$E^{\overline{U}}(Y \mid X)$
(*Joe*, *no*, −)	.09	*Joe*	0	0	.70	.45	.40	.70	.80	.45
(*Joe*, *no*, +)	.21	*Joe*	0	1	.70	.45	.40	.70	.80	.45
(*Joe*, *yes*, −)	.04	*Joe*	1	0	.80	.60	.40	.70	.80	.60
(*Joe*, *yes*, +)	.16	*Joe*	1	1	.80	.60	.40	.70	.80	.60
(*Ann*, *no*, −)	.24	*Ann*	0	0	.20	.45	.40	.20	.40	.45
(*Ann*, *no*, +)	.06	*Ann*	0	1	.20	.45	.40	.20	.40	.45
(*Ann*, *yes*, −)	.12	*Ann*	1	0	.40	.60	.40	.20	.40	.60
(*Ann*, *yes*, +)	.08	*Ann*	1	1	.40	.60	.40	.20	.40	.60

Table 4: Joe and Ann Homogeneous

Outcomes ω				Observables			Regressions		
Unit	Treatment	Success	$P(\{\omega\})$	Person variable U	Treatment variable X	Outcome variable Y	$E(Y \mid X,U) = P(Y=1 \mid X,U)$	$E(Y \mid X) = P(Y=1 \mid X)$	$E(X \mid U) = P(X=1 \mid U)$
Joe	*no*	−	.03	*Joe*	0	0	.70	.70	.80
Joe	*no*	+	.07	*Joe*	0	1	.70	.70	.80
Joe	*yes*	−	.08	*Joe*	1	0	.80	.80	.80
Joe	*yes*	+	.32	*Joe*	1	1	.80	.80	.80
Ann	*no*	−	.09	*Ann*	0	0	.70	.70	.40
Ann	*no*	+	.21	*Ann*	0	1	.70	.70	.40
Ann	*yes*	−	.04	*Ann*	1	0	.80	.80	.40
Ann	*yes*	+	.16	*Ann*	1	1	.80	.80	.40

Table 5: Joe and Ann Homogeneous With Partial Regressions

Outcomes ω				Observables			Regressions					
Unit	Treatment	Success	$P(\{\omega\})$	Person variable U	Treatment variable X	Outcome variable Y	$E(Y \mid X,U) = P(Y=1 \mid X,U)$	$E(Y \mid X) = P(Y=1 \mid X)$	$E(X \mid U) = P(X=1 \mid U)$	$E^{X=0}(Y \mid U) = \tau_0$	$E^{X=1}(Y \mid U) = \tau_1$	$E^{\overline{U}}(Y \mid X=x)$ Adjusted Regression
(*Joe*,	*no*,	−)	.03	*Joe*	0	0	.70	.70	.80	.70	.80	.70
(*Joe*,	*no*,	+)	.07	*Joe*	0	1	.70	.70	.80	.70	.80	.70
(*Joe*,	*yes*,	−)	.08	*Joe*	1	0	.80	.80	.80	.70	.80	.80
(*Joe*,	*yes*,	+)	.32	*Joe*	1	1	.80	.80	.80	.70	.80	.80
(*Ann*,	*no*,	−)	.09	*Ann*	0	0	.70	.70	.40	.70	.80	.70
(*Ann*,	*no*,	+)	.21	*Ann*	0	1	.70	.70	.40	.70	.80	.70
(*Ann*,	*yes*,	−)	.04	*Ann*	1	0	.80	.80	.40	.70	.80	.80
(*Ann*,	*yes*,	+)	.16	*Ann*	1	1	.80	.80	.40	.70	.80	.80

Table 6: Expectations in three treatment conditions

treatment	expectation of Y in the treatment conditions $E(Y \mid X{=}x)$	treatment probabilities $P(X{=}x)$
$X{=}0$ (control)	111.25	1/3
$X{=}1$ (treatment 1)	100.00	1/3
$X{=}2$ (treatment 2)	114.25	1/3
$E(Y)$	108.50	

