Tools to Ascend the Ladder of Causation

Causal reasoning

Charles Assaad charles.assaad@inserm.fr

L'Institut Pierre Louis d'Epidémiologie et de Santé Publique,
Inserm, Sorbonne Université

June 2026

SANTÉ SORBONNE UNIVERSITÉ

SANTÉ SORBONNE UNIVERSITÉ

1

Causal hierarchy

Some health-related questions

Some health-related questions

- Do people who exercise more tend to have lower rates of heart disease?

Some health-related questions

- Do people who exercise more tend to have lower rates of heart disease?
- Can we predict who is at higher risk of developing diabetes using BIG DATA?

Some health-related questions

- Do people who exercise more tend to have lower rates of heart disease?
- Can we predict who is at higher risk of developing diabetes using BIG DATA?
- Does exercising reduce cholesterol?

Some health-related questions

- Do people who exercise more tend to have lower rates of heart disease?
- Can we predict who is at higher risk of developing diabetes using BIG DATA?
- Does exercising reduce cholesterol?
- What happens to headache intensity if we prescribe aspirin?

Some health-related questions

- Do people who exercise more tend to have lower rates of heart disease?
- Can we predict who is at higher risk of developing diabetes using BIG DATA?
- Does exercising reduce cholesterol?
- What happens to headache intensity if we prescribe aspirin?
- Would COVID-19 outcomes in France have changed with a different policy?

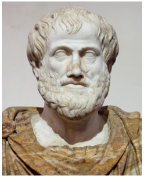
Some health-related questions

- Do people who exercise more tend to have lower rates of heart disease?
- Can we predict who is at higher risk of developing diabetes using BIG DATA?
- Does exercising reduce cholesterol?
- What happens to headache intensity if we prescribe aspirin?
- Would COVID-19 outcomes in France have changed with a different policy?
- Was it the aspirin that stopped by headache?

A history of attempts ...

−384 1748 1954 1970

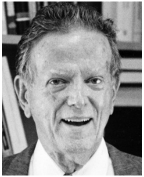
Aristotle



Hume

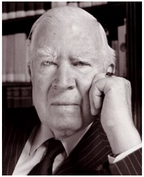
Reichenbach

Suppes

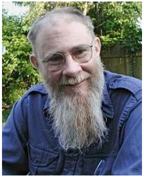


1971 1972 1973 2011

von Wright



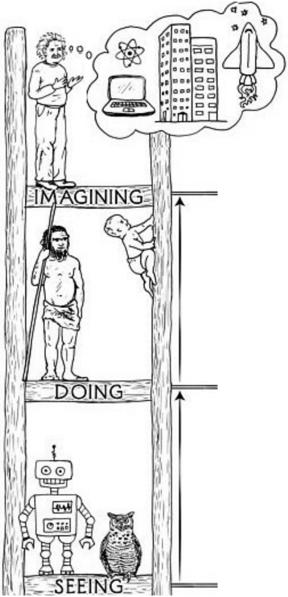
Lewis



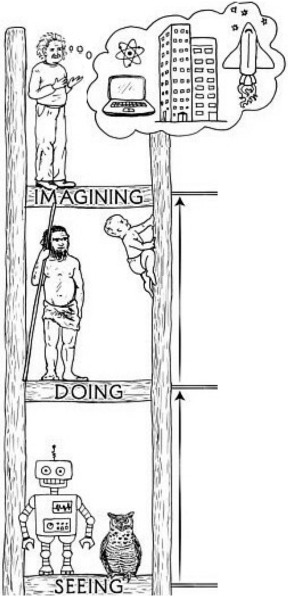
Anscombe

Pearl

Pearl causal hierarchy

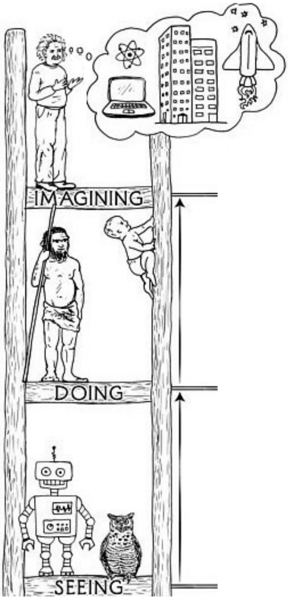


Pearl causal hierarchy



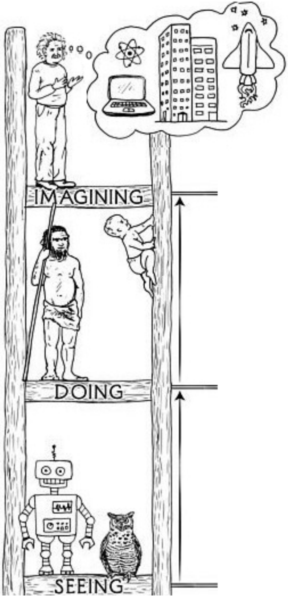
- Associations
  - Questions : What if I see ...?
  - Do people who exercise more tend to have lower rates of heart disease?
  - Can we predict who is at higher risk of developing diabetes using BIG DATA?

Pearl causal hierarchy



- Interventions
  - Questions: What if I do ...? How?
  - Does exercising reduce cholesterol?
  - What happens to headache intensity if we prescribe aspirin?
- Associations
  - Questions : What if I see ...?
  - Do people who exercise more tend to have lower rates of heart disease?
  - Can we predict who is at higher risk of developing diabetes using BIG DATA?

Pearl causal hierarchy



- Counterfactuals
  - Questions: What if I had done ...? Why?
  - Would COVID-19 outcomes in France have changed with a different policy?
  - Was it the aspirin that stopped by headache?
- Interventions
  - Questions: What if I do ...? How?
  - Does exercising reduce cholesterol?
  - What happens to headache intensity if we prescribe aspirin?
- Associations
  - Questions : What if I see ...?
  - Do people who exercise more tend to have lower rates of heart disease?
  - Can we predict who is at higher risk of developing diabetes using BIG DATA?

SANTÉ SORBONNE UNIVERSITÉ

2

Associations

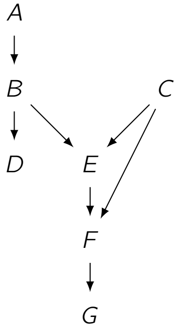
Notations

- Sets: $\mathbb{A} = \{X, Y, Z\}$
- Statistical independence: $\perp\!\!\!\perp_P$
- Statistical dependence: $\not\!\perp\!\!\!\perp_P$

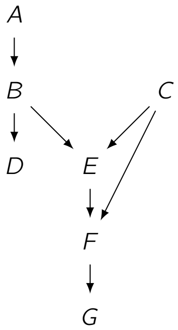
Directed acyclic graphs

A graph $\mathcal{G} = (\mathbb{V}, \mathbb{E})$ is said to be a directed acyclic graph (DAG) if

- $\mathbb{V}$ is the set of nodes (usually each node corresponds to a random variable),
- $\mathbb{E}$ is the set of edges,
- $\forall (X, Y) \in \mathbb{E}$,
  there is an arrow pointing from X to Y,
- there are no directed cycle in $\mathcal{G}$.



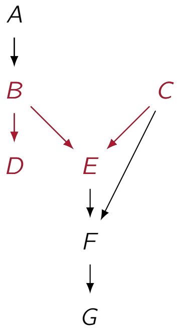
Consider the following DAG $\mathcal{G} = (\mathbb{V}, \mathbb{E})$:



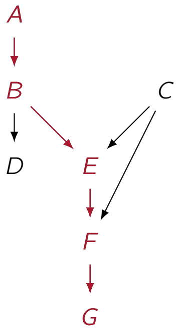
DAGs: basic concepts

Consider the following DAG $\mathcal{G} = (\mathbb{V}, \mathbb{E})$:

Path: $D \leftarrow B \rightarrow E \leftarrow C$

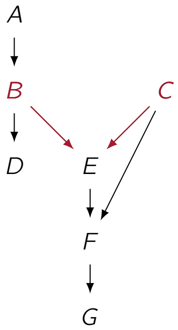


Consider the following DAG $\mathcal{G} = (\mathbb{V}, \mathbb{E})$:



Path: $D \leftarrow B \rightarrow E \leftarrow C$
Directed path: $A \rightarrow B \rightarrow E \rightarrow F \rightarrow G$

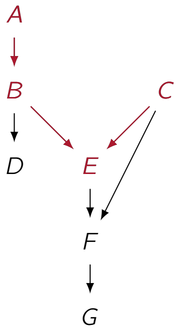
Consider the following DAG $\mathcal{G} = (\mathbb{V}, \mathbb{E})$:



Path: $D \leftarrow B \rightarrow E \leftarrow C$
Directed path: $A \rightarrow B \rightarrow E \rightarrow F \rightarrow G$
Parents: $Pa(E) = \{B, C\}$

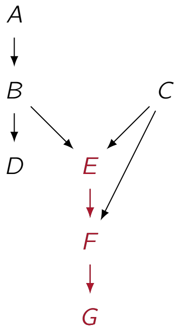
DAGs: basic concepts

Consider the following DAG $\mathcal{G} = (\mathbb{V}, \mathbb{E})$:



Path: $D \leftarrow B \rightarrow E \leftarrow C$
Directed path: $A \rightarrow B \rightarrow E \rightarrow F \rightarrow G$
Parents: $Pa(E) = \{B, C\}$
Ancestors: $An(E) = \{A, B, C, E\}$

Consider the following DAG $\mathcal{G} = (\mathbb{V}, \mathbb{E})$:



Path: $D \leftarrow B \rightarrow E \leftarrow C$

Directed path: $A \rightarrow B \rightarrow E \rightarrow F \rightarrow G$

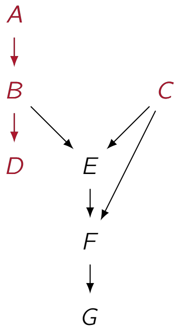
Parents: $Pa(E) = \{B, C\}$

Ancestors: $An(E) = \{A, B, C, E\}$

Descendants: $De(E) = \{E, F, G\}$

DAGs: basic concepts

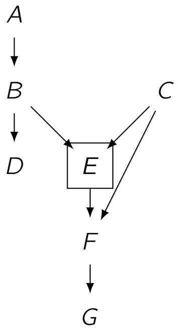
Consider the following DAG $\mathcal{G} = (\mathbb{V}, \mathbb{E})$:



Path: $D \leftarrow B \rightarrow E \leftarrow C$
Directed path: $A \rightarrow B \rightarrow E \rightarrow F \rightarrow G$
Parents: $Pa(E) = \{B, C\}$
Ancestors: $An(E) = \{A, B, C, E\}$
Descendants: $De(E) = \{E, F, G\}$
Non-descendants: $Nd(E) = \{A, B, C, D\}$

DAGs: basic concepts

Consider the following DAG $\mathcal{G} = (\mathbb{V}, \mathbb{E})$:



Path: $D \leftarrow B \rightarrow E \leftarrow C$
Directed path: $A \rightarrow B \rightarrow E \rightarrow F \rightarrow G$
Parents: $Pa(E) = \{B, C\}$
Ancestors: $An(E) = \{A, B, C, E\}$
Descendants: $De(E) = \{E, F, G\}$
Non-descendants: $Nd(E) = \{A, B, C, D\}$

Conditioning: represented by a square

Python code: DAGs using pyCIPHOD package

```
#Load the pyCIPHOD package
from pyciphod.graphs import DirectedAcyclicGraph
#Define the Directed graph
g = DirectedAcyclicGraph()
g.add_vertices(["A","B","C"])
g.add_directed_edge('B', 'A')
g.add_directed_edge('B', 'C')
g.remove_directed_edge("A", "C")
#Print nodes and edges
print(g.get_vertices())
print(g.get_directed_edges())
#Print parents and children
print(g.get_parents("A"))
print(g.get_children("B"))
```

Bayesian network

A distribution $P(\mathbb{V})$ is compatible with a DAG $\mathcal{G} = (\mathbb{V}, \mathbb{E})$ if

A distribution $P(\mathbb{V})$ is compatible with a DAG $\mathcal{G} = (\mathbb{V}, \mathbb{E})$ if

$$\forall X \in \mathbb{V}, X \perp\!\!\!\perp_P Nd(X) \mid Pa(X) \text{ (Markov condition)}$$

Bayesian network

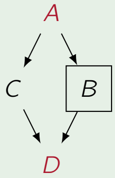
A distribution $P(\mathbb{V})$ is compatible with a DAG $\mathcal{G} = (\mathbb{V}, \mathbb{E})$ if

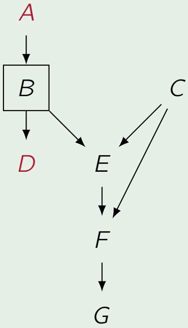
$$\forall X \in \mathbb{V}, X \perp\!\!\!\perp_P Nd(X) \mid Pa(X) \text{ (Markov condition)}$$

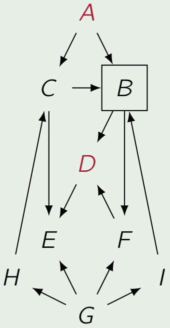
A DAG $\mathcal{G} = (\mathbb{V}, \mathbb{E})$ is a Bayesian network if there exists a joint distribution $P(\mathbb{V})$ that is compatible with $\mathcal{G}$.

