Methods of Evaluation Research: Lectures 3 and 4: An outline of the core of the theory

Rolf Steyer

Friedrich Schiller University Jena
Institute of Psychology
Chair of Methodology and Evaluation Research

November 2 and 9, 2015

Introduction

In Lecture 2 we treated three causality conditions for the conditional expectation $E(Y|X)$, where we contended ourselves with a dichotomous treatment variable X with values 0 and 1:

- $E(Y|X) = E^{\overline{U}}(Y|X)$ causal unbiasedness of $E(Y|X)$, which is equivalent to: For $x = 0, 1$: $E(Y|X{=}x) = E(\tau_x)$, where $\tau_x := E^{X=x}(Y|U)$.
- $P(X{=}1|U) = P(X{=}1)$ independence of X and U
- $E(Y|X, U) = E(Y|X)$ X-conditional mean independence of Y from U

The last two conditions imply the first, and they can be tested empirically.

However, these causality conditions require more than we actually need for causal inference, *provided that there are covariates.*

If we consider only the total effect of a treatment variable, then a *covariate* is any pre-treatment variable such as gender, educational status, a pretest (measured before the onset of the treatment), etc. Note that gender and educational status are (deterministic) functions of U, whereas a pretest (e.g., depression before treatment) is not, because it can be measured only with measurement error.

Which type of *empirical phenomenon* do we consider from now on?

- 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 of (a possibly multivariate qualitative or quantitative and possibly fallible) covariate Z_{all} of the unit
- 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.

Table 1: Four Persons With Self-Selection to Treatment

Outcomes ω		Observables				Conditional expectations			
Unit, Treatment, Success	$P(\{\omega\})$	Person variable U	Sex Z	Treatment variable X	Outcome variable Y	$E(Y\|X,U) = P(Y=1\|X,U)$	$E(Y\|X,Z) = P(Y=1\|X,Z)$	$E(Y\|X) = P(Y=1\|X)$	$P(X=1\|U)$
(*Joe*, *no*, −)	.0675	*Joe*	*m*	0	0	.7	.66	.63	.1
(*Joe*, *no*, +)	.1575	*Joe*	*m*	0	1	.7	.66	.63	.1
(*Joe*, *yes*, −)	.0050	*Joe*	*m*	1	0	.8	.44	.44	.1
(*Joe*, *yes*, +)	.0200	*Joe*	*m*	1	1	.8	.44	.44	.1
(*Jim*, *no*, −)	.0175	*Jim*	*m*	0	0	.3	.66	.63	.9
(*Jim*, *no*, +)	.0075	*Jim*	*m*	0	1	.3	.66	.63	.9
(*Jim*, *yes*, −)	.1350	*Jim*	*m*	1	0	.4	.44	.44	.9
(*Jim*, *yes*, +)	.0900	*Jim*	*m*	1	1	.4	.44	.44	.9
(*Sue*, *no*, −)	.0600	*Sue*	*f*	0	0	.7	.60	.63	.2
(*Sue*, *no*, +)	.1400	*Sue*	*f*	0	1	.7	.60	.63	.2
(*Sue*, *yes*, −)	.0200	*Sue*	*f*	1	0	.6	.44	.44	.2
(*Sue*, *yes*, +)	.0300	*Sue*	*f*	1	1	.6	.44	.44	.2
(*Ann*, *no*, −)	.0400	*Ann*	*f*	0	0	.2	.60	.63	.8
(*Ann*, *no*, +)	.0100	*Ann*	*f*	0	1	.2	.60	.63	.8
(*Ann*, *yes*, −)	.1200	*Ann*	*f*	1	0	.4	.44	.44	.8
(*Ann*, *yes*, +)	.0800	*Ann*	*f*	1	1	.4	.44	.44	.8

www.causal-effects.de

Kassel0.tab: self-selection and biased conditional expectations $E(Y \mid X)$ and $E(Y \mid X, Z)$

Kassel1.tab: Z-conditional independence between X and U

Kassel2.tab: Z-conditional mean independence of Y from U

In the last two examples, the conditional expectations $E(Y \mid X, Z)$ are unbiased, where Z denotes sex.

