EXERCISES FOR RANDOMIZATION AND CAUSATION (MATH-336)

Exercise Sheet 4

Exercise 1 (Balancing property of propensity scores)**.** (from [1]) Let $A = 0, 1$ be a binary treatment variable, L be a discrete baseline covariate and $Y = 0, 1$ be a binary outcome variable. The propensity score is defined as $\pi(L) := P(A = 1 \mid L)$. Prove that

(a) $P(A = 1 \mid \pi(L), L) = \pi(L)$.
(b) $P(A = 1 \mid \pi(L)) = \pi(L)$ and thus deduce that $A \perp\!\!\!\perp L \mid \pi(L)$.

Solution:

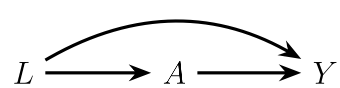
(a) The event defined by $\{L = l, \pi(L) = \pi(l)\}$ is identical[1] to the event $\{L = l\}$ for all l, and therefore we can remove $\pi(L)$ from the conditioning set. Note that the converse is not true, that is, the event $\{L = l, \pi(L) = \pi(l)\}$ is not equal to $\{\pi(L) = \pi(l)\}$ for all l, because $\pi(L)$ may not be a bijective function of L.
(b) We start by noticing that $P(A = 1 \mid \pi(L)) = E[A \mid \pi(L)]$. Next, we have from the law of total expectation that $E[A] = E[E[A \mid L]]$. When we further condition on $\pi(L)$, then $E[A \mid \pi(L)] = E\{E[A \mid L, \pi(L)] \mid \pi(L)\}$, which allows us to write

$$\begin{aligned} P(A = 1 \mid \pi(L)) &= E[A \mid \pi(L)] \\ &= E\{ \underbrace{E[A \mid L, \pi(L)]}_{= P(A=1 \mid \pi(L), L) \overset{\text{(a)}}{=} \pi(L)} \mid \pi(L)\} \\ &= E\{\pi(L) \mid \pi(L)\} \\ &= \pi(L) \ . \end{aligned}$$

From (a) and (b), we have the equality $\pi(L) = P(A = 1 \mid \pi(L), L) = P(A = 1 \mid \pi(L))$, which allows us to conclude that $A \perp\!\!\!\perp L \mid \pi(L)$.

Exercise 2 (Propensity scores)**.** Let A, Y denote treatment and outcome respectively. Furthermore, let L be a set of baseline covariates (a common cause of A and Y) and denote by $\pi(L)$ the propensity score $P(A = 1 \mid L)$, a deterministic function of L.

(i) Assume that A, L, Y satisfy the causal model $\mathcal{G}$:



$\mathcal{G}$

Draw the causal DAG $\mathcal{G}^\star$ containing nodes $A, L, \pi(L), Y$ which satisfies $A \perp\!\!\!\perp L \mid \pi(L)$ (a fact about $\pi(L)$ which we proved in Exercise 1).

[1]To see this more formally, suppose we have a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We then have that $\{\omega \in \Omega : L(\omega) = l, \pi \circ L(\omega) = \pi(l)\} = \{\omega \in \Omega : L(\omega) = l\}$ for all l.

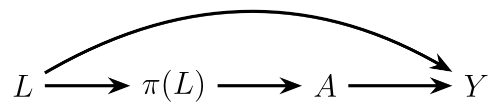
(ii) By drawing the SWIG $\mathcal{G}^\star(a)$ and using the rules of d-separation, show that

$$Y^a \perp\!\!\!\perp A \mid \pi(L) \quad \text{whenever} \quad Y^a \perp\!\!\!\perp A \mid L\ .$$

In other words, we want to show that the propensity score is sufficient to adjust for confounding whenever L is sufficient to adjust for confounding.

Solution:

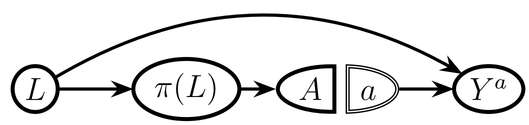
(i) The desired graph is



$\mathcal{G}^\star(a)$

The node $\pi(L)$ must lie on the path $L \to A$ because $\pi(L)$ is a descendant of L and $A \perp\!\!\!\perp L \mid \pi(L)$.