Example

$A \overset{?}{\perp\!\!\!\perp}_P D \mid B$







Basic structures

Basic structures

Fork: contains a common parent

$$A \leftarrow C \rightarrow B \qquad\qquad A \leftarrow \boxed{C} \rightarrow B$$

$$A \overset{?}{\perp\!\!\!\perp}_P B \qquad\qquad A \perp\!\!\!\perp_P B \mid C$$

Basic structures

Fork: contains a common parent

$A \longleftarrow C \longrightarrow B$

$A \overset{?}{\perp\!\!\!\perp}_P B$

$A \longleftarrow \boxed{C} \longrightarrow B$

$A \perp\!\!\!\perp_P B \mid C$

Chain: directed path

$A \longrightarrow C \longrightarrow B$

$A \overset{?}{\perp\!\!\!\perp}_P B$

$A \longrightarrow \boxed{C} \longrightarrow B$

$A \perp\!\!\!\perp_P B \mid C$

Basic structures

Fork: contains a common parent

$$A \longleftarrow C \longrightarrow B \qquad\qquad A \longleftarrow \boxed{C} \longrightarrow B$$

$$A \overset{?}{\perp\!\!\!\perp}_P B \qquad\qquad A \perp\!\!\!\perp_P B \mid C$$

Chain: directed path

$$A \longrightarrow C \longrightarrow B \qquad\qquad A \longrightarrow \boxed{C} \longrightarrow B$$

$$A \overset{?}{\perp\!\!\!\perp}_P B \qquad\qquad A \perp\!\!\!\perp_P B \mid C$$

Collider : contains a common child

$$A \longrightarrow C \longleftarrow B \qquad\qquad A \longrightarrow \boxed{C} \longleftarrow B$$

$$A \perp\!\!\!\perp_P B \qquad\qquad A \overset{?}{\perp\!\!\!\perp}_P B \mid C$$

Basic structures

Fork: contains a common parent

$$A \longleftarrow C \longrightarrow B \qquad\qquad A \longleftarrow \boxed{C} \longrightarrow B$$

$$A \overset{?}{\perp\!\!\!\perp}_P B \qquad\qquad A \perp\!\!\!\perp_P B \mid C$$

Chain: directed path

$$A \longrightarrow C \longrightarrow B \qquad\qquad A \longrightarrow \boxed{C} \longrightarrow B$$

$$A \overset{?}{\perp\!\!\!\perp}_P B \qquad\qquad A \perp\!\!\!\perp_P B \mid C$$

Collider : contains a common child

$$A \longrightarrow C \longleftarrow B$$

$$A \perp\!\!\!\perp_P B$$

$$\begin{tikzcd} A \arrow[r] & C \arrow[d] & B \arrow[l] \\ & \boxed{D} & \end{tikzcd}$$

$$A \overset{?}{\perp\!\!\!\perp}_P B \mid D$$

Artificial correlation in colliders: example 1

Example

$$A \longrightarrow \boxed{C} \longleftarrow B$$

Artificial correlation in colliders: example 1

Example

$$A \longrightarrow \boxed{C} \longleftarrow B$$

$$A = \begin{cases} \text{Mother carrier} \\ \text{Mother not carrier} \end{cases} \qquad B = \begin{cases} \text{Father carrier} \\ \text{Father not carrier} \end{cases}$$

$$C = (A \text{ or } B) = \begin{cases} \text{Child carrier} \\ \text{Child not carrier} \end{cases}$$

Artificial correlation in colliders: example 1

Example

$$A \longrightarrow \boxed{C} \longleftarrow B$$

$$A = \begin{cases} \text{Mother carrier} \\ \text{Mother not carrier} \end{cases} \qquad B = \begin{cases} \text{Father carrier} \\ \text{Father not carrier} \end{cases}$$

$$C = (A \text{ or } B) = \begin{cases} \text{Child carrier} \\ \text{Child not carrier} \end{cases}$$

If $C = \text{Child carrier} \implies$

$$\begin{cases} \text{If } A = \text{Mother not carrier then } B = \text{Father carrier} \\ \text{If } B = \text{Father not carrier then } A = \text{Mother carrier} \end{cases}$$

Artificial correlation in colliders: example 2

Example

$$A \longrightarrow \boxed{C} \longleftarrow B$$

Artificial correlation in colliders: example 2

Example

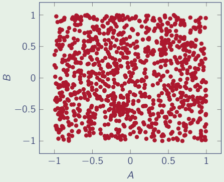
$$A \longrightarrow \boxed{C} \longleftarrow B$$

$A, B \sim U(-1, 1)$ $\quad U_c \sim N(0, \frac{1}{2})$ $\quad C = 2AB + U_c$

Example

$$A \longrightarrow \boxed{C} \longleftarrow B$$

$A, B \sim U(-1, 1)$ $\quad U_c \sim N(0, \frac{1}{2})$ $\quad C = 2AB + U_c$



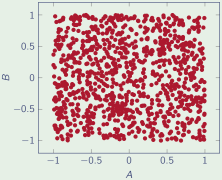
$Corr(A; B) = 0.002$

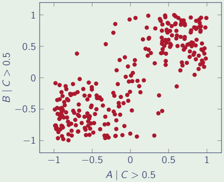
Artificial correlation in colliders: example 2

Example

$$A \longrightarrow \boxed{C} \longleftarrow B$$

$A, B \sim U(-1, 1)$ $\quad U_c \sim N(0, \frac{1}{2})$ $\quad C = 2AB + U_c$





$Corr(A; B) = 0.002$ $\quad Corr(A; B \mid C > 0.5) = 0.8$

Blocked paths

A path is said to be blocked by a set of nodes $\mathbb{Z} \in \mathbb{V}$ if:

- it contains a chain $A \to B \to C$ or a fork $A \leftarrow B \to C$ and $B \in \mathbb{Z}$; or
- it contains a collider $A \to B \leftarrow C$ such that no descendant of B is in $\mathbb{Z}$.

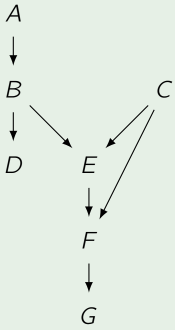
Blocked paths

A path is said to be blocked by a set of nodes $\mathbb{Z} \in \mathbb{V}$ if:

- it contains a chain $A \rightarrow B \rightarrow C$ or a fork $A \leftarrow B \rightarrow C$ and $B \in \mathbb{Z}$; or
- it contains a collider $A \rightarrow B \leftarrow C$ such that no descendant of B is in $\mathbb{Z}$.

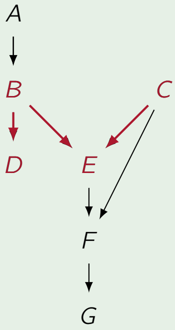
A path that is not blocked is active.

Example



Blocked path: examples

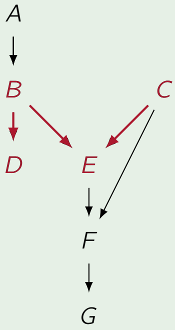
Example



Is the path $< D, B, E, C >$ blocked?

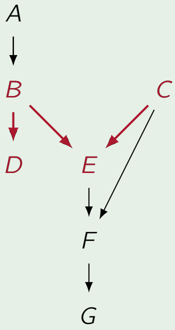
Blocked path: examples

Example



Is the path $< D, B, E, C >$ blocked? Yes

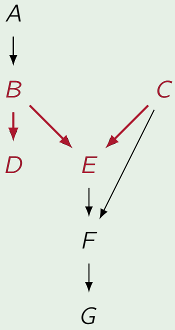
Example



Is the path $< D, B, E, C >$ blocked? Yes
Is the path $< D, B, E, C >$ blocked by E?

Blocked path: examples

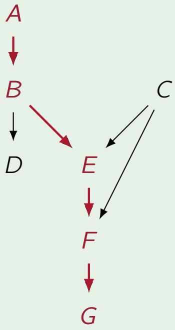
Example



Is the path $< D, B, E, C >$ blocked? Yes

Is the path $< D, B, E, C >$ blocked by E? No

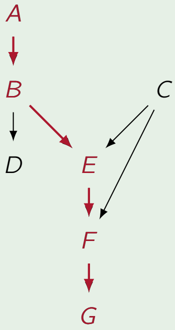
Example



Is the path $< D, B, E, C >$ blocked? Yes
Is the path $< D, B, E, C >$ blocked by E? No
Is the path $< A, B, E, F, G >$ blocked?

Blocked path: examples

Example



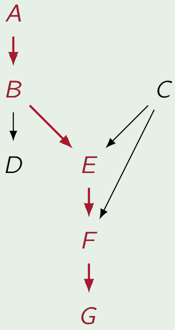
Is the path $< D, B, E, C >$ blocked? Yes

Is the path $< D, B, E, C >$ blocked by E? No

Is the path $< A, B, E, F, G >$ blocked? No

Blocked path: examples

Example



Is the path $< D, B, E, C >$ blocked? Yes

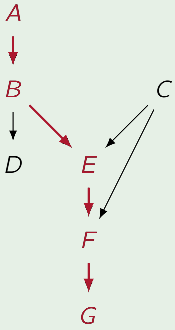
Is the path $< D, B, E, C >$ blocked by E? No

Is the path $< A, B, E, F, G >$ blocked? No

Is the path $< A, B, E, F, G >$ blocked by E?

Blocked path: examples

Example



Is the path $< D, B, E, C >$ blocked? Yes

Is the path $< D, B, E, C >$ blocked by E? No

Is the path $< A, B, E, F, G >$ blocked? No

Is the path $< A, B, E, F, G >$ blocked by E? Yes

d-separation

Given disjoint sets $\mathbb{X}, \mathbb{Y}, \mathbb{Z} \subseteq \mathbb{V}$, we say that $\mathbb{X}$ and $\mathbb{Y}$ are d-separated by $\mathbb{Z}$ if every path between a node in $\mathbb{X}$ and a node in $\mathbb{Y}$ is blocked by $\mathbb{Z}$ and we write $\mathbb{X} \perp\!\!\!\perp_{\mathcal{G}} \mathbb{Y} \mid \mathbb{Z}$.

Given disjoint sets $\mathbb{X}, \mathbb{Y}, \mathbb{Z} \subseteq \mathbb{V}$, we say that $\mathbb{X}$ and $\mathbb{Y}$ are d-separated by $\mathbb{Z}$ if every path between a node in $\mathbb{X}$ and a node in $\mathbb{Y}$ is blocked by $\mathbb{Z}$ and we write $\mathbb{X} \perp\!\!\!\perp_{\mathcal{G}} \mathbb{Y} \mid \mathbb{Z}$.

If one path is not blocked, we say that $\mathbb{X}$ and $\mathbb{Y}$ are d-connected given $\mathbb{Z}$ and we write $\mathbb{X} \not\perp\!\!\!\perp_{\mathcal{G}} \mathbb{Y} \mid \mathbb{Z}$.

d-separation

Given disjoint sets $\mathbb{X}, \mathbb{Y}, \mathbb{Z} \subseteq \mathbb{V}$, we say that $\mathbb{X}$ and $\mathbb{Y}$ are d-separated by $\mathbb{Z}$ if every path between a node in $\mathbb{X}$ and a node in $\mathbb{Y}$ is blocked by $\mathbb{Z}$ and we write $\mathbb{X} \perp\!\!\!\perp_{\mathcal{G}} \mathbb{Y} \mid \mathbb{Z}$.

If one path is not blocked, we say that $\mathbb{X}$ and $\mathbb{Y}$ are d-connected given $\mathbb{Z}$ and we write $\mathbb{X} \not\perp\!\!\!\perp_{\mathcal{G}} \mathbb{Y} \mid \mathbb{Z}$.

Theorem

$\mathbb{X} \perp\!\!\!\perp_{\mathcal{G}} \mathbb{Y} \mid \mathbb{Z} \Rightarrow \forall P$ compatible with $\mathcal{G}, \mathbb{X} \perp\!\!\!\perp_{P} \mathbb{Y} \mid \mathbb{Z}$

Given disjoint sets $\mathbb{X}, \mathbb{Y}, \mathbb{Z} \subseteq \mathbb{V}$, we say that $\mathbb{X}$ and $\mathbb{Y}$ are d-separated by $\mathbb{Z}$ if every path between a node in $\mathbb{X}$ and a node in $\mathbb{Y}$ is blocked by $\mathbb{Z}$ and we write $\mathbb{X} \perp\!\!\!\perp_{\mathcal{G}} \mathbb{Y} \mid \mathbb{Z}$.

If one path is not blocked, we say that $\mathbb{X}$ and $\mathbb{Y}$ are d-connected given $\mathbb{Z}$ and we write $\mathbb{X} \not\perp\!\!\!\perp_{\mathcal{G}} \mathbb{Y} \mid \mathbb{Z}$.

Theorem

$\mathbb{X} \perp\!\!\!\perp_{\mathcal{G}} \mathbb{Y} \mid \mathbb{Z} \Rightarrow \forall P$ compatible with $\mathcal{G}, \mathbb{X} \perp\!\!\!\perp_{P} \mathbb{Y} \mid \mathbb{Z}$
but $\mathbb{X} \not\perp\!\!\!\perp_{\mathcal{G}} \mathbb{Y} \mid \mathbb{Z} \not\Rightarrow \forall P$ compatible with $\mathcal{G}, \mathbb{X} \not\perp\!\!\!\perp_{P} \mathbb{Y} \mid \mathbb{Z}$

d-separation

Given disjoint sets $\mathbb{X}, \mathbb{Y}, \mathbb{Z} \subseteq \mathbb{V}$, we say that $\mathbb{X}$ and $\mathbb{Y}$ are d-separated by $\mathbb{Z}$ if every path between a node in $\mathbb{X}$ and a node in $\mathbb{Y}$ is blocked by $\mathbb{Z}$ and we write $\mathbb{X} \perp\!\!\!\perp_{\mathcal{G}} \mathbb{Y} \mid \mathbb{Z}$.

If one path is not blocked, we say that $\mathbb{X}$ and $\mathbb{Y}$ are d-connected given $\mathbb{Z}$ and we write $\mathbb{X} \not\perp\!\!\!\perp_{\mathcal{G}} \mathbb{Y} \mid \mathbb{Z}$.

Theorem

$\mathbb{X} \perp\!\!\!\perp_{\mathcal{G}} \mathbb{Y} \mid \mathbb{Z} \Rightarrow \forall P$ compatible with $\mathcal{G}, \mathbb{X} \perp\!\!\!\perp_{P} \mathbb{Y} \mid \mathbb{Z}$
but $\mathbb{X} \not\perp\!\!\!\perp_{\mathcal{G}} \mathbb{Y} \mid \mathbb{Z} \not\Rightarrow \forall P$ compatible with $\mathcal{G}, \mathbb{X} \not\perp\!\!\!\perp_{P} \mathbb{Y} \mid \mathbb{Z}$ instead
$\mathbb{X} \not\perp\!\!\!\perp_{\mathcal{G}} \mathbb{Y} \mid \mathbb{Z} \Rightarrow \exists P$ compatible with $\mathcal{G}, \mathbb{X} \not\perp\!\!\!\perp_{P} \mathbb{Y} \mid \mathbb{Z}$

Given disjoint sets $\mathbb{X}, \mathbb{Y}, \mathbb{Z} \subseteq \mathbb{V}$, we say that $\mathbb{X}$ and $\mathbb{Y}$ are d-separated by $\mathbb{Z}$ if every path between a node in $\mathbb{X}$ and a node in $\mathbb{Y}$ is blocked by $\mathbb{Z}$ and we write $\mathbb{X} \perp\!\!\!\perp_{\mathcal{G}} \mathbb{Y} \mid \mathbb{Z}$.

If one path is not blocked, we say that $\mathbb{X}$ and $\mathbb{Y}$ are d-connected given $\mathbb{Z}$ and we write $\mathbb{X} \not\!\perp\!\!\!\perp_{\mathcal{G}} \mathbb{Y} \mid \mathbb{Z}$.

