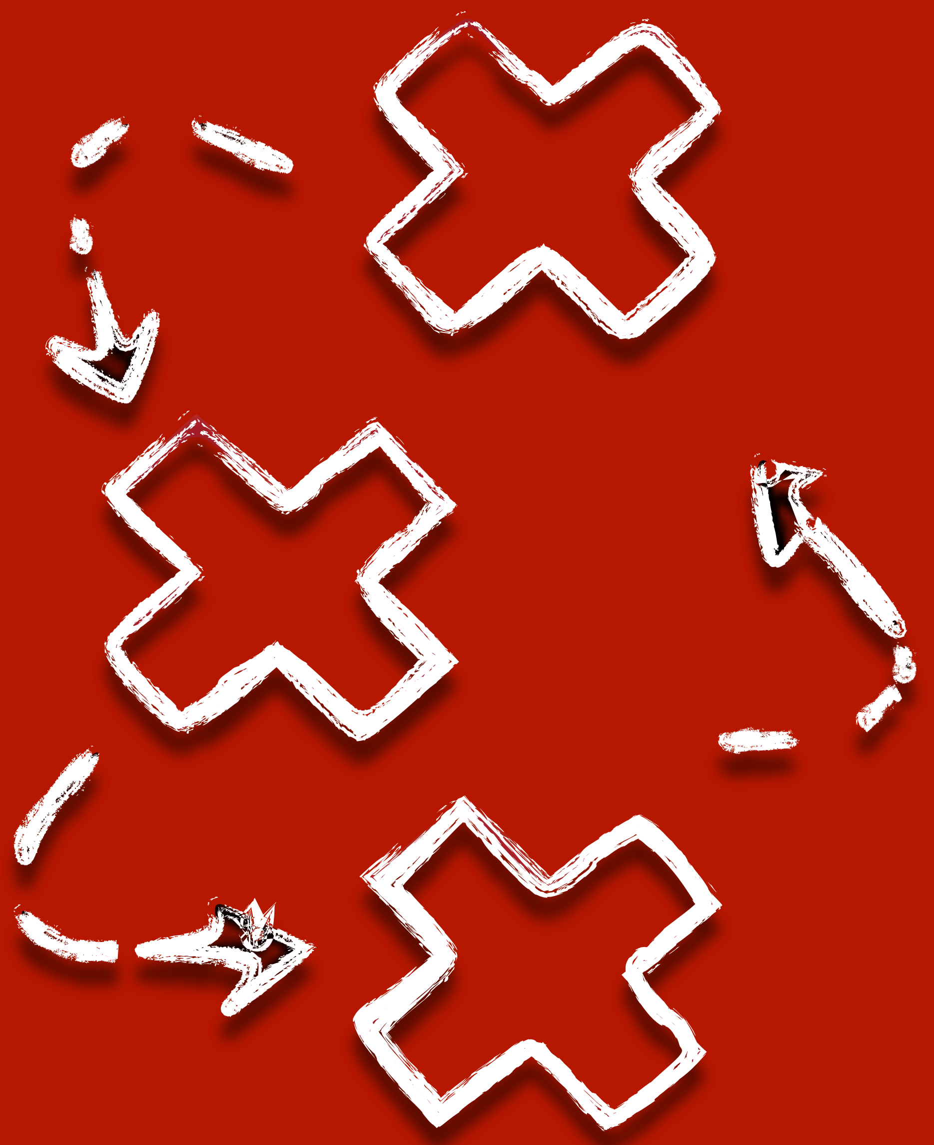


Causal Data Science

Lecture 10.0: Recap of 9 and solutions

Lecturer: Sara Magliacane

UvA - Spring 2022

Overview on where we are in the course

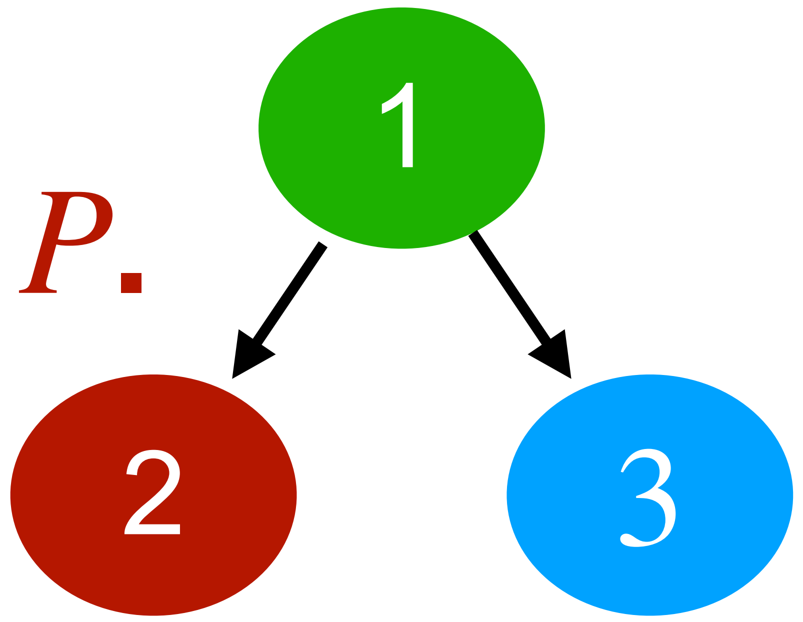
7/02/2022	Introduction
10/02/2022	Probability recap
14/02/2022	Causal graphs
17/02/2022	Interventions
21/02/2022	Covariate adjustment
24/02/2022	Frontdoor criterion, Instrumental variables
