Directed Acyclic Graphs: Marginal Independence

STSCI / INFO / ILRST 3900: Causal Inference

18 Sep 2025

Logistics

- Quiz 1 today
- PSET 2 due tomorrow 11:59pm
- PSET 2 peer review due Sep 26

Quiz

- Don't forget your name and Net ID
- 15 minutes

Learning goals for today

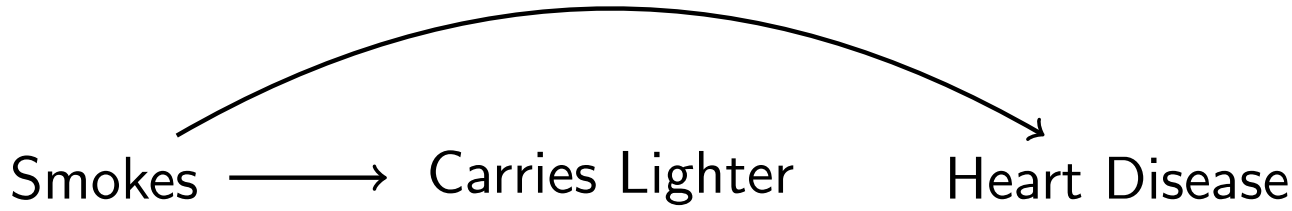
At the end of class, you will be able to

- draw a causal Directed Acyclic Graph
- enumerate edges in the graph
- read statistical dependence of nodes in the graph
- determine marginal exchangeability in the graph

After class:

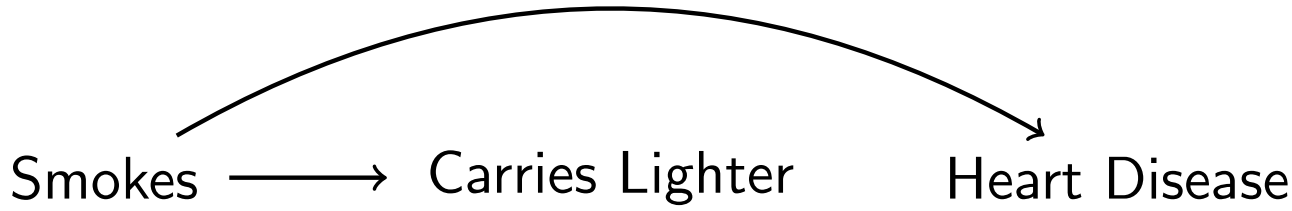
- Hernán and Robins 2020 Chapter 6.1 and 6.2

Directed Acyclic Graphs (DAGs) formalize **causal beliefs**



Causal beliefs:
1) Smoking may cause you to carry a lighter
2) Smoking may cause Heart Disease
3) Carrying a lighter does not cause Heart Disease