Theorem

$\mathbb{X} \perp\!\!\!\perp_{\mathcal{G}} \mathbb{Y} \mid \mathbb{Z} \Rightarrow \forall P$ compatible with $\mathcal{G}, \mathbb{X} \perp\!\!\!\perp_{P} \mathbb{Y} \mid \mathbb{Z}$
but $\mathbb{X} \not\!\perp\!\!\!\perp_{\mathcal{G}} \mathbb{Y} \mid \mathbb{Z} \not\Rightarrow \forall P$ compatible with $\mathcal{G}, \mathbb{X} \not\!\perp\!\!\!\perp_{P} \mathbb{Y} \mid \mathbb{Z}$ instead
$\mathbb{X} \not\!\perp\!\!\!\perp_{\mathcal{G}} \mathbb{Y} \mid \mathbb{Z} \Rightarrow \exists P$ compatible with $\mathcal{G}, \mathbb{X} \not\!\perp\!\!\!\perp_{P} \mathbb{Y} \mid \mathbb{Z}$

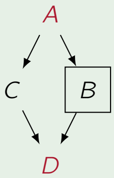
$$\mathbb{X} \perp\!\!\!\perp_{\mathcal{G}} \mathbb{Y} \mid \mathbb{Z} \Rightarrow \mathbb{X} \perp\!\!\!\perp_{P} \mathbb{Y} \mid \mathbb{Z}$$

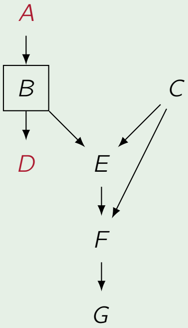
$$\mathbb{X} \not\!\perp\!\!\!\perp_{\mathcal{G}} \mathbb{Y} \mid \mathbb{Z} \Rightarrow \mathbb{X} \overset{?}{\perp\!\!\!\perp}_{P} \mathbb{Y} \mid \mathbb{Z}$$

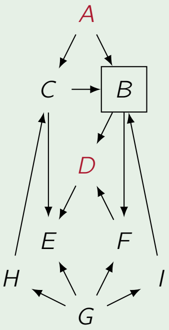
Reading conditional independencies using d-separation

Example

Is $A \perp\!\!\!\perp_P D \mid B$?

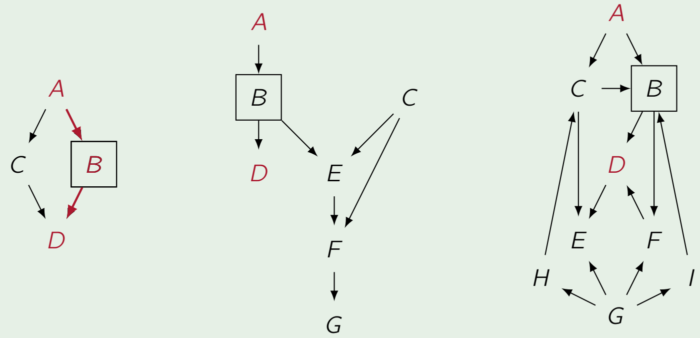






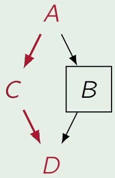
Example

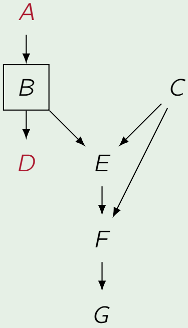
Is $A \perp\!\!\!\perp_P D \mid B$?

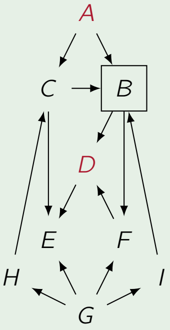


Example

Is $A \perp\!\!\!\perp_P D \mid B$?



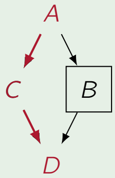




Reading conditional independencies using d-separation

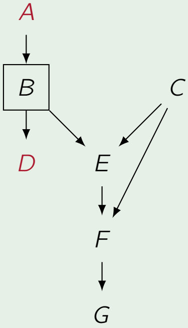
Example

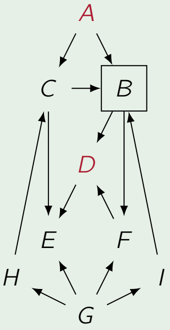
Is $A \perp\!\!\!\perp_P D \mid B$?



$< A, C, D >$ is not blocked

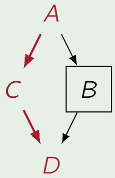
$\overset{?}{\Longrightarrow} A \perp\!\!\!\perp_P D \mid B$





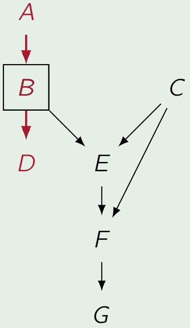
Example

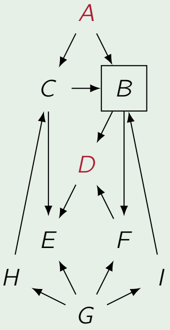
Is $A \perp\!\!\!\perp_P D \mid B$?



$< A, C, D >$ is not blocked

$\overset{?}{\Longrightarrow} A \perp\!\!\!\perp_P D \mid B$

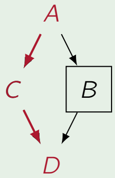




Reading conditional independencies using d-separation

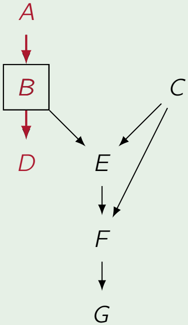
Example

Is $A \perp\!\!\!\perp_P D \mid B$?



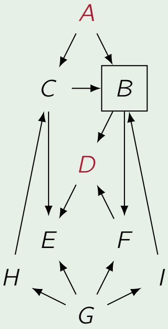
$< A, C, D >$ is not blocked

$\implies A \overset{?}{\perp\!\!\!\perp}_P D \mid B$



All paths are blocked

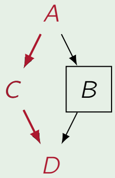
$\implies A \perp\!\!\!\perp_P D \mid B$



Reading conditional independencies using d-separation

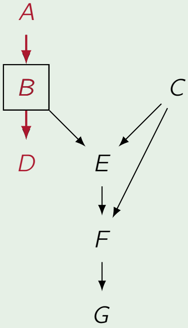
Example

Is $A \perp\!\!\!\perp_P D \mid B$?



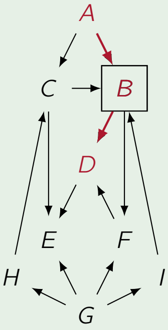
$< A, C, D >$ is not blocked

$\overset{?}{\Longrightarrow} A \perp\!\!\!\perp_P D \mid B$



All paths are blocked

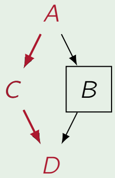
$\Longrightarrow A \perp\!\!\!\perp_P D \mid B$



Reading conditional independencies using d-separation

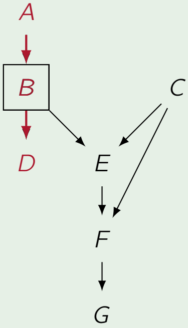
Example

Is $A \perp\!\!\!\perp_P D \mid B$?



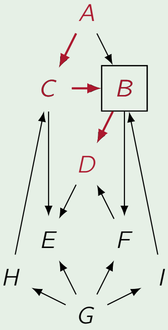
$< A, C, D >$ is not blocked

$\Longrightarrow A \overset{?}{\perp\!\!\!\perp}_P D \mid B$



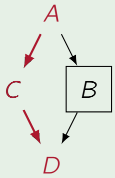
All paths are blocked

$\Longrightarrow A \perp\!\!\!\perp_P D \mid B$



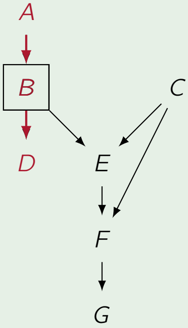
Example

Is $A \perp\!\!\!\perp_P D \mid B$?



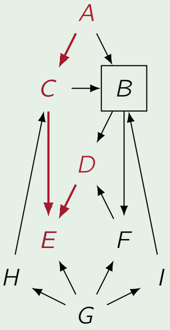
$< A, C, D >$ is not blocked

$\implies A \overset{?}{\perp\!\!\!\perp}_P D \mid B$



All paths are blocked

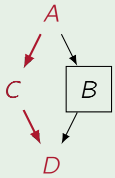
$\implies A \perp\!\!\!\perp_P D \mid B$



Reading conditional independencies using d-separation

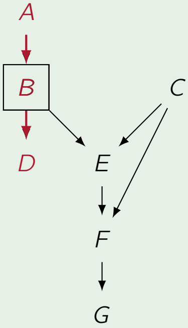
Example

Is $A \perp\!\!\!\perp_P D \mid B$?



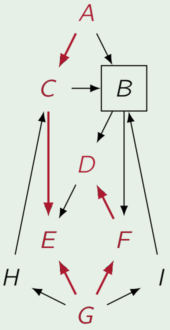
$< A, C, D >$ is not blocked

$\implies A \overset{?}{\perp\!\!\!\perp}_P D \mid B$



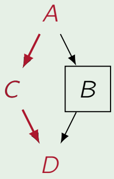
All paths are blocked

$\implies A \perp\!\!\!\perp_P D \mid B$



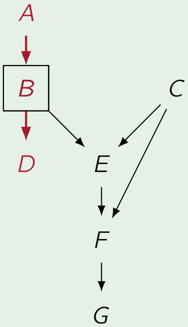
Example

Is $A \perp\!\!\!\perp_P D \mid B$?



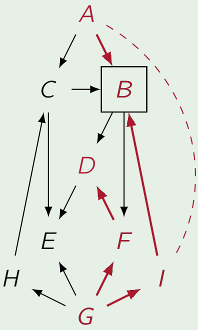
$< A, C, D >$ is not blocked

$\implies A \overset{?}{\perp\!\!\!\perp}_P D \mid B$



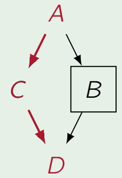
All paths are blocked

$\implies A \perp\!\!\!\perp_P D \mid B$



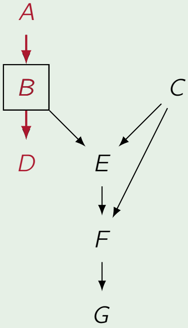
Example

Is $A \perp\!\!\!\perp_P D \mid B$?



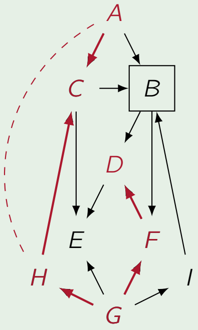
$< A, C, D >$ is not blocked

$\implies A \overset{?}{\perp\!\!\!\perp}_P D \mid B$



All paths are blocked

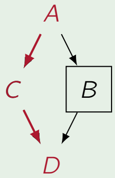
$\implies A \perp\!\!\!\perp_P D \mid B$



Reading conditional independencies using d-separation

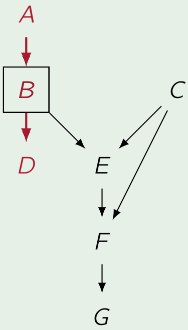
Example

Is $A \perp\!\!\!\perp_P D \mid B$?



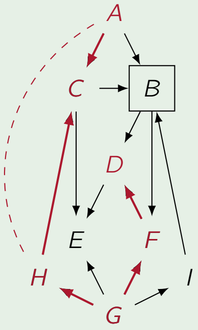
$< A, C, D >$ is not blocked

$\implies A \overset{?}{\perp\!\!\!\perp}_P D \mid B$



All paths are blocked

$\implies A \perp\!\!\!\perp_P D \mid B$



$< A, B, I, G, F, D >$ is not blocked

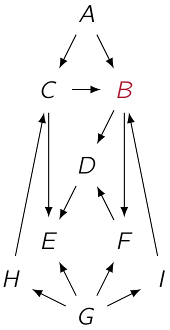
$\implies A \overset{?}{\perp\!\!\!\perp}_P D \mid B$

Python code: d-separation using pyCIPHOD package

```
#Load the pyCIPHOD package
from pyciphod.graphs import DirectedAcyclicGraph
from pyciphod.separation import d_separated
#Define the DAG
g = DirectedAcyclicGraph()
g.add_vertices(["A","B","C"])
g.add_directed_edge('B', 'A')
g.add_directed_edge('B', 'C')
#Check if B d-separates A and C
ds = g.d_separated(["A"], ["C"], ["B"])
print(ds)
```

Exercise 1

What are the relevant variables for predicting B?



3

Interventions

Causal effect

$$\text{The causal effect of } \mathbb{A} \text{ on } \mathbb{Y}$$
$$= \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a})) - \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a}')).$$

The operator $do()$ represents interventions.

Causal effect

The causal effect of $\mathbb{A}$ on $\mathbb{Y}$

$$= \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a})) - \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a}'))\text{.}$$

The operator $do()$ represents interventions.

Population

Causal effect

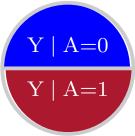
The causal effect of $\mathbb{A}$ on $\mathbb{Y}$

$$= \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a})) - \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a}')).$$

The operator $do()$ represents interventions.

Population

Sub-populations



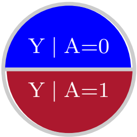
Causal effect

$$\text{The causal effect of } \mathbb{A} \text{ on } \mathbb{Y}$$
$$= \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a})) - \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a}')).$$

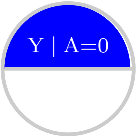
The operator $do()$ represents interventions.

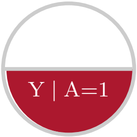
Population

Sub-populations



Conditioning





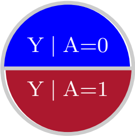
The causal effect of $\mathbb{A}$ on $\mathbb{Y}$

$$= \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a})) - \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a}')).$$

The operator $do()$ represents interventions.

