Directed Acyclic Graphs

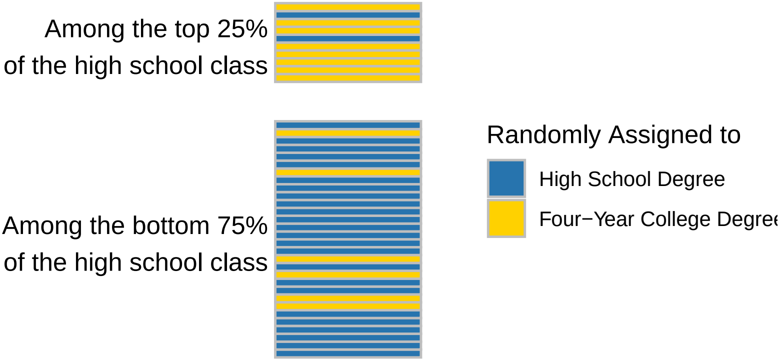
Sociol 114

Learning goals for today

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

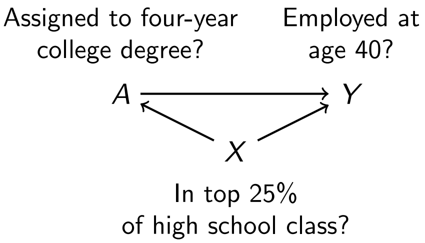
1. Read a Directed Acyclic Graph
2. Recognize causal paths
3. Understand two key structures
   - Fork structures ($\bullet \leftarrow \bullet \rightarrow \bullet$)
   - Collider structures ($\bullet \rightarrow \bullet \leftarrow \bullet$)
4. List all paths in a DAG
5. Determine which paths are blocked under a particular adjustment set
6. Select a sufficient adjustment set to isolate causal paths

A hypothetical experiment:
Conditional randomization



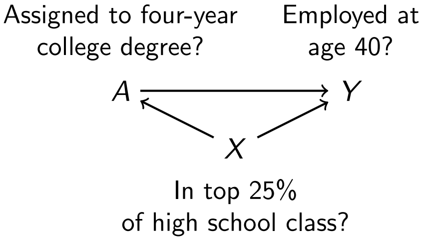
Outcome: Employed at age 40

Elements of a Directed Acyclic Graph (DAG)



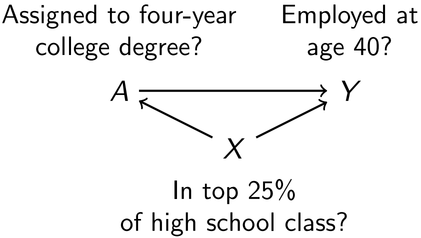
- **Nodes** (X, A, Y) are random variables
- **Edges** $(\rightarrow)$ are causal relationships.
  - X has a causal effect on A
  - X has a causal effect on Y
  - A has a causal effect on Y

Elements of a Directed Acyclic Graph (DAG)



A **path** is a sequence of edges connecting two nodes.

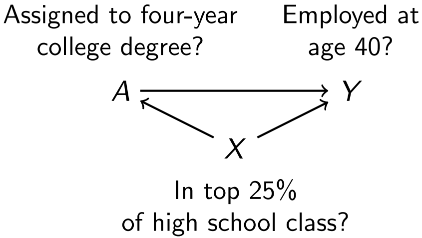
Elements of a Directed Acyclic Graph (DAG)



A **path** is a sequence of edges connecting two nodes.

Between A and Y, what are the two paths?

Elements of a Directed Acyclic Graph (DAG)



A **path** is a sequence of edges connecting two nodes.

Between A and Y, what are the two paths?

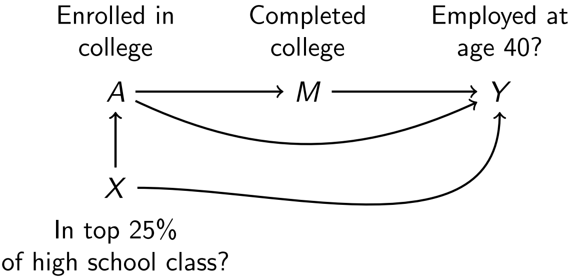
- $A \rightarrow Y$
- $A \leftarrow X \rightarrow Y$

Causal path: A path with arrows pointing one way

$\bullet \rightarrow \bullet \rightarrow \bullet$

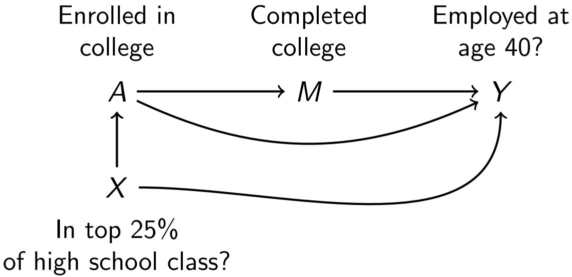
Causal path: A path with arrows pointing one way

$\bullet \rightarrow \bullet \rightarrow \bullet$



Causal path: A path with arrows pointing one way

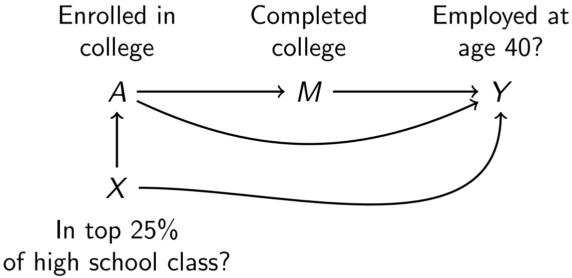
$\bullet \to \bullet \to \bullet$



What three paths connect A and Y?
Which two are causal paths?

Causal path: A path with arrows pointing one way

$\bullet \rightarrow \bullet \rightarrow \bullet$

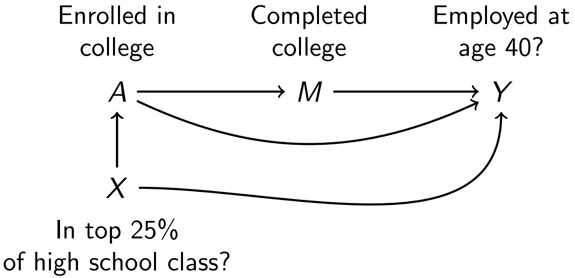


What three paths connect A and Y?
Which two are causal paths?