Directed Acyclic Graphs (DAGs) formalize **causal beliefs**



Directed Acyclic Graphs (DAGs) formalize **causal beliefs**

Nodes represent random variables

Directed Acyclic Graphs (DAGs) formalize **causal beliefs**

Nodes represent random variables

Edges represent direct causal effects

Directed Acyclic Graphs (DAGs) formalize **causal beliefs**

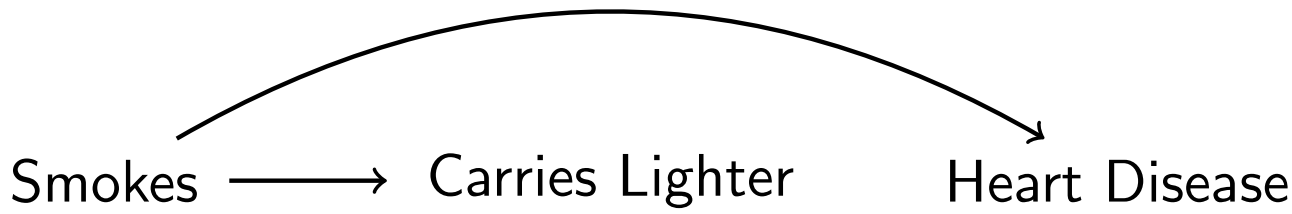
Nodes represent random variables

Edges represent direct causal effects

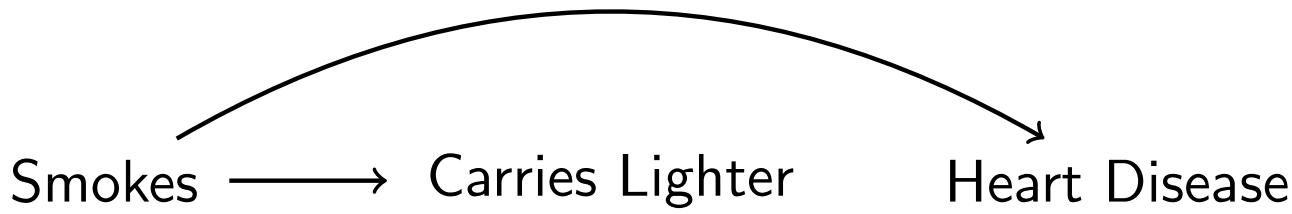
Additional Requirements

- In this class we will think about **acylic** graphs
- Nodes with edges to at least two other nodes should be included

Directed Acyclic Graphs (DAGs) formalize **statistical dependence**



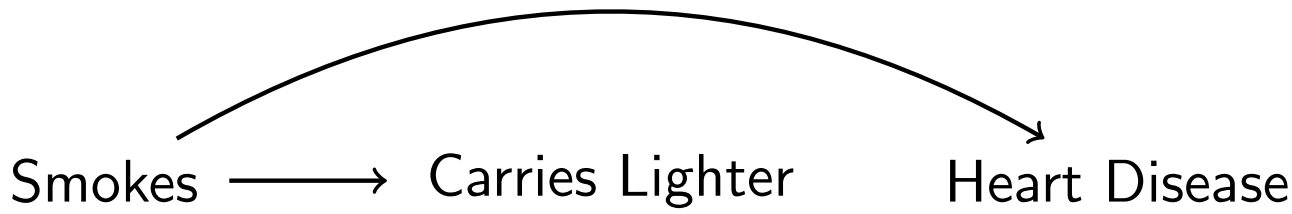
Directed Acyclic Graphs (DAGs) formalize **statistical dependence**



(Smokes, Carries Lighter) are statistically dependent

— because (Smokes) causes (Carries Lighter)

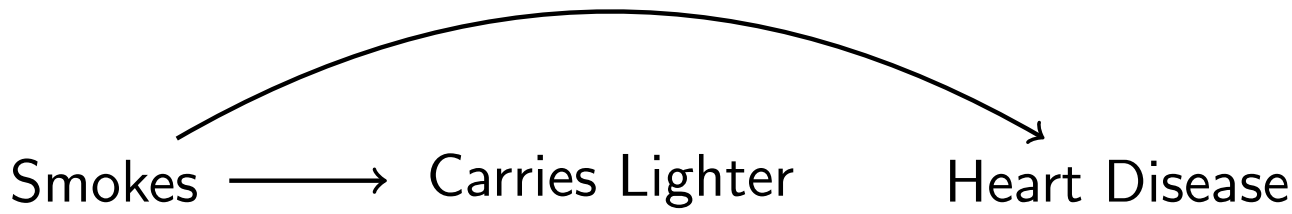
Directed Acyclic Graphs (DAGs) formalize **statistical dependence**



(Smokes, Carries Lighter) are statistically dependent
— because (Smokes) causes (Carries Lighter)

(Smokes, Heart Disease) are statistically dependent
— because (Smokes) causes (Heart Disease)

Directed Acyclic Graphs (DAGs) formalize **statistical dependence**



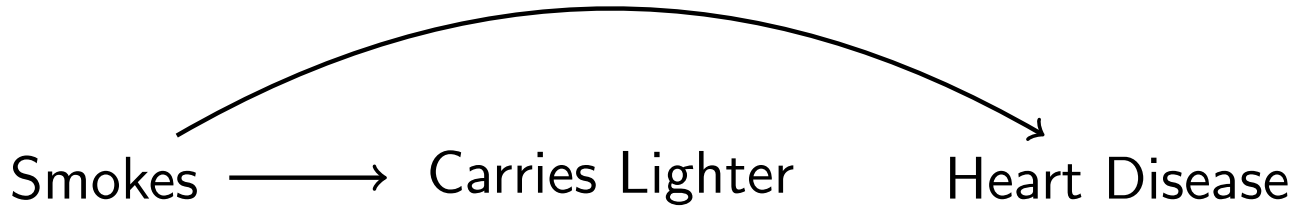
(Smokes, Carries Lighter) are statistically dependent
— because (Smokes) causes (Carries Lighter)