(ii) The desired SWIG is



$\mathcal{G}^\star(a)$

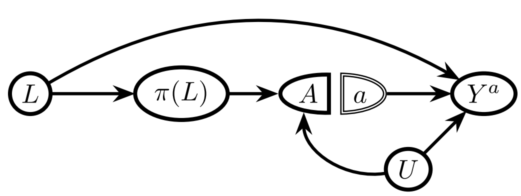
Let $G = \{\mathcal{G}_i : (Y^a \perp\!\!\!\perp A \mid L)_{\mathcal{G}_i(a)}\}$ be the set of causal models satisfying conditional exchangeability.[2] By definition, every $\mathcal{G}_i \in G$ has in common that the path $(A \leftarrow \cdots \leftarrow \pi(L) \leftarrow L \to \cdots \to Y^a)_{\mathcal{G}_i}$ is closed by conditioning on L. Because $\pi(L)$ always intersects the path $(L \to \cdots \to A)_{\mathcal{G}_i}$, we can therefore block this path for all $\mathcal{G}_i \in G$ by conditioning on $\pi(L)$, and thus $Y^a \perp\!\!\!\perp A \mid \pi(L)$ whenever $Y^a \perp\!\!\!\perp A \mid L$.

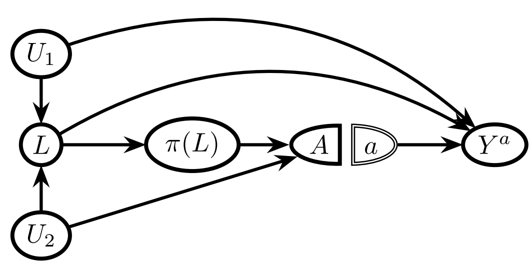
Exercise 3 (SWIGS and independencies)**.** (from Robins' EPI 207, Homework 2 [2])

(a) Given the graph in Fig. 1, draw SWIGs corresponding to
  (i) Intervening on A_0 alone,
  (ii) Intervening on A_1 alone,
  (iii) Intervening on both A_0, A_1.

[2] The causal model $\mathcal{G}^\star$ is an example of a model satisfying $Y^a \perp\!\!\!\perp A \mid L$ (i.e. $\mathcal{G}^\star \in G$) because the path $(A \leftarrow \pi(L) \leftarrow L \to Y^a)_{\mathcal{G}^\star}$ is closed by conditioning on L.

Examples of other causal models where $Y^a \perp\!\!\!\perp A \mid L$ is *not* satisfied include the following:





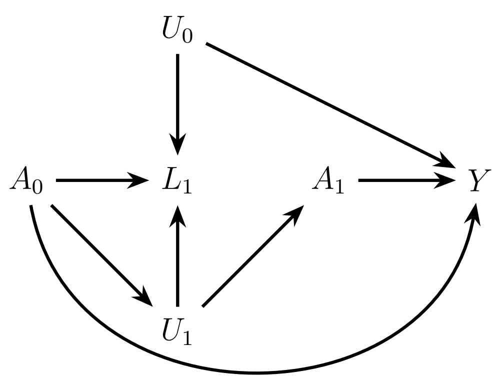


FIGURE 1

(b) Use your SWIGs from part (a) to determine whether the following statements are true or false and explain why:
(i) $Y^{a_0} \perp\!\!\!\perp A_0$
(ii) $Y^{a_0} \perp\!\!\!\perp A_0 \mid L_1^{a_0}$
(iii) $Y^{a_1} \perp\!\!\!\perp A_0$
(iv) $Y^{a_1} \perp\!\!\!\perp A_1$
(v) $Y^{a_1} \perp\!\!\!\perp A_1 \mid L_1, A_0$
(vi) $Y^{a_1} \perp\!\!\!\perp A_1 \mid A_0$
(vii) $Y^{a_0,a_1} \perp\!\!\!\perp A_0$
(viii) $Y^{a_0,a_1} \perp\!\!\!\perp A_1^{a_0}$
(ix) $Y^{a_0,a_1} \perp\!\!\!\perp A_1^{a_0} \mid L_1^{a_0}, A_0$
(x) $Y^{a_0,a_1} \perp\!\!\!\perp A_1^{a_0} \mid A_0$
(xi) $L_1^{a_0} = L_1$
(xii) $A_1^{a_0} = A_1$

Solution:

(a) The desired SWIGS are shown in Fig. 2.