28/02/2022	Counterfactuals and potential outcomes
3/03/2022	Estimating causal effects, Missing data
7/03/2022	Constraint based structure learning
10/03/2022	Score based structure learning
14/03/2022	Advanced structure learning and transportability
17/03/2022	Causality-inspired ML

Background on causal graphs

We know the causal graph, how do we estimate causal effects?

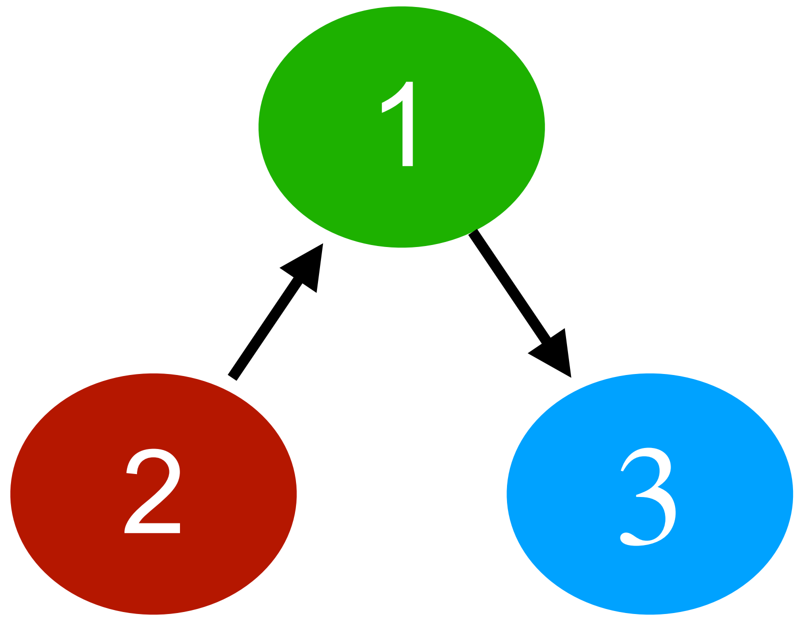
What happens if the graph is unknown?