$A \rightarrow Y$
$A \rightarrow M \rightarrow Y$
$A \leftarrow X \rightarrow Y$

Causal path: A path with arrows pointing one way

$\bullet \to \bullet \to \bullet$

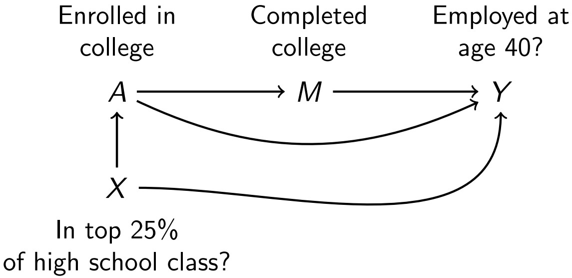


What three paths connect A and Y?
Which two are causal paths?

$A \to Y$ causal path
$A \to M \to Y$
$A \leftarrow X \to Y$

Causal path: A path with arrows pointing one way

$\bullet \rightarrow \bullet \rightarrow \bullet$

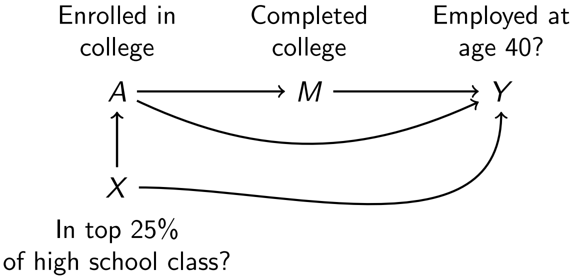


What three paths connect A and Y?
Which two are causal paths?

$A \rightarrow Y$	causal path
$A \rightarrow M \rightarrow Y$	causal path
$A \leftarrow X \rightarrow Y$	

Causal path: A path with arrows pointing one way

$\bullet \to \bullet \to \bullet$



What three paths connect A and Y?
Which two are causal paths?

$A \to Y$	causal path
$A \to M \to Y$	causal path
$A \leftarrow X \to Y$	not a causal path

Causal path: Marginal dependence

$\bullet \to \bullet \to \bullet$

A causal path $A \to \cdots \to B$ will make the variables A and B statistically dependent

Example:

(visits grocery store) $\to$ (buys ice cream) $\to$ (eats ice cream)

Causal path: Marginal dependence

$\bullet \rightarrow \bullet \rightarrow \bullet$

A causal path $A \rightarrow \cdots \rightarrow B$ will make the variables A and B statistically dependent

Example:

(visits grocery store) $\rightarrow$ (buys ice cream) $\rightarrow$ (eats ice cream)

What if we condition:
filter to those with (buys ice cream $=$ `FALSE`)?

Causal path: Conditional independence

$\bullet \rightarrow \bullet \rightarrow \bullet$

A causal path $A \rightarrow \cdots \rightarrow B$ will not make the variables A and B statistically dependent if we condition on a variable along the path

Example:

$$\text{(visits grocery store)} \rightarrow \boxed{\text{(buys ice cream)}} \rightarrow \text{(eats ice cream)}$$

Causal path: Conditional independence

$\bullet \to \bullet \to \bullet$

A causal path $A \to \cdots \to B$ will not make the variables A and B statistically dependent if we condition on a variable along the path

Example:

$$\text{(visits grocery store)} \to \boxed{\text{(buys ice cream)}} \to \text{(eats ice cream)}$$

Among people who didn't buy ice cream today,
those who went to the store and didn't
are equally likely to be eating ice cream.

Causal path: Conditional independence

$\bullet \to \bullet \to \bullet$

A causal path $A \to \cdots \to B$ will not make the variables A and B statistically dependent if we condition on a variable along the path

Example:

$$\text{(visits grocery store)} \to \boxed{\text{(buys ice cream)}} \to \text{(eats ice cream)}$$

Among people who didn't buy ice cream today,
those who went to the store and didn't
are equally likely to be eating ice cream.

Conditioning on (buys ice cream $=$ `FALSE`) **blocks** this path.

Fork structure

$\bullet \leftarrow \bullet \rightarrow \bullet$

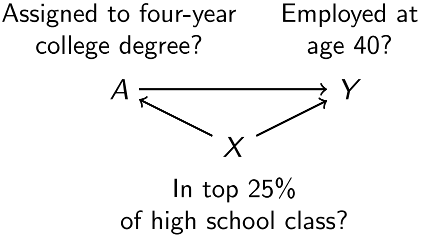
A sequence of edges within a path in which two variables are both caused by a third variable: $A \leftarrow C \rightarrow B$

Fork structure

$\bullet \leftarrow \bullet \rightarrow \bullet$

A sequence of edges within a path in which two variables are both caused by a third variable: $A \leftarrow C \rightarrow B$

In our initial graph, what path contains a fork structure?



Recall that there are two paths:

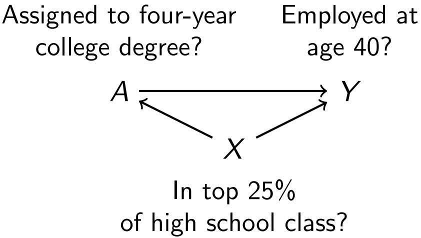
1. $A \rightarrow Y$
2. $A \leftarrow X \rightarrow Y$

Fork structure

$\bullet \leftarrow \bullet \rightarrow \bullet$

A sequence of edges within a path in which two variables are both caused by a third variable: $A \leftarrow C \rightarrow B$

In our initial graph, what path contains a fork structure?



Recall that there are two paths:

1. $A \rightarrow Y$
2. $A \leftarrow X \rightarrow Y$ (this path contains a fork structure)

Fork structure: Marginal dependence

$\bullet \leftarrow \bullet \rightarrow \bullet$

A fork structure $A \leftarrow C \rightarrow B$ will make A and B statistically dependent (because C causes both).

