RANDOMIZATION AND CAUSATION (MATH-336)

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Mock exam - 2025 - solutions

Date: 26th of June, 2025
Time: 08:15–10:00

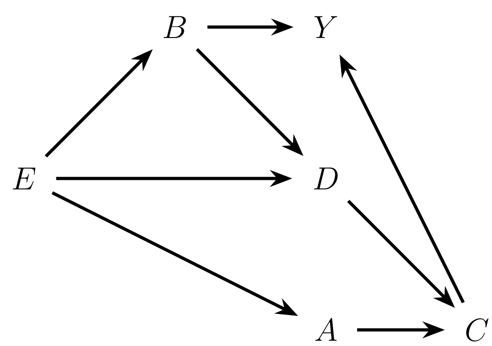
Name: ________________________

SCIPER: ________________________

INSTRUCTIONS TO CANDIDATES

- This is a practice exam, so it does not contribute to your final grade.

Question 1.
Consider the DAG below.



Use the rules of d-separation to decide whether the following independencies hold. Justify your answer:

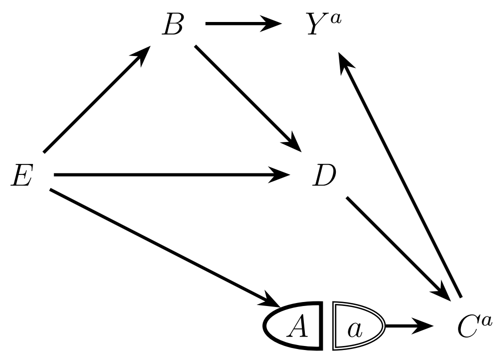
(1) $D \perp\!\!\!\perp A | E$,
(2) $Y \perp\!\!\!\perp E | B, C$
(3) $Y \perp\!\!\!\perp A | C, E$

Suppose we are interested in the causal effect of A on Y and C. Are the following statements true or false? Justify your answer.

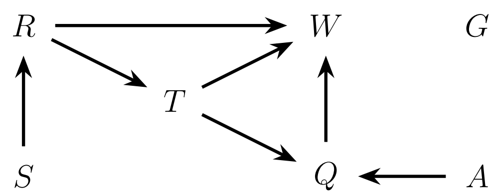
(4) $C^a \perp\!\!\!\perp A | D, Y^a$
(5) $Y^a \perp\!\!\!\perp A | E$

Solution:

(1) True, all paths from D to A are blocked conditional on E.
(2) True, all paths from Y to E are blocked conditional on B and C.
(3) False, $Y \leftarrow B \rightarrow D \rightarrow C \leftarrow A$ is an open path from Y to A conditional on C and E.
(4) False, $A \leftarrow E \rightarrow D \leftarrow B \rightarrow Y^a \leftarrow C^a$ is an open path in the SWIG below.



(5) True, all paths from Y^a to A are blocked conditional on E in the SWIG.



Question 2.

In the DAG above, R represents area of residence, S represents socioeconomic status, T represents traffic noise, W represents subjective well-being, Q represents sleep quality, G represents sex and A represents age.

(1) Which statements are correct, supposing that all regression models are correctly specified and all relationships between variables are linear, without interactions? Justify your answers.

(a) In a regression of ‘subjective well-being’ on ‘age’ and ‘sex’, we expect both coefficients to be zero.

(b) The coefficient for ‘traffic noise’ in a linear regression of ‘sleep quality’ on ‘socioeconomic status’, ‘traffic noise’, and ‘subjective well-being’ has a causal interpretation. *Note: We say the coefficient has a "causal interpretation" if it can be interpreted as the difference between the expectation of two potential outcomes in the same population.*

(c) If we regress ‘sleep quality’ on ‘age’ and obtain a coefficient of about zero, we know that the DAG is wrong.

(2) Suppose we are interested in the total causal effect of ‘sleep quality’ on ‘subjective well-being’. For what confounders do we need to adjust? How could the desired effect be obtained?

Solution:

(1) (a) This statement is false. W is d-connected to A, both with and without conditioning on G. The coefficient for A can therefore not be expected to be zero.

(b) This statement is false. Conditioning on W opens a collider on the path $T \leftarrow R \rightarrow W \leftarrow Q$.

(c) This statement is false, unless faithfulness has been assumed. A DAG with an open path such as $A \rightarrow Q$ does not require A and Q to be independent, but neither does it rule out such independence. If a faithfulness assumption is made, the DAG can be falsified by this empirical finding of marginal independence.

(2) We would need to adjust for T in order to block all backdoor paths between W and Q. The desired effect can then be obtained using the g-formula, $E(W^q) = \sum_t E(W|Q = q, T = t) \times P(T = t)$

Question 3.