(b) (i) True. There are no arrows into or out of A_0 in the SWIG, so it is d-separated from every other node.
(ii) True. There are no arrows into or out of A_0 in the SWIG, so it is d-separated from every other node, whether or not we condition on $L_1^{a_0}$.
(iii) False. The path $A_0 \to Y^{a_1}$ is an open directed path from A_0 to Y^{a_1}.
(iv) False. $A_1 \leftarrow U_1 \leftarrow A_0 \to Y^{a_1}$ is an open directed path from A_1 to Y^{a_1}.
(v) False. Conditioning on L_1 unblocks the path $A_1 \leftarrow U_1 \to L_1 \leftarrow U_0 \to Y^{a_1}$.
(vi) True. Conditioning on A_0 blocks all paths through it and all paths through L_1 are blocked because it is a collider.
(vii) True. There are no arrows into or out of A_0 in the SWIG so it is d-separated from every other node.
(viii) True. All paths from $A_1^{a_0}$ to Y^{a_0,a_1} are blocked by either the intervention at A_0 or the collider at $L_1^{a_0}$.
(ix) False. The path $A_1^{a_0} \leftarrow U_1^{a_0} \to L_1^{a_0} \leftarrow U_0 \to Y^{a_0,a_1}$ from $A_1^{a_0}$ to Y^{a_0,a_1} is unblocked by conditioning on $L_1^{a_0}$.

