

Practice Question 1

Causal Inference

E P Swens

A. Sketch the Directed Acyclic Graph (DAG)

Sketch (on paper or in R) the following DAG, representing our beliefs that:

- x_1 causes x_2
- x_1 causes x_3
- x_2 causes x_3

B. Identify the number of open path(s)

How many open path(s) are there?

- A) 0
- B) 1
- C) 2
- D) 3

C. Identify the number of backdoor path(s)

How many backdoor path(s) are there?

- A) 0
- B) 1
- C) 2
- D) 3

D. What is the valid adjustment set?

What are the valid adjustment set(s)? (Multiple solutions are possible)

- A) \emptyset
- B) $\{X_2\}$
- C) $\{X_1, X_2\}$
- D) $\{X_2, X_3\}$

E. Simulate the Data

Simulate the data ($n = 1000$ with `set.seed(1)`) from the structural equations:

$$\begin{aligned}X_1 &\sim \mathcal{N}(0, 1) \\X_2 &\sim 2X_1 + \mathcal{N}(0, 1) \\X_3 &\sim X_1 - 0.5X_2 + \mathcal{N}(0, 1)\end{aligned}$$

where $\epsilon_1, \epsilon_2, \epsilon_3 \sim \mathcal{N}(0, 1)$ (i.i.d.)

F. Test the Correlations

Test the correlations between the variables use $\alpha = 0.05$. Tip: You can use `cor.test` from (`stats`) to test marginal dependencies and `pcor.test` from (`ppcor`) to test conditional dependencies.

G. Recover the DAG

Which correlation(s) are (in)significant from part C? What structure can you infer from the results?

- A) Mediator
- B) Collider
- C) Confounder

H. Recovered DAG from PC Algorithm

Figure 1 (next page) shows that the `PC-algorithm` did not recover the DAG. Which model assumption is not met?

- A) Positivity
- B) Exchangeability
- C) Faithfulness
- D) Sufficiency
- E) Strong ignorability

I. Why does the algorithm require this assumption?

J. In general, we do not worry about violation of this assumption, why?

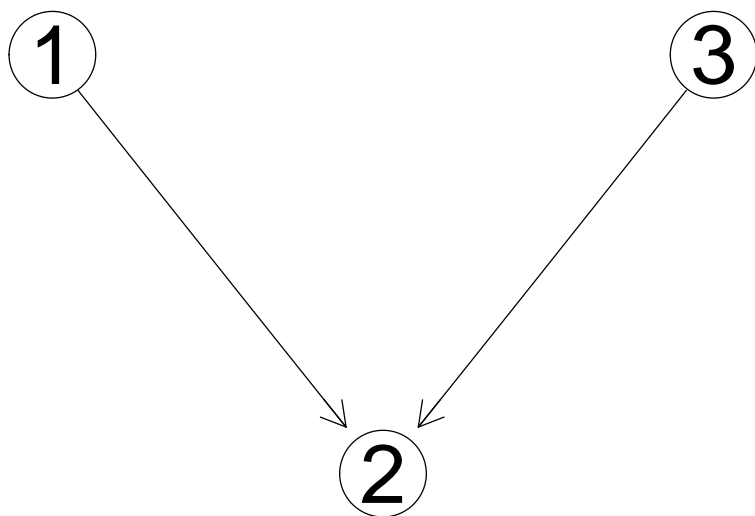


Figure 1: DAG Recovered by PC Algorithm