(Smokes, Heart Disease) are statistically dependent
— because (Smokes) causes (Heart Disease)

(Carries Lighter, Heart Disease) are statistically dependent
— because (Smokes) causes both

Task. Propose a rule for when two nodes are dependent

Two nodes are dependent if and only if _____



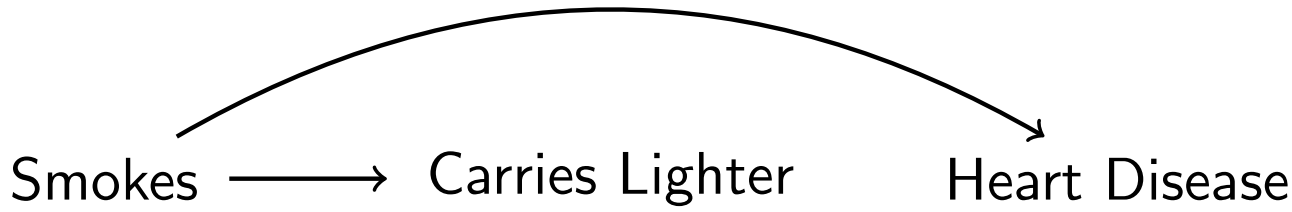
(Smokes, Carries Lighter) are statistically dependent
— because (Smokes) causes (Carries Lighter)

(Smokes, Heart Disease) are statistically dependent
— because (Smokes) causes (Heart Disease)

(Carries Lighter, Heart Disease) are statistically dependent
— because (Smokes) causes both

Task. Propose a rule for when two nodes are dependent

Two nodes are dependent if and only if _____

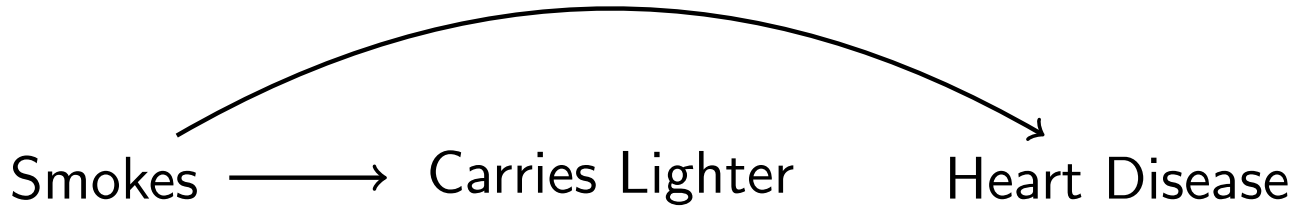


Possible rule (not yet correct)

Two nodes are dependent if and only if they are connected by a path

Task. Propose a rule for when two nodes are dependent

Two nodes are dependent if and only if _____



Possible rule (not yet correct)

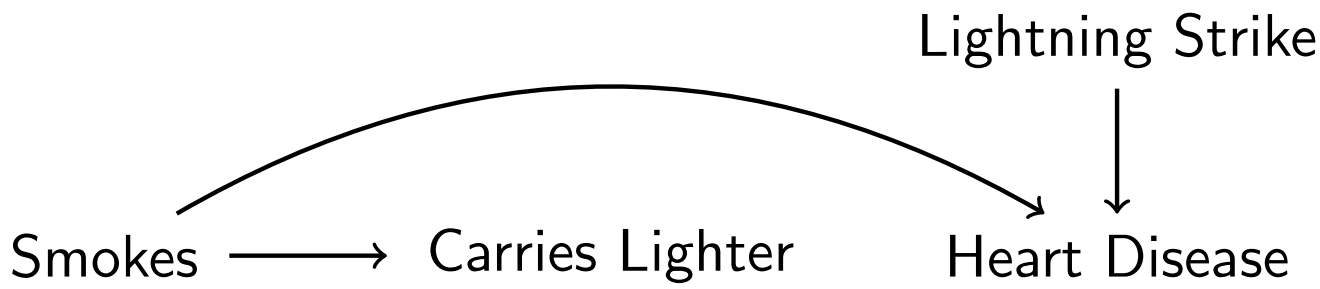
Two nodes are dependent if and only if they are connected by a path

