EXERCISES FOR RANDOMIZATION AND CAUSATION (MATH-336)

Exercise Sheet 5

Exercise 1. For each of the following, state whether the effect is identified. If not, explain why. If so, provide a formula in terms of the observed variables and evaluate that formula for $(a_0 = 1, a_1 = 0)$, using Table 1 below (if you need to choose a value for l_1, use $l_1 = 0$).

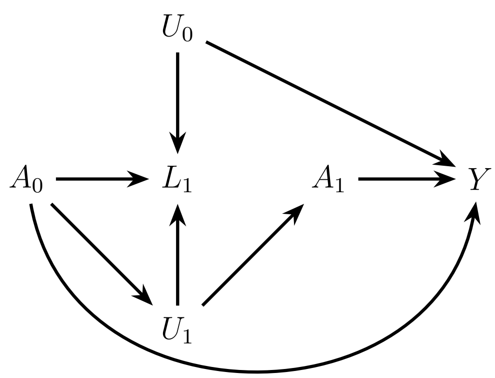


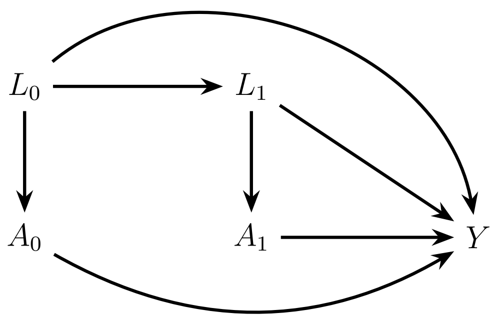
Figure 1

Table 1

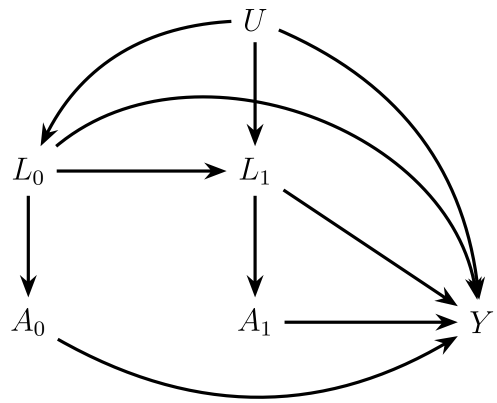
Row	A_0	L_1	A_1	N	$E[Y \mid A_0, L_1, A_1]$
1	0	0	0	6000	50
2	0	0	1	2000	70
3	0	1	0	2000	200
4	0	1	1	6000	220
5	1	0	0	3000	230
6	1	0	1	1000	250
7	1	1	0	3000	130
8	1	1	1	9000	110

Hint: Use your answers from part (b) and consistency.

(i) $E[Y^{a_0}]$
(ii) $E[Y^{a_0,a_1}]$
(iii) $E[Y^{a_1}]$
(iv) $E[Y^{a_1} \mid L_1 = l_1, A_0 = a_0]$
(v) $E[Y^{a_0} \mid A_1 = a_1, L_1 = l_1, A_0 = a_0]$
(vi) $E[Y^{a_1} \mid A_1 = a_1, L_1 = l_1, A_0 = a_0]$



$\mathcal{G}$



$\mathcal{G}_U$

(vii) $E[Y^{a_0,a_1} \mid A_1 = a_1, L_1 = l_1, A_0 = a_0]$
(viii) $E[Y^{a_0} \mid L_1^{a_0}]$
(ix) $E[Y^{a_0,a_1} \mid A_1^{a_0}]$
(x)* $E[Y^{a_1} \mid L_1^{a_0}]$
(xi)* $E[Y^{a_0} \mid L_1 = l_1, A_1 = a_1]$

Exercise 2 (Two point treatments)**.** Consider a conditionally randomized study with binary variables, in which two point treatments A_t at times $t \in \{0, 1\}$ are assigned by randomizing conditional on (possibly) time-varying covariates L_0 and L_1 (i.e. A_0 can possibly depend on L_0, and A_1 can possibly depend on L_0 and L_1).