Example:

(completed college) $\leftarrow$ (top 25% of high school) $\rightarrow$ (employed at 40)

Fork structure: Marginal dependence

$\bullet \leftarrow \bullet \rightarrow \bullet$

A fork structure $A \leftarrow C \rightarrow B$ will make A and B statistically dependent (because C causes both).

Example:

(lifeguard rescues) $\leftarrow$ (temperature) $\rightarrow$ (ice cream sales)

Fork structure: Marginal dependence

$\bullet \leftarrow \bullet \rightarrow \bullet$

A fork structure $A \leftarrow C \rightarrow B$ will make A and B statistically dependent (because C causes both).

Example:

$$\text{(lifeguard rescues)} \leftarrow \text{(temperature)} \rightarrow \text{(ice cream sales)}$$

On days with many lifeguard rescues,
there are also many ice cream sales.
Warm temperature causes both.

Fork structure: Marginal dependence

$\bullet \leftarrow \bullet \rightarrow \bullet$

A fork structure $A \leftarrow C \rightarrow B$ will make A and B statistically dependent (because C causes both).

Example:

$$\text{(lifeguard rescues)} \leftarrow \text{(temperature)} \rightarrow \text{(ice cream sales)}$$

On days with many lifeguard rescues,
there are also many ice cream sales.
Warm temperature causes both.

What if we look only at days with a given temperature?

Fork structure: Conditional independence

$\bullet \leftarrow \bullet \rightarrow \bullet$

A fork structure $A \leftarrow \boxed{C} \rightarrow B$ does not make A and B statistically dependent if we condition on C.

Example:

$$\text{(lifeguard rescues)} \leftarrow \boxed{\text{(temperature)}} \rightarrow \text{(ice cream sales)}$$

Among days with a given temperature,
lifeguard rescues and ice cream sales are unrelated.

Conditioning on (temperature) blocks this path.

Collider structure

$\bullet \rightarrow \bullet \leftarrow \bullet$

A sequence of edges within a path in which two variables both cause a third variable: $A \rightarrow C \leftarrow B$

Collider structure

$\bullet \rightarrow \bullet \leftarrow \bullet$

A sequence of edges within a path in which two variables both cause a third variable: $A \rightarrow C \leftarrow B$

Example:

- sprinklers on a timer
- rain on random days
- either one can make the grass wet

$$\text{(sprinklers on)} \rightarrow \text{(grass wet)} \leftarrow \text{(raining)}$$

Collider structure

$\bullet \rightarrow \bullet \leftarrow \bullet$

A sequence of edges within a path in which two variables both cause a third variable: $A \rightarrow C \leftarrow B$

Example:

- sprinklers on a timer
- rain on random days
- either one can make the grass wet

$$\text{(sprinklers on)} \rightarrow \text{(grass wet)} \leftarrow \text{(raining)}$$

Are (sprinklers on) and (raining) statistically related?

Collider structure: Marginal independence

$\bullet \rightarrow \bullet \leftarrow \bullet$

In a collider structure $A \rightarrow C \leftarrow B$,
A and B are marginally independent.

$$\text{(sprinklers on)} \rightarrow \text{(grass wet)} \leftarrow \text{(raining)}$$

Knowing (sprinklers on $=$ `TRUE`) tells me nothing about whether (raining $=$ `TRUE`)

Collider structure: Marginal independence

$\bullet \rightarrow \bullet \leftarrow \bullet$

In a collider structure $A \rightarrow C \leftarrow B$,
A and B are marginally independent.

$$\text{(sprinklers on)} \rightarrow \text{(grass wet)} \leftarrow \text{(raining)}$$

Knowing (sprinklers on $=$ `TRUE`) tells me nothing about whether (raining $=$ `TRUE`)

What if I condition: look only at days when the grass is wet?

Collider structure: Conditional dependence

$\bullet \rightarrow \bullet \leftarrow \bullet$

$$\text{(sprinklers on)} \rightarrow \boxed{\text{(grass wet)}} \leftarrow \text{(raining)}$$

Collider structure: Conditional dependence

$\bullet \rightarrow \bullet \leftarrow \bullet$

$$\text{(sprinklers on)} \rightarrow \boxed{\text{(grass wet)}} \leftarrow \text{(raining)}$$

Among days when (grass wet = `TRUE`),
if (sprinklers on = `FALSE`)
then it must be (raining = `TRUE`)

(grass had to get wet somehow!)

Collider structure: Conditional dependence

$\bullet \rightarrow \bullet \leftarrow \bullet$

$$\text{(sprinklers on)} \rightarrow \boxed{\text{(grass wet)}} \leftarrow \text{(raining)}$$

Among days when (grass wet = `TRUE`),
if (sprinklers on = `FALSE`)
then it must be (raining = `TRUE`)

(grass had to get wet somehow!)

In a collider structure $A \rightarrow \boxed{C} \leftarrow B$,
A and B are conditionally dependent.

Review: Three structures

Name	Structure	A and B marginally dependent?	A and B conditionally dependent given C?
Causal path	$A \rightarrow C \rightarrow B$	Yes	No
Fork	$A \leftarrow C \rightarrow B$	Yes	No
Collider	$A \rightarrow C \leftarrow B$	No	Yes

A path can involve forks, colliders, and causal paths

(timer displays clock) $\leftarrow$ (timer works) $\rightarrow$ (sprinklers on) $\rightarrow$ (grass wet) $\leftarrow$ (raining)

A path can involve forks, colliders, and causal paths

(timer displays clock) $\leftarrow$ (timer works) $\rightarrow$ (sprinklers on) $\rightarrow$ (grass wet) $\leftarrow$ (raining)

(timer displays clock) is statistically related to which variables?

timer works
sprinklers on
grass wet
raining

A path can involve forks, colliders, and causal paths

(timer displays clock) $\leftarrow$ (timer works) $\rightarrow$ (sprinklers on) $\rightarrow$ (grass wet) $\leftarrow$ (raining)

(timer displays clock) is statistically related to which variables?