In other words, the conditional expectations $E^{Z=1}(Y \mid X)$ (for males) and $E^{Z=2}(Y \mid X)$ (for females) are unbiased.

How do you compute the average total effect in this case?

Identification of Causal Effects

Notation

X — treatment variable Z — covariate, i.e., $Z = f(U, Z_{all})$

Y — outcome variable U — person variable

Z_{all} — vector of all observed covariates

Notation

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

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

Z_{all} — vector of all observed covariates

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

Notation

X — treatment variable Z — covariate, i.e., $Z = f(U, Z_{all})$
Y — outcome variable U — person variable
Z_{all} — vector of all observed covariates

$\tau_x := E^{X=x}(Y \mid U, Z_{all}),\quad x = 0, 1,$ — 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 conditional expectation $E^{X=x}(Y \mid Z)$

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

Notation

X — treatment variable Z — covariate, i.e., $Z = f(U, Z_{all})$
Y — outcome variable U — person variable
Z_{all} — vector of all observed covariates

$\tau_x := E^{X=x}(Y \mid U, Z_{all}),\quad x = 0, 1,$ — 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 conditional expectation $E^{X=x}(Y \mid Z)$

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

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

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

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

Notation

X — treatment variable Z — covariate, i.e., $Z = f(U, Z_{all})$
Y — outcome variable U — person variable
Z_{all} — vector of all observed covariates

$\tau_x := E^{X=x}(Y \mid U, Z_{all}),\;\; x = 0, 1,$ — 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 conditional expectation $E^{X=x}(Y \mid Z)$

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

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

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

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

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

(3) $E\big[E^{X=x}(Y \mid Z) \,\big|\, V\big] = E\big[E(\tau_x \,\big|\, Z) \mid V\big] = E(\tau_x \mid V)$

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

Let Z_i, $i = a, \ldots, d$, denote (possibly multivariate) covariates, i.e., $Z_i = f(U, Z_{all})$. Then each of the following conditions is sufficient for unbiasedness with $Z = Z_i$:

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

$$P(X{=}x | U, Z_{all}) = P(X{=}x | Z_a), \quad \forall x$$

Let Z_i, $i = a, \ldots, d$, denote (possibly multivariate) covariates, i.e., $Z_i = f(U, Z_{all})$. Then each of the following conditions is sufficient for unbiasedness with $Z = Z_i$:

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

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

- Complete cause condition $(Y \vdash U, Z_{all} \mid X, Z_b)$

$$E(Y \mid X, U, Z_{all}) \; = \; E(Y \mid X, Z_b)$$

Let Z_i, $i = a, \ldots, d$, denote (possibly multivariate) covariates, i.e., $Z_i = f(U, Z_{all})$. Then each of the following conditions is sufficient for unbiasedness with $Z = Z_i$:

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

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

- Complete cause condition ($Y \vdash U, Z_{all} \mid X, Z_b$)

$$E(Y \mid X, U, Z_{all}) \;=\; E(Y \mid X, Z_b)$$

- Z_c-Conditional Strong Causality

  if W is (U, Z_{all})-measurable, then there is a real-valued function h such that
  $E(Y \mid X, Z_c, W) \;=\; E(Y \mid X, Z_c) \;+\; h(Z_c, W)$ and $P(X{=}x \mid U, Z_{all}) > 0, \quad \forall x$

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

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

Note that there are still more causality conditions (see, e.g., ?, ?).

Table 2: 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 3: Expectations $E(Y \mid X=x, Z=z)$ in treatment × neediness conditions

treat-ment	neediness: 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.

Mayer, Thoemmes, Rose, Steyer, and West (2014)
Steyer, Mayer, and Fiege (2014)

Mayer, A., Thoemmes, F., Rose, N., Steyer, R., & West, S. G. (2014). Theory and analysis of total, direct, and indirect causal effects. *Multivariate Behavioral Research*, *49*(5), 425–442.

Steyer, R., Mayer, A., & Fiege, C. (2014). Causal inference on total, direct, and indirect effects. In Michalos, A. C. (Ed.), *Encyclopedia of Quality of Life and Well-Being Research* (pp. 606–631). Dordrecht, Netherlands: Springer.