Correction to Proposition 6.13 in 'Elements of Causal Inference'

Jonas Peters*

September 16, 2024

Remark 1 *I am omitting some measure theoretic details in the arguments and only refer to the relevant literature when the statements are non-trivial. I hope that the arguments are clear enough, so that you could add the details if you wanted to. In that case, the definition of SCMs by Bongers et al. [2021] might be a good starting point.*

Definition 2 *Let $\mathfrak{C}$ be an SCM over variables $\mathbf{X}$ with $X, Y \in \mathbf{X}$. We say that there is a* (total) causal effect *from X to Y if there exists a random variable $\tilde{N}_X$ s.t.*

$$X \not\perp\!\!\!\perp Y \qquad \text{in } P^{\mathfrak{C};do\left(X:=\tilde{N}_X\right)}.$$

Proposition 3 *Let $\mathfrak{C}$ be an SCM over variables $\mathbf{X}$ with $X, Y \in \mathbf{X}$. Consider*

(i) There is a total causal effect from X to Y.

(ii) There are $x^\triangle$ and $x^\square$ such that $P_Y^{\mathfrak{C};\,do\left(X:=x^\triangle\right)} \neq P_Y^{\mathfrak{C};\,do\left(X:=x^\square\right)}$.

(iii) There is $x^\triangle$ such that $P_Y^{\mathfrak{C};\,do\left(X:=x^\triangle\right)} \neq P_Y^{\mathfrak{C}}$.

(iv) $X \not\perp\!\!\!\perp Y$ in $P_{X,Y}^{\mathfrak{C};\,do\left(X:=\tilde{N}_X\right)}$ for any $\tilde{N}_X$ whose distribution has full support.

We then have $(iv) \Rightarrow (i) \Leftrightarrow (ii) \Rightarrow (iii)$.

To prove this proposition, we first consider the following lemma.

Lemma 4 *Let $\mathfrak{C}$ be an SCM over variables $\mathbf{X}$ with $X, Y \in \mathbf{X}$ and consider a noise variable $\tilde{N}_X$. Then,*

$$(P^{\mathfrak{C};do(X:=x)}(Y \in \cdot))_x \text{ is a cond. distribution for } Y \text{ given } X \text{ in } P^{\mathfrak{C};do\left(X:=\tilde{N}_X\right)}.$$

*Many thanks to Joris Mooij and Michael Law for helpful discussions and providing ideas for Example 5(*ii*) and (*i*), respectively.

Proof. Let $\mathcal{G}$ be the graph induced by the SCM. Substituting, iteratively, the structural assignments into each other (but not the one for X), we obtain that there is a function g s.t.

$$Y = g(X, N_{A_Y^X}),$$

where $A_Y^X := \mathbf{AN}_Y^{\mathcal{G}^*} \setminus \{X\}$ and $\mathcal{G}^*$ equals $\mathcal{G}$ after removing the edges incoming into X. We have that $X \perp\!\!\!\perp N_{A_Y^X}$ in $P^{\mathfrak{C};do\left(X:=\tilde{N}_X\right)}$. Therefore, the substitution theorem of conditional probabilities [Rønn-Nielsen and Hansen, 2014, Theorem 2.1.1] states that

$$(P^{\mathfrak{C};do\left(X:=\tilde{N}_X\right)}(g(x, N_{A_Y^X}) \in \cdot)_x \tag{1}$$

is a conditional probability for Y given X in $P^{\mathfrak{C};do\left(X:=\tilde{N}_X\right)}$. But (1) equals

$$(P^{\mathfrak{C};do(X:=x)}(Y \in \cdot))_x,$$

which concludes the proof of Lemma 4. □

We are now ready to prove Proposition 3.

Proof. The implications $(iv) \Rightarrow (i)$ and $(ii) \Rightarrow (iii)$ are trivial.