(a) Let $\overline{L}_1 = (L_0, L_1)$. Suppose that $Y^{a_0,a_1} = Y$ for those with $A_0 = a_0$ and $A_1 = a_1$ (consistency). Also, suppose $P(A_1 = a_1 \mid \overline{L}_1 = \bar{l}_1, A_0 = a_0) > 0$ for all $a_1, a_0, \bar{l}_1$ whenever $P(\overline{L}_1 = \bar{l}_1, A_0 = a_0) > 0$ and that $P(A_0 = a_0 \mid L_0 = l_0) > 0$ for all a_0, l_0 whenever $P(L_0 = l_0) > 0$ (positivity). Explain why the following sequential exchangeability assumptions are expected to hold for all a_0, a_1 in a conditionally randomized trial:

(1) $$Y^{a_0,a_1} \perp\!\!\!\perp A_0 \mid L_0 \ ,$$

(2) $$Y^{a_0,a_1} \perp\!\!\!\perp A_1 \mid A_0, \overline{L}_1 \ .$$

Prove that under the above assumptions,

$$E[Y^{a_0,a_1}] = E\left[\frac{I(A_0 = a_0, A_1 = a_1)Y}{P(A_1 = a_1 \mid \overline{L}_1, A_0 = a_0)P(A_0 = a_0 \mid L_0)}\right] \ .$$

(b) Draw the SWIG $\mathcal{G}(a_0, a_1)$, corresponding to the intervention where we assign treatment a_0 at time 0 and a_1 at time 1 for the causal DAG $\mathcal{G}$, and show that $\mathcal{G}(a_0, a_1)$ is an example of a SWIG satisfying the exchangeability conditions in part (a).

Show that under the causal model $\mathcal{G}$, we can further simplify the identification formulas as:

$$E[Y^{a_0,a_1}] = E\left[\frac{I(A_0 = a_0, A_1 = a_1)Y}{P(A_1 = a_1 \mid L_1)P(A_0 = a_0 \mid L_0)}\right] \ .$$

(c) Draw the SWIG $\mathcal{G}_U(a_0, a_1)$ corresponding to the DAG $\mathcal{G}_U$. Does $\mathcal{G}_U(a_0, a_1)$ satisfy the exchangeability conditions in part (a) (and is $E[Y^{a_0,a_1}]$ thus identified by the formula in part (a))? Is $E[Y^{a_0,a_1}]$ identified by the formula in part (b)?

We will now study the more general setting of an observational study, where a clinician assigns treatment A_t according to a decision rule based on A_{t-1} and $\overline{L}_t$ (take $A_{-1} = \emptyset$, so A_0 is only assigned based on L_0). As as example, consider survival Y (0 indicates survival, 1 indicates death) after randomizing patients to standard antibiotics versus broad-spectrum antibiotics following hospitalization with a bacterial infection.[1] In the observed data, the clinician decides whether to start a standard antibiotic treatment regime ($A_0 = 0$) or use broad-spectrum antibiotics ($A_0 = 1$) depending on the medical history, clinical examination and test results (L_0) on admission at time 0.[2] Based on updated findings L_1 at a later time 1, the clinician decides whether to continue with standard antibiotics or switch to broad-spectrum antibiotics ($A_1 = 0$ versus $A_1 = 1$).

(d) Suppose that consistency and positivity hold, as in part (a). Next, *assume* that the following conditional exchangeability conditions hold:[3]

$$Y^{a_0,a_1} \perp\!\!\!\perp A_0 \mid L_0^{a_0} \ , \tag{3}$$

$$Y^{a_0,a_1} \perp\!\!\!\perp A_1^{a_0} \mid A_0, \overline{L}_1^{a_0} \ . \tag{4}$$

Show that the identification formula for $E[Y^{a_0,a_1}]$ is the same as in part (a).

(e) Using consistency, positivity and conditional exchangeability (Eqs. 3-4), prove that $E[Y^{a_0,a_1}]$ is identified by

$$E[Y^{a_0,a_1}] = \sum_{l_0,l_1,y} y \cdot p(y \mid a_0, a_1, l_0, l_1) p(l_1 \mid a_0, l_0) p(l_0) \ .$$

The right hand side is an example of a g-formula. A g-formula is defined as a functional of the observed data distributions, but it is a valid identification formula for a causal estimand if conditional exchangeability, consistency and positivity hold.

Hint: Use conditional exchangeability and the law of total probability in alternation to add A's and L's to the conditioning set of $P(Y^{a_0,a_1})$, starting at time 0.

(f) Show that the complete SWIG $\mathcal{G}_c(a_0, a_1)$ is an example of a causal model satisfying Eqs. 3-4.

Finally, we will consider a type of dynamic intervention. Consider the same story as in parts (d)–(f), except that we suppose A_0 to be randomized at time 0 (and thus we can take $L_0 = \emptyset$). Suppose for simplicity that clinicians decide on treatment A_1 based on the patient's blood test result for C-reactive protein (a biomarker which is positively correlated with the

[1] If the infection is widespread across the population, we would no longer have independent Y_i across individuals i, i.e. a violation of the no interference assumption. In this case, we need to use a different approach.