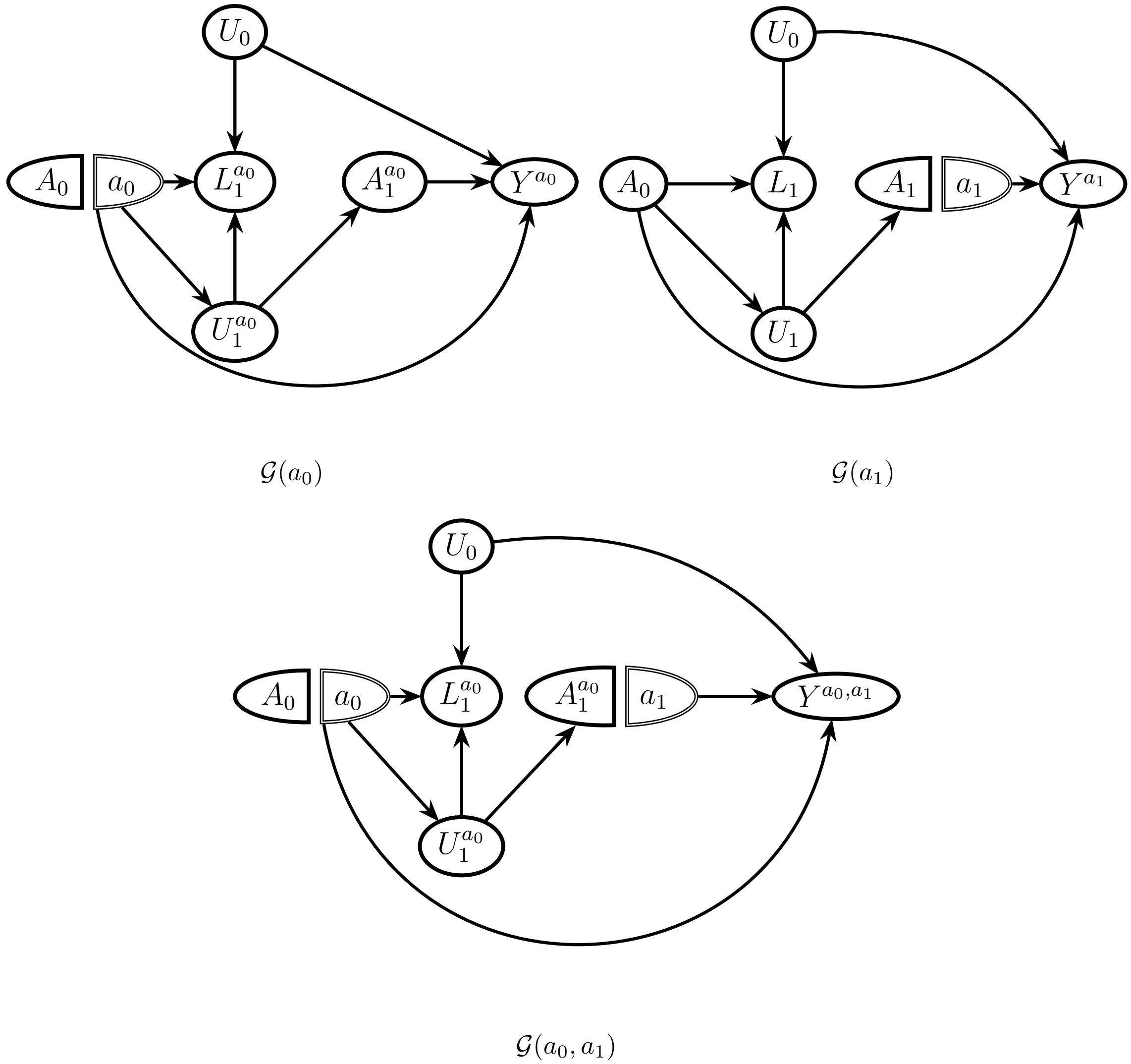


FIGURE 2. Showing the SWIGs in (a) (i)-(iii)

(x) True. All paths from $A_1^{a_0}$ to Y^{a_0,a_1} are blocked by either the intervention at A_0 or the collider at $L_1^{a_0}$.
(xi) False. There is an arrow directly from a_0 to $L_1^{a_0}$.
(xii) False. $a_0 \to U_1^{a_0} \to A_1^{a_0}$ is an open path from a_0 to $A_1^{a_0}$.

Exercise 4 (Evaluating the causal assumptions)**.** Consider again the study investigating whether GRE test scores can be used to predict future performance [3], discussed in Exercise Sheet 3. As suggested at the end of the exercise solution, suppose we conduct a modified version of the original study where all applicants to graduate school are admitted regardless of their GRE score. We will now consider whether the contrast $E[Y \mid G = 1] - E[Y \mid G = 0]$ can be interpreted as a causal effect.

(a) State the identification conditions required for the following equality to hold

$$E[Y^g] = E[Y \mid G = g] \ .$$

(b) Evaluate whether the assumptions hold in this study.

(c) Deduce whether the contrast $E[Y \mid G = 1] - E[Y \mid G = 0] \neq E[Y^{g=1} - Y^{g=0}]$ can be interpreted as a causal effect.

Solution:

(a)
- Consistency
- Exchangeability: $Y^g \perp\!\!\!\perp G$
- Positivity: $P(G = g) > 0$ for all $g \in \{0, 1\}$.

(b)
- Consistency holds, that is, the intervention on GRE test score is well-defined. For example changing the numerical value of the GRE score given to the admissions officers (in other words, an intervention on GRE test score can be thought of as cheating on the test, without changing the candidates underlying skills). By contrast, an intervention on A_1 would not be well-defined, because it is not clear how to change skill set A_1 without altering skill set A_2.
- Exchangeability: $Y^g \not\perp\!\!\!\perp G$ because there is an open path $G \leftarrow A_1 \rightarrow Y^g$. Intuitively, GRE is associated with performance due to the common cause A_1 of GRE scores and Y^g, and an intervention on GRE test score (changing the score displayed to admission officers) does not affect the candidates future performance, given that all candidates are admitted regardless of GRE score.
- Positivity can be checked in the data, and would be expected to hold in this case.

(c) We deduce that $E[Y \mid G = 1] - E[Y \mid G = 0] \neq E[Y^{g=1} - Y^{g=0}]$ because exchangeability fails, and thus the contrast on LHS cannot be interpreted as a causal effect.

References

[1] Proving the balancing score property of propensity score.

[2] J. M. Robins. EPI 207 (Harvard T.H. Chan School of Public Health).

[3] Liane Moneta-Koehler, Abigail M. Brown, Kimberly A. Petrie, Brent J. Evans, and Roger Chalkley. The Limitations of the GRE in Predicting Success in Biomedical Graduate School. *PLOS ONE*, 12(1):e0166742, January 2017. Publisher: Public Library of Science.