Last class: Perfect maps - Markov equivalence



- If P is Markov and faithful to G, we say that **G is a perfect map of P.** Then, for any disjoint $\mathbf{A}, \mathbf{B}, \mathbf{C} \subseteq \mathbf{V}$:

$$\mathbf{A} \perp \mathbf{B} \,|\, \mathbf{C} \iff X_{\mathbf{A}} \perp\!\!\!\perp X_{\mathbf{B}} \,|\, X_{\mathbf{C}}$$



- In general there are multiple DAGs that can describe the same d-separations (and independences)

- We call these DAGs **Markov equivalent** and we **cannot distinguish them from observational data alone** (or without further assumptions)

Last class: Markov equivalence class and CPDAGs

- (Verma and Pearl 1990) show that all DAGs in a Markov equivalence class have the **same skeleton** and the **same v-structures**
- We can represent the skeleton and the orientations (edge marks) all DAGs in a Markov equivalence class (MEC) have in common with a **Complete Partially Directed Acyclic Graph (CPDAG):**
  - We have a directed edge $i \to j$ if all DAGs in the MEC have $i \to j$
  - We have an undirected edge $i - j$ if some DAGs in the MEC have $i \to j$ and others have $j \to i$

Last class: Constraint-based causal discovery

- **Idea:** we perform conditional independence tests on **observational** data and use them to constrain the possible graphs using d-separation
- In general, we can narrow down the possible graphs only up to their **Markov equivalence class (MEC)**
- The output of the algorithms we will see (e.g. SGS, PC) is a **CPDAG**, a mixed graph in which directed edges represent causal relations on which all DAGs in the MEC agree - these relations are **identifiable**

Last class: SGS algorithm (Spirtes, Glymour, Scheines)

- Assuming P is Markov and faithful to an unknown graph G

- We can estimate a CPDAG from samples of P in three steps:
  1. Determine the **skeleton**
  2. Determine the **v-structures**
  3. Direct as many remaining edges as possible

- **Note**: the directed parts of the CPDAG will agree with G, but some parts might stay undirected

Last class: SGS step 1: Skeleton learning

1. Start with **completely connected undirected graph** U

2. For each pair $i, j \in \mathbf{V}, i \neq j$, and for any subset $\mathbf{S} \subseteq \mathbf{V} \backslash \{i, j\}$
   - Check **if** $X_i \perp\!\!\!\perp X_j | X_{\mathbf{S}}$ for any $\mathbf{S}$ in data
     - If this is true, by faithfulness $i \perp_G j | \mathbf{S}$, so we can **remove** $i - j$ **in** U

Last class: SGS step 2: Determine v-structures

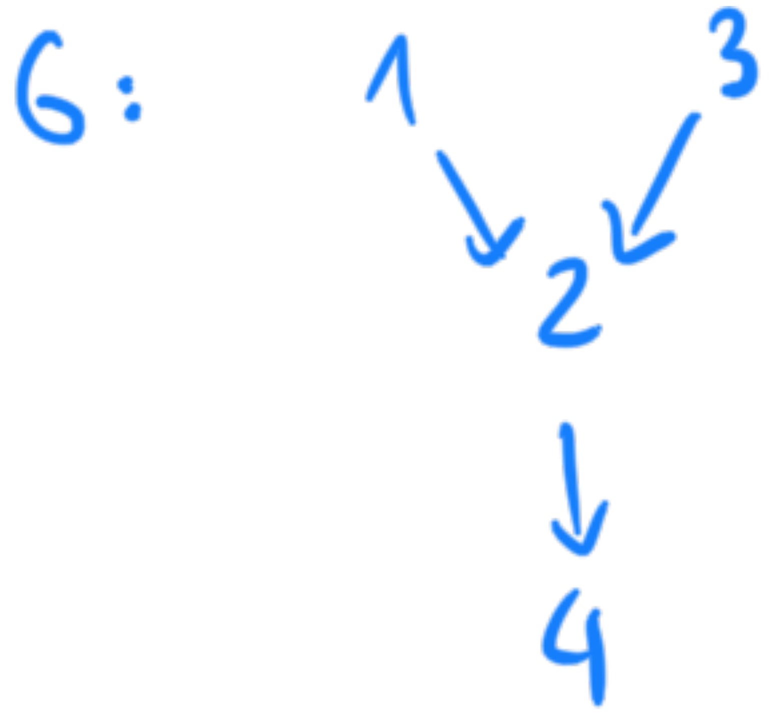
1. Start from the skeleton U from previous step