Path

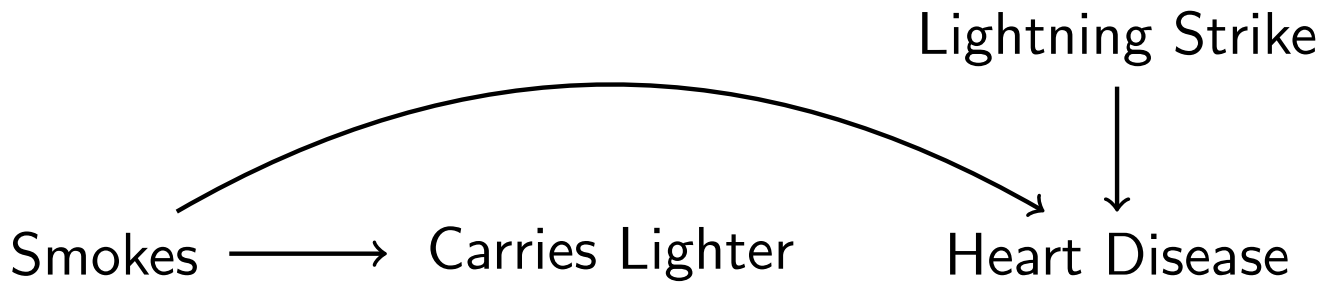
A sequence of edges connecting two nodes

Smokes $\rightarrow$ Carries Lighter
Smokes $\rightarrow$ Heart Disease
Carries Lighter $\leftarrow$ Smokes $\rightarrow$ Heart Disease

Task. Propose a rule for when two nodes are dependent

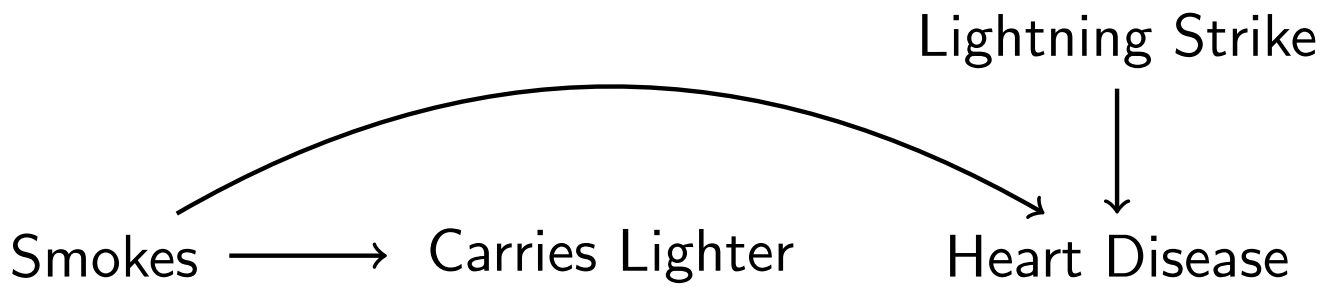


Task. Propose a rule for when two nodes are dependent



(Lightning Strike) causes (Heart Disease)

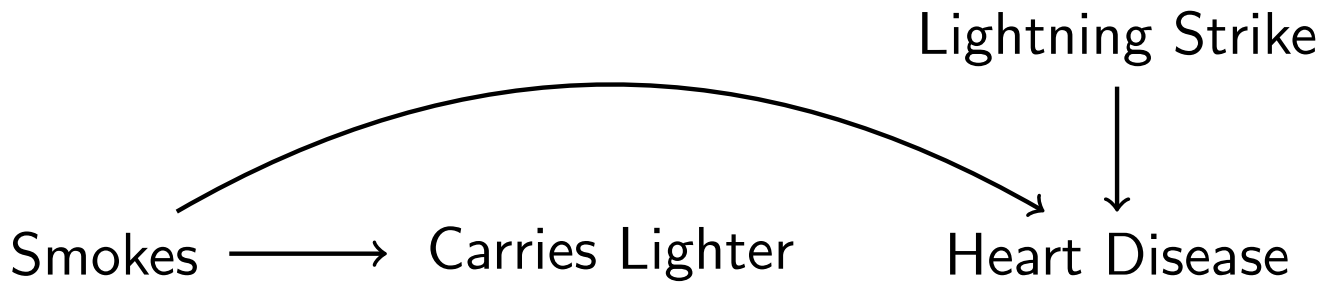
Task. Propose a rule for when two nodes are dependent



