Sufficient adjustment sets in DAGs

INFO/STSCI/ILRST 3900: Causal Inference

1 Oct 2024

Learning goals for today

At the end of class, you will be able to:

1. Identify a sufficient adjustment set using the backdoor criterion
2. Assess whether selection bias may hold in a gathered sample

Logistics

- Project Task 1 due Thursday
- Read project description on Course Project Page
- Pset 3 will be posted Thursday

Big picture

- Fundamental problem of causal inference:
  We want to know

  $$\mathrm{E}(Y^a) \qquad \text{Causal Quantity}$$

  but only observe

  $$\mathrm{E}(Y \mid A = a) \qquad \text{Observational Quantity}$$

- When exchangeability holds, we can estimate causal quantities using observed quantities

  $$A \perp\!\!\!\perp (Y^1, Y^0) \qquad \Rightarrow \qquad \mathrm{E}(Y \mid A = a) = \mathrm{E}(Y^a)$$

Big picture

- Fundamental problem of causal inference:
  We want to know

  $$\mathrm{E}(Y^a) \qquad \text{Causal Quantity}$$

  but only observe

  $$\mathrm{E}(Y \mid A = a) \qquad \text{Observational Quantity}$$

- When exchangeability holds, we can estimate causal quantities using observed quantities

  $$A \perp\!\!\!\perp (Y^1, Y^0) \quad \Rightarrow \quad \mathrm{E}(Y \mid A = a) = \mathrm{E}(Y^a)$$

- Exchangeability typically doesn't hold in observational data, but conditional exchangeability is somtimes plausible

  $$A \perp\!\!\!\perp (Y^1, Y^0) \mid L \quad \Rightarrow \quad \mathrm{E}(Y \mid A = a, L = \ell) = \mathrm{E}(Y^a \mid L = \ell)$$

- Use standardization or inverse probability weighting to estimate the average causal effect

Big picture

- (Conditional) exchangeability is not so easy to reason about

Big picture

- (Conditional) exchangeability is not so easy to reason about
- DAGs communicate causal assumptions clearly and intuitively

Big picture

- (Conditional) exchangeability is not so easy to reason about
- DAGs communicate causal assumptions clearly and intuitively
- Given a graph and a conditioning set L, we can determine if conditional exchangeability holds or not
  1. List out all paths from A to Y
  2. Check if each path is open or blocked conditional on L
  3. Conditional exchangeability holds if all open paths are causal paths

Big picture

- (Conditional) exchangeability is not so easy to reason about
- DAGs communicate causal assumptions clearly and intuitively
- Given a graph and a conditioning set L, we can determine if conditional exchangeability holds or not
  1. List out all paths from A to Y
  2. Check if each path is open or blocked conditional on L
  3. Conditional exchangeability holds if all open paths are causal paths

If we believe the causal DAG, then we have reason to believe the estimated causal effect

Sufficient adjustment set

To conduct a causal analysis, we need to find a conditioning set of variables L which give us conditional exchangeability

- Find a set of variables L that blocks all non-causal paths from A and Y
- L is called **sufficient adjustment set**

Sufficient adjustment set

To conduct a causal analysis, we need to find a conditioning set of variables L which give us conditional exchangeability

- Find a set of variables L that blocks all non-causal paths from A and Y
- L is called **sufficient adjustment set**
- Various ways to do this, we will talk about one criterion

Open or blocked?

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

Open or blocked?

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 non-collider:
  - Blocked if it is in the conditioning set
  - Otherwise it is open

Open or blocked?

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 non-collider:
  - Blocked if it is in the conditioning set
  - Otherwise it is open
- If collider:
  - Open if it or any of its descendants are in the conditioning set
  - Otherwise it is blocked

Open or blocked?