timer works	yes
sprinklers on	
grass wet	
raining	

A path can involve forks, colliders, and causal paths

(timer displays clock) $\leftarrow$ (timer works) $\rightarrow$ (sprinklers on) $\rightarrow$ (grass wet) $\leftarrow$ (raining)

(timer displays clock) is statistically related to which variables?

timer works	yes
sprinklers on	yes
grass wet	
raining	

A path can involve forks, colliders, and causal paths

(timer displays clock) $\leftarrow$ (timer works) $\rightarrow$ (sprinklers on) $\rightarrow$ (grass wet) $\leftarrow$ (raining)

(timer displays clock) is statistically related to which variables?

timer works	yes
sprinklers on	yes
grass wet	yes
raining	

A path can involve forks, colliders, and causal paths

(timer displays clock) $\leftarrow$ (timer works) $\rightarrow$ (sprinklers on) $\rightarrow$ (grass wet) $\leftarrow$ (raining)

(timer displays clock) is statistically related to which variables?

timer works	yes
sprinklers on	yes
grass wet	yes
raining	no

A path can involve forks, colliders, and causal paths

(timer displays clock) $\leftarrow$ (timer works) $\rightarrow$ (sprinklers on) $\rightarrow$ (grass wet) $\leftarrow$ (raining)

(timer displays clock) is statistically related to which variables?

timer works	yes
sprinklers on	yes
grass wet	yes
raining	no

We just learned: One collider can block an entire path

A path can involve forks, colliders, and causal paths

(timer displays clock) $\leftarrow$ (timer works) $\rightarrow$ $\boxed{\text{(sprinklers on)}}$ $\rightarrow$ (grass wet) $\leftarrow$ (raining)

A path can involve forks, colliders, and causal paths

(timer displays clock) $\leftarrow$ (timer works) $\rightarrow$ $\boxed{\text{(sprinklers on)}}$ $\rightarrow$ (grass wet) $\leftarrow$ (raining)

(timer displays clock) is statistically related to which variables?

timer works
grass wet
raining

A path can involve forks, colliders, and causal paths

(timer displays clock) $\leftarrow$ (timer works) $\rightarrow$ $\boxed{\text{(sprinklers on)}}$ $\rightarrow$ (grass wet) $\leftarrow$ (raining)

(timer displays clock) is statistically related to which variables?

timer works	yes
grass wet	
raining	

A path can involve forks, colliders, and causal paths

(timer displays clock) $\leftarrow$ (timer works) $\rightarrow$ $\boxed{\text{(sprinklers on)}}$ $\rightarrow$ (grass wet) $\leftarrow$ (raining)

(timer displays clock) is statistically related to which variables?

timer works	yes
grass wet	no
raining	

A path can involve forks, colliders, and causal paths

(timer displays clock) $\leftarrow$ (timer works) $\rightarrow$ $\boxed{\text{(sprinklers on)}}$ $\rightarrow$ (grass wet) $\leftarrow$ (raining)

(timer displays clock) is statistically related to which variables?

timer works	yes
grass wet	no
raining	no

A path can involve forks, colliders, and causal paths

(timer displays clock) $\leftarrow$ (timer works) $\rightarrow$ $\boxed{\text{(sprinklers on)}}$ $\rightarrow$ (grass wet) $\leftarrow$ (raining)

(timer displays clock) is statistically related to which variables?

timer works	yes
grass wet	no
raining	no

We just learned: One conditioned non-collider can block an entire path

Rules for whether paths are open or blocked

- If a path contains an unconditioned collider, it is blocked
- If a path contains a conditioned non-collider, it is blocked
- Otherwise, the path is open

Open paths create statistical dependence.
Blocked paths do not.

Determining statistical dependence: A procedure

How do you know if two nodes (e.g., A and B are dependent?

Determining statistical dependence: A procedure

How do you know if two nodes (e.g., A and B are dependent?

$$\begin{tikzcd}
 & C \arrow[dl] \arrow[dr] & \\
A \arrow[dr] & & B \arrow[dl] \\
 & D &
\end{tikzcd}$$

Determining statistical dependence: A procedure

How do you know if two nodes (e.g., A and B are dependent?

$$\begin{tikzcd} & C \arrow[dl] \arrow[dr] & \\ A \arrow[dr] & & B \arrow[dl] \\ & D & \end{tikzcd}$$

1. List all paths between the two nodes
   - $A \leftarrow C \rightarrow B$
   - $A \rightarrow D \leftarrow B$

Determining statistical dependence: A procedure

How do you know if two nodes (e.g., A and B are dependent?

$$\begin{tikzcd}
 & C \arrow[dl] \arrow[dr] & \\
A \arrow[dr] & & B \arrow[dl] \\
 & D &
\end{tikzcd}$$

1. List all paths between the two nodes
   - $A \leftarrow C \rightarrow B$
   - $A \rightarrow D \leftarrow B$
2. Cross out any blocked paths that are blocked
   - $A \leftarrow C \rightarrow B$
   - $\cancel{A \rightarrow D \leftarrow B}$

Determining statistical dependence: A procedure

How do you know if two nodes (e.g., A and B are dependent?

$$\begin{tikzcd} & C \arrow[dl] \arrow[dr] & \\ A \arrow[dr] & & B \arrow[dl] \\ & D & \end{tikzcd}$$

1. List all paths between the two nodes
   - $A \leftarrow C \rightarrow B$
   - $A \rightarrow D \leftarrow B$
2. Cross out any blocked paths that are blocked
   - $A \leftarrow C \rightarrow B$
   - $\cancel{A \rightarrow D \leftarrow B}$
3. If any paths remain, the two nodes are dependent
   - Dependent!

Determining statistical dependence: A procedure

1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.

Determining statistical dependence: A procedure

1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.

Power outage throughout town X

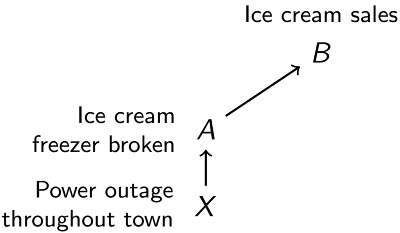
Determining statistical dependence: A procedure

1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.