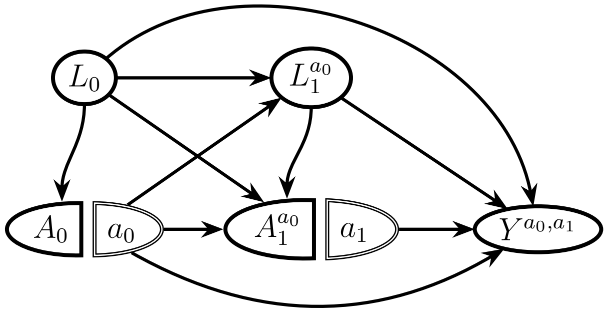
[2] Typically, clinicians avoid starting with broad-spectrum antibiotics immediately if possible, because frequent use of such antibiotics leads to antibiotic resistance. The idea is to reserve such treatments for 'backup' use, in case standard treatments fail.

[3] Using consistency, we can rewrite the Eqs. 3-4 as

$$Y^{a_0,a_1} \perp\!\!\!\perp A_0 \mid L_0 \ ,$$

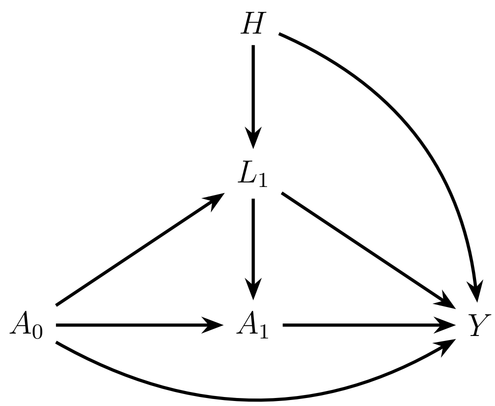
$$Y^{a_0,a_1} \perp\!\!\!\perp A_1 \mid A_0 = a_0, \overline{L}_1 \ ,$$

which differ subtly from Eqs. 1-2 in that A_0 is instantiated to the value a_0.



$\mathcal{G}_c(a_0, a_1)$

severity of the infection), which we denote by L_1. Let $L_1 = 0$ indicate low score and 1 indicate a high score, and suppose our system can be described by the causal model $\mathcal{G}_d$.



$\mathcal{G}_d$

We want to study a hypothetical intervention where the initial treatment A_0 is deterministically set to a_0 (for example, always start with standard antibiotics) and the subsequent treatment A_1 is assigned according to the decision rule g given by $A_1^{g+} = g(L_1^{a_0}) := L_1^{a_0}$.

(g) Draw the d-SWIG $\mathcal{G}_d(g)$ corresponding to the above intervention on $\mathcal{G}_d$.
(h) For the dynamic intervention considered here, the conditional exchangeability condition becomes

$$Y^g \perp\!\!\!\perp A_0 \ ,$$
$$Y^g \perp\!\!\!\perp A_1^{a_0} \mid L_1^{a_0}, A_0 \ ,$$

and the consistency condition becomes

$$Y^g = Y \quad \text{whenever} \quad \overline{A}_1 = \overline{A}_1^{g+} \ .$$

Use this to show that the expected outcome under the the intervention g is given by[4]

$$E[Y^g] = \sum_y \sum_{l_1} \sum_{a_0'} \sum_{a_1'} y \cdot p(y \mid a_0', a_1', l_1) p(l_1 \mid a_0') p^g(a_1' \mid l_1) p^g(a_0') \ ,$$

where $p^g(a_1' \mid l_1) = I(a_1' = l_1)$ and $p^g(a_0') = I(a_0' = a_0)$ are the interventional distributions.

Hint: The consistency condition can be written as[5]

$$Y^g = Y \quad \text{whenever} \quad A_0 = a_0 \text{ and } A_1 = L_1 \ .$$

(i) Using part (h), show further that

$$E[Y^g] = \sum_{l_1} E[Y \mid A_1 = l_1, L_1 = l_1, A_0 = a_0] P(L_1 = l_1 \mid A_0 = a_0) \ .$$

(j) The optimal regime g (which maximizes survival) subject to $\mathcal{G}_d(g)$ is given by

$$\arg\min_{g \in \mathbb{G}} E[Y^g] \ ,$$

where $\mathbb{G} = \{g : \{0,1\} \to \{0,1\}\}$ is the set of functions from $\{0,1\}$ to $\{0,1\}$. List the members of $\mathbb{G}$ and determine how many decision rules g we must decide between to find the optimal regime in this example.

References

[4] This is an example of a (marginal) *extended* g-formula.

[5] Let $C = \{\omega : A_0 = a_0, A_1(\omega) = L_1(\omega)\}$. Consistency can be written more elaborately as $Y^g(\omega) = Y(\omega)$ whenever $\omega \in C$ for some unit or individual $\omega \in \Omega$.