Back-door and front-door criterions

Charles K. Assaad, Emilie Devijver, Eric Gaussier

charles.assaad@ens-lyon.fr

Table of content

Table of content

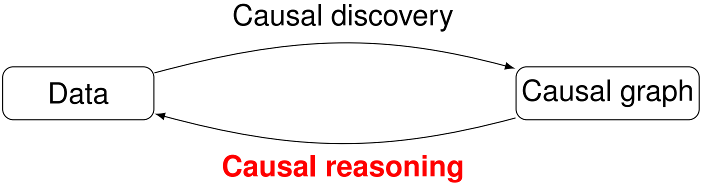
Causal reasoning (1/2)

Counterfactuals

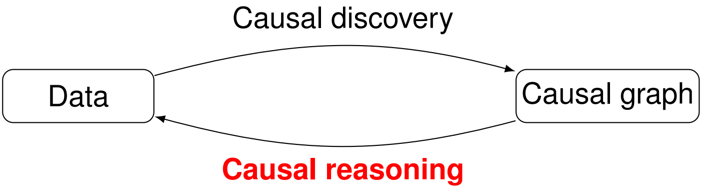
Interventions

Associations

Causal reasoning (2/2)

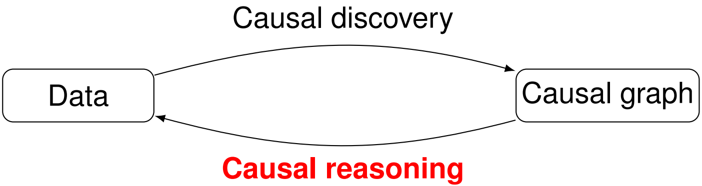


Causal reasoning (2/2)



Goal: Estimate the causal effect or effect of an intervention.

Causal reasoning (2/2)



Goal: Estimate the causal effect or effect of an intervention.

It is not always possible.

Recap about causal graphical models (1/4)

Active and blocked paths A path is said to be *blocked* by a set of vertices $\mathcal{Z} \in \mathcal{V}$ if:

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

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

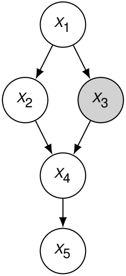
Recap about causal graphical models (1/4)

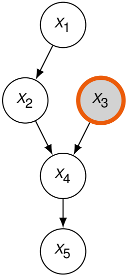
Active and blocked paths A path is said to be *blocked* by a set of vertices $\mathcal{Z} \in \mathcal{V}$ if:

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

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

Conditioning vs intervention





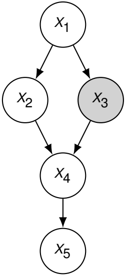
$\Pr(X_1, X_2, X_4, X_5 \mid X_3 = off)$ vs $\Pr_{X_3=off}(X_1, X_2, X_4, X_5)$

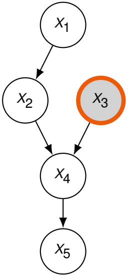
$\Pr(X_1, X_2, X_4, X_5 \mid X_3 = off)$ vs $\Pr(X_1, X_2, X_4, X_5 \mid do(X_3 = off))$

The do() operator allows to represent interventions in equations.

Recap about causal graphical models (2/4)

Conditioning vs intervention





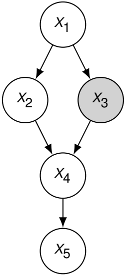
$\Pr(X_1, X_2, X_4, X_5 \mid X_3 = \mathit{off})$ vs $\Pr_{X_3=\mathit{off}}(X_1, X_2, X_4, X_5)$

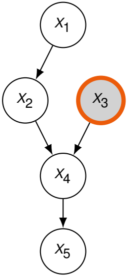
$\Pr(X_1, X_2, X_4, X_5 \mid X_3 = \mathit{off})$ vs $\Pr(X_1, X_2, X_4, X_5 \mid do(X_3 = \mathit{off}))$

The do() operator allows to represent interventions in equations.

Recap about causal graphical models (2/4)

Conditioning vs intervention





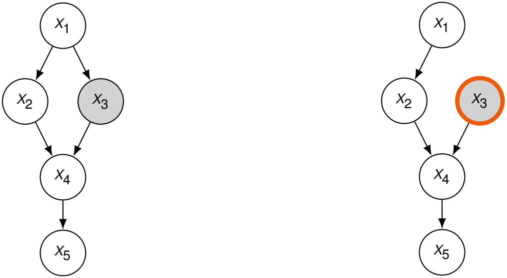
$$\Pr(X_1, X_2, X_4, X_5 \mid X_3 = \mathit{off}) \text{ vs } \Pr_{X_3=\mathit{off}}(X_1, X_2, X_4, X_5)$$

$$\Pr(X_1, X_2, X_4, X_5 \mid X_3 = \mathit{off}) \text{ vs } \Pr(X_1, X_2, X_4, X_5 \mid do(X_3 = \mathit{off}))$$

The do() operator allows to represent interventions in equations.