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 non-collider:
  - Blocked if it is in the conditioning set
  - Otherwise it is open
- If collider:
  - Open if it or any of its descendants are in the conditioning set
  - Otherwise it is blocked

Backdoor criterion

$$\begin{tikzcd} L \arrow[r] \arrow[rr, bend left] & A \arrow[r] & Y \end{tikzcd}$$

Backdoor criterion

$$\begin{tikzcd} L \arrow[r] \arrow[rr, bend left] & A \arrow[r] & Y \end{tikzcd}$$

Backdoor path starts with an edge pointing in to A and ends at Y

Backdoor criterion

$$
\begin{tikzcd}
L \arrow[r] \arrow[rr, bend left] & A \arrow[r] & Y
\end{tikzcd}
$$

Backdoor path starts with an edge pointing in to A and ends at Y

A set of variables satisfies the backdoor criterion if

1. Blocks all backdoor paths
2. Does not contain any descendant of A

Backdoor criterion

$$\begin{tikzcd} L \arrow[r] \arrow[rr, bend left] & A \arrow[r] & Y \end{tikzcd}$$

Backdoor path starts with an edge pointing in to A and ends at Y

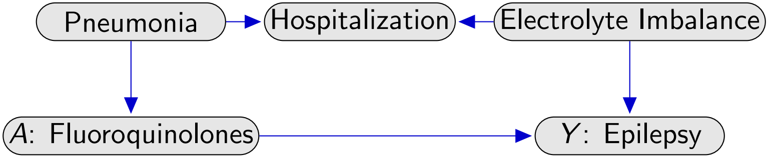
A set of variables satisfies the backdoor criterion if

1. Blocks all backdoor paths
2. Does not contain any descendant of A

Sets that satisfy the backdoor criterion are sufficient adjustment sets!

Exercise

Researchers may be interested in the effect of fluoroquinolones, a class of antibiotics, on epilepsy[1]

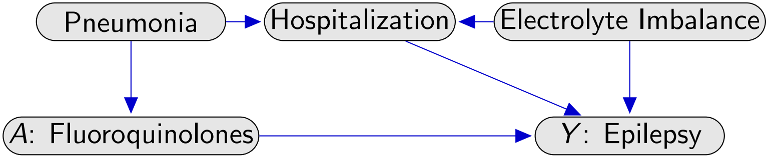


- Does a sufficient adjustment set exist? If so, what is it?

[1] Example from "Using Causal Diagrams to Improve the Design and Interpretation of Medical Research" (Etminan et. al. 2020, Chest)

Exercise

Researchers may be interested in the effect of fluoroquinolones, a class of antibiotics, on epilepsy[2]

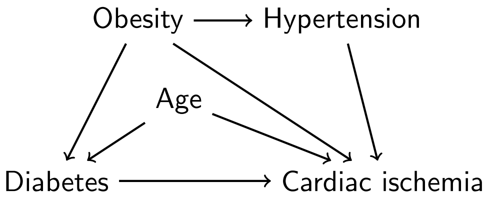


- Does a sufficient adjustment set exist? If so, what is it?

[2] Example from "Using Causal Diagrams to Improve the Design and Interpretation of Medical Research" (Etminan et. al. 2020, Chest)

Exercise

Researchers may be interested in the effect of diabetes on cardiac Ischemia[3]



- Does a sufficient adjustment set exist? If so, what is it?

[3] Example from "Using Causal Diagrams for Biomedical Research" (Kyriacou et. al. 2023, Annals of Emergency Medicine)

Selection Bias

- Sufficient adjustment set to close backdoor paths

Selection Bias

- Sufficient adjustment set to close backdoor paths
- Does not always mean conditioning on more things

$$
\begin{tikzcd}
Z_1 \arrow[dr] \arrow[dd] & & Z_3 \arrow[dl] \arrow[dd] \\
& Z_2 & \\
A \arrow[rr] & & Y
\end{tikzcd}
$$

Selection Bias

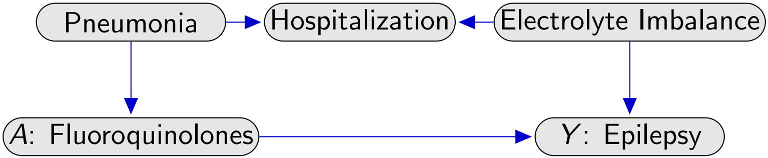
In some settings, certain variables may already be “conditioned on”