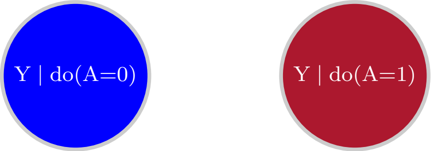
Population

Sub-populations

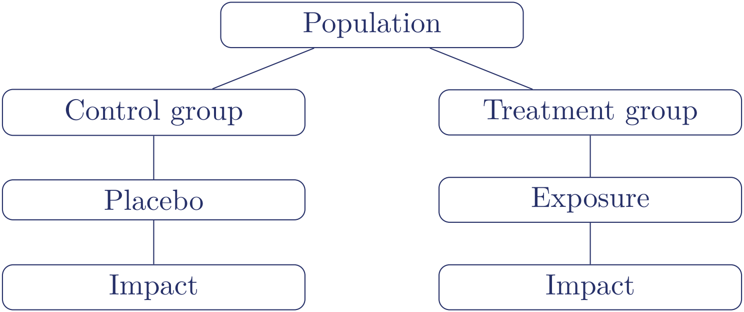


Conditioning

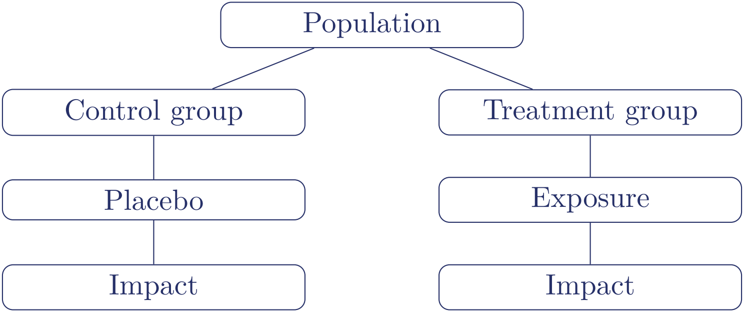
Intervening



Gold standard: Randomized controlled trials



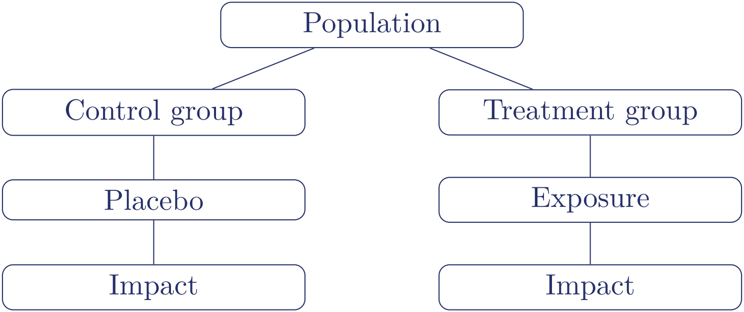
Gold standard: Randomized controlled trials



Limitations:

- Unethical
- Infeasible
- Costly

Gold standard: Randomized controlled trials



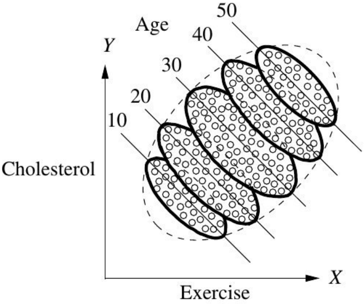
Limitations:

- Unethical
- Infeasible
- Costly

$\Longrightarrow$ Sometimes we have to rely on observational studies.

Simpson paradox

In a study, we measure weekly exercise and cholesterol levels across different age groups. What is the causal effect of exercise on cholesterol?



Causal inference in observational studies

Estimating causal effects from observational data necessarily requires strong assumptions.

Estimating causal effects from observational data necessarily requires strong assumptions.

Two main frameworks

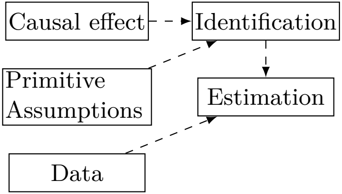
Potential outcomes

Structural causal models

Estimating causal effects from observational data necessarily requires strong assumptions.

Two main frameworks

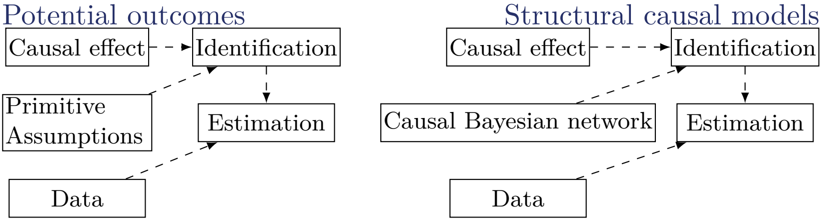
Potential outcomes



Structural causal models

Estimating causal effects from observational data necessarily requires strong assumptions.

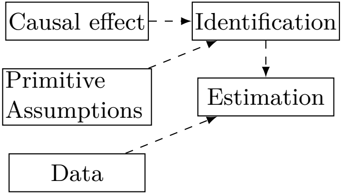
Two main frameworks



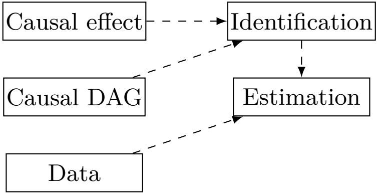
Estimating causal effects from observational data necessarily requires strong assumptions.

Two main frameworks

Potential outcomes

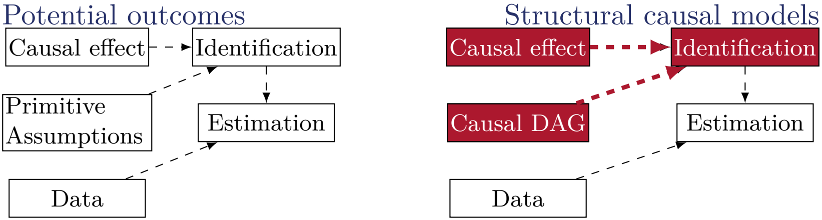


Structural causal models



Estimating causal effects from observational data necessarily requires strong assumptions.

Two main frameworks



Estimating causal effects from observational data necessarily requires strong assumptions.

Two main frameworks

Potential outcomes

Causal effect → Identification

Primitive Assumptions → Identification

Identification → Estimation

Data → Estimation

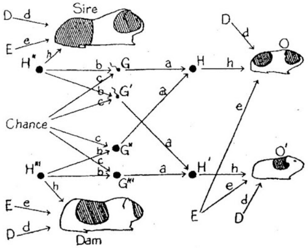
Structural causal models

Causal effect → Identification

Causal DAG → Identification

Identification → Estimation

Data → Estimation



Sewal Wright

Causal inference in observational studies

Estimating causal effects from observational data necessarily requires strong assumptions.

Two main frameworks

Potential outcomes

Causal effect → Identification

Primitive Assumptions → Identification

Identification → Estimation

Data → Estimation

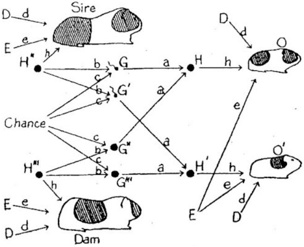
Structural causal models

Causal effect → Identification

Causal DAG → Identification

Identification → Estimation

Data → Estimation



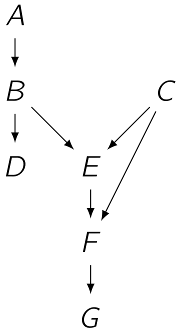
Sewal Wright

Judea Pearl

Causal DAGs

A causal DAG $\mathcal{G} = (\mathbb{V}, \mathbb{E})$ is a bayesian network where

- $\mathbb{E}$ represents direct causal relations between variables in $\mathbb{V}$.

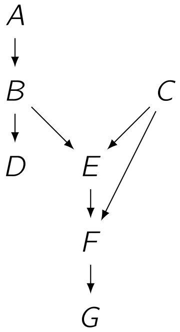


Causal DAGs

A causal DAG $\mathcal{G} = (\mathbb{V}, \mathbb{E})$ is a bayesian network where

- $\mathbb{E}$ represents direct causal relations between variables in $\mathbb{V}$.

$$\text{Structural causal model : } \begin{cases} A := f_a(U_a) \\ B := f_b(A, U_b) \\ C := f_c(U_c) \\ D := f_d(B, U_d) \\ E := f_e(B, C, U_e) \\ F := f_f(C, E, U_f) \\ G := f_g(F, U_g) \end{cases}$$

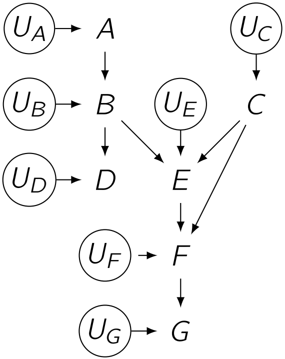


Causal DAGs

A causal DAG $\mathcal{G} = (\mathbb{V}, \mathbb{E})$ is a bayesian network where

- $\mathbb{E}$ represents direct causal relations between variables in $\mathbb{V}$.

Structural causal model : $\begin{cases} A := f_a(U_a) \\ B := f_b(A, U_b) \\ C := f_c(U_c) \\ D := f_d(B, U_d) \\ E := f_e(B, C, U_e) \\ F := f_f(C, E, U_f) \\ G := f_g(F, U_g) \end{cases}$

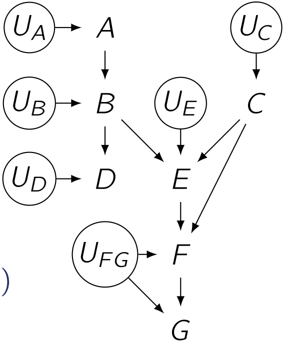


Causal DAGs

A causal DAG $\mathcal{G} = (\mathbb{V}, \mathbb{E})$ is a bayesian network where

- $\mathbb{E}$ represents direct causal relations between variables in $\mathbb{V}$.

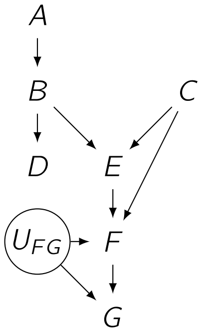
Structural causal model :
$$\begin{cases} A := f_a(U_a) \\ B := f_b(A, U_b) \\ C := f_c(U_c) \\ D := f_d(B, U_d) \\ E := f_e(B, C, U_e) \\ F := f_f(C, E, U_{fg}) \\ G := f_g(F, U_{fg}) \end{cases}$$



A causal DAG $\mathcal{G} = (\mathbb{V}, \mathbb{E})$ is a bayesian network where

- $\mathbb{E}$ represents direct causal relations between variables in $\mathbb{V}$.

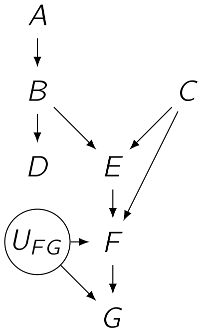
Structural causal model : $\begin{cases} A := f_a(U_a) \\ B := f_b(A, U_b) \\ C := f_c(U_c) \\ D := f_d(B, U_d) \\ E := f_e(B, C, U_e) \\ F := f_f(C, E, U_{fg}) \\ G := f_g(F, U_{fg}) \end{cases}$



A causal DAG $\mathcal{G} = (\mathbb{V}, \mathbb{E})$ is a bayesian network where

- $\mathbb{E}$ represents direct causal relations between variables in $\mathbb{V}$.
- Unmeasured confounders should be represented graphically (represented by a circle).

Structural causal model : $\begin{cases} A := f_a(U_a) \\ B := f_b(A, U_b) \\ C := f_c(U_c) \\ D := f_d(B, U_d) \\ E := f_e(B, C, U_e) \\ F := f_f(C, E, U_{fg}) \\ G := f_g(F, U_{fg}) \end{cases}$

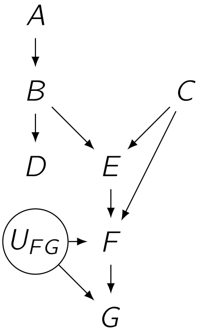


Causal DAGs

A causal DAG $\mathcal{G} = (\mathbb{V}, \mathbb{E})$ is a bayesian network where

- $\mathbb{E}$ represents direct causal relations between variables in $\mathbb{V}$.
- Unmeasured confounders should be represented graphically (represented by a circle).

$$\text{Structural causal model : } \begin{cases} A := f_a(U_a) \\ B := f_b(A, U_b) \\ C := f_c(U_c) \\ D := f_d(B, U_d) \\ E := f_e(B, C, U_e) \\ F := f_f(C, E, U_{fg}) \\ G := f_g(F, U_{fg}) \end{cases}$$

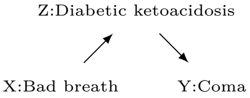


In a structural causal model, the intervention $\mathrm{do}(F = f)$ is defined by replacing the structural equation of F with $F := f$, leaving all other equations unchanged.

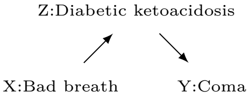
Bayesian networks vs causal DAGs

Not all Bayesian networks are necessarily causal DAGs.

Not all Bayesian networks are necessarily causal DAGs.



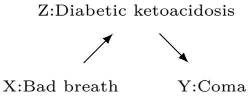
Not all Bayesian networks are necessarily causal DAGs.



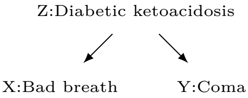
Bayesian network

Bayesian networks vs causal DAGs

Not all Bayesian networks are necessarily causal DAGs.



Bayesian network



Bayesian networks vs causal DAGs

Not all Bayesian networks are necessarily causal DAGs.

$$\begin{tikzcd} & \text{Z:Diabetic ketoacidosis} \arrow[dr] & \\ \text{X:Bad breath} \arrow[ur] & & \text{Y:Coma} \end{tikzcd}$$

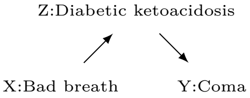
Bayesian network

$$\begin{tikzcd} & \text{Z:Diabetic ketoacidosis} \arrow[dl] \arrow[dr] & \\ \text{X:Bad breath} & & \text{Y:Coma} \end{tikzcd}$$

Causal DAG

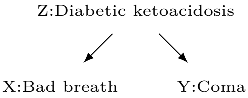
Bayesian networks vs causal DAGs

Not all Bayesian networks are necessarily causal DAGs.



Bayesian network

Oracle for conditional independencies

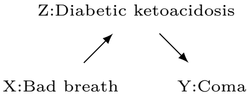


Causal DAG

Oracle for conditional independencies
+
Oracle for interventions

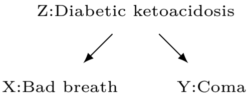
Bayesian networks vs causal DAGs

Not all Bayesian networks are necessarily causal DAGs.



Bayesian network

Oracle for conditional independencies

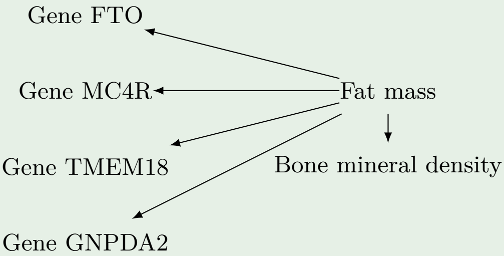


Causal DAG

Oracle for conditional independencies
+
Oracle for interventions

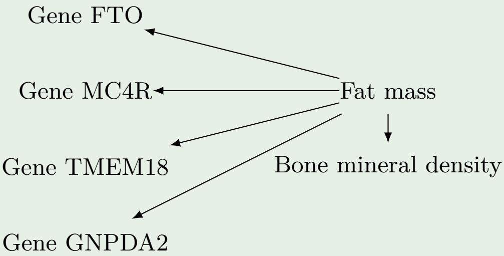
It is impossible to determine, without additional assumptions, which of the two DAGs is causal based solely on the observed distribution!

