RANDOMIZATION AND CAUSATION (MATH-336)

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Mock exam - 2025 - questions

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

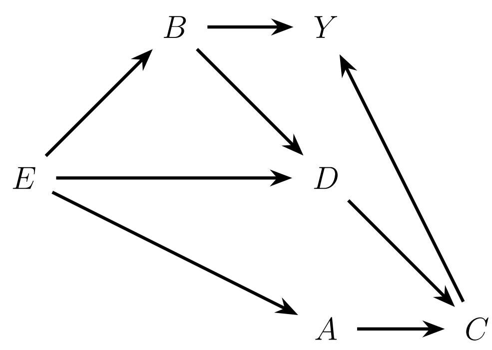
Name: ________________________

SCIPER: ________________________

INSTRUCTIONS TO CANDIDATES

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

Question 1.
Consider the DAG below.

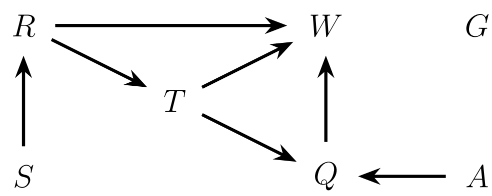


Use the rules of d-separation to decide whether the following independencies hold. Justify your answer:

(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) $Y^a \perp\!\!\!\perp A | E$



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?

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.