Ice cream freezer broken	A
	$\uparrow$
Power outage throughout town	X

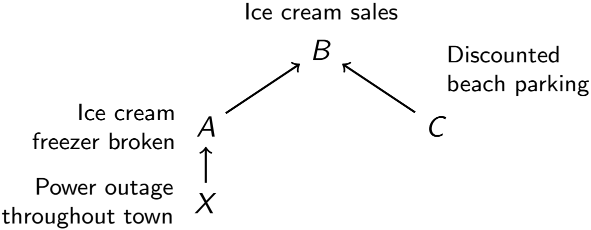
Determining statistical dependence: A procedure

1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.



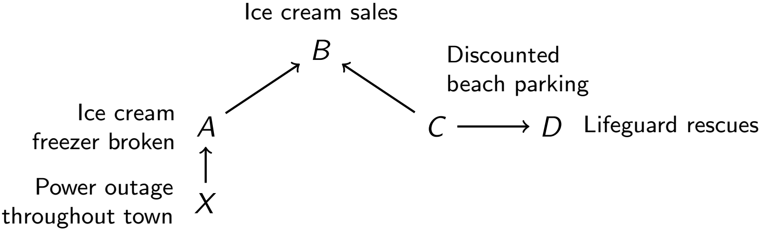
Determining statistical dependence: A procedure

1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.



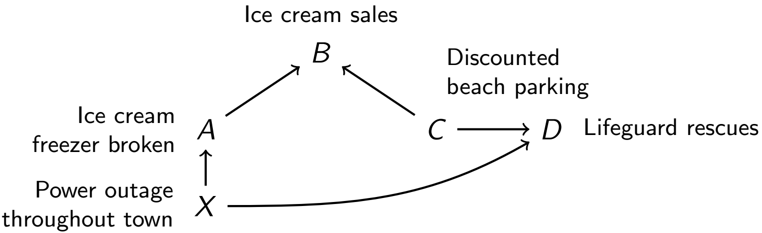
Determining statistical dependence: A procedure

1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.



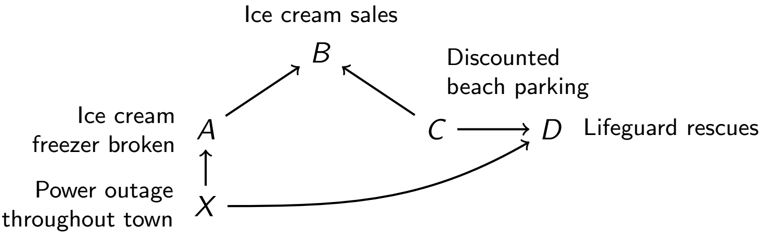
Determining statistical dependence: A procedure

1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.



Determining statistical dependence: A procedure

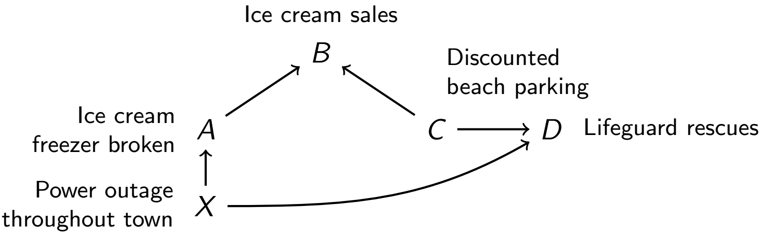
1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.



Are A and C statistically independent or dependent?

Determining statistical dependence: A procedure

1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.

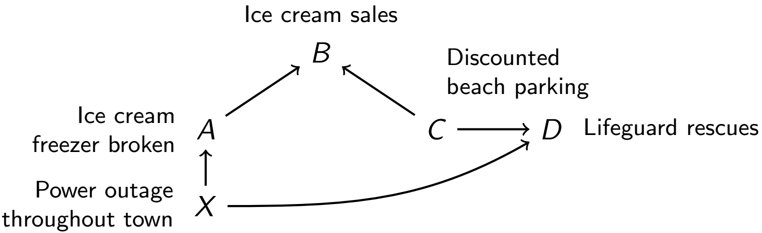


Are A and C statistically independent or dependent?

- $A \rightarrow B \leftarrow C$
- $A \leftarrow X \rightarrow D \leftarrow C$

Determining statistical dependence: A procedure

1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.

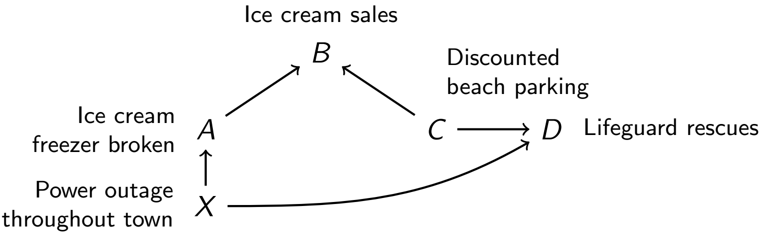


Are A and C statistically independent or dependent?

- $\cancel{A \rightarrow B \leftarrow C}$
- $A \leftarrow X \rightarrow D \leftarrow C$

Determining statistical dependence: A procedure

1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.

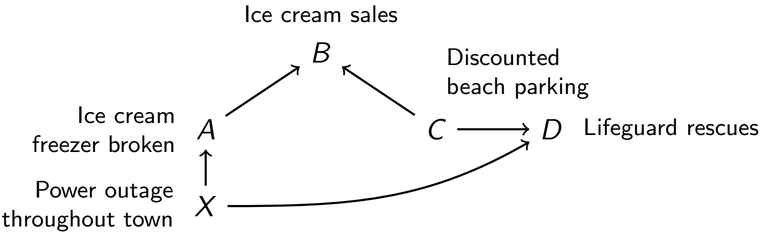


Are A and C statistically independent or dependent?

- $\cancel{A \rightarrow B \leftarrow C}$
- $\cancel{A \leftarrow X \rightarrow D \leftarrow C}$

Determining statistical dependence: A procedure

1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.