Example



Is this DAG causal?

Example

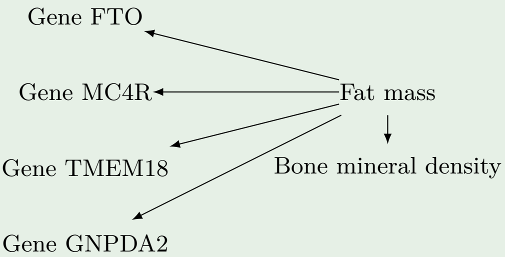


Is this DAG causal?

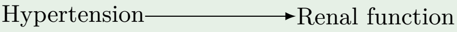
No

Bayesian networks vs causal DAGs: examples

Example



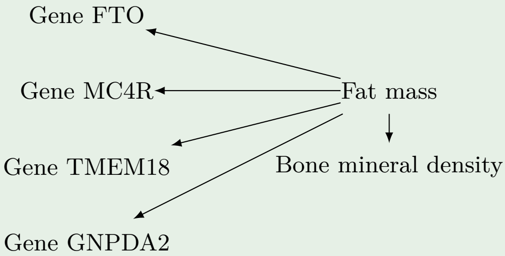
Is this DAG causal?
No



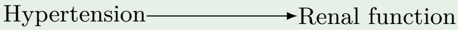
Is this DAG causal?

Bayesian networks vs causal DAGs: examples

Example



Is this DAG causal?
No



Is this DAG causal?
Yes

In causal DAGs, the information lies in the absence of edges and the direction of arrows, rather than merely in the presence of edges.

Guidlines for constructing causal DAGs

In causal DAGs, the information lies in the absence of edges and the direction of arrows, rather than merely in the presence of edges.

- Think carrefully about the orientation

Guidlines for constructing causal DAGs

In causal DAGs, the information lies in the absence of edges and the direction of arrows, rather than merely in the presence of edges.

- Think carrefully about the orientation
- When you are not sure if you need to add or not an edge (for example $Z \to A$) to the graph, ADD IT! (as long as you keep the graph acyclic)

Causal paths vs non-causal paths

The causal effect of $\mathbb{A}$ on $\mathbb{Y} = \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a})) - \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a}'))$.

The causal effect of $\mathbb{A}$ on $\mathbb{Y} = \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a})) - \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a}'))$.

It is identifiable if $P(\mathbb{Y} = \mathbb{y} \mid do(\mathbb{A} = \mathbb{a}))$ is identifiable.

Causal paths vs non-causal paths

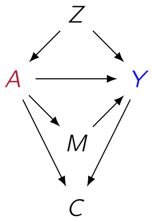
The causal effect of $\mathbb{A}$ on $\mathbb{Y} = \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a})) - \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a}'))$.

It is identifiable if $P(\mathbb{Y} = \mathbb{y} \mid do(\mathbb{A} = \mathbb{a}))$ is identifiable.
Equivalently, if $P(\mathbb{Y} = \mathbb{y} \mid do(\mathbb{A} = \mathbb{a}))$ can be expressed without the do() operator.

Causal paths vs non-causal paths

The causal effect of $\mathbb{A}$ on $\mathbb{Y} = \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a})) - \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a}'))$.

It is identifiable if $P(\mathbb{Y} = \mathbb{y} \mid do(\mathbb{A} = \mathbb{a}))$ is identifiable.
Equivalently, if $P(\mathbb{Y} = \mathbb{y} \mid do(\mathbb{A} = \mathbb{a}))$ can be expressed without the do() operator.

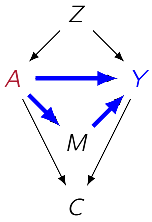


Causal paths vs non-causal paths

The causal effect of $\mathbb{A}$ on $\mathbb{Y} = \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a})) - \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a}'))$.

It is identifiable if $P(\mathbb{Y} = \mathbb{y} \mid do(\mathbb{A} = \mathbb{a}))$ is identifiable.
Equivalently, if $P(\mathbb{Y} = \mathbb{y} \mid do(\mathbb{A} = \mathbb{a}))$ can be expressed without the do() operator.

- Causal paths: All directed paths A to Y.

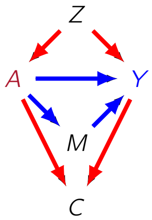


Causal paths vs non-causal paths

The causal effect of $\mathbb{A}$ on $\mathbb{Y} = \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a})) - \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a}')).$

It is identifiable if $P(\mathbb{Y} = \mathbb{y} \mid do(\mathbb{A} = \mathbb{a}))$ is identifiable.
Equivalently, if $P(\mathbb{Y} = \mathbb{y} \mid do(\mathbb{A} = \mathbb{a}))$ can be expressed without the do() operator.

- Causal paths: All directed paths A to Y.
- Non-causal paths: All non-directed paths A to Y.

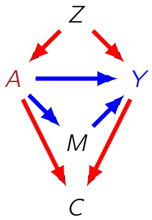


Causal paths vs non-causal paths

The causal effect of $\mathbb{A}$ on $\mathbb{Y} = \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a})) - \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a}'))$.

It is identifiable if $P(\mathbb{Y} = \mathbb{y} \mid do(\mathbb{A} = \mathbb{a}))$ is identifiable.
Equivalently, if $P(\mathbb{Y} = \mathbb{y} \mid do(\mathbb{A} = \mathbb{a}))$ can be expressed without the do() operator.

- Causal paths: All directed paths A to Y.
- Non-causal paths: All non-directed paths A to Y.

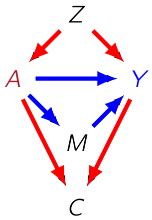


Identification involves finding a way of keeping all associations induced by causal path while removing all association induced by non-causal paths.

The causal effect of $\mathbb{A}$ on $\mathbb{Y} = \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a})) - \mathbb{E}(\mathbb{Y} \mid do(\mathbb{A} = \mathbb{a}'))$.

It is identifiable if $P(\mathbb{Y} = \mathbb{y} \mid do(\mathbb{A} = \mathbb{a}))$ is identifiable.
Equivalently, if $P(\mathbb{Y} = \mathbb{y} \mid do(\mathbb{A} = \mathbb{a}))$ can be expressed without the do() operator.

- Causal paths: All directed paths A to Y.
- Non-causal paths: All non-directed paths A to Y.



Identification involves finding a way of keeping all associations induced by causal path while removing all association induced by non-causal paths.

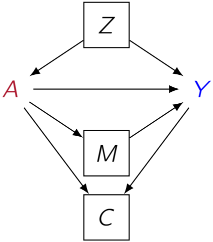
In what follows, we will simplify the notation by using $P(\mathbb{y} \mid do(\mathbb{a}))$ instead of $P(\mathbb{Y} = \mathbb{y} \mid do(\mathbb{A} = \mathbb{a}))$.

Non-graphical solutions (1/2)

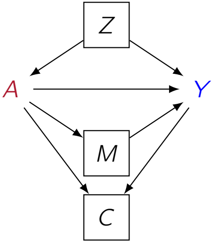
Solution 1: Adjust for all available variables.

Non-graphical solutions (1/2)

Solution 1: Adjust for all available variables.



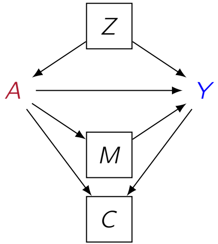
Solution 1: Adjust for all available variables.



- Can block non-causal paths (remove confounding bias if all confounders are measured)

Non-graphical solutions (1/2)

Solution 1: Adjust for all available variables.



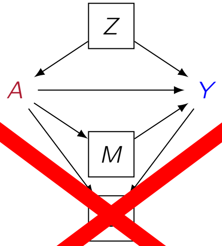
- Can block non-causal paths (remove confounding bias if all confounders are measured)

But

- Can block causal paths
- Can activate non-causal paths (collider bias and selection bias)

Non-graphical solutions (1/2)

Solution 1: Adjust for all available variables.



- Can block non-causal paths (removes confounding bias if all confounders are measured)

But

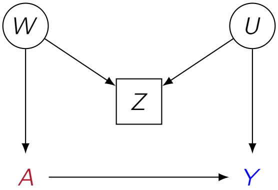
- Can block causal paths
- Can activate non-causal paths (collider bias and selection bias)

Non-graphical solutions (2/2)

Solution 2: Adjust for all available variables that are measured before the exposure.

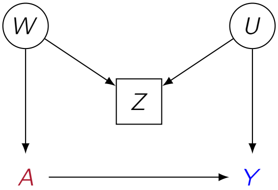
Non-graphical solutions (2/2)

Solution 2: Adjust for all available variables that are measured before the exposure.



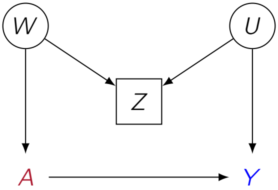
Non-graphical solutions (2/2)

Solution 2: Adjust for all available variables that are measured before the exposure.



- Can block non-causal paths (remove confounding bias if all confounders are measured)
- Cannot block causal paths

Solution 2: Adjust for all available variables that are measured before the exposure.



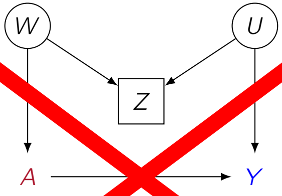
- Can block non-causal paths (remove confounding bias if all confounders are measured)
- Cannot block causal paths

But

- Can activate non-causal paths (collider bias and artificial confounding bias)

Non-graphical solutions (2/2)

Solution 2: Adjust for all available variables that are measured before the exposure.



- Can block non-causal paths (remove confounding bias if all confounders are measured)
- Cannot block causal paths

But

- Can activate non-causal paths (collider bias and artificial confounding bias)

The back-door criterion

A set of measured variables $\mathbb{Z}$ satisfies the back-door criterion relative to an ordered pair of variables (A, Y) in causal DAG $\mathcal{G}$ if:

- No node in $\mathbb{Z}$ is a descendant of A; and
- $\mathbb{Z}$ blocks all non-causal paths between A and Y.

The back-door criterion

A set of measured variables $\mathbb{Z}$ satisfies the back-door criterion relative to an ordered pair of variables (A, Y) in causal DAG $\mathcal{G}$ if:

- No node in $\mathbb{Z}$ is a descendant of A; and
- $\mathbb{Z}$ blocks all non-causal paths between A and Y.

Theorem

If $\mathbb{Z}$ satisfies the back-door criterion with respect to (A, Y) and if $P(a, \mathbb{z}) > 0$, then $P(y \mid do(a))$ is identifiable and is given by:

$$P(y \mid do(a)) = \sum_{\mathbb{z}} P(y \mid a, \mathbb{z}) P(\mathbb{z}).$$

The back-door criterion

A set of measured variables $\mathbb{Z}$ satisfies the back-door criterion relative to an ordered pair of variables (A, Y) in causal DAG $\mathcal{G}$ if:

- No node in $\mathbb{Z}$ is a descendant of A; and
- $\mathbb{Z}$ blocks all non-causal paths between A and Y.

Theorem

If $\mathbb{Z}$ satisfies the back-door criterion with respect to (A, Y) and if $P(a, \mathbb{z}) > 0$, then $P(y \mid do(a))$ is identifiable and is given by:

$$P(y \mid do(a)) = \sum_{\mathbb{z}} P(y \mid a, \mathbb{z}) P(\mathbb{z}).$$

This criterion can be extended to a set $\mathbb{A}$ and a set $\mathbb{Y}$.

The back-door criterion: intuition

- Why "no node in $\mathbb{Z}$ is a descendant of A"?

- Why "$\mathbb{Z}$ blocks all non-causal paths between A and Y"?

The back-door criterion: intuition

- Why "no node in $\mathbb{Z}$ is a descendant of A"?

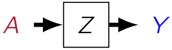
$$A \rightarrow \boxed{Z} \rightarrow Y$$

To avoid blocking causal paths.

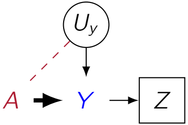
- Why "$\mathbb{Z}$ blocks all non-causal paths between A and Y"?

The back-door criterion: intuition

- Why "no node in $\mathbb{Z}$ is a descendant of A"?



To avoid blocking causal paths.

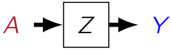


To avoid introducing artificial confounding bias resulting from conditioning on a collider.

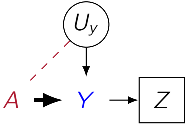
- Why "$\mathbb{Z}$ blocks all non-causal paths between A and Y"?

The back-door criterion: intuition

- Why "no node in $\mathbb{Z}$ is a descendant of A"?



To avoid blocking causal paths.

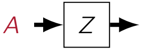


To avoid introducing artificial confounding bias resulting from conditioning on a collider.
To avoid selection bias!

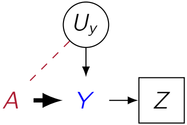
- Why "$\mathbb{Z}$ blocks all non-causal paths between A and Y"?

The back-door criterion: intuition

- Why "no node in $\mathbb{Z}$ is a descendant of A"?



To avoid blocking causal paths.



To avoid introducing artificial confounding bias resulting from conditioning on a collider.
To avoid selection bias!

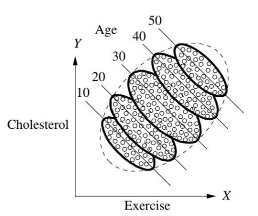
- Why "$\mathbb{Z}$ blocks all non-causal paths between A and Y"?

This one is the "obvious".

Simpson Paradox is no longer a paradox

Example

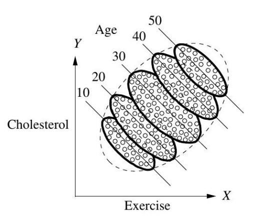
In a study, we measure weekly exercise and cholesterol levels across different age groups. What is the causal effect of exercise on cholesterol $P(c \mid do(e))$?



Simpson Paradox is no longer a paradox

Example

In a study, we measure weekly exercise and cholesterol levels across different age groups. What is the causal effect of exercise on cholesterol $P(c \mid do(e))$?

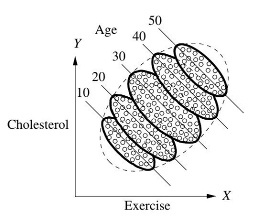


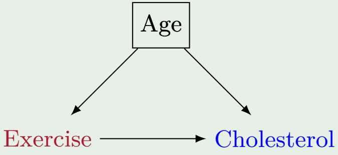
$$
\begin{tikzcd}
 & \text{Age} \arrow[dl] \arrow[dr] & \\
\text{Exercise} \arrow[rr] & & \text{Cholesterol}
\end{tikzcd}
$$

Simpson Paradox is no longer a paradox

Example

In a study, we measure weekly exercise and cholesterol levels across different age groups. What is the causal effect of exercise on cholesterol $P(c \mid do(e))$?