Table 7: Expectations $E(Y \mid X{=}x, Z{=}z)$ in treatment × neediness conditions

	neediness						
treat-ment	low ($Z{=}0$)		medium ($Z{=}1$)		high ($Z{=}2$)		$P(X{=}x)$
$X{=}0$	120	(20/120)	110	(17/120)	60	(3/120)	(40/120)
$X{=}1$	100	(7/120)	100	(26/120)	100	(7/120)	(40/120)
$X{=}2$	80	(3/120)	90	(17/120)	140	(20/120)	(40/120)
$P(Z{=}z)$		(30/120)		(60/120)		(30/120)	

Note. Probabilities $P(X{=}x, Z{=}z)$, $P(Z{=}z)$, and $P(X{=}x)$ in parentheses.

Components of the Theory of Causal Effects

Which type of *empirical phenomenon* does the theory refer to?

- Drawing a person u out of a set of persons. This value u is a value of the random variable U.
- observing the value z_{all} of (a possibly multivariate qualitative and/or quantitative) fallible covariate Z_{all}. The covariate Z_{all} contains *all* fallible covariates. Non-fallible covariates are mappings $f(U)$. Examples are sex, race, true pretest variables $E(Z_1 \mid U)$ if Z_1 denotes a pretest variable.
- assigning the unit or observing its assignment to one of several experimental conditions (represented by the value x of the treatment variable X),
- recording the numerical value y of the outcome variable Y.

A *probability space* $(\Omega, \mathscr{A}, P)$ consists of

- a set Ω of possible outcomes (of the random experiment)
- a σ-algebra $\mathscr{A}$ of possible events
- a probability measure $P : \mathscr{A} \to [0,1]$

On this space we can consider random variables $U, X, Y, Z, \ldots$, all of which are measurable w. r. t. $\mathscr{A}$.

A *filtered space* $\langle (\Omega, \mathscr{A}, P), (\mathscr{F}_t)_{t \in T} \rangle$ consists of:

- a probability space $(\Omega, \mathscr{A}, P)$
- a filtration $(\mathscr{F}_t)_{t \in T}$ (w. r. t. which random variables and events can be ordered)

On this space we can consider random variables $U, X, Y, Z, \ldots$, all of which are measurable w. r. t. $\mathscr{A}$. Some are measurable w. r. t. some of the σ-algebras $\mathscr{F}_t$, some are not.

Let $(\Omega, \mathscr{A})$ be a measurable space, let T be a set on which there are the relations $=$, $<$, and $\leq$, and let $s, t \in T$. A family $(\mathscr{F}_t)_{t \in T}$ of sub-σ-algebras $\mathscr{F}_t$ of $\mathscr{A}$ is called a *filtration* in $\mathscr{A}$, if $\mathscr{F}_s \subset \mathscr{F}_t$ for all $s \leq t$.