Motivation (not strictly needed to answer the question). Consider a physician who meets a patient diagnosed with localized breast cancer. The tumor is small and it grows slowly, so the physician believes the patient has a good chance of never relapsing if treated with standard chemotherapy.

However, the doctor also knows that a new treatment, immunotherapy, will be launched to treat this type of tumor. At a medical conference, the doctor heard that this new treatment is extremely effective and likely outperforms standard chemotherapy. However, the doctor heard that there is a small subset of patients in whom, for unknown reasons, the use of immunotherapy could actually increase tumor growth and prevent recovery. Remembering the oath he took at the end of medical school, "First, do no harm" the physician hesitates to prescribe the new drug to the patient.

To help guide his decision, the physician conducts a literature review and finds a randomized controlled experiment study that examined the effect of immunotherapy *versus* standard chemotherapy on breast cancer relapse. The physician then visits two of the most respected statisticians in the country, presents the results of the study, and asks the statisticians to quantify the "likelihood" that his patient will be harmed by immunotherapy. However, the two statisticians used two different mathematical definitions of harm and apparently came up with two opposite guidelines for the physician. What should the physician do?

Notation and assumptions. This question will concern data from a randomized controlled experiment. We let A be treatment received, binary ($A = 1$ is immunotherapy, $A = 0$ is standard chemotherapy), Y is the binary outcome ($Y = 1$ is relapse or death after five years, $Y = 0$ is no relapse or death after five years). As usual, we let Y^a be the potential outcome under treatment $A = a$.

We make the following assumptions:

$$E[Y^a|A = a] = E[Y|A = a] \text{ for } a \in \{0, 1\} \tag{A1}$$

$$Y^a \perp\!\!\!\perp A \quad \text{for } a \in \{0, 1\} \tag{A2}$$

$$Pr(A = a) > 0 \text{ for } a \in \{0, 1\} \tag{A3}$$

In this question, you will study two common mathematical definitions of harm and their properties: these are called the interventionist and the counterfactual definition of harm.

In the interventionist approach, we say that assigning treatment $A = 1$ does harm if

$$Pr(Y^{a=1} = 1) - Pr(Y^{a=0} = 1) > 0. \tag{D1}$$

In the counterfactual approach, we say that assigning treatment $A = 1$ does harm if

$$Pr(Y^{a=1} = 1, Y^{a=0} = 0) > 0. \tag{D2}$$

(1) Using the assumptions listed above, express $Pr(Y^{a=1} = 1) - Pr(Y^{a=0} = 1)$ as a function of observed variables only.

(2) In general, can you point identify $Pr(Y^{a=1} = 1, Y^{a=0} = 0)$ with the assumptions above? If yes, give an expression; if no, explain why.

(3) For the rest of the question, let $Y^{a=1} := Y^1$ and $Y^{a=0} := Y^0$ and denote

$$\tau := Pr(Y^1 = 1) - Pr(Y^0 = 1) \tag{D3}$$
$$\rho := Pr(Y^1 = 1) - Pr(Y^0 = 0). \tag{D4}$$

Express all entries of the so-called transition matrix

$$T = \begin{bmatrix} Pr(Y^1 = 1) & Pr(Y^1 = 0) \\ Pr(Y^0 = 1) & Pr(Y^0 = 0) \end{bmatrix}$$

as functions of τ and ρ only.

(4) Give a necessary and sufficient condition involving only τ and ρ that ensures all entries in T are non-negative.

(5) Next, denote ξ as

$$\xi := Pr(Y^0 = 0, Y^1 = 1) + Pr(Y^0 = 1, Y^1 = 0). \tag{D5}$$

Express the joint distribution (Y^1, Y^0) as a function of τ, ρ, and ξ; that is, express every entry in the following matrix

$$P = \begin{bmatrix} Pr(Y^0 = 1, Y^1 = 1) & Pr(Y^0 = 0, Y^1 = 1) \\ Pr(Y^0 = 1, Y^1 = 0) & Pr(Y^0 = 0, Y^1 = 0) \end{bmatrix}$$

as a function of τ, ρ, and ξ only.

(6) Show that there exist two functions $l(\tau, \rho)$ and $u(\tau, \rho)$, which are functions of only τ and ρ, such that the following condition is satisfied:
"All entries of the matrix P are non-negative if and only if $l(\tau, \rho) \leq \xi \leq u(\tau, \rho)$."
You need to provide explicit formulations of $l(\tau, \rho)$ and $u(\tau, \rho)$, as functions of τ and ρ only.

(7) Deduce lower and upper bounds for $Pr(Y^{a=1} = 1, Y^{a=0} = 0)$ as functions of τ and ρ. We will denote these lower and upper bounds as $\mathcal{L}_{R=1}$ and $\mathcal{U}_{R=1}$, respectively, because they were computed using experimental data only.