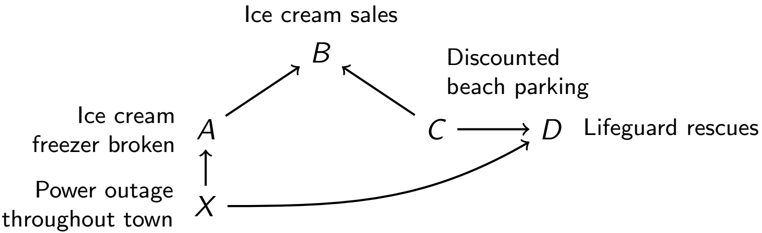
Are A and C statistically independent or dependent?

- $\cancel{A \rightarrow B \leftarrow C}$
- $\cancel{A \leftarrow X \rightarrow D \leftarrow C}$

No unblocked paths. Independent!

Determining statistical dependence: A procedure

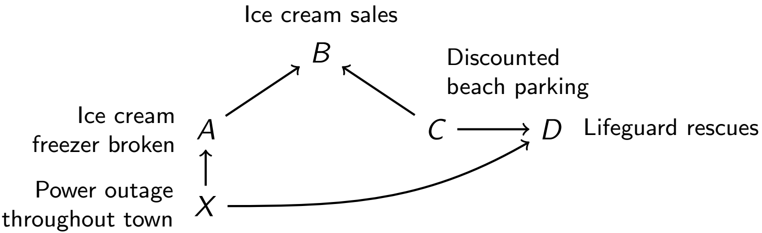
1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.



Are A and D statistically independent or dependent?

Determining statistical dependence: A procedure

1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.

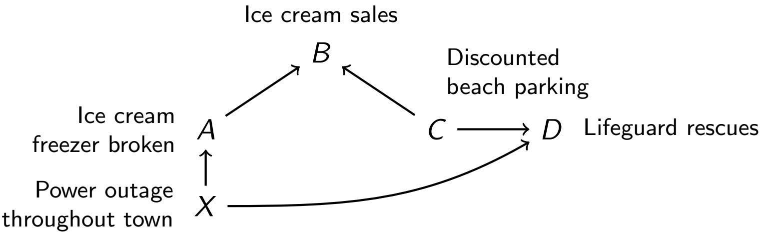


Are A and D statistically independent or dependent?

- $A \rightarrow B \leftarrow C \rightarrow D$
- $A \leftarrow X \rightarrow D$

Determining statistical dependence: A procedure

1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.

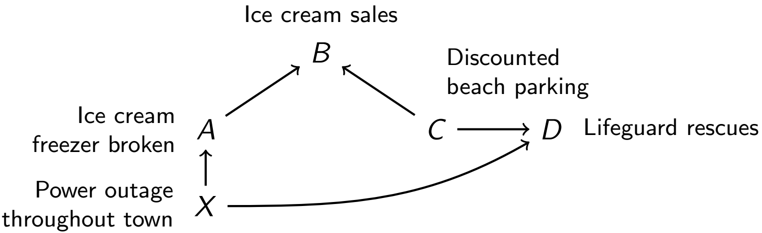


Are A and D statistically independent or dependent?

- ~~$A \rightarrow B \leftarrow C \rightarrow D$~~
- $A \leftarrow X \rightarrow D$

Determining statistical dependence: A procedure

1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.



Are A and D statistically independent or dependent?

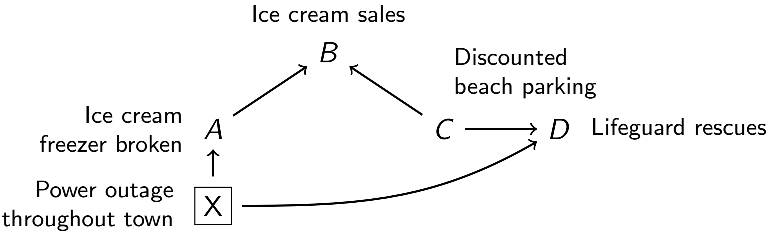
- $\cancel{A \rightarrow B \leftarrow C \rightarrow D}$
- $A \leftarrow X \rightarrow D$

A path remains unblocked. Dependent!

Practice with **conditional** dependence
(holding something constant)

Determining statistical dependence: A procedure

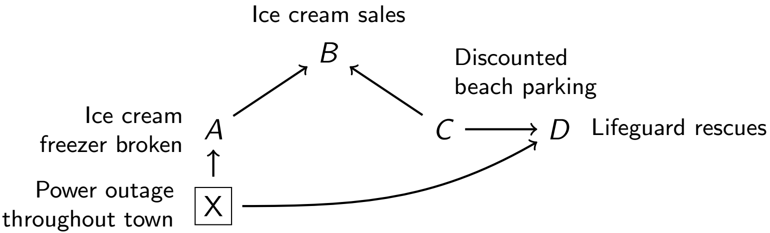
1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.



Practice: Are A and D statistically independent or dependent, conditional on $X = \texttt{FALSE}$?

Determining statistical dependence: A procedure

1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.

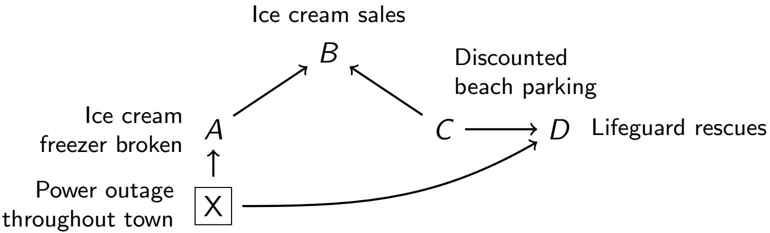


Practice: Are A and D statistically independent or dependent, conditional on $X = \texttt{FALSE}$?

- $A \rightarrow B \leftarrow C \rightarrow D$
- $A \leftarrow \boxed{X} \rightarrow D$

Determining statistical dependence: A procedure

1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.