2. For each unshielded triple (i, j, k) in U, i.e. $\forall i, j, k \in \mathbf{V}$ such that $i - j, j - k$ **and** $i \not- k$ **in** U

- For all $\mathbf{S} \subseteq \mathbf{V} \setminus \{i, j, k\}$ check if $X_i \not\perp\!\!\!\perp X_k \mid X_j \cup X_{\mathbf{S}}$ in data
  - If this is true, $i \rightarrow j \leftarrow k$ **is a v-structure**

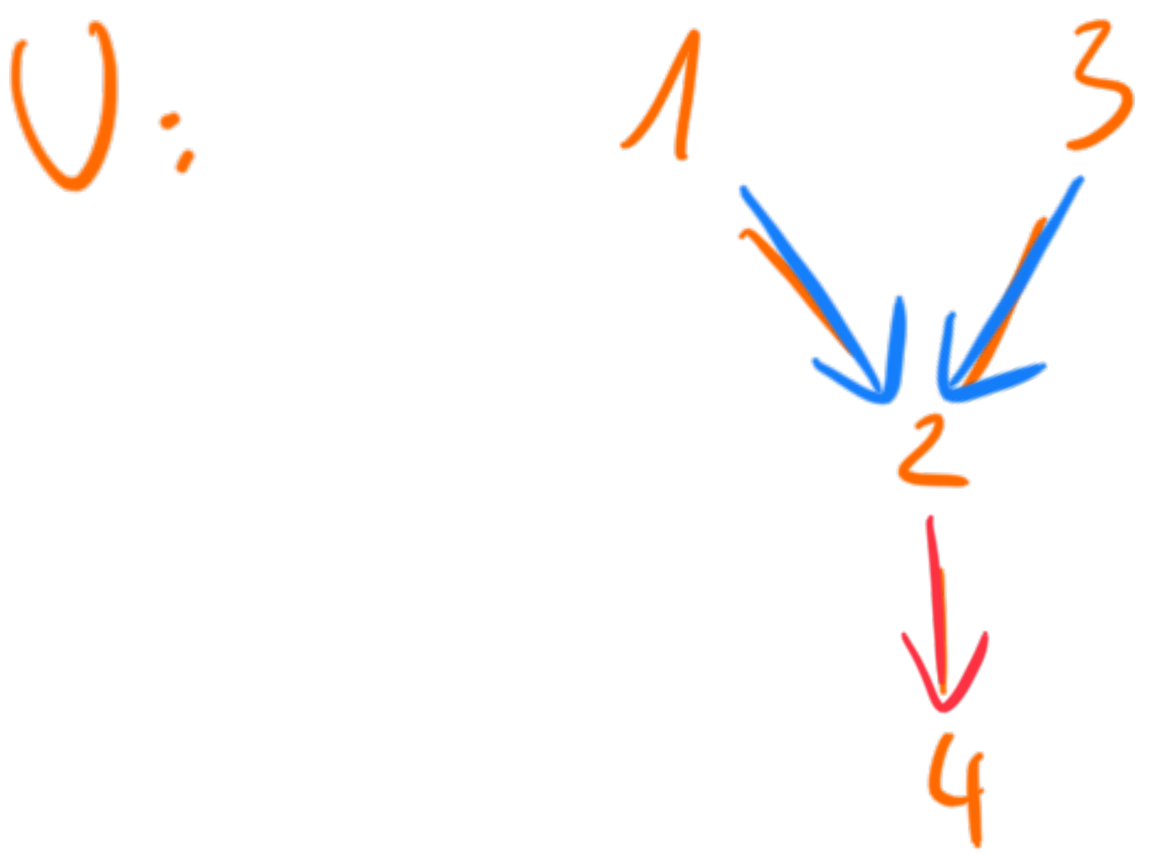
Last class: SGS step 3: Direct as many edges as possible

