

DAGs Review

INFO/STSCI/ILRST 3900: Causal Inference

20 Sep 2023

Reminders and Announcements

- ▶ Peer reviews for HW2 are **due tomorrow** (Sept 21st) by 5pm
- ▶ HW 3 is up & due Thursday, Sept 28th by 5pm
- ▶ Office hours throughout the week (check the syllabus!)
- ▶ Check Ed for announcements and use for HW help!

Agenda

- ▶ Reminders and Announcements
- ▶ Review: open/closed paths and independence in DAGs
- ▶ Icebreaker: practice with DAGs
- ▶ Causal Discovery

Review

How to check if a *path* is **open** or **blocked**:

1. Traverse the path node by node
2. If any node is blocked, the entire path is blocked
3. If all nodes are open, then entire path is open

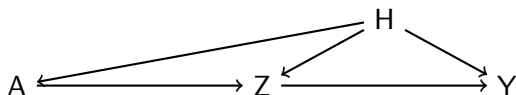
How to check if a *node* is **open** or **blocked**:

- ▶ If **collider** ($\rightarrow X \leftarrow$):
 - ▶ Open if it or any of its descendants are in the conditioning set
 - ▶ Otherwise it is blocked
- ▶ If **non-collider** ($\rightarrow X \rightarrow$ or $\leftarrow X \leftarrow$ or $\leftarrow X \rightarrow$):
 - ▶ Open if it is not in the conditioning set
 - ▶ Blocked if it is in the conditioning set

Two variables are dependent (statistically associated) if there exists an unblocked path between them. They are independent if all paths between them are blocked.

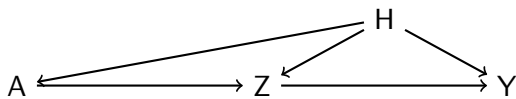
Icebreaker: Practice Identifying Paths

Go through this exercise in groups. Don't forget to introduce yourselves!

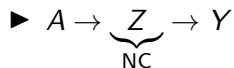
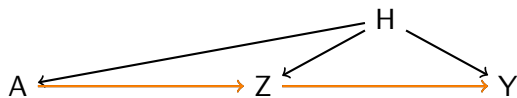


- ▶ What are the paths from A to Y ?
- ▶ Determine if each of the paths is causal or non-causal
- ▶ Determine whether each node on each path is a collider or non-collider

Practice

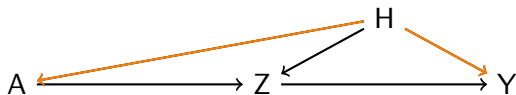


Practice



causal path

Practice



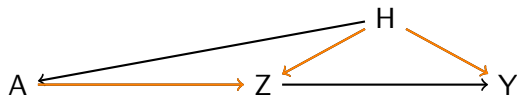
▶ $A \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$

causal path

▶ $A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$

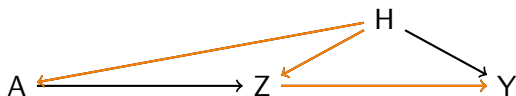
non-causal

Practice



- ▶ $A \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$ causal path
- ▶ $A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$ non-causal
- ▶ $A \rightarrow \underbrace{Z}_{\text{Col}} \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$ non-causal

Practice



► $A \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$

causal path

► $A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$

non-causal

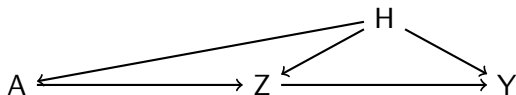
► $A \rightarrow \underbrace{Z}_{\text{Col}} \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$

non-causal

► $A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$

non-causal

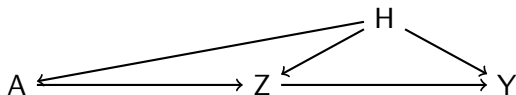
Practice: Is the path open or blocked?



If we condition on $L = \emptyset$, which paths are open? Which paths are blocked?

- ▶ $A \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$
- ▶ $A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$
- ▶ $A \rightarrow \underbrace{Z}_{\text{CoI}} \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$
- ▶ $A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$

Practice: Is the path open or blocked?



If we condition on $L = \emptyset$, which paths are open? Which paths are blocked?

$$\blacktriangleright A \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$$

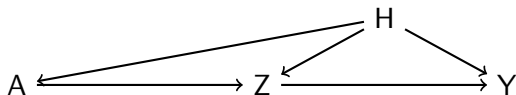
Open

$$\blacktriangleright A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$$

$$\blacktriangleright A \rightarrow \underbrace{Z}_{\text{CoI}} \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$$

$$\blacktriangleright A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$$

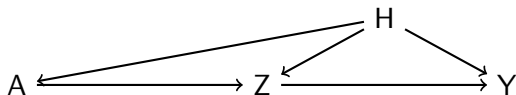
Practice: Is the path open or blocked?



If we condition on $L = \emptyset$, which paths are open? Which paths are blocked?

- ▶ $A \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$ Open
- ▶ $A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$ Open
- ▶ $A \rightarrow \underbrace{Z}_{\text{CoI}} \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$
- ▶ $A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$

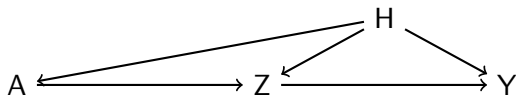
Practice: Is the path open or blocked?



If we condition on $L = \emptyset$, which paths are open? Which paths are blocked?