(Lightning Strike) causes (Heart Disease)

There are no common causes of (Smokes, Lightning Strike)

Task. Propose a rule for when two nodes are dependent

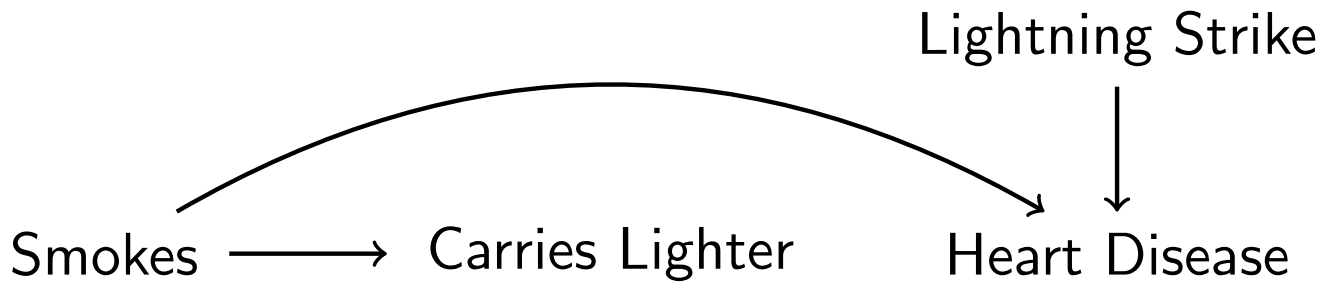


(Lightning Strike) causes (Heart Disease)

There are no common causes of (Smokes, Lightning Strike)

There is a path: (Smokes) $\rightarrow$ (Heart Disease) $\leftarrow$ (Lightning Strike)

Task. Propose a rule for when two nodes are dependent



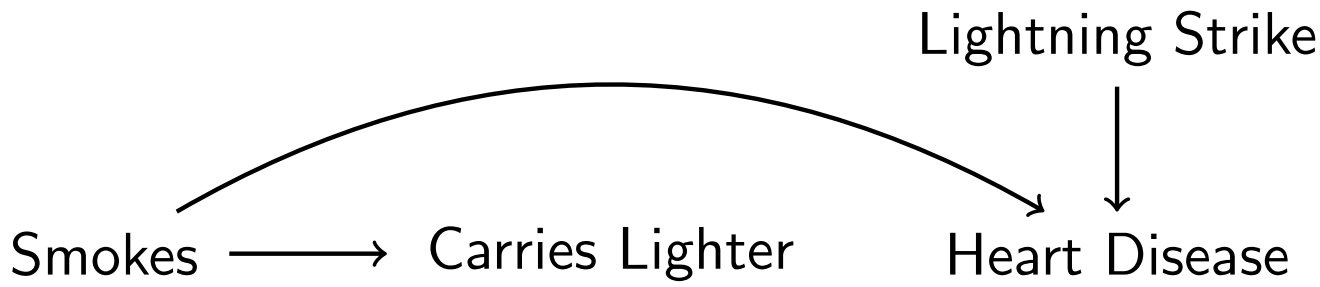
(Lightning Strike) causes (Heart Disease)

There are no common causes of (Smokes, Lightning Strike)

There is a path: (Smokes) $\rightarrow$ (Heart Disease) $\leftarrow$ (Lightning Strike)

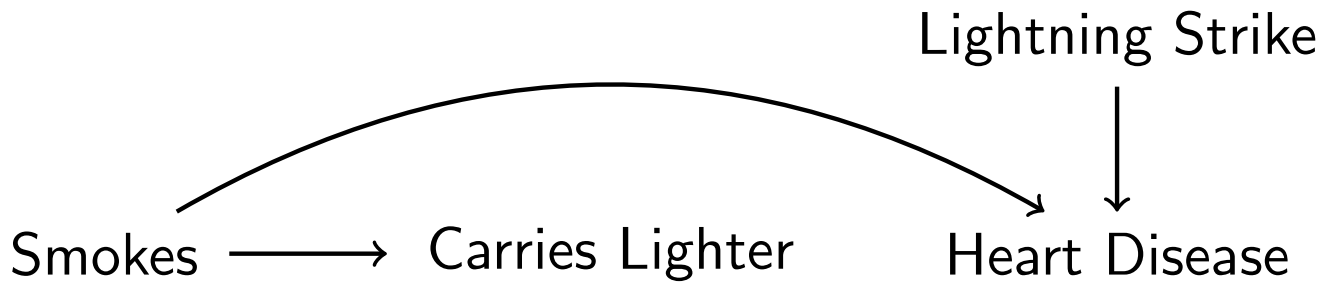
Is (Smokes) statistically related to (Lightning Strike)?