- Cannot create cycles or new v-structures
- Some of the edges can be oriented to disallow these situations to happen



P: $P(X_1)\cdot P(X_3)\cdot P(X_2|X_1,X_3)P(X_4|X_2)$

X_1	X_2	X_3	X_4
1	1	1	0
1	0	0	1
0	1	0	1
1	0	1	0
1	0	0	0



Last class: PC algorithm (Peter Spirtes, Clark Glymour)

- We can estimate a CPDAG from samples of P in three steps:
  1. Determine the **skeleton in an optimised way**
  2. Determine the **v-structures**
  3. Direct as many remaining edges as possible

1. Start with **completely connected undirected graph U**
2. For $k = 0,1,2\ldots, p-2$
   - If $i - j$ in U and there exists a set $\mathbf{S} \subseteq \text{Adj}(i) \cup \text{Adj}(j)$ of size k for which $X_i \perp\!\!\!\perp X_j \mid \mathbf{S}$, **remove $i - j$ in U**

Last class: PC algorithm - when does it fail?

- If the conditional independence tests give the wrong result
  - Too few samples
  - A very weak dependence
  - Wrong parametric assumption (e.g. partial correlation on nonlinear data)
- If there are unmeasured confounders or selection bias **(causal sufficiency)**

Main advantage of constraint-based methods

- More advanced algorithms like Fast Causal Inference (FCI)
  - Chapter 6 in [SGS] Causation Prediction and Search
    - https://www.researchgate.net/publication/242448131_Causation_Prediction_and_Search

Canvas exercise - PC algorithm

1 Multiple choice 1 point

What is the skeleton of the graph over nodes {1,2,3,4} that one can learn using the above conditional independences (and no other independence)?
If you forgot, here is the pseudo-code for the skeleton learning phase

1. Start with **completely connected undirected graph** U
2. For $k = 0,1,2\ldots,p-2$
   - If $i - j$ in U and there exists a set $\mathbf{S} \subseteq \mathrm{Adj}(i) \cup \mathrm{Adj}(j)$ of size k for which $X_i \perp\!\!\!\perp X_j \mid \mathbf{S}$, **remove** $i - j$ **in** U