- ▶ $A \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$ Open
- ▶ $A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$ Open
- ▶ $A \rightarrow \underbrace{Z}_{\text{Col}} \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$ Blocked
- ▶ $A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$

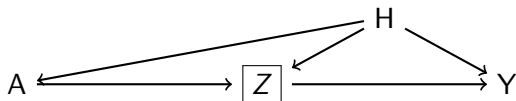
Practice: Is the path open or blocked?



If we condition on $L = \emptyset$, which paths are open? Which paths are blocked?

- ▶ $A \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$ Open
- ▶ $A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$ Open
- ▶ $A \rightarrow \underbrace{Z}_{\text{Col}} \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$ Blocked
- ▶ $A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$ Open

Practice: Is the path open or blocked?



If we condition on $L = \{Z\}$, which paths are open? Which paths are blocked?

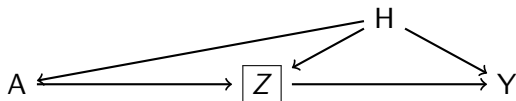
$$\blacktriangleright A \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$$

$$\blacktriangleright A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$$

$$\blacktriangleright A \rightarrow \underbrace{Z}_{\text{Col}} \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$$

$$\blacktriangleright A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$$

Practice: Is the path open or blocked?



If we condition on $L = \{Z\}$, which paths are open? Which paths are blocked?

▶ $A \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$

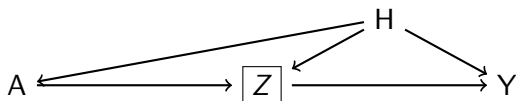
Blocked

▶ $A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$

▶ $A \rightarrow \underbrace{Z}_{\text{Col}} \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$

▶ $A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$

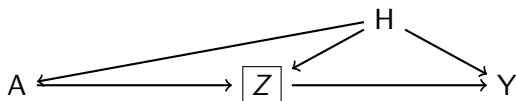
Practice: Is the path open or blocked?



If we condition on $L = \{Z\}$, which paths are open? Which paths are blocked?

- ▶ $A \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$ Blocked
- ▶ $A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$ Open
- ▶ $A \rightarrow \underbrace{Z}_{\text{Col}} \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$
- ▶ $A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$

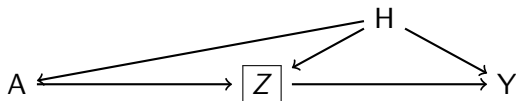
Practice: Is the path open or blocked?



If we condition on $L = \{Z\}$, which paths are open? Which paths are blocked?

- ▶ $A \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$ Blocked
- ▶ $A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$ Open
- ▶ $A \rightarrow \underbrace{Z}_{\text{Col}} \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$ Open
- ▶ $A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$

Practice: Is the path open or blocked?



If we condition on $L = \{Z\}$, which paths are open? Which paths are blocked?

- ▶ $A \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$ Blocked
- ▶ $A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$ Open
- ▶ $A \rightarrow \underbrace{Z}_{\text{Col}} \leftarrow \underbrace{H}_{\text{NC}} \rightarrow Y$ Open
- ▶ $A \leftarrow \underbrace{H}_{\text{NC}} \rightarrow \underbrace{Z}_{\text{NC}} \rightarrow Y$ Blocked

Causal Discovery

- ▶ So far, we have assumed the DAG is known
- ▶ DAG tells us about conditional independence we would observe in data

DAG \Rightarrow Conditional independence in data

Causal Discovery

- ▶ So far, we have assumed the DAG is known
- ▶ DAG tells us about conditional independence we would observe in data

DAG \Rightarrow Conditional independence in data

- ▶ Conditional independence is an observational quantity (i.e., not causal)
- ▶ Can be tested in observed data
- ▶ Can we go in the opposite direction? Given data, can we construct the DAG it came from? Can we tell which nodes are/aren't connected by an edge and the edge's direction?

Conditional independence in data $\stackrel{?}{\Rightarrow}$ DAG

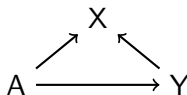
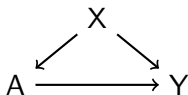
Create your own DAG

Say that we want to estimate the causal effect of asthma (A) on mortality (Y). We have some data and from this data, we know that smoking (X) is statistically associated with both asthma as well as mortality. How might we represent this as a DAG?

Create your own DAG

Say that we want to estimate the causal effect of asthma (A) on mortality (Y). We have some data and from this data, we know that smoking (X) is statistically associated with both asthma as well as mortality. How might we represent this as a DAG?

Potential Solutions:



This serves as a reminder that association and correlation are not the same as causation!

Create your own DAG

Two variables are dependent (statistically associated) if there exists an unblocked path between them. They are independent if all paths between them are blocked.

Create three DAGs, one for each of the following situations.

1. The variables A and Y are statistically associated
2. The variables A and Y are independent conditioned on X
3. The variables A and Y are dependent conditioned on X , but independent otherwise

Create your own DAG

Two variables are dependent (statistically associated) if there exists an unblocked path between them. They are independent if all paths between them are blocked.

Create three DAGs, one for each of the following situations.

1. The variables A and Y are statistically associated
2. The variables A and Y are independent conditioned on X
3. The variables A and Y are dependent conditioned on X , but independent otherwise

$$A \longrightarrow Y$$

$$A \longleftarrow Y$$

$$A \longleftarrow X \longrightarrow Y$$

$$A \longrightarrow X \longrightarrow Y$$

$$A \longleftarrow X \longleftarrow Y$$

$$A \longleftarrow X \longrightarrow Y$$