(8) Under which condition on the entries of T is $Pr(Y^{a=1} = 1, Y^{a=0} = 0)$ point identified, that is $\mathcal{L}_{R=1} = \mathcal{U}_{R=1}$?

(9) Prove that $\mathcal{L}_{R=1} > 0$ if and only if $\tau > 0$.

(10) Suppose that, under the assumptions A1-A3, we detect counterfactual harm. Do we then also detect interventionist harm? Conversely, suppose we detect interventionist harm. Do we then also detect counterfactual harm? Justify your answers.

Solutions:

(1) We have

$$\begin{aligned} Pr(Y^{a=1}=1) - Pr(Y^{a=0}=1) &= Pr(Y^{a=1}=1|A=1) - Pr(Y^{a=0}=1|A=0) && \text{(A2, A3)} \\ &= Pr(Y=1|A=1) - Pr(Y=1|A=0,), && \text{(A1)} \end{aligned}$$

which concludes the identification proof.

(2) No, in general we cannot identify this probability. Indeed, $Y^{a=1}$ and $Y^{a=0}$ can never be observed simultaneously for a patient : there are cross-world. In order to identify quantities related to the joint of the two counterfactuals, we require independence assumptions relating counterfactuals under different interventions, i.e. in this case relating $Y^{a=0}$ and $Y^{a=1}$. Such assumptions are strong, cannot be verified empirically in a real life experiment, and were not made here.

(3) We have

$$\tau + \rho = 2 \cdot Pr(Y^1 = 1) - Pr(Y^0 = 1) - Pr(Y^0 = 0) = 2 \cdot Pr(Y^1 = 1) - 1,$$

so that $Pr(Y^1 = 1) = \frac{1}{2}(1 + \tau + \rho)$, from which we deduce

$$Pr(Y^1 = 0) = 1 - Pr(Y^1 = 1) = \frac{1}{2}(1 - \tau - \rho)$$

Similary

$$\tau - \rho = -Pr(Y^0 = 1) + Pr(Y^0 = 0) = 1 - 2 \cdot Pr(Y^0 = 1),$$

so that $Pr(Y^0 = 1) = \frac{1}{2}(1 - \tau + \rho)$, from which we deduce

$$Pr(Y^0 = 0) = 1 - Pr(Y^0 = 1) = \frac{1}{2}(1 + \tau - \rho)$$

Finally,

$$T = \begin{bmatrix} Pr(Y^1 = 1) & Pr(Y^1 = 0) \\ Pr(Y^0 = 1) & Pr(Y^0 = 0) \end{bmatrix} = \begin{bmatrix} \frac{1}{2}(1 + \tau + \rho) & \frac{1}{2}(1 - \tau - \rho) \\ \frac{1}{2}(1 - \tau + \rho) & \frac{1}{2}(1 + \tau - \rho) \end{bmatrix}$$

(4) **Necessary condition:** We need all of $\begin{cases} 1 \geq -\tau - \rho \\ 1 \geq \tau - \rho \\ 1 \geq \tau + \rho \\ 1 \geq -\tau + \rho \end{cases}$ which implies that

$$1 \geq |\tau| + |\rho|.$$

Let's prove that this is a sufficient condition. **Sufficient condition:** Since $\forall x \in \mathbb{R}, |x| \geq x$ and $|x| \geq -x$
it is straighforward to show that if $|\tau| + |\rho| \leq 1$, then, *a fortiori* :

$$\tau + \rho \leq 1; -\tau + \rho \leq 1; -\tau - \rho \leq 1; \tau - \rho \leq 1.$$

Finally, all entries of T to be non-negative if and only if $|\tau| + |\rho| \leq 1$

(5) We have :

$$Pr(Y^1 = 1) = Pr(Y^0 = 0, Y^1 = 1) + Pr(Y^0 = 1, Y^1 = 1) \qquad \text{(LOTP)}$$

and

$$Pr(Y^0 = 1) = Pr(Y^0 = 1, Y^1 = 0) + Pr(Y^0 = 1, Y^1 = 1) \qquad \text{(LOTP)}$$

so that

$$\begin{aligned}
\tau &= Pr(Y^1 = 1) - Pr(Y^0 = 1) \\
&= Pr(Y^0 = 0, Y^1 = 1) + Pr(Y^0 = 1, Y^1 = 1) - Pr(Y^0 = 1) \\
&= \xi - Pr(Y^0 = 1, Y^1 = 0) + Pr(Y^0 = 1, Y^1 = 1) - Pr(Y^0 = 1) \\
&= \xi - Pr(Y^0 = 1, Y^1 = 0) + Pr(Y^0 = 1) - Pr(Y^0 = 1, Y^1 = 0) - Pr(Y^0 = 1) \\
&= \xi - 2 \cdot Pr(Y^0 = 1, Y^1 = 0),
\end{aligned}$$