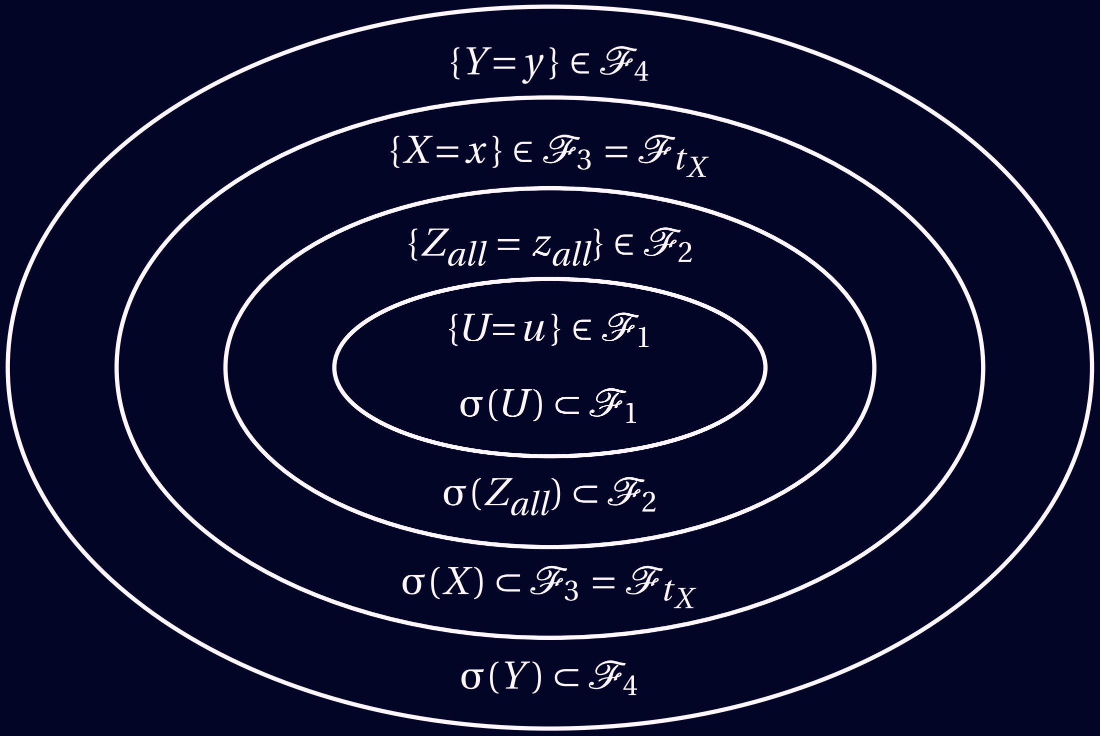

Figure 1: Venn-diagram of a filtration with $T = \{1, \ldots, 4\}$

In the kind of random experiment considered in this presentation the global covariate is (U, Z_{all}).

A random variable Z on $(\Omega, \mathscr{A}, P)$ is called a *covariate* of X if $\sigma(Z) \subset \sigma(U, Z_{all})$.

For $x = 0, 1$, let $E^{X=x}(Y \mid U, Z_{all})$ denote the (U, Z_{all})-conditional expectation of Y with respect to the measure $P^{X=x}$ defined by $P^{X=x}(A) = P(A \mid X=x)$, for all $A \in \mathscr{A}$. $E^{X=x}(Y \mid U, Z_{all})$ is called *P-unique* if all versions of $E^{X=x}(Y \mid U, Z_{all})$ are P-equivalent to each other. $[P^{X=x}(A) = 0] \Rightarrow [P(A) = 0], \ \forall A \in \sigma(U, Z_{all})$, implies that $E^{X=x}(Y \mid U, Z_{all})$ is P-unique.

- Using the conditional expectation $E^{X=x}(Y \mid U, Z_{all})$ of Y given (U, Z_{all}) in treatment x, we can define the *(total-effect) true outcome variables* $\tau_x := E^{X=x}(Y \mid U, Z_{all})$ and the *atomic total-effect variables* $\delta_{10} := \tau_1 - \tau_0$.
- These variables τ_x and δ_{10} are, by definition, unconfounded, because we condition on all covariates (all potential confounfers).
- τ_x and δ_{10} are random variables on the same probability space as the original random variables X and Y. Under the assumption of *P-uniqueness* they have uniquely defined expectations, conditional expectations, variances, covariances, etc.
- The true outcome variables τ_x play the same role as Rubin's potential outcome variables Y_x. In contrast to Y_x, the τ_x allow for effects of variables that are in between X and Y. (This includes mediators.)