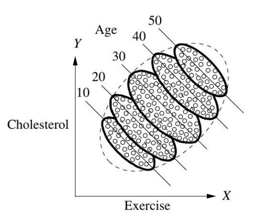


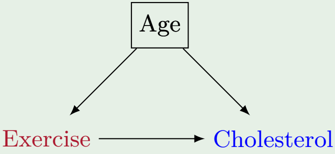


Simpson Paradox is no longer a paradox

Example

In a study, we measure weekly exercise and cholesterol levels across different age groups. What is the causal effect of exercise on cholesterol $P(c \mid do(e))$?





$$P(c \mid do(e)) = \sum_a P(c \mid a, e) P(a)$$

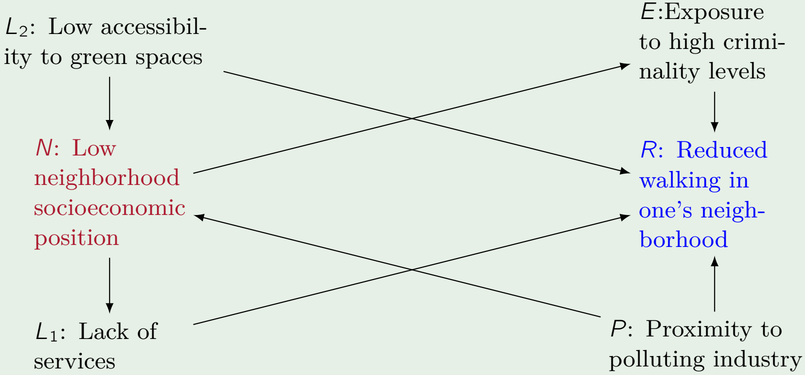
Example

In this study, we aim to estimate the effect of the neighborhood's socioeconomic status (N) on the reduction of walking within the neighborhood (R), $P(r \mid do(n))$?

Back-door criterion in action 1

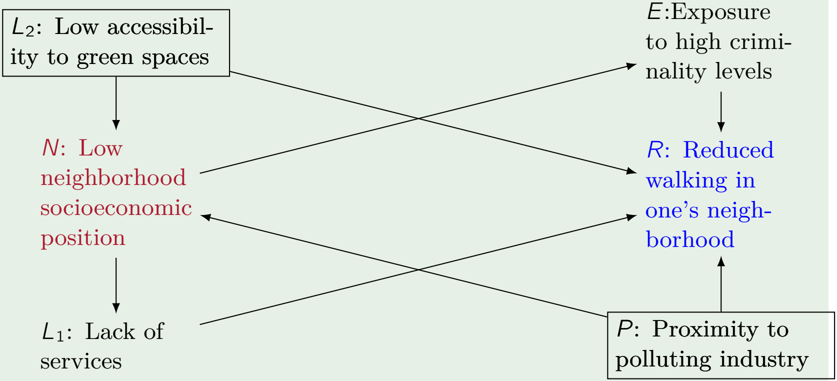
Example

In this study, we aim to estimate the effect of the neighborhood's socioeconomic status (N) on the reduction of walking within the neighborhood (R), $P(r \mid do(n))$?



Example

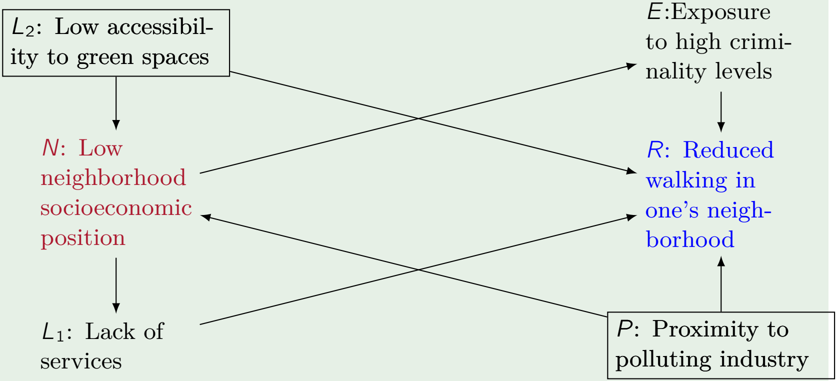
In this study, we aim to estimate the effect of the neighborhood's socioeconomic status (N) on the reduction of walking within the neighborhood (R), $P(r \mid do(n))$?



Back-door criterion in action 1

Example

In this study, we aim to estimate the effect of the neighborhood's socioeconomic status (N) on the reduction of walking within the neighborhood (R), $P(r \mid do(n))$?

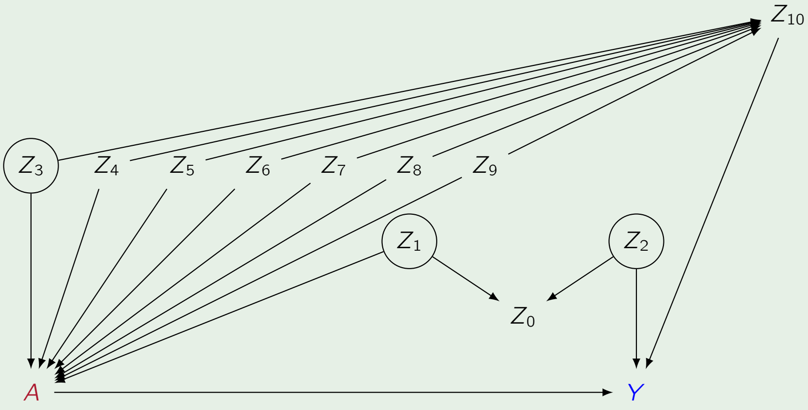


$$P(r \mid do(n)) = \sum P(r|n, l_2, p)P(l_2, p)$$

Back-door criterion in action 2

Example

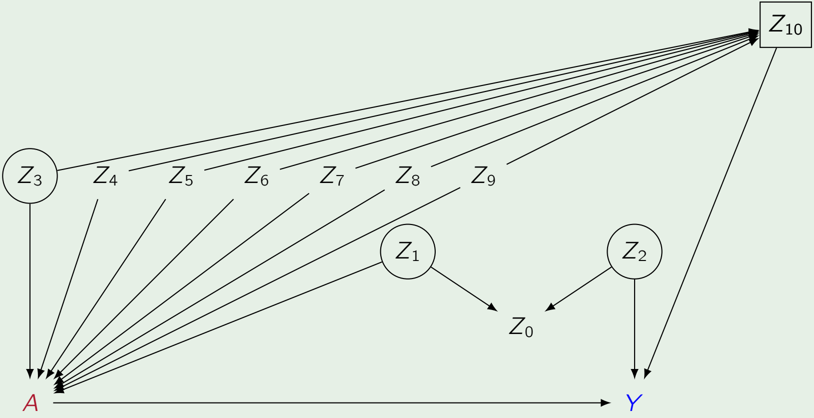
Which adjustment sets allow us to identify $\Pr(y \mid do(a))$?



Back-door criterion in action 2

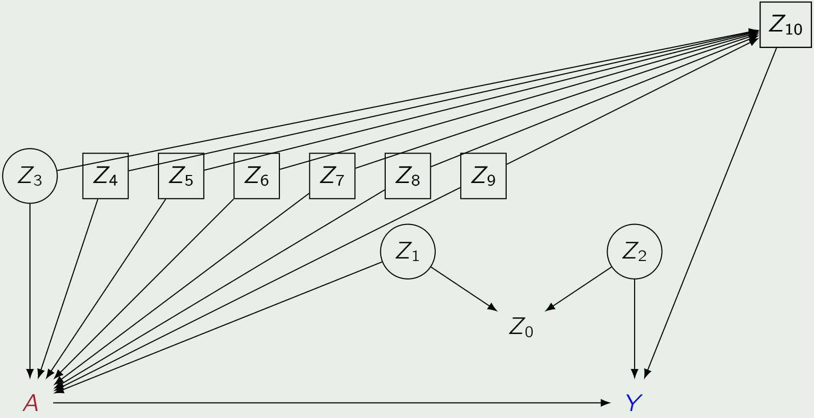
Example

Which adjustment sets allow us to identify $\Pr(y \mid do(a))$?



Example

Which adjustment sets allow us to identify $\Pr(y \mid do(a))$?



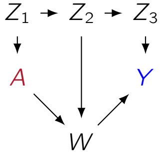
Finding the optimal set satisfying the back-door criterion

Many sets of variables can satisfy the back-door criterion:

- All these sets successfully remove bias.
- In practice, they do not all yield the same variance.

Many sets of variables can satisfy the back-door criterion:

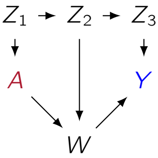
- All these sets successfully remove bias.
- In practice, they do not all yield the same variance.

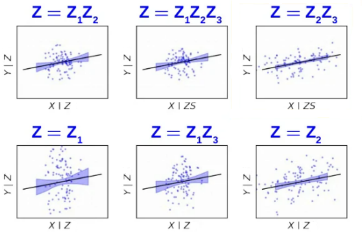


Finding the optimal set satisfying the back-door criterion

Many sets of variables can satisfy the back-door criterion:

- All these sets successfully remove bias.
- In practice, they do not all yield the same variance.

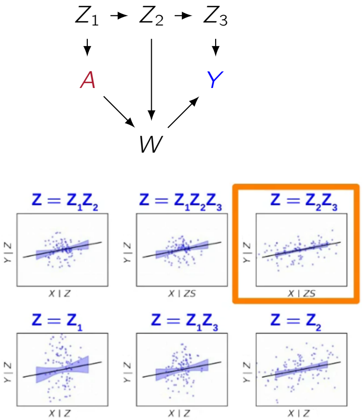




Finding the optimal set satisfying the back-door criterion

Many sets of variables can satisfy the back-door criterion:

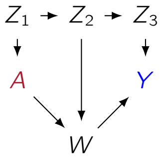
- All these sets successfully remove bias.
- In practice, they do not all yield the same variance.

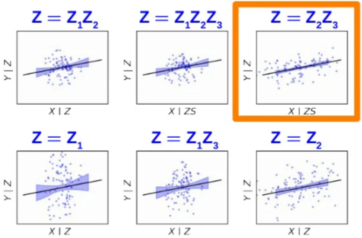


Finding the optimal set satisfying the back-door criterion

Many sets of variables can satisfy the back-door criterion:

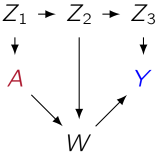
- All these sets successfully remove bias.
- In practice, they do not all yield the same variance.
- Is it possible to find the optimal set that yields the smallest variance?





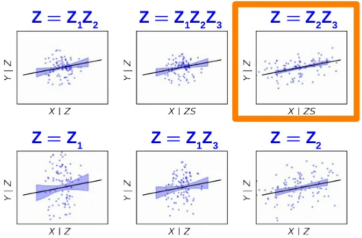
Many sets of variables can satisfy the back-door criterion:

- All these sets successfully remove bias.
- In practice, they do not all yield the same variance.
- Is it possible to find the optimal set that yields the smallest variance?



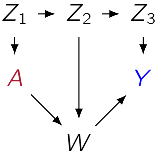
A set $\mathbb{Z}$ satisfying the back-door criterion relative to (A, Y) is optimal if:

- $\mathbb{Z}$ includes all parents of Y that are not mediators between A and Y, and



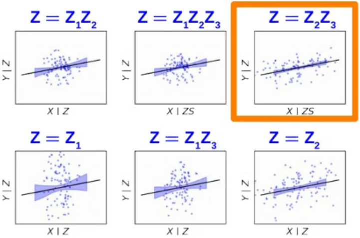
Many sets of variables can satisfy the back-door criterion:

- All these sets successfully remove bias.
- In practice, they do not all yield the same variance.
- Is it possible to find the optimal set that yields the smallest variance?



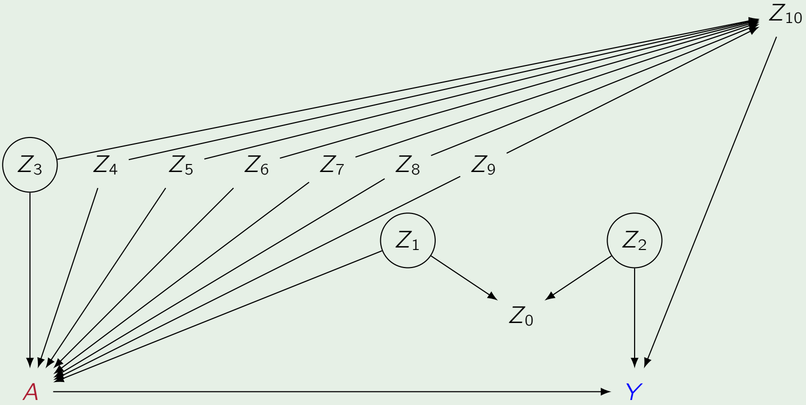
A set $\mathbb{Z}$ satisfying the back-door criterion relative to (A, Y) is optimal if:

- $\mathbb{Z}$ includes all parents of Y that are not mediators between A and Y, and
- $\mathbb{Z}$ includes all parents of any mediator between A and Y, that are not themselves mediators between A and Y.



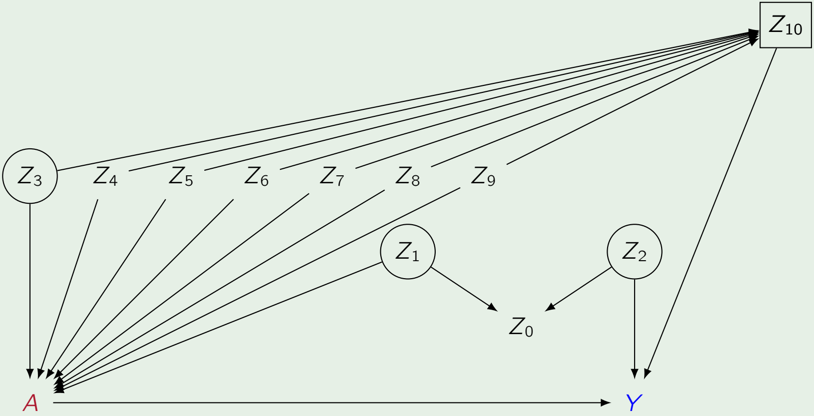
Example

What is the optimal adjustment set that allow us to identify $\Pr(y \mid do(a))$?



