Methods of Evaluation Research
Lectures 1 and 2: Basic ideas of causal effects

Rolf Steyer

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

October 26, 2015

Motivation

Many theories in psychology involve theoretical variables that cannot directly be observed and many theories use causal concepts such as "leads to", "determines", "is a function of", etc. Obviously, these theories can only be tested in a strict sense if we succeed in translating these concepts of colloquial language into a language that is directly related to our statistical data analysis. So what is the language that is directly related to our statistical data analysis? The answer is obvious: it is the language in which probability, conditional probability, expectation, conditional expectation, variance, correlation, distribution, etc. are defined, i.e., probability theory.

- How can we define latent variables and causal dependencies and effects in terms of probability theory?
- What is the difference between a stochastic dependency and a causal stochastic dependency?
- What is the difference between a manifest random variable and a latent random variable?

In this presentation I introduce the basic ideas using simple examples.

Joe and Ann With Self-Selection – Compressed

Table 1: Joe and Ann With Self-Selection – Compressed

U	$P(U{=}u)$	$P(X{=}1\,\vert\,U)$	$P^{X=0}(Y{=}1\,\vert\,U) =: \tau_0$	$P^{X=1}(Y{=}1\,\vert\,U) =: \tau_1$	$\delta_{10} := \tau_1 - \tau_0$	$P(U{=}u\,\vert\,X{=}0)$	$P(U{=}u\,\vert\,X{=}1)$
Joe	1/2	.04	.70	.80	.10	.80	.05
Ann	1/2	.76	.20	.40	.20	.20	.95
			$E(\tau_0) = .45$	$E(\tau_1) = .60$			

Individual causal total effect function: $\delta_{10} := \tau_1 - \tau_0$.

Average causal total effect: $E(\delta_{10}) = E(\tau_1 - \tau_0) = E(\tau_1) - E(\tau_0) = .15$.

For the definition and properties of the concepts used in this table see ? (?).

Joe and Ann With Self-Selection

Table 2: Joe and Ann With Self-Selection

Outcomes ω_i		Observables			Conditional Expectations					
Unit, Treatment, Success	$P(\{\omega_i\})$	Person variable U	Treatment variable X	Outcome variable Y	$P(Y{=}1 \mid X, U) = E(Y \mid X, U)$	$P(Y{=}1 \mid X) = E(Y \mid X)$	$P(X{=}1 \mid U)$	$P^{X=0}(Y{=}1 \mid U) = \tau_0$	$P^{X=1}(Y{=}1 \mid U) = \tau_1$	$E^{\overline{U}}(Y \mid X)$
$\omega_1 = (Joe,\ no,\ -)$	.144	*Joe*	0	0	.70	.60	.04	.70	.80	.45
$\omega_2 = (Joe,\ no,\ +)$	.336	*Joe*	0	1	.70	.60	.04	.70	.80	.45
$\omega_3 = (Joe,\ yes,\ -)$	.004	*Joe*	1	0	.80	.42	.04	.70	.80	.60
$\omega_4 = (Joe,\ yes,\ +)$	.016	*Joe*	1	1	.80	.42	.04	.70	.80	.60
$\omega_5 = (Ann,\ no,\ -)$	.096	*Ann*	0	0	.20	.60	.76	.20	.40	.45
$\omega_6 = (Ann,\ no,\ +)$	.024	*Ann*	0	1	.20	.60	.76	.20	.40	.45
$\omega_7 = (Ann,\ yes,\ -)$	.228	*Ann*	1	0	.40	.42	.76	.20	.40	.60
$\omega_8 = (Ann,\ yes,\ +)$	.152	*Ann*	1	1	.40	.42	.76	.20	.40	.60

$$E^{\overline{U}}(Y \mid X)(\omega) := \begin{cases} E(\tau_0), & \text{if } X(\omega) = 0 \\ E(\tau_1), & \text{if } X(\omega) = 1 \end{cases} \quad (U\text{-adjusted } X\text{-conditional expectation of } Y)$$

Causality condition: Causal unbiasedness of $E(Y \mid X)$: $E(Y \mid X) = E^{\overline{U}}(Y \mid X)$ Individual causal total effect function: $\tau_1 - \tau_0$.
Average causal total effect: $E(\tau_1 - \tau_0) = E(\tau_1) - E(\tau_0) = .15$.