Primitives

$\Omega = \Omega_U \times \Omega_{Z_{all}} \times \Omega_X \times \Omega_Y$	The *set of possible outcomes*
$U : \Omega \to \Omega_U$	*Person variable*
$Z : \Omega \to \Omega'_Z$	*Covariate* if $\sigma(Z) \subset \sigma(U, Z_{all})$
$X : \Omega \to \{0, 1\}$	*Treatment variable*
$Y : \Omega \to \mathbb{R}$	*Outcome variable*

Theoretical Concepts the Theory of Causal Effects

$\tau_x := E^{X=x}(Y \mid U, Z_{all})$	*True outcome variables in treatment conditions x with respect to total effects*
$\delta_{10} := \tau_1 - \tau_0$	*Atomic total-effect variables*
$E(\delta_{10})$	*Average total effect*
$E(\delta_{10} \mid Z{=}z)$	*Conditional total effect given* $Z{=}z$
$E(\delta_{10} \mid X{=}x)$	*Conditional total effect given* $X{=}x$

Identification of Causal Effects

Notation

X — treatment variable $\quad Z$ — covariate, i.e., $\sigma(Z) \subset \sigma(U, Z_{all})$

Y — outcome variable $\quad U$ — person variable

$\tau_x := E^{X=x}(Y \mid U, Z_{all}), \quad x = 0, 1,$ — true outcome variables

$\delta_{10} := \tau_1 - \tau_0$ — atomic total-effect variable

$E(\delta_{10})$ — average total effect

$E(\delta_{10} \mid Z{=}z)$ — conditional total effect given $Z{=}z$

Unbiasedness

...of the regression $E^{X=x}(Y \mid Z)$: $\quad E^{X=x}(Y \mid Z) \underset{P}{=} E(\tau_x \mid Z)$

Unbiasedness of the regression $E^{X=x}(Y \mid Z)$ implies:

(1) $E\big[E^{X=x}(Y \mid Z)\big] = E\big[E(\tau_x \mid Z)\big] = E(\tau_x)$

Remember : $E(\delta_{10}) = E(\tau_1) - E(\tau_0)$

and, if $V = f(Z)$,

(2) $E\big[E^{X=x}(Y \mid Z) \mid V\big] \underset{P}{=} E\big[E(\tau_x \mid Z) \mid V\big] \underset{P}{=} E(\tau_x \mid V)$

Remember : $E(\delta_{10} \mid V) \underset{P}{=} E(\tau_1 \mid V) - E(\tau_0 \mid V)$

Let Z_i, $i = 1, \dots, 4$, denote (possibly multivariate) covariates, i.e., $\sigma(Z_i) \subset \sigma(U, Z_{all})$. Then each of the following conditions is sufficient for unbiasedness with $Z = Z_i$:

- Z_1-conditional independence of (U, Z_{all}) and treatments $(X \perp\!\!\!\perp U, Z_{all} \mid Z_1)$

$$P(X{=}x \mid U, Z_{all}) \underset{P}{=} P(X{=}x \mid Z_1), \quad \forall x$$

- Completeness of the regression $(Y \vdash U, Z_{all} \mid X, Z_2)$

$$E(Y \mid X, U, Z_{all}) \underset{P}{=} E(Y \mid X, Z_2)$$

- Z_3-Conditional Strong Causality

  if W is (U, Z_{all})-measurable, then there is a real-valued function h such that

$$E(Y \mid X, Z_3, W) \underset{P}{=} E(Y \mid X, Z_3) + h(Z_3, W) \quad \text{and} \quad P(X{=}x \mid U, Z_{all}) \underset{P}{>} 0, \quad \forall x$$

- Z_4-conditional independence of true outcomes and treatments ($\tau_0, \tau_1 \perp\!\!\!\perp X \mid Z_4$ "strong ignorability")

$$P(X{=}x \mid Z_4, \tau_0, \tau_1) \underset{P}{=} P(X{=}x \mid Z_4) \quad \text{and} \quad P(X{=}x \mid U, Z_{all}) \underset{P}{>} 0, \quad \forall x$$

Note that there are still more causality conditions.

Steyer, R., Mayer, A., & Fiege, C. (2014). Causal inference on total, direct, and indirect effects. In: A.C. Michalos (Ed.) *Encyclopedia of quality of Life and well-being research* (pp. 606-631). Dordrecht, Netherlands: Springer.