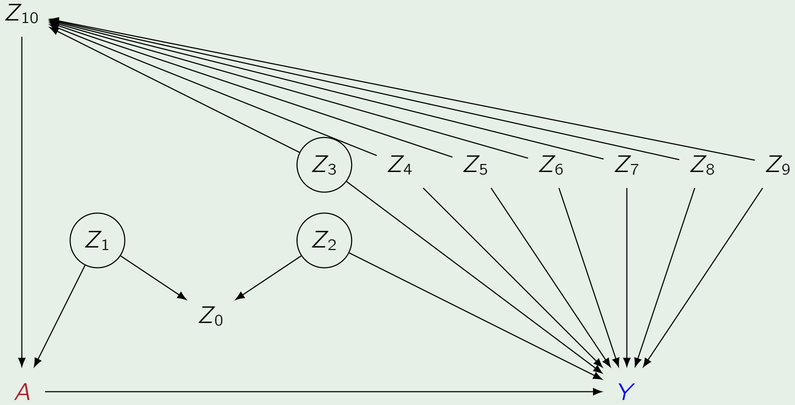
Example

What is the optimal adjustment set that allow us to identify $\Pr(y \mid do(a))$?



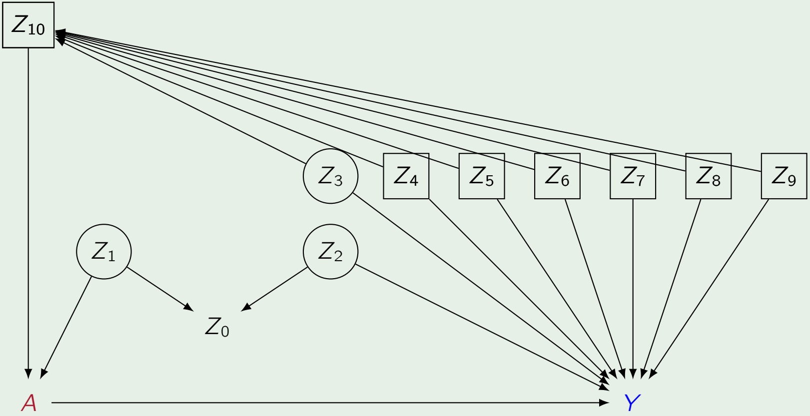
Example

What is the optimal adjustment set that allow us to identify $\Pr(y \mid do(a))$?



Example

What is the optimal adjustment set that allow us to identify $\Pr(y \mid do(a))$?



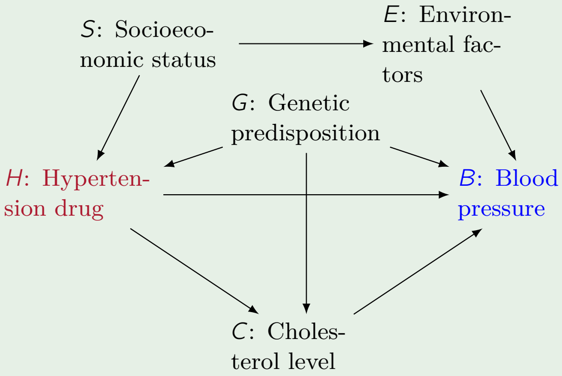
Optimal set in action 3

Example

How to optimally estimate the effect of new hypertension drug (H) on blood pressure (B) $P(b \mid do(h))$?

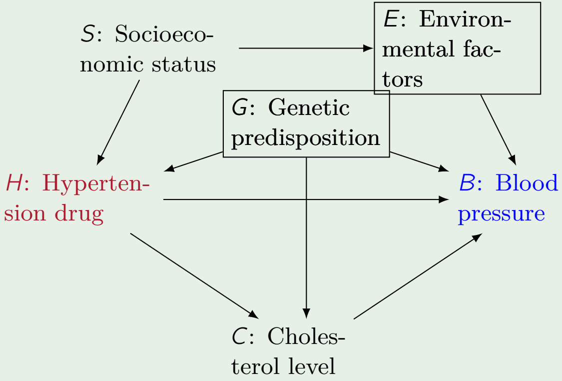
Example

How to optimally estimate the effect of new hypertension drug (H) on blood pressure (B) $P(b \mid do(h))$?



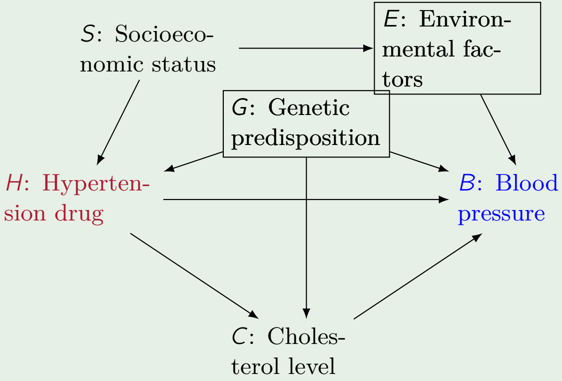
Example

How to optimally estimate the effect of new hypertension drug (H) on blood pressure (B) $P(b \mid do(h))$?



Example

How to optimally estimate the effect of new hypertension drug (H) on blood pressure (B) $P(b \mid do(h))$?

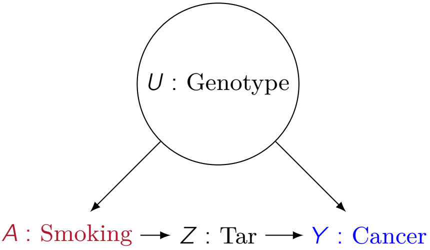


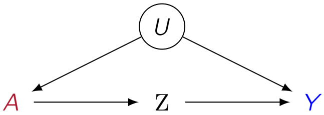
$$P(b \mid do(h)) = \sum_{g,e} P(b|h,g,e)P(g,e)$$

The back-door criterion is not complete:

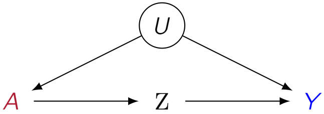
- If no set satisfies the back-door criterion for $P(y \mid do(x))$, this does not necessarily mean that $P(y \mid do(x))$ is not identifiable.

Going beyond the back-door



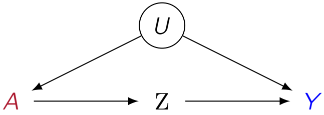

U
A
Z
Y

Going beyond the back-door



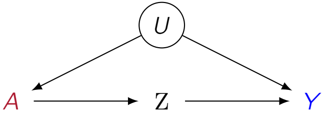
- $P(z \mid do(a)) = P(z \mid a)$ (No non-causal path)

Going beyond the back-door



- $P(z \mid do(a)) = P(z \mid a)$ (No non-causal path)
- $P(y \mid do(z)) = \sum_a P(y \mid z, a) P(a)$ (A blocks non-causal paths)

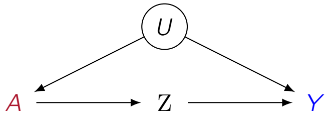
Going beyond the back-door



- $P(z \mid do(a)) = P(z \mid a)$ (No non-causal path)
- $P(y \mid do(z)) = \sum_a P(y \mid z, a)P(a)$ (A blocks non-causal paths)

$$P(y \mid do(a)) = \sum_z P(y \mid do(z))P(z \mid do(a))$$

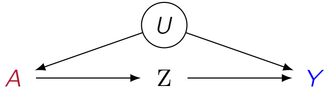
Going beyond the back-door



- $P(z \mid do(a)) = P(z \mid a)$ (No non-causal path)
- $P(y \mid do(z)) = \sum_a P(y \mid z, a)P(a)$ (A blocks non-causal paths)

$$P(y \mid do(a)) = \sum_z P(y \mid do(z))P(z \mid do(a))$$
$$= \sum_z P(z \mid a) \sum_{a'} P(y \mid z, a')P(a')$$

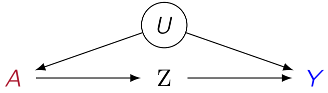
The front-door criterion



A set of measured variables $\mathbb{Z}$ satisfies the front-door criterion relative to an ordered pair of variables (A, Y) in causal DAG $\mathcal{G}$ if:

- $\mathbb{Z}$ intercepts all causal paths from A to Y;
- There is no non-causal path from A to $\mathbb{Z}$;
- All non-causal paths from $\mathbb{Z}$ to Y are blocked by A.

The front-door criterion



A set of measured variables $\mathbb{Z}$ satisfies the front-door criterion relative to an ordered pair of variables (A, Y) in causal DAG $\mathcal{G}$ if:

- $\mathbb{Z}$ intercepts all causal paths from A to Y;
- There is no non-causal path from A to $\mathbb{Z}$;
- All non-causal paths from $\mathbb{Z}$ to Y are blocked by A.

Theorem

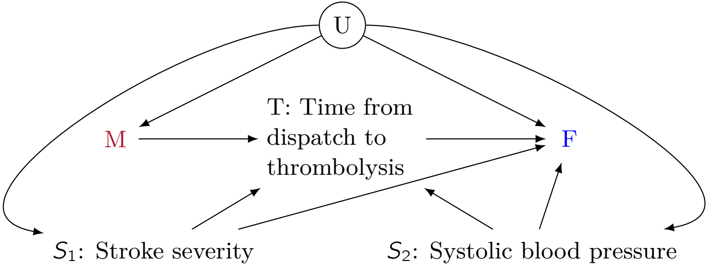
If $\mathbb{Z}$ satisfies the front-door criterion relative to (A, Y) and if $P(a, \mathbb{z}) > 0$, then the causal effect of A on Y is identifiable and is given by

$$P(y \mid do(a)) = \sum_{\mathbb{z}} P(\mathbb{z} \mid a) \sum_{a'} P(y \mid a', \mathbb{z}) P(a').$$

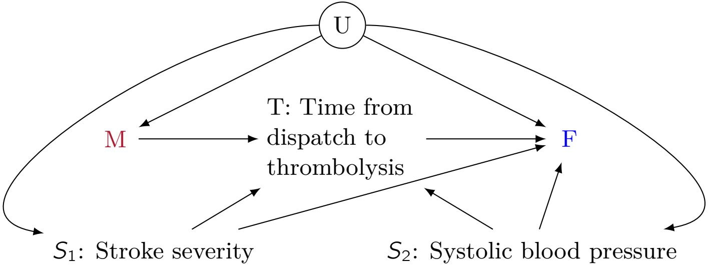
Back-door and Front-door criteria in a real application

In this study, Piccininni et al. were interested in estimating the effect of Mobile Stroke Unit dispatch (M) on functional outcomes (F). In other words, $P(f \mid do(m))$?

In this study, Piccininni et al. were interested in estimating the effect of Mobile Stroke Unit dispatch (M) on functional outcomes (F). In other words, $P(f \mid do(m))$?



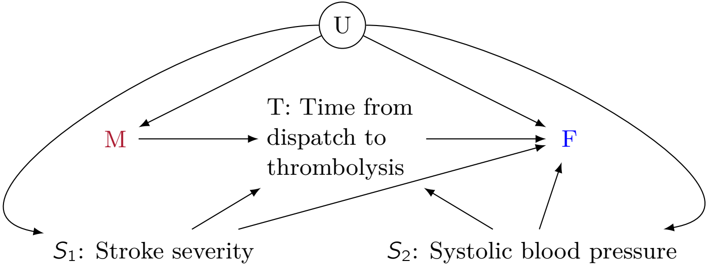
In this study, Piccininni et al. were interested in estimating the effect of Mobile Stroke Unit dispatch (M) on functional outcomes (F). In other words, $P(f \mid do(m))$?



The back-door criterion is not satisfied for any set of variables.
The front-door criterion is not satisfied for any set of variables.
However, it is possible to identify $P(f \mid do(m))$ by leveraging a combination of the back-door and front-door criteria!

Back-door and Front-door criteria in a real application

In this study, Piccininni et al. were interested in estimating the effect of Mobile Stroke Unit dispatch (M) on functional outcomes (F). In other words, $P(f \mid do(m))$?



The back-door criterion is not satisfied for any set of variables.
The front-door criterion is not satisfied for any set of variables.
However, it is possible to identify $P(f \mid do(m))$ by leveraging a combination of the back-door and front-door criteria!

Best paper in Epidemiology in 2024!

The front-door criterion is not complete:

- If no set satisfies the front-door criterion for $P(y \mid do(x))$, this does not necessarily mean that $P(y \mid do(x))$ is not identifiable.

Uncompletness of the front-door criterion

The front-door criterion is not complete:

- If no set satisfies the front-door criterion for $P(y \mid do(x))$, this does not necessarily mean that $P(y \mid do(x))$ is not identifiable.

The combination of the back-door and the front-door criteria is also not complete.

The do-calculus

Mutilated graphs:

$$\begin{tikzcd}
B \arrow[r] \arrow[d] & Z \arrow[d] & C \arrow[l] \arrow[d] \\
D & A \arrow[l] \arrow[r] & E
\end{tikzcd}$$

$\mathcal{G}$

Mutilated graphs:

$$\begin{tikzcd} B \arrow[r] \arrow[d] & Z \arrow[d] & C \arrow[l] \arrow[d] \\ D & A \arrow[l] \arrow[r] & E \end{tikzcd} \qquad \begin{tikzcd} B \arrow[d] & Z \arrow[d] & C \arrow[d] \\ D & A \arrow[l] \arrow[r] & E \end{tikzcd}$$

$\mathcal{G}$ $\qquad\qquad$ $\mathcal{G}_{\overline{Z}}$

The do-calculus

Mutilated graphs:

$$\begin{tikzcd} B \arrow[r] \arrow[d] & Z \arrow[d] & C \arrow[l] \arrow[d] \\ D & A \arrow[l] \arrow[r] & E \end{tikzcd}$$

$\mathcal{G}$

$$\begin{tikzcd} B \arrow[d] & Z \arrow[d] & C \arrow[d] \\ D & A \arrow[l] \arrow[r] & E \end{tikzcd}$$

$\mathcal{G}_{\overline{Z}}$

$$\begin{tikzcd} B \arrow[r] \arrow[d] & Z \arrow[d] & C \arrow[l] \arrow[d] \\ D & A & E \end{tikzcd}$$

$\mathcal{G}_{\underline{A}}$

The do-calculus

Mutilated graphs:

$$\begin{tikzcd} B \arrow[r] \arrow[d] & Z \arrow[d] & C \arrow[l] \arrow[d] \\ D & A \arrow[l] \arrow[r] & E \end{tikzcd}$$

$\mathcal{G}$

$$\begin{tikzcd} B \arrow[d] & Z \arrow[d] & C \arrow[d] \\ D & A \arrow[l] \arrow[r] & E \end{tikzcd}$$

$\mathcal{G}_{\overline{Z}}$

$$\begin{tikzcd} B \arrow[r] \arrow[d] & Z \arrow[d] & C \arrow[l] \arrow[d] \\ D & A & E \end{tikzcd}$$

$\mathcal{G}_{\underline{A}}$

The do-calclulus:

Rule 1: $P(\mathbb{y} \mid do(\mathbb{z}), \mathbb{a}, \mathbb{w}) = P(\mathbb{y} \mid do(\mathbb{z}), \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp \mathbb{A} \mid \mathbb{Z}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{Z}}}}$

Rule 2: $P(\mathbb{y} \mid do(\mathbb{z}, \mathbb{a}), \mathbb{w}) = P(\mathbb{y} \mid do(\mathbb{z}), \mathbb{a}, \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp \mathbb{A} \mid \mathbb{Z}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{Z}}\underline{\mathbb{A}}}}$

Rule 3: $P(\mathbb{y} \mid do(\mathbb{z}, \mathbb{a}), \mathbb{w}) = P(\mathbb{y} \mid do(\mathbb{z}), \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp \mathbb{A} \mid \mathbb{Z}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{Z}}\overline{\mathbb{A}(\mathbb{W})}}}$

The do-calculus

Mutilated graphs:

$$\begin{tikzcd} B \arrow[r] \arrow[d] & Z \arrow[d] & C \arrow[l] \arrow[d] \\ D & A \arrow[l] \arrow[r] & E \end{tikzcd} \qquad \begin{tikzcd} B \arrow[d] & Z \arrow[d] & C \arrow[d] \\ D & A \arrow[l] \arrow[r] & E \end{tikzcd} \qquad \begin{tikzcd} B \arrow[r] \arrow[d] & Z \arrow[d] & C \arrow[l] \arrow[d] \\ D & A & E \end{tikzcd}$$

$$\mathcal{G} \qquad\qquad\qquad \mathcal{G}_{\overline{Z}} \qquad\qquad\qquad \mathcal{G}_{\underline{A}}$$

The do-calclulus:

Rule 1: $P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}), \mathbb{a}, \mathbb{w}) = P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}), \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp \mathbb{A} \mid \mathbb{Z}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{Z}}}}$

Rule 2: $P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}, \mathbb{a}), \mathbb{w}) = P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}), \mathbb{a}, \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp \mathbb{A} \mid \mathbb{Z}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{Z}}\underline{\mathbb{A}}}}$

Rule 3: $P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}, \mathbb{a}), \mathbb{w}) = P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}), \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp \mathbb{A} \mid \mathbb{Z}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{Z}}\,\overline{\mathbb{A}(\mathbb{W})}}}$

Theorem

$P(\mathbb{y} \mid do(\mathbb{a}))$ is identifiable if and only if there exists a finite sequence of transformations, each conforming to either one of the Rules 1-3 or some standard probability manipulations, that reduces $P(\mathbb{y} \mid do(\mathbb{a}))$ into a do-free formula.

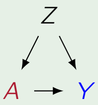
The do-calculus in action

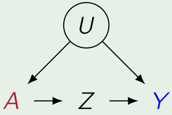
Rule 1: $P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}), \mathbb{x}, \mathbb{w}) = P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}), \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp \mathbb{X} \mid \mathbb{Z}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{Z}}}}$

Rule 2: $P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}, \mathbb{x}), \mathbb{w}) = P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}), \mathbb{x}, \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp \mathbb{X} \mid \mathbb{Z}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{Z}}\underline{\mathbb{X}}}}$

Rule 3: $P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}, \mathbb{x}), \mathbb{w}) = P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}), \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp \mathbb{X} \mid \mathbb{Z}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{Z}}\overline{\mathbb{X}(\mathbb{W})}}}$

Example





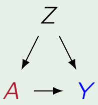
The do-calculus in action

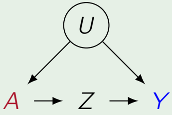
Rule 1: $P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}), \mathbb{x}, \mathbb{w}) = P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}), \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp \mathbb{X} \mid \mathbb{Z}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{Z}}}}$

Rule 2: $P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}, \mathbb{x}), \mathbb{w}) = P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}), \mathbb{x}, \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp \mathbb{X} \mid \mathbb{Z}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{Z}}\underline{\mathbb{X}}}}$

Rule 3: $P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}, \mathbb{x}), \mathbb{w}) = P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}), \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp \mathbb{X} \mid \mathbb{Z}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{Z}}\overline{\mathbb{X}(\mathbb{W})}}}$

Example





$P(y \mid do(a))$

$= \sum_z P(\mathbb{y} \mid do(a), z) P(z \mid do(a))$

$= \sum_z P(y \mid a, z) P(z \mid do(a))$ (Rule 2)

$= \sum_z P(y \mid a, z) P(z)$ (Rule 3)

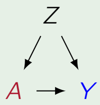
The do-calculus in action

Rule 1: $P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}), \mathbb{x}, \mathbb{w}) = P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}), \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp \mathbb{X} \mid \mathbb{Z}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{Z}}}}$

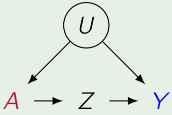
Rule 2: $P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}, \mathbb{x}), \mathbb{w}) = P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}), \mathbb{x}, \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp \mathbb{X} \mid \mathbb{Z}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{Z}}\underline{\mathbb{X}}}}$

Rule 3: $P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}, \mathbb{x}), \mathbb{w}) = P(\mathbb{y} \mid \mathrm{do}(\mathbb{z}), \mathbb{w})$ if $(\mathbb{Y} \perp\!\!\!\perp \mathbb{X} \mid \mathbb{Z}, \mathbb{W})_{\mathcal{G}_{\overline{\mathbb{Z}}\overline{\mathbb{X}(\mathbb{W})}}}$

Example



$$
\begin{aligned}
&P(y \mid do(a)) \\
&= \sum_z P(\mathbb{y} \mid do(a), z) P(z \mid do(a)) \\
&= \sum_z P(y \mid a, z) P(z \mid do(a)) \quad \text{(Rule 2)} \\
&= \sum_z P(y \mid a, z) P(z) \quad \text{(Rule 3)}
\end{aligned}
$$

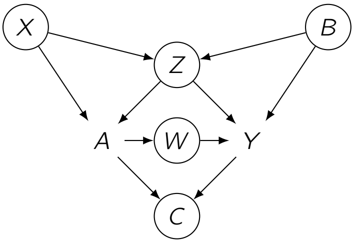


$$
\begin{aligned}
&P(y \mid do(a)) \\
&= \sum_z P(y \mid do(a), z) P(z \mid do(a)) \\
&= \sum_z P(y \mid do(a), do(z)) P(z \mid a) && \text{(Rule 2) x2} \\
&= \sum_z P(y \mid do(z)) P(z \mid a) && \text{(Rule 3)} \\
&= \sum_{z,a'} P(y \mid do(z), a') P(a' \mid do(z)) P(z \mid a) \\
&= \sum_{z,a'} P(y \mid z, a') P(a' \mid do(z)) P(z \mid a) && \text{(Rule 2)}
\end{aligned}
$$

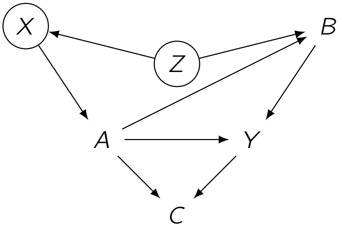
```
#Load the pyCIPHOD package
from pyciphod.graphs import DirectedAcyclicGraph
from pyciphod.causal_reasoning import back_door_criterio
from pyciphod.causal_reasoning import front_door_criteri
#Define the DAG
g = DirectedAcyclicGraph()
g.add_vertices(["A","B","C", "D"])
g.add_directed_edge('B', 'A')
g.add_directed_edge('B', 'C')
g.add_directed_edge('A', 'D')
g.add_directed_edge('D', 'C')
#Check if B satifies the back-door criterion for (A,C)
bd = back_door_criterion(g, ["A"], ["C"], ["B"])
print(bd)
#Check if D satifies the front-door criterion for (A,C)
fd = front_door_criterion(g, ["A"], ["C"], ["D"])
print(fd)
```

Exercise 2

Consider that in the following causal DAG, only A and Y, and one additional variable can be measured. Which variable would allow the identification of $P(y \mid do(a))$?



- Consider the following causal DAG. List all sets of variables that satisfy the back-door criterion for $P(y \mid do(a))$;

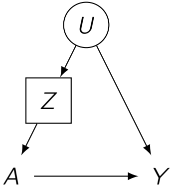


Exercise 4

Is $\{Z\}$ a good, bad or neutral adjustment set for $P(y \mid do(a))$?

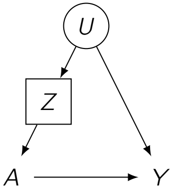
Exercise 4

Is $\{Z\}$ a good, bad or neutral adjustment set for $P(y \mid do(a))$?



Exercise 4

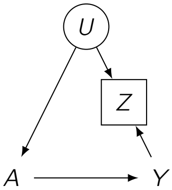
Is $\{Z\}$ a good, bad or neutral adjustment set for $P(y \mid do(a))$?



- Z blocks all non-causal paths and it is not a descendant of A
  $\implies \{Z\}$ is a good adjustment set.

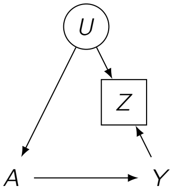
Exercise 4

Is $\{Z\}$ a good, bad or neutral adjustment set for $P(y \mid do(a))$?



Exercise 4

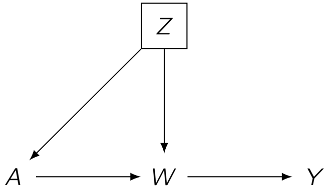
Is $\{Z\}$ a good, bad or neutral adjustment set for $P(y \mid do(a))$?



- Z activates a back-door path
  $\implies \{Z\}$ is a bad adjustment set.

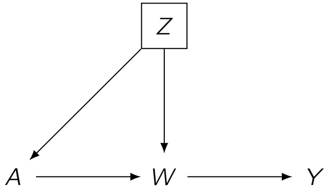
Exercise 4

Is $\{Z\}$ a good, bad or neutral adjustment set for $P(y \mid do(a))$?



Exercise 4

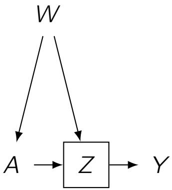
Is $\{Z\}$ a good, bad or neutral adjustment set for $P(y \mid do(a))$?



- Z blocks all non-causal paths and it is not a descendant of A
  $\implies$ $\{Z\}$ is a good adjustment set.

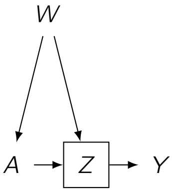
Exercise 4

Is $\{Z\}$ a good, bad or neutral adjustment set for $P(y \mid do(a))$?



Exercise 4

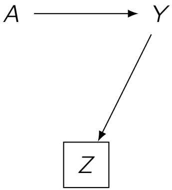
Is $\{Z\}$ a good, bad or neutral adjustment set for $P(y \mid do(a))$?



- Z d-separates A from Y
  $\implies$ $\{Z\}$ is a bad adjustment set.

Exercise 4

Is $\{Z\}$ a good, bad or neutral adjustment set for $P(y \mid do(a))$?



Exercise 4

Is $\{Z\}$ a good, bad or neutral adjustment set for $P(y \mid do(a))$?

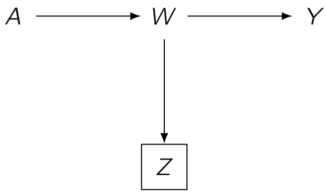
$$
\begin{tikzcd}
A \arrow[r] & Y \arrow[dl] \\
\boxed{Z} &
\end{tikzcd}
$$

- Selection bias

  $\Longrightarrow$ $\{Z\}$ is a bad control.

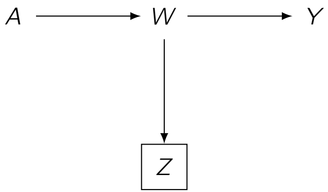
Exercise 4

Is $\{Z\}$ a good, bad or neutral adjustment set for $P(y \mid do(a))$?



Exercise 4

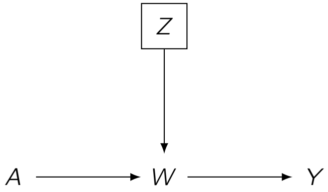
Is $\{Z\}$ a good, bad or neutral adjustment set for $P(y \mid do(a))$?



- Z is a descendant of A
  $\implies \{Z\}$ is a bad adjustment set.

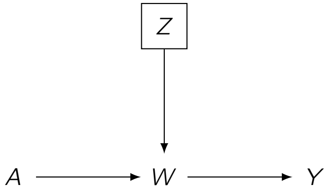
Exercise 4

Is $\{Z\}$ a good, bad or neutral adjustment set for $P(y \mid do(a))$?



Exercise 4

Is $\{Z\}$ a good, bad or neutral adjustment set for $P(y \mid do(a))$?



- Z is a parent of a mediator W
  $\implies$ $\{Z\}$ is a neutral adjustment set but good for optimality.

SANTÉ SORBONNE UNIVERSITÉ

4

Conclusion

Conclusion

Counterfactuals

Interventions

Associations

Counterfactuals

Interventions

Bayesian network

Associations

Counterfactuals

Causal DAG

Interventions

Bayesian network

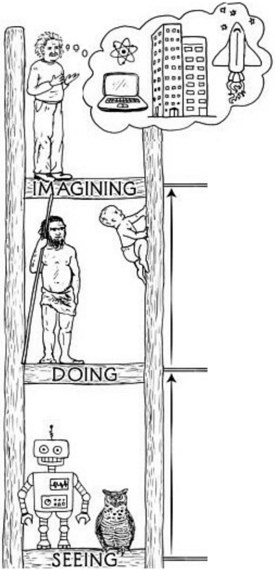
Associations

Conclusion

? | Counterfactuals

Causal DAG | Interventions

Bayesian network | Associations



5

References

References

- Pearl J. Causal diagrams for emperical studies. Biometrika. 1995.
- Pearl J. Causality: Models, Reasoning and Inference. Cambridge University Press. 2009.
- Chaix B, Leal C, Evans D. Neighborhood-level Con- founding in Epidemiologic Studies Unavoidable Challenges, Uncertain Solutions. Epidemiology (Cambridge, Mass.) 2009.
- Assaad C. Causal discovery between time series (Doctoral dissertation), Chap 2. UGA, 2021.
- Henckel L, Perkovič E, Maathuis M. Graphical Criteria for Efficient Total Effect Estimation via Adjustment in Causal Linear Models. Journal of the Royal Statistical Society. 2022.
- Piccininni M, Kurth T, Audebert HJ, Rohmann JL. The Effect of Mobile Stroke Unit Care on Functional Outcomes: An Application of the Front-door Formula. Epidemiology. 2023.
- Cinelli C, Forney A, Pearl J. A Crash Course in Good and Bad Controls. Sociological Methods & Research. 2024.

THANK YOU!

SANTÉ SORBONNE UNIVERSITÉ