Joe and Ann With Random Assignment

Table 3: Joe and Ann With Random Assignment

Outcomes ω_i		Observables			Conditional expectations					
Unit, Treatment, Success	$P(\{\omega_i\})$	Person variable U	Treatment variable X	Outcome variable Y	$P(Y{=}1 \mid X, U) = E(Y \mid X, U)$	$P(Y{=}1 \mid X) = E(Y \mid X)$	$P(X{=}1 \mid U) = E(X \mid U)$	$P^{X=0}(Y{=}1 \mid U) =: \tau_0$	$P^{X=1}(Y{=}1 \mid U) =: \tau_1$	$E^{\overline{U}}(Y \mid X)$
$\omega_1 = (Joe,\ no,\ -)$	.09	*Joe*	0	0	.70	.45	.40	.70	.80	.45
$\omega_2 = (Joe,\ no,\ +)$	.21	*Joe*	0	1	.70	.45	.40	.70	.80	.45
$\omega_3 = (Joe,\ yes,\ -)$	.04	*Joe*	1	0	.80	.60	.40	.70	.80	.60
$\omega_4 = (Joe,\ yes,\ +)$	.16	*Joe*	1	1	.80	.60	.40	.70	.80	.60
$\omega_5 = (Ann,\ no,\ -)$	.24	*Ann*	0	0	.20	.45	.40	.20	.40	.45
$\omega_6 = (Ann,\ no,\ +)$	.06	*Ann*	0	1	.20	.45	.40	.20	.40	.45
$\omega_7 = (Ann,\ yes,\ -)$	.12	*Ann*	1	0	.40	.60	.40	.20	.40	.60
$\omega_8 = (Ann,\ yes,\ +)$	.08	*Ann*	1	1	.40	.60	.40	.20	.40	.60

Causality condition of $E(Y \mid X)$: Independence of X and U

$$P(X{=}1 \mid U) = P(X{=}1).$$

This condition implies unbiasedness of $E(Y \mid X)$.

Joe and Ann Homogeneous

Table 4: Joe and Ann Homogeneous

Outcomes ω_i (Unit, Treatment, Success)	$P(\{\omega_i\})$	Person variable U	Treatment variable X	Outcome variable Y	$P(Y=1 \mid X,U) = E(Y \mid X,U)$	$P(Y=1 \mid X) = E(Y \mid X)$	$P(X=1 \mid U) = E(X \mid U)$	$P^{X=0}(Y=1 \mid U) = \tau_0$	$P^{X=1}(Y=1 \mid U) = \tau_1$	$E^{\overline{U}}(Y \mid X)$
$\omega_1 = (Joe,\ no,\ -)$	.03	*Joe*	0	0	.70	.70	.80	.70	.80	.70
$\omega_2 = (Joe,\ no,\ +)$	.07	*Joe*	0	1	.70	.70	.80	.70	.80	.70
$\omega_3 = (Joe,\ yes,\ -)$	.08	*Joe*	1	0	.80	.80	.80	.70	.80	.80
$\omega_4 = (Joe,\ yes,\ +)$	.32	*Joe*	1	1	.80	.80	.80	.70	.80	.80
$\omega_5 = (Ann,\ no,\ -)$	.09	*Ann*	0	0	.70	.70	.40	.70	.80	.70
$\omega_6 = (Ann,\ no,\ +)$	.21	*Ann*	0	1	.70	.70	.40	.70	.80	.70
$\omega_7 = (Ann,\ yes,\ -)$	.04	*Ann*	1	0	.80	.80	.40	.70	.80	.80
$\omega_8 = (Ann,\ yes,\ +)$	.16	*Ann*	1	1	.80	.80	.40	.70	.80	.80

Causality condition for $E(Y \mid X)$: X-conditional mean independence of Y from U

$$E(Y \mid X, U) = E(Y \mid X).$$

This condition also implies unbiasedness of $E(Y \mid X)$.

What Joe and Ann teach us about causal total effects

- When we talk about conditional and average causal effects, then **we refer to a random experiment**. In the Joe-Ann example, this random experiment consists of drawing a person from a set of persons, observing whether or not the person is treated, and observing the value of the outcome variable. In this example the set of persons is $\Omega_U = \{Joe, Ann\}$.
- **The random experiment is represented by a probability space** $(\Omega, \mathscr{A}, P)$. In this example, the set of (possible) outcomes is $\Omega = \{\omega_1, \ldots, \omega_8\}$, the σ-algebra $\mathscr{A}$, i.e., the set of (possible) events, is the power set of Ω, and the probability measure P is completely determined by the probabilities $P(\{\omega_i\})$ of the elementary events $\{\omega_i\}$, $i = 1, \ldots, 8$.
- The **person variable** U is an ordinary random variable. Therefore, it has a distribution and a joint distribution with the other random variables on this probability space, such as X and Y.
- **The individual total causal effects and the average total causal effect** are perfectly defined exclusively using the joint probability distribution of the random variables U, X, and Y.
- **Without the person variable U causal effects cannot be defined.**

Conclusion

- The theory of stochastic causality that I outlined has many practical implications. A very useful tool for data analysis is EffectLiteR, an R-program written by Axel Mayer.
- Treatments usually do not deterministically determine the outcomes. Treatments just change distributions or expectations of random variables. Usually there are many other causes aside from the treatment, including those that are intermediate between treatment and outcome. (In contrast, compare Rubin's approach to causal effects. In this context, see also ? (?) on direct effects and the limitations of the randomized experiment.)
- There is no need for do-operators in stochastic causality. (In contrast, compare e.g. ?, ?).

Conclusion

- Causal effects can be defined exclusively using well-known and well-defined concepts of probability theory. In this case, the implications of our theories, assumptions, and hypotheses can be tested directly in appropriate samples. Statistical inference aims at aspects of the (joint) distributions of random variables. Only if theories refer to the same (joint) distributions can we learn from empirical findings about the validity of our theories.
- In the definitions of of causal effects we do not refer to substantive or philosophical theories. It is straight-forward mathematics and we can use logical instead of plausibility arguments in the derivation of further properties and implications.

References

Books