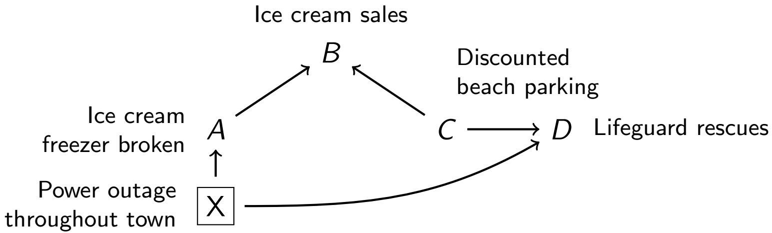


Practice: Are A and D statistically independent or dependent, conditional on $X = \texttt{FALSE}$?

- $\cancel{A \rightarrow B \leftarrow C \rightarrow D}$
- $A \leftarrow \boxed{X} \rightarrow D$

Determining statistical dependence: A procedure

1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.

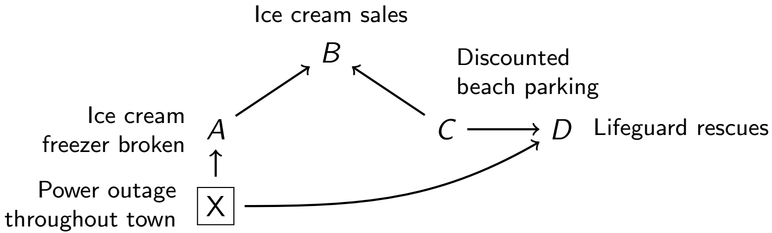


Practice: Are A and D statistically independent or dependent, conditional on $X = \texttt{FALSE}$?

- $\cancel{A \rightarrow B \leftarrow C \rightarrow D}$
- $\cancel{A \leftarrow \boxed{X} \rightarrow D}$

Determining statistical dependence: A procedure

1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.



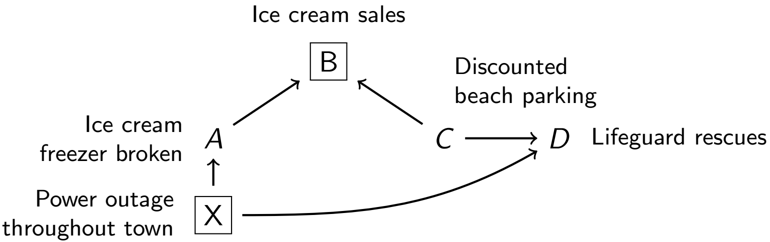
Practice: Are A and D statistically independent or dependent, conditional on $X = \texttt{FALSE}$?

- $\cancel{A \rightarrow B \leftarrow C \rightarrow D}$
- $\cancel{A \leftarrow \boxed{X} \rightarrow D}$

No unblocked paths. Independent!

Determining statistical dependence: A procedure

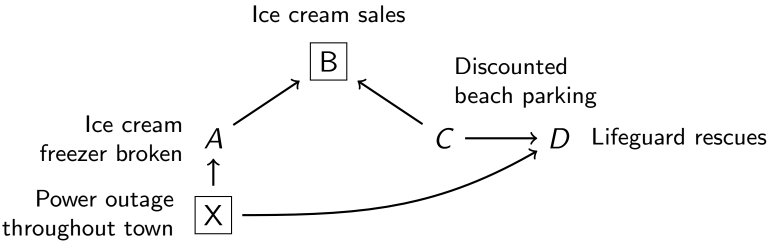
1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.



Practice: Are A and D statistically independent or dependent, conditional on $X = \texttt{FALSE}$ and $B = 0$?

Determining statistical dependence: A procedure

1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.

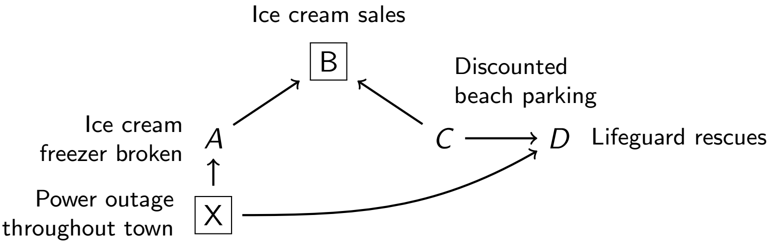


Practice: Are A and D statistically independent or dependent, conditional on $X = \texttt{FALSE}$ and $B = 0$?

- $A \rightarrow \boxed{B} \leftarrow C \rightarrow D$
- $A \leftarrow \boxed{X} \rightarrow D$

Determining statistical dependence: A procedure

1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.

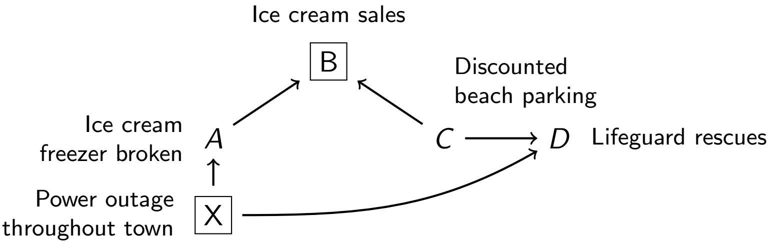


Practice: Are A and D statistically independent or dependent, conditional on $X = \texttt{FALSE}$ and $B = 0$?

- $A \rightarrow \boxed{B} \leftarrow C \rightarrow D$
- $\cancel{A \leftarrow \boxed{X} \rightarrow D}$

Determining statistical dependence: A procedure

1. List all paths. 2. Cross out blocked paths. 3. Dependent if any paths remain.



Practice: Are A and D statistically independent or dependent, conditional on $X = \texttt{FALSE}$ and $B = 0$?

- $A \rightarrow \boxed{B} \leftarrow C \rightarrow D$
- $\cancel{A \leftarrow \boxed{X} \rightarrow D}$

A path remains. Dependent!

DAGs and conditional exchangeability

When studying the effect of A on Y, conditional exchangeability holds if the only unblocked paths between A and Y are causal paths from A to Y.