To prove $(i) \Rightarrow (ii)$, assume that (ii) does not hold. Then, there exists a function c s.t. for all x and for all measurable B

$$P^{\mathfrak{C};do(X:=x)}(Y \in B) = c(B).$$

Thus, for all measurable A and B, we have

$$\begin{aligned} P^{\mathfrak{C};do\left(X:=\tilde{N}_X\right)}(X \in A, Y \in B) &= \int_A P^{\mathfrak{C};do(X:=x)}(Y \in B)\, dP_X^{\mathfrak{C};do\left(X:=\tilde{N}_X\right)}(x) \\ &= c(B) P^{\mathfrak{C};do\left(X:=\tilde{N}_X\right)}(X \in A), \end{aligned}$$

where for the first equality we have made use of Lemma 4. But this implies

$$c(B) = P^{\mathfrak{C};do\left(X:=\tilde{N}_X\right)}(Y \in B),$$

which shows that X and Y are independent in $P^{\mathfrak{C};do\left(X:=\tilde{N}_X\right)}$.

To prove $(ii) \Rightarrow (i)$, assume that (ii) holds and let $x^{\triangle}$ and $x^{\square}$ be values and B be a measurable set such that

$$P_Y^{\mathfrak{C};do\left(X:=x^{\triangle}\right)}(B) \neq P_Y^{\mathfrak{C};do\left(X:=x^{\square}\right)}(B).$$

Consider a random variable $\tilde{N}_X$ with $P(\tilde{N}_X = x^{\triangle}) = 1/2 = P(\tilde{N}_X = x^{\square})$. We then have

$$\begin{aligned} P^{\mathfrak{C};do\left(X:=\tilde{N}_X\right)}(X = x^{\triangle}, Y \in B) &= \int_{\{x^{\triangle}\}} P^{\mathfrak{C};do(X:=x)}(Y \in B)\, dP_X^{\mathfrak{C};do\left(X:=\tilde{N}_X\right)}(x) \\ &= \frac{1}{2} P^{\mathfrak{C};do\left(X:=x^{\triangle}\right)}(Y \in B) \\ &\neq \frac{1}{2} P^{\mathfrak{C};do\left(X:=x^{\square}\right)}(Y \in B) \\ &= P^{\mathfrak{C};do\left(X:=\tilde{N}_X\right)}(X = x^{\square}, Y \in B), \end{aligned}$$

where, again, we have made use of Lemma 4. This implies that X and Y are not independent in $P^{\mathfrak{C};do(X:=\tilde{N}_X)}$ (if they were, then the first and last term should both equal $1/2 \cdot P^{\mathfrak{C};do(X:=\tilde{N}_X)}(Y \in B)$. This concludes the proof of Proposition 3. $\square$

Example 5 *(i) Let $\mathfrak{C}$ be an SCM over (X, Y) with*

$$X := N_X \tag{2}$$
$$Y := 1_0(X) + N_Y, \tag{3}$$

where N_X and N_Y are i.i.d. $\mathcal{N}(0,1)$ and $1_0(0) = 1$ and $1_0(x) = 0$ for all $x \neq 0$. Then, (ii) of Proposition 3 is satisfied but not (iv).

(ii) Let $\mathfrak{C}$ be an SCM over (X, Y) with

$$Z := N_Z \tag{4}$$
$$X := Z \tag{5}$$
$$Y := \mathrm{sg}(X) \cdot \mathrm{sg}(Z) + N_Y, \tag{6}$$

where N_Z and N_Y are i.i.d. $\mathcal{N}(0,1)$ and $\mathrm{sg}(x) = 1$ for all $x \geq 0$ and $\mathrm{sg}(x) = -1$ for all $x < 0$. Then, (iii) of Proposition 3 is satisfied but not (ii).

References

S. Bongers, P. Forre, J. Peters, and J. M. Mooij. Foundations of structural causal models with cycles and latent variables. *Annals of Statistics*, 49(5): 2885–2915, 2021.

A. Rønn-Nielsen and E. Hansen. Conditioning and Markov properties, 2014.