$X_1 \perp\!\!\!\perp X_4$

$X_2 \perp\!\!\!\perp X_4$

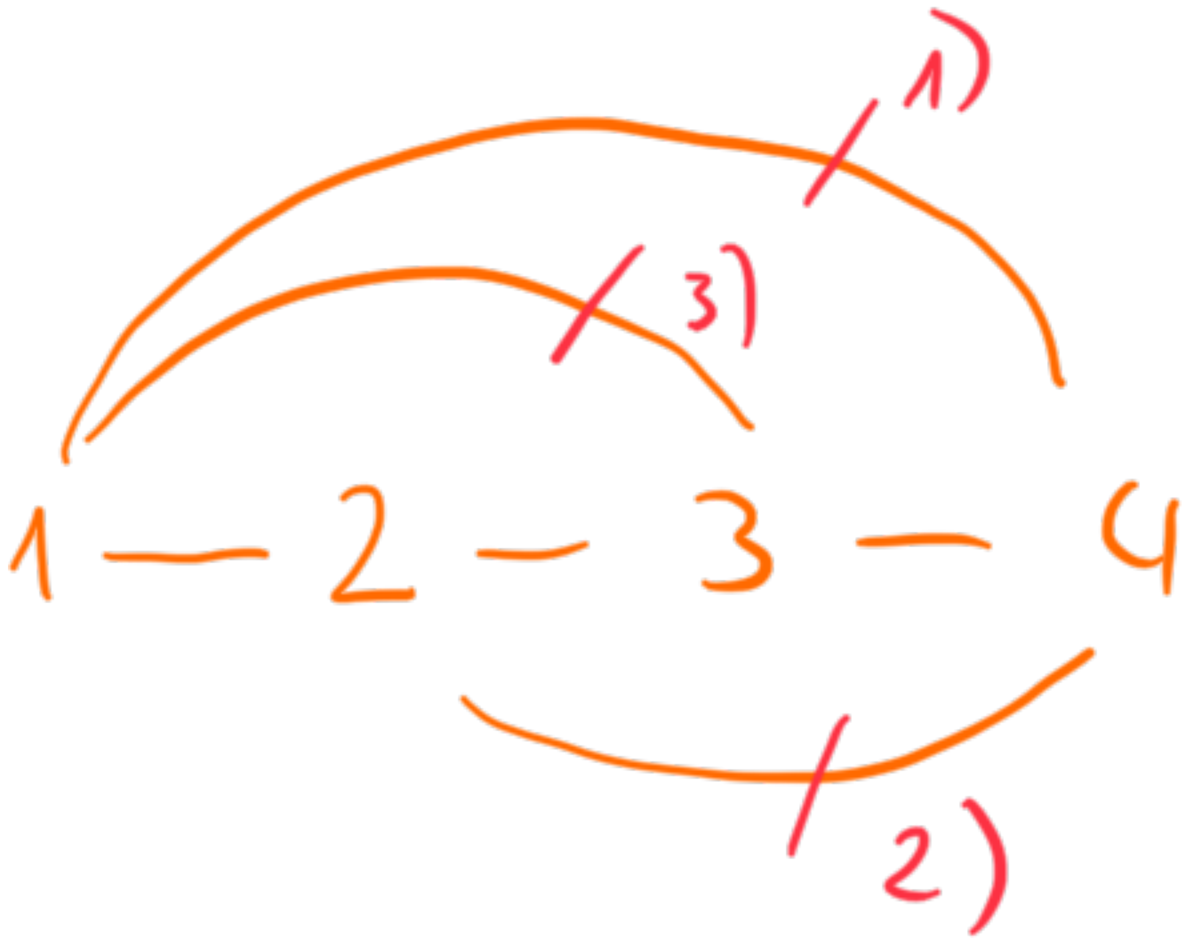
$X_1 \perp\!\!\!\perp X_3 \mid X_2$

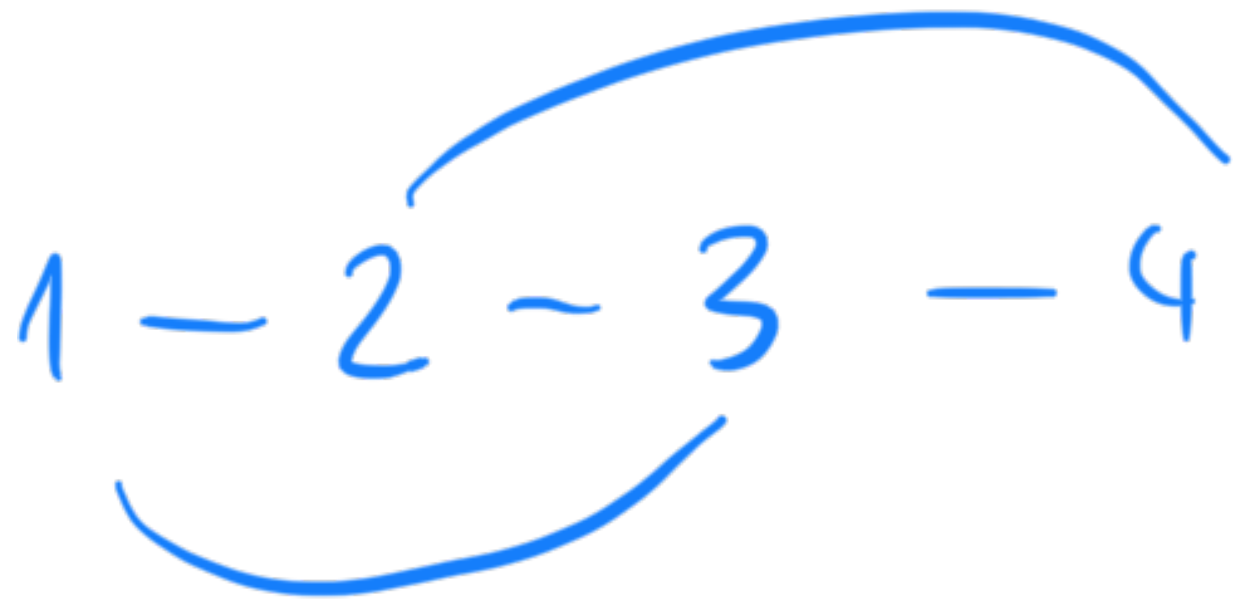
$X_1 \perp\!\!\!\perp X_4 \mid X_2$

$X_2 \perp\!\!\!\perp X_4 \mid X_1$

$X_1 \perp\!\!\!\perp X_4 \mid X_3, X_2$

1 — 2 — 3 — 4





Canvas exercise - PC algorithm

$X_1 \perp\!\!\!\perp X_4$

$X_2 \perp\!\!\!\perp X_4$

$X_1 \perp\!\!\!\perp X_3 \mid X_2$

$X_1 \perp\!\!\!\perp X_4 \mid X_2$

$X_2 \perp\!\!\!\perp X_4 \mid X_1$

$X_1 \perp\!\!\!\perp X_4 \mid X_3, X_2$

2 Multiple choice 1 point

Which are the unshielded triples in the skeleton (note that the definition is the same for undirected graphs):

- A triple of nodes (i, j, k) in a DAG G is a **an unshielded triple** if $i - j, j - k$ and **i is not adjacent to k**, i.e. $i \not- k$, in G

- [x] (1,2,3), (2,3,4)
- [ ] (1,4,3), (4,3,2)
- [ ] none
- [ ] (1,3,4)

$1 - 2 - 3 - 4$

(1,4,3) ✗ 1 is not adjacent to 4 or 3

(1,3,4) ✗

Canvas exercise - PC algorithm

$X_1 \perp\!\!\!\perp X_4$

$X_2 \perp\!\!\!\perp X_4$

$X_1 \perp\!\!\!\perp X_3 \mid X_2$

$X_1 \perp\!\!\!\perp X_4 \mid X_2$

$X_2 \perp\!\!\!\perp X_4 \mid X_1$