so that $Pr(Y^0 = 1, Y^1 = 0) = \frac{1}{2}(\xi - \tau)$ From that, and from the expressions in matrix T, we deduce easily that :

$$\begin{aligned}
Pr(Y^0 = 0, Y^1 = 0) &= Pr(Y^1 = 0) - Pr(Y^0 = 1, Y^1 = 0) \\
&= \frac{1}{2}(1 - \tau - \rho) - \frac{1}{2}(\xi - \tau) \\
&= \frac{1}{2}(1 - \rho - \xi)
\end{aligned}$$

$$\begin{aligned}
Pr(Y^0 = 1, Y^1 = 1) &= Pr(Y^1 = 1) - Pr(Y^0 = 0, Y^1 = 1) \\
&= Pr(Y^1 = 1) - Pr(Y^0 = 0) + Pr(Y^0 = 0, Y^1 = 0) \\
&= \rho + \frac{1}{2}(1 - \rho - \xi) \\
&= \frac{1}{2}(1 + \rho - \xi)
\end{aligned}$$

and

$$\begin{aligned}
Pr(Y^0 = 0, Y^1 = 1) &= Pr(Y^1 = 1) - Pr(Y^0 = 1, Y^1 = 1) \\
&= \frac{1}{2}(1 + \tau + \rho) - \frac{1}{2}(1 + \rho - \xi) \\
&= \frac{1}{2}(\tau + \xi)
\end{aligned}$$

Finally,

$$P = \begin{bmatrix} Pr(Y^0 = 1, Y^1 = 1) & Pr(Y^0 = 0, Y^1 = 1) \\ Pr(Y^0 = 1, Y^1 = 0) & Pr(Y^0 = 0, Y^1 = 0) \end{bmatrix} = \begin{bmatrix} \frac{1}{2}(1 + \rho - \xi) & \frac{1}{2}(\xi + \tau) \\ \frac{1}{2}(\xi - \tau) & \frac{1}{2}(1 - \rho - \xi) \end{bmatrix}$$

(6) **Necessary condition:** We need all of $\begin{cases} \xi \leq 1-\rho \\ \xi \leq 1+\rho \\ \xi \geq \tau \\ \xi \geq -\tau \end{cases}$ so that we need

$$\xi \leq \min(1-\rho, 1+\rho) = 1-|\rho|$$

and

$$\xi \geq \max(-\tau, \tau) = |\tau|$$

Finally, we need

$$|\tau| \leq \xi \leq 1-|\rho|.$$

Let's prove that this is a sufficient condition.

Sufficient condition: Straightforward since $\forall x \in \mathbb{R}, |x| \geq x$ and $|x| \geq -x$.

Finally, all entries of P to be non-negative if and only if $|\tau| \leq \xi \leq 1-|\rho|$.

(7) Since $Pr(Y^0 = 0, Y^1 = 1) = \frac{1}{2}(\xi + \tau)$, and it must be that $|\tau| \leq \xi \leq 1-|\rho|$, we have that

$$\mathcal{L}_{R=1} := \frac{1}{2}(|\tau| + \tau) \leq Pr(Y^0 = 0, Y^1 = 1) \leq \frac{1}{2}(1-|\rho| + \tau) := \mathcal{U}_{R=1}$$

(8)

$$\mathcal{L}_{R=1} = \mathcal{U}_{R=1} \Leftrightarrow |\tau| + \tau = 1-|\rho| + \tau \Leftrightarrow |\tau| = 1-|\rho| \Leftrightarrow |\rho| + |\tau| = 1$$

This is true if and only if one of $1+\tau+\rho$, $1+\tau-\rho$, $1-\tau+\rho$, or $1-\tau-\rho$ is zero, which means that one of $Pr(Y^1 = 1)$ or $Pr(Y^0 = 1)$ should be either 0 or 1. In other words, for at least one of the interventions ($A = 1$ or $A = 0$), the outcome Y can be predicted with certainty (which is very unlikely).

(9) The lower bound could be re-written as

$$\mathcal{L}_{R=1} = \frac{1}{2}(|\tau| + \tau) = \max(0, \tau).$$

From that, it is straightforward to show that :

$$\mathcal{L}_{R=1} > 0 \Leftrightarrow \max(0, \tau) > 0 \Leftrightarrow \tau > 0$$

(10) When only experimental data is available, the criterion for interventional harm is point identified, while, in general, we can only get bounds of the criterion for counterfactual harm. Furthermore, we are able to detect harm with the counterfactual approach, if and only if the lower bound we get is strictly positive, which is the case if an only if the ATE was strictly positive, that if we were already able to detect harm with the interventionist approach. That is, when only experimental data is available, counterfactual approach is unnecessary.