Task. Propose a rule for when two nodes are dependent



Heart Disease is a **collider** on the path
(Smokes) $\rightarrow$ (Heart Disease) $\leftarrow$ (Lightning Strike)

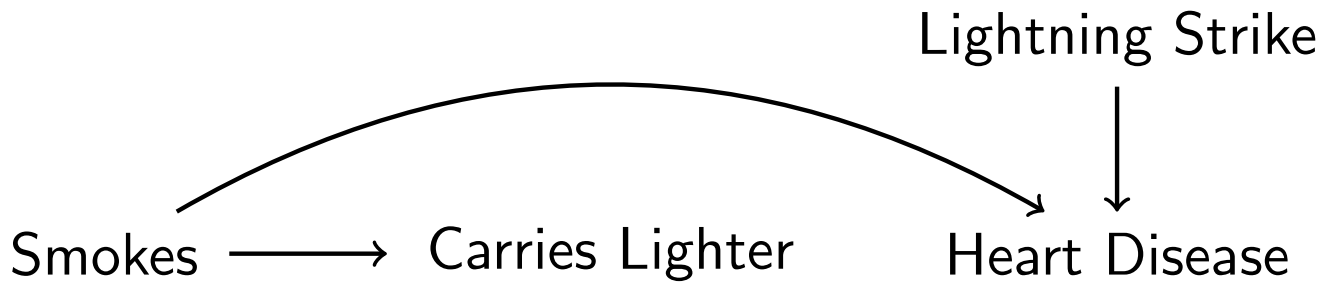
Task. Propose a rule for when two nodes are dependent



Heart Disease is a **collider** on the path
(Smokes) $\rightarrow$ (Heart Disease) $\leftarrow$ (Lightning Strike)

Collider A node on a path where two edges collide $\rightarrow \bullet \leftarrow$

Task. Propose a rule for when two nodes are dependent

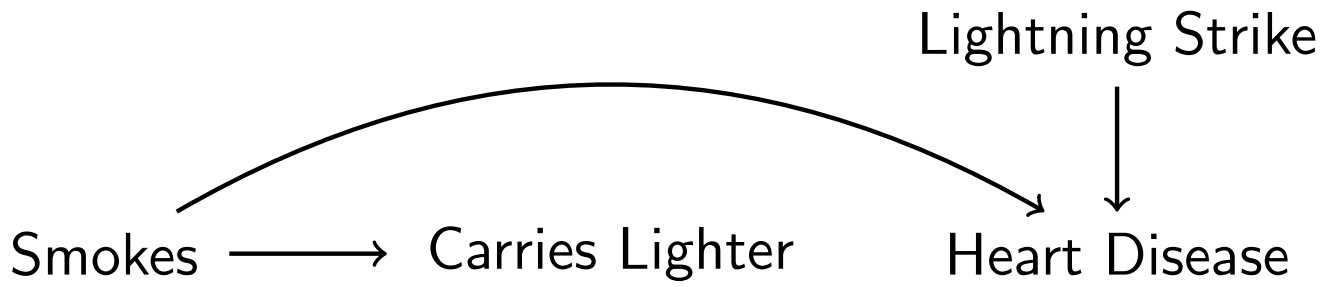


Heart Disease is a **collider** on the path
(Smokes) $\rightarrow$ (Heart Disease) $\leftarrow$ (Lightning Strike)

Collider A node on a path where two edges collide $\rightarrow \bullet \leftarrow$

A collider **blocks the path**.
A blocked path does not create statistical dependence.