Recap about causal graphical models (2/4)

Conditioning vs intervention



$$\Pr(X_1, X_2, X_4, X_5 \mid X_3 = off) \text{ vs } \Pr_{X_3=off}(X_1, X_2, X_4, X_5)$$

$$\Pr(X_1, X_2, X_4, X_5 \mid X_3 = off) \text{ vs } \Pr(X_1, X_2, X_4, X_5 \mid do(X_3 = off))$$

The do() operator allows to represent interventions in equations.

Recap about causal graphical models (3/4)

Bayesian network factorization:

$$\Pr(V_1 = v_1, \cdots, V_d = v_d) = \prod_i \Pr(V_i = v_i \mid Parents(V_i))$$

Truncated factorization: if we intervene on a subset $S \subset \mathbf{V}$, then

$$\Pr(V_1 = v_1, \cdots, V_d = v_d \mid do(S = s)) = \prod_{i \notin S} \Pr(V_i = v_i \mid Parents(V_i))$$

if $v_1, \cdots, v_d$ are values consistant with the intervention,
else,

$$\Pr(V_1 = v_1, \cdots, V_d = v_d \mid do(S = s)) = 0$$

Recap about causal graphical models (3/4)

Bayesian network factorization:

$$\Pr(V_1 = v_1, \cdots, V_d = v_d) = \prod_i \Pr(V_i = v_i \mid Parents(V_i))$$

Truncated factorization: if we intervene on a subset $S \subset \mathbf{V}$, then

$$\Pr(V_1 = v_1, \cdots, V_d = v_d \mid do(S = s)) = \prod_{i \notin S} \Pr(V_i = v_i \mid Parents(V_i))$$

if $v_1, \cdots, v_d$ are values consistant with the intervention,
else,

$$\Pr(V_1 = v_1, \cdots, V_d = v_d \mid do(S = s)) = 0$$

Recap about causal graphical models (3/4)

Bayesian network factorization:

$$\Pr(v_1, \cdots, v_d) = \prod_i \Pr(v_i \mid Parents(v_i))$$

Truncated factorization: if we intervene on a subset $S \subset \mathbf{V}$, then

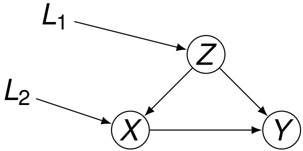
$$\Pr(v_1, \cdots, v_d \mid do(S = s)) = \prod_{i \notin S} \Pr(v_i \mid Parents(V_i))$$

if $v_1, \cdots, v_d$ are values consistant with the intervention, else,

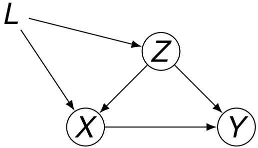
$$\Pr(v_1, \cdots, v_d \mid do(s)) = 0$$

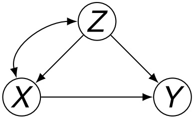
Recap about causal graphical models (4/4)

Markovian models: A model M is Markovian if the graph induced by M contains no bidirected edges (the graph is a DAG).



Semi-Markovian models: A model M is semi-Markovian if the graph induced by M contains bidirected edges (the graph is a ADMG).





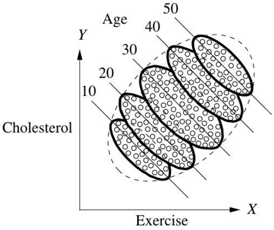
Causal effect identifiability

The causal effect $\Pr(y \mid do(x))$ from a causal graph $\mathcal{G}$ is identifiable if $\Pr(y \mid do(x))$ can be computed uniquely from observational data.

Table of content

Simpson paradox 1

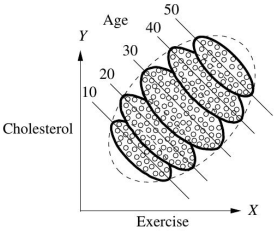
In a study, we measure weekly exercise and cholesterol levels for various age groups.



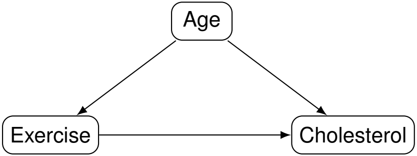
What is the effect of exercise on cholesterol $\Pr(c \mid do(e))$?

Simpson paradox 1

In a study, we measure weekly exercise and cholesterol levels for various age groups.

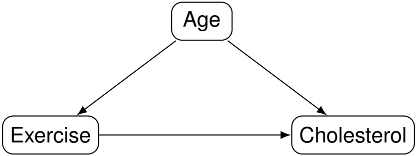


What is the effect of exercise on cholesterol $\Pr(c \mid do(e))$?



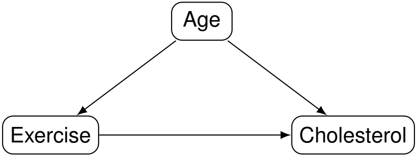
Simpson paradox 1: a simple solution

$\Pr(c \mid do(e))$?



Simpson paradox 1: a simple solution

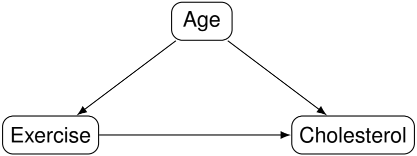
$\Pr(c \mid do(e))$?



$$\Pr(a, e, c) = \Pr(a)\Pr(e \mid a)\Pr(c \mid a, e) \qquad \text{(BN fact.)}$$

Simpson paradox 1: a simple solution