DAGs and conditional exchangeability

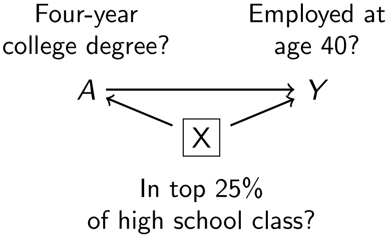
When studying the effect of A on Y, conditional exchangeability holds if the only unblocked paths between A and Y are causal paths from A to Y.

- Why? Because then any association between A and Y must be due to the causal effect

DAGs and conditional exchangeability

When studying the effect of A on Y, conditional exchangeability holds if the only unblocked paths between A and Y are causal paths from A to Y.

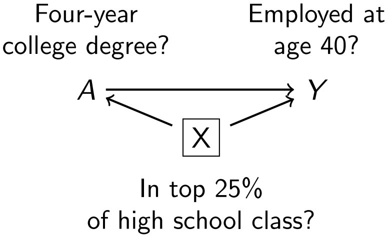
- Why? Because then any association between A and Y must be due to the causal effect



DAGs and conditional exchangeability

When studying the effect of A on Y, conditional exchangeability holds if the only unblocked paths between A and Y are causal paths from A to Y.

- Why? Because then any association between A and Y must be due to the causal effect



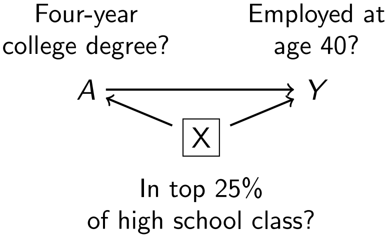
- $A \rightarrow Y$
- $A \leftarrow \boxed{X} \rightarrow Y$ (blocked by conditioning on X)

DAGs and conditional exchangeability

When studying the effect of A on Y, conditional exchangeability holds if the only unblocked paths between A and Y are causal paths from A to Y.

- Why? Because then any association between A and Y must be due to the causal effect

X is a **sufficient adjustment set** to identify the effect of A on Y

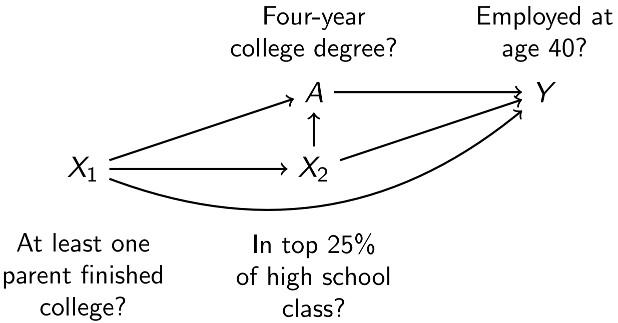


- $A \to Y$
- $A \leftarrow \boxed{X} \to Y$ (blocked by conditioning on X)

DAGs and conditional exchangeability: Practice

1. List all paths. 2. Choose adjustment set. 3. Only causal paths remain unblocked.

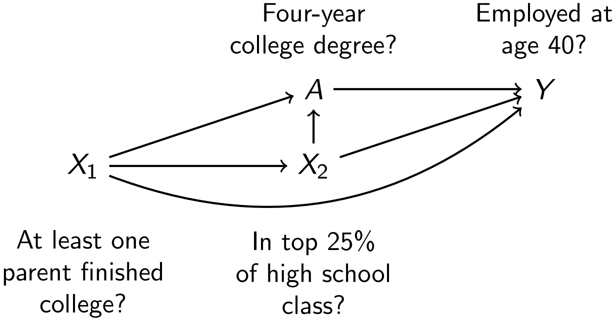
Find a sufficient adjustment set to identify the effect of A on Y.



DAGs and conditional exchangeability: Practice

1. List all paths. 2. Choose adjustment set. 3. Only causal paths remain unblocked.

Find a sufficient adjustment set to identify the effect of A on Y.

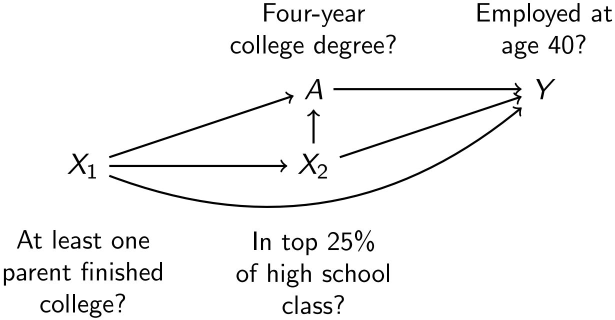


Paths: $(A \rightarrow Y)$, $(A \leftarrow X_2 \rightarrow Y)$, $(A \leftarrow X_1 \rightarrow Y)$,
$(A \leftarrow X_1 \rightarrow X_2 \rightarrow Y)$,$(A \leftarrow X_2 \leftarrow X_1 \rightarrow Y)$

DAGs and conditional exchangeability: Practice

1. List all paths. 2. Choose adjustment set. 3. Only causal paths remain unblocked.

Find a sufficient adjustment set to identify the effect of A on Y.



Paths: $(A \rightarrow Y)$, $(A \leftarrow X_2 \rightarrow Y)$, $(A \leftarrow X_1 \rightarrow Y)$, $(A \leftarrow X_1 \rightarrow X_2 \rightarrow Y)$,$(A \leftarrow X_2 \leftarrow X_1 \rightarrow Y)$

Adjust for $\{X_1, X_2\}$

How to draw a DAG

1. Begin with treatment A and outcome Y
2. Add any variable that affects both
3. Add any variable that affects any two variables in the DAG.

Assumptions are about nodes and edges that you omit.

Exercise: Draw a DAG

Treatment is college degree. Outcome is employment at age 40.

Learning goals for today

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

1. Read a Directed Acyclic Graph
2. Recognize causal paths
3. Understand two key structures
   - Fork structures ($\bullet \leftarrow \bullet \rightarrow \bullet$)
   - Collider structures ($\bullet \rightarrow \bullet \leftarrow \bullet$)
4. List all paths in a DAG
5. Determine which paths are blocked under a particular adjustment set
6. Select a sufficient adjustment set to isolate causal paths