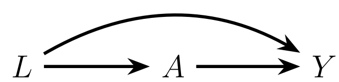
EXERCISES FOR RANDOMIZATION AND CAUSATION (MATH-336)

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 $A \perp\!\!\!\perp L \mid \pi(L)$.

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).

(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.

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

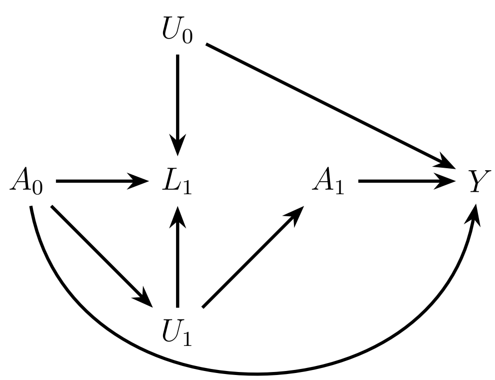


Figure 1

(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.

(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) $A_1^{a_0} = A_1$

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.

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.