2 Tutorial exercises

2.1 Seeing vs. Doing in SCMs

Consider a simple causal model of a car. The endogenous variables are (all binary):

X: the battery is charged
Y: the start engine is operational
S: the car starts

The exogenous variables (latent, independent, binary) have Bernoulli distributions:

$$E_X \sim \text{Ber}(0.95)$$
$$E_Y \sim \text{Ber}(0.99)$$
$$E_Z \sim \text{Ber}(0.999)$$

The structural equations are specified by:

$$X = E_X$$
$$Y = E_Y$$
$$S = X \wedge Y \wedge E_S$$

where "$\wedge$" is the logical AND.

a) *Draw the corresponding augmented causal graph and the causal graph.*

b) *Write pseudocode to draw a sample (x, y, s) from this model.*

c) *How does the model change under a perfect intervention* $\text{do}(S = 0)$*? Write down the intervened model. How does the pseudocode to sample from the model change?*

d) *Is* $p(X = x|S = s) = p(X = x|\,\text{do}(S = s))$*? Motivate your answer.*

e) *Is* $p(S = s|X = x) = p(S = s|\,\text{do}(X = x))$*? Motivate your answer.*

f) *Calculate:*

i) $p(S = 1)$ *(the probability that the car starts)*
ii) $p(S = 1|X = 1)$ *(the probability that the car starts, given that the battery is charged)*
iii) $p(S = 1|\,\text{do}(X = 1))$ *(the probability that the car starts when we charge the battery)*

g) *Calculate:*

i) $p(X = 0)$ *(the probability that the battery is empty)*
ii) $p(X = 0|S = 0)$ *(the probability that the battery is empty given that the car fails to start)*
iii) $p(X = 0|\,\text{do}(S = 0))$ *(the probability that the battery is empty if we lost the key)*

2.2 The Back-Door Criterion

Theorem 1 (Back-Door Criterion (Pearl, 2000)) *For an acyclic SCM $\mathcal{M}$, variables X, Y and set of variables $\boldsymbol{H}$: Let $\hat{\mathcal{G}}$ be $\mathcal{G}(\mathcal{M})$ extended with an intervention node $I_X \to X$. If*

1. $X, Y \notin \boldsymbol{H}$*;*
2. $H \perp_{\hat{\mathcal{G}}} I_X$*;*
3. $Y \perp_{\hat{\mathcal{G}}} I_X \mid X, H$*,*

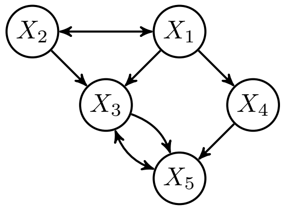
then $\boldsymbol{H}$ *is called* admissible for adjustment *to find the causal effect of* X *on* Y, *and this causal effect is given by:*

$$p_{\mathcal{M}}\big(y \mid \mathrm{do}(X=x)\big) = \int p_{\mathcal{M}}(y \mid x, \boldsymbol{h}) p_{\mathcal{M}}(\boldsymbol{h})\, d\boldsymbol{h} \left(= \sum_{\boldsymbol{h}} p_{\mathcal{M}}(y \mid x, \boldsymbol{h}) p_{\mathcal{M}}(\boldsymbol{h}) \right).$$

where the last equation only applies if $\boldsymbol{H}$ *is discrete-valued. For the special case* $\boldsymbol{H} = \emptyset$, *this should be read as:*

$$p_{\mathcal{M}}\big(y \mid \mathrm{do}(X=x)\big) = p_{\mathcal{M}}(y \mid x).$$

Consider an SCM $\mathcal{M}$ with the following causal graph $\mathcal{G}(\mathcal{M})$:



a) *Give a set that is admissible for adjustment to find the causal effect of* X_4 *on* X_5.

b) *Provide an expression for this causal effect in terms of the observational distribution.*

c) *Give a set that is admissible for adjustment to find the causal effect of* X_1 *on* X_5.

d) *Provide an expression for this causal effect in terms of the observational distribution.*

e) *Is* $\emptyset$ *admissible for adjustment to find the causal effect of* X_1 *on* X_4? *If so, provide an expression for this causal effect in terms of the observational distribution.*

f) *Which sets are admissible for adjustment to find the causal effect of* X_3 *on* X_5?

g) *Which sets are admissible for adjustment to find the causal effect of* X_5 *on* X_4?

2.3 Simpson's Paradox: resolution

This exercise continues where exercise 1 ended. We will make use of causal reasoning with SCMs to resolve Simpson's paradox.