Task. Propose a rule for when two nodes are dependent



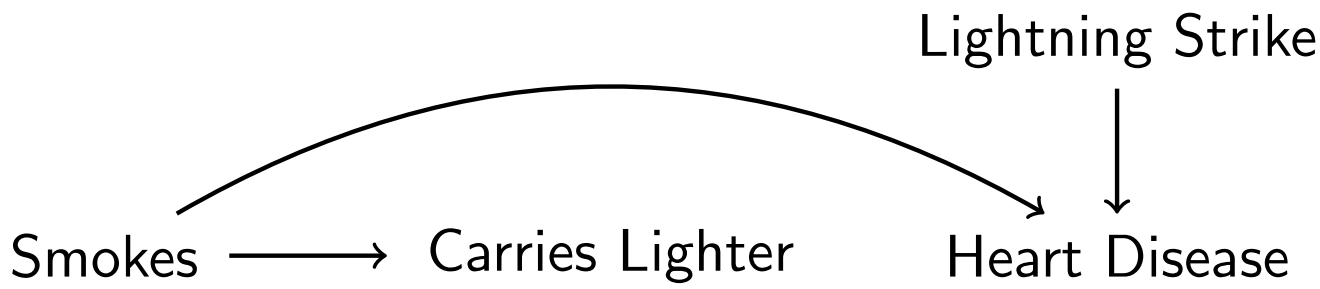
Heart Disease is a **collider** on the path
(Smokes) $\rightarrow$ (Heart Disease) $\leftarrow$ (Lightning Strike)

Collider A node on a path where two edges collide $\rightarrow \bullet \leftarrow$

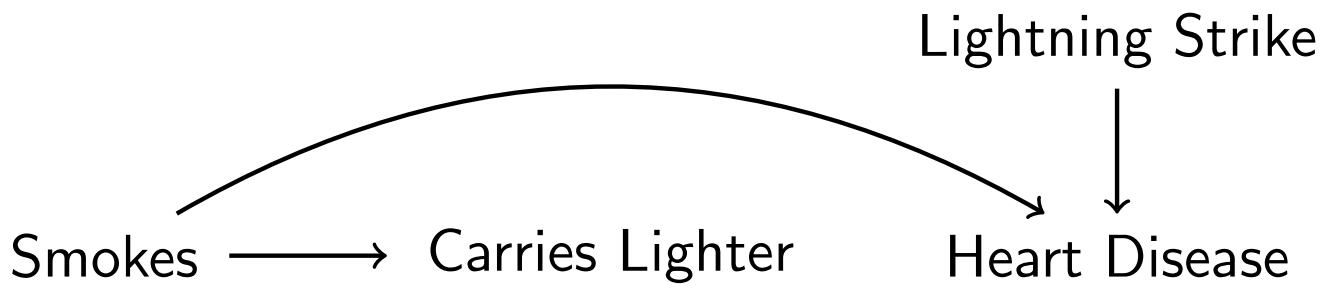
A collider **blocks the path**.
A blocked path does not create statistical dependence.

Intuition: If two variables affect one outcome,
that does not make those two variables related

Task. Propose a rule for when two nodes are dependent



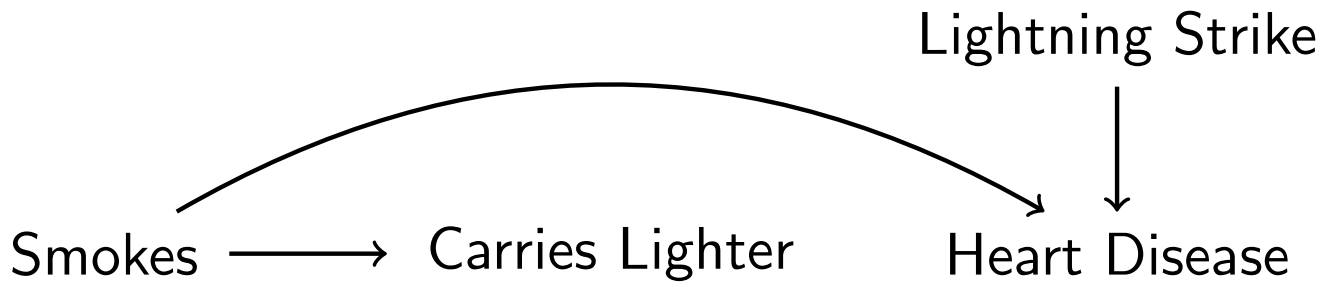
Task. Propose a rule for when two nodes are dependent