[4] Example from “Using Causal Diagrams to Improve the Design and Interpretation of Medical Research” (Etminan et. al. 2020, Chest)

Selection Bias

In some settings, certain variables may already be "conditioned on"

Researchers may be interested in the effect of fluoroquinolones, a class of antibiotics, on epilepsy[4]

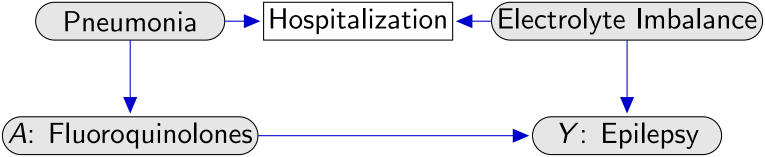


[4]Example from "Using Causal Diagrams to Improve the Design and Interpretation of Medical Research" (Etminan et. al. 2020, Chest)

Selection Bias

In some settings, certain variables may already be “conditioned on”

Researchers may be interested in the effect of fluoroquinolones, a class of antibiotics, on epilepsy[4]

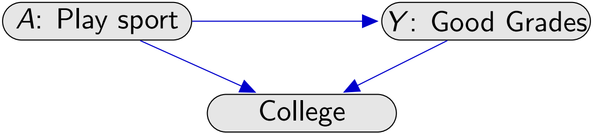


[4] Example from “Using Causal Diagrams to Improve the Design and Interpretation of Medical Research” (Etminan et. al. 2020, Chest)

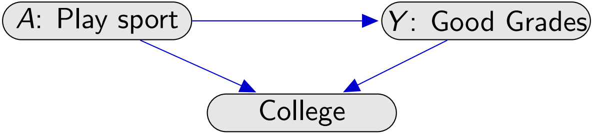
Selection Bias

Selection Bias

Selection Bias

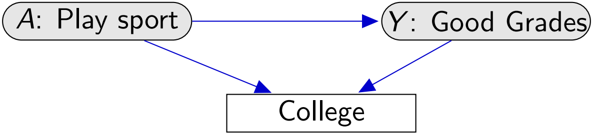


Selection Bias

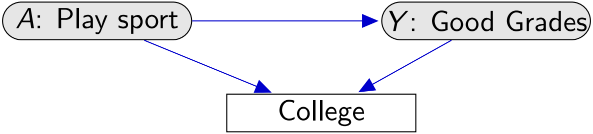


$$A \perp\!\!\!\perp Y^a$$

Selection Bias



Selection Bias



$$A \not\perp\!\!\!\perp Y^a \mid \text{College}$$

Selection Bias

- Gathered data is typically restricted to some group

Selection Bias

- Gathered data is typically restricted to some group
- This may implicitly condition on a variable

Selection Bias

- Gathered data is typically restricted to some group
- This may implicitly condition on a variable
- May open non-causal paths

Takeaways

- Drawing the DAG makes your causal assumptions clear to yourself and others

Takeaways

- Drawing the DAG makes your causal assumptions clear to yourself and others
- Clear rules for what to adjust for **if the DAG is true**
  - **Sufficient adjustment set** blocks all non-causal paths
  - If it exists, use standardization or IPW to estimate causal effect
  - If it does not exist, consider gathering more variables
- Carefully consider the data gathering process

Takeaways

- Drawing the DAG makes your causal assumptions clear to yourself and others
- Clear rules for what to adjust for **if the DAG is true**
  - **Sufficient adjustment set** blocks all non-causal paths
  - If it exists, use standardization or IPW to estimate causal effect
  - If it does not exist, consider gathering more variables
- Carefully consider the data gathering process
- Causal claims come from assumptions + data

Learning goals for today

At the end of class, you will be able to:

1. Identify a sufficient adjustment set using the backdoor criterion
2. Assess whether selection bias may hold in a gathered sample