$X_1 \perp\!\!\!\perp X_4 \mid X_3, X_2$

3 Multiple choice 1 point

Which are the v-structures that we can determine using the following rules and the original conditional independences? written in the original list, then the variables are dependent. Here are the steps to determine v-structures:

1. Start from the skeleton U from previous step
2. For each unshielded triple (i, j, k) in U, i.e. $i - j, j - k$ **and** $i \not- k$ **in** U
   - For all $\mathbf{S} \subseteq \mathbf{V} \setminus \{i, j, k\}$ check if $X_i \not\perp\!\!\!\perp X_k | X_j \cup X_\mathbf{S}$ in data
     - If this is true, $i \rightarrow j \leftarrow k$ **is a v-structure**

- (x) (2,3,4)
- ( ) (1,2,3)
- ( ) (1,4,3)
- ( ) none

(1,2,3) ✗ $X_1 \perp\!\!\!\perp X_3 \mid X_2$ ✗

(2,3,4) $X_2 \not\perp\!\!\!\perp X_4 \mid X_3$

(not in list)

$X_2 \not\perp\!\!\!\perp X_4 \mid X_3, X_1$

Canvas exercise - PC algorithm

4 True or False 1 point

Consider the graph you have obtained up to now, after phase 1: skeleton learning, and phase 2: determining the v-structures. In this graph, in phase 3, we can orient one or more edges by using the acyclicity or "no new v-structures" constraint.