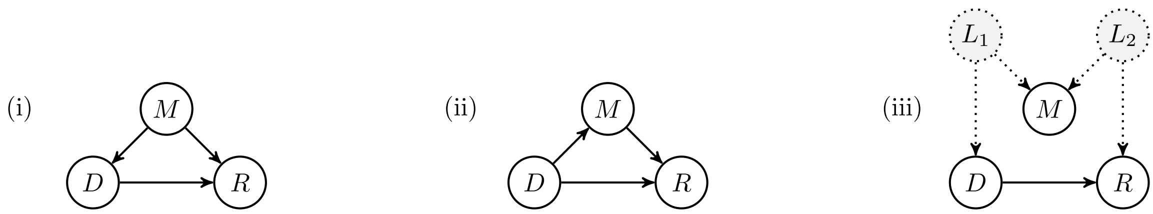


Figure 1: Different hypothetical causal graphs, where R stands for *Recovery*, D for taking the *Drug*, and M has different interpretations in cases (i), (ii) and (iii).

Suppose you believe that an SCM with causal graph as in Figure 1(i) applies, where M denotes gender of the patient (male/female).

a) *Apply the back-door criterion to obtain a formula that expresses* $p(r \mid \mathrm{do}(D=d))$ *in terms of observable quantities (i.e., in terms of marginal or conditional distributions where the* do*-operator does not appear).*

b) *Is* $p(r \mid \mathrm{do}(D=d)) = p(r \mid d)$ *in this case?*

c) *What would be your advice for a patient with unknown gender?*

Now suppose that instead, you believe that an SCM with causal graph as in Figure 1(ii) applies. Intuitively, this would be quite unlikely, as we know that most drugs don't change gender, but we could have used a slightly different story where the variable M has a different interpretation (for example, "blood pressure"), and then this causal structure would also be a plausible one.

d) *Again, use the back-door criterion to express* $p(r \mid \mathrm{do}(D = d))$ *in terms of observable quantities.*

e) *Is* $p(r \mid \mathrm{do}(D = d)) = p(r \mid d)$ *in this case?*

f) *What would be your advice for a patient with unknown* M *(say, blood pressure) in this case?*

Finally, suppose that you believe that the SCM has the causal graph of Figure 1(iii).

g) *Invent an interpretation of* M *and the two latent variables* L_1, L_2 *yourself that could match the causal model depicted in Figure 1(iii).*

h) *Express* $p(r \mid \mathrm{do}(D = d))$ *in terms of* observable *quantities.*

i) *Is* $p(r \mid \mathrm{do}(D = d)) = p(r \mid d)$ *in this case?*

j) *Again, what would be your advice for a patient with unknown* M *in this case?*

Conclusion: whether or not you should prescribe the drug depends on which causal model you believe to apply to this situation. The fact that different causal models will lead to different conclusions should not be paradoxical, it is another illustration that "correlation does not imply causation".

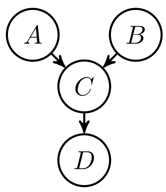
2.4 Y-structure

Theorem 2 (Global Markov Property) *For an acyclic SCM, the following* Global Markov Property *holds:*

$$\boldsymbol{X}, \boldsymbol{Y} \underset{\mathcal{G}(\mathcal{M})}{\perp} \boldsymbol{Z} \qquad \Longrightarrow \qquad \boldsymbol{X} \underset{p_{\mathcal{M}}}{\perp\!\!\!\perp} \boldsymbol{Y} \mid \boldsymbol{Z}$$

for all subsets $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}$ *of nodes.*

Given an SCM $\mathcal{M}$ with the following causal graph:



Which (conditional) independences in $p_{\mathcal{M}}$ *are implied by the Global Markov Property?*

2.5 LCD

Theorem 3 (Cooper, 1997) *Given an acyclic SCM* $\mathcal{M}$ *with three endogenous variables* C, X, Y*. If:*

1. $X \to C \notin \mathcal{G}(\mathcal{M})$,
2. $Y \to C \notin \mathcal{G}(\mathcal{M})$,
3. $C \not\perp\!\!\!\perp_{p_{\mathcal{M}}} X$,
4. $X \not\perp\!\!\!\perp_{p_{\mathcal{M}}} Y$,
5. $C \perp\!\!\!\perp_{p_{\mathcal{M}}} Y \mid X$,
6. *Faithfulness holds, i.e., the Global Markov Property gives* all *(conditional) independences in* $p_{\mathcal{M}}$.

Then $X \to Y \in \mathcal{G}(\mathcal{M})$, $C \to Y \notin \mathcal{G}(\mathcal{M})$, $X \leftrightarrow Y \notin \mathcal{G}(\mathcal{M})$, $C \leftrightarrow Y \notin \mathcal{G}(\mathcal{M})$.

Prove this theorem by considering for each possible ADMG $\mathcal{G}(\mathcal{M})$ *whether it satisfies the assumptions. (Hint: could there be an edge between* C *and* Y*?)*