Possible rule
(not yet correct)

Two nodes are dependent if and only if they are connected by a path

Task. Propose a rule for when two nodes are dependent



Rule Two nodes are dependent if and only if they are connected by an unblocked path (path with no colliders)

DAGs help us reason about exchangeability

DAGs help us reason about exchangeability

DAGs tell us why two variables are statistically dependent

- A set of unblocked paths

DAGs help us reason about exchangeability

DAGs tell us why two variables are statistically dependent

- A set of unblocked paths

Exchangeability requires statistical independence: $A \perp\!\!\!\perp Y^a$

- Exchangeability holds if the only reason A and Y are related is the causal effect of A on Y

DAGs help us reason about exchangeability

DAGs tell us why two variables are statistically dependent

- A set of unblocked paths

Exchangeability requires statistical independence: $A \perp\!\!\!\perp Y^a$

- Exchangeability holds if the only reason A and Y are related is the causal effect of A on Y

Exchangeability holds if all unblocked paths between A and Y are causal paths that point from A to Y

DAGs help us reason about exchangeability

DAGs help us reason about exchangeability

Procedure

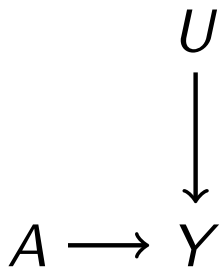
1) List all paths between A to Y
2) Cross out the blocked paths
3) Exchangeability holds if all remaining paths are causal

DAGs help us reason about exchangeability

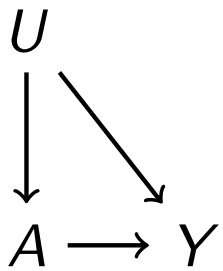
Procedure

1) List all paths between A to Y
2) Cross out the blocked paths
3) Exchangeability holds if all remaining paths are causal

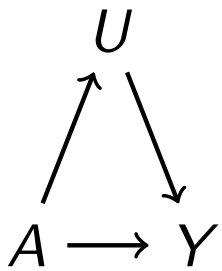
DAG 1



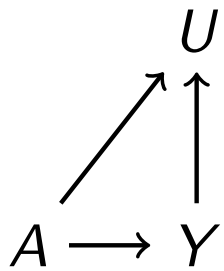
DAG 2



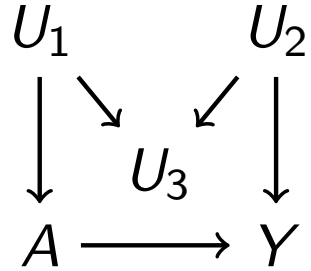
DAG 3



DAG 4



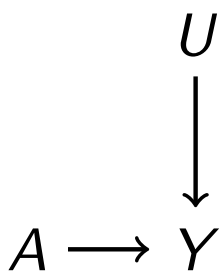
DAG 5



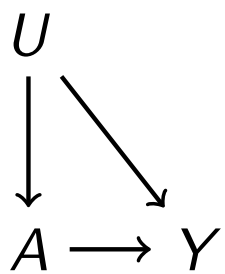
DAGs help us reason about exchangeability

DAGs help us reason about exchangeability

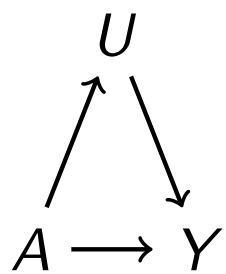
DAG 1



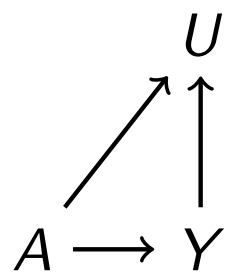
DAG 2



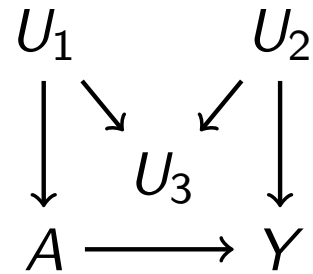
DAG 3



DAG 4



DAG 5



Learning goals for today

At the end of class, you will be able to

- draw a causal Directed Acyclic Graph
- enumerate edges in the graph
- read statistical dependence of nodes in the graph
- determine marginal exchangeability in the graph

After class:

- Hernán and Robins 2020 Chapter 6.1 and 6.2