- ○ True
- ● False

$X_1 \perp\!\!\!\perp X_4$

$X_2 \perp\!\!\!\perp X_4$

$X_1 \perp\!\!\!\perp X_3 \mid X_2$

$X_1 \perp\!\!\!\perp X_4 \mid X_2$

$X_2 \perp\!\!\!\perp X_4 \mid X_1$

$X_1 \perp\!\!\!\perp X_4 \mid X_3, X_2$

$1 - 2 \rightarrow 3 \leftarrow 4$

$1 \rightarrow 2 \rightarrow 3 \rightarrow 4$ ✓

$1 \leftarrow 2 \rightarrow 3 \rightarrow 4$ ✓

Canvas exercise - PC algorithm

$X_1 \perp\!\!\!\perp X_4$

$X_2 \perp\!\!\!\perp X_4$

$X_1 \perp\!\!\!\perp X_3 \mid X_2$

$X_1 \perp\!\!\!\perp X_4 \mid X_2$

$X_2 \perp\!\!\!\perp X_4 \mid X_1$

$X_1 \perp\!\!\!\perp X_4 \mid X_3, X_2$

4 True or False 1 point

Consider the graph you have obtained up to now, after phase 1: skeleton learning, and phase 2: determining the v-structures. In this graph, in phase 3, we can orient one or more edges by using the acyclicity or "no new v-structures" constraint.

- ◯ True
- ◉ False

$1 - 2 \rightarrow 3 \leftarrow 4$

CPDAG

5 True or False 1 point

The final graph (the CPDAG we obtain after all phases) is fully oriented (there are no undirected edges, only directed ones)

- ◯ True
- ◉ False

Canvas exercise - PC algorithm

6 Multiple choice 1 point

In the true causal graph, what is the relationship between 1 and 2 (based on what you can learn from the CPDAG)?

We can represent the skeleton and the orientations (edge marks) all DAGs in a Markov equivalence class (MEC) have in common with a **Complete Partially Directed Acyclic Graph (CPDAG):**

- We have a directed edge $i \to j$ if all DAGs in the MEC have $i \to j$
- We have an undirected edge $i - j$ if some DAGs in the MEC have $i \to j$ and others have $j \to i$

- [x] either 1 -> 2 or 1 <- 2
- [ ] 1 -> 2 (and there are no other options)
- [ ] 1 <- 2 (and there are no other options)
- [ ] 1 is not adjacent to 2

$X_1 \perp\!\!\!\perp X_4$

$X_2 \perp\!\!\!\perp X_4$

$X_1 \perp\!\!\!\perp X_3 \mid X_2$

$X_1 \perp\!\!\!\perp X_4 \mid X_2$

$X_2 \perp\!\!\!\perp X_4 \mid X_1$

$X_1 \perp\!\!\!\perp X_4 \mid X_3, X_2$

$1 \to 2 \to 3 \to 4$ ✓

$1 \leftarrow 2 \to 3 \to 4$ ✓

$1 - 2 \to 3 \leftarrow 4$ CPDAG

Canvas exercise - PC algorithm

$X_1 \perp\!\!\!\perp X_4$

$X_2 \perp\!\!\!\perp X_4$

$X_1 \perp\!\!\!\perp X_3 \mid X_2$

$X_1 \perp\!\!\!\perp X_4 \mid X_2$

$X_2 \perp\!\!\!\perp X_4 \mid X_1$

$X_1 \perp\!\!\!\perp X_4 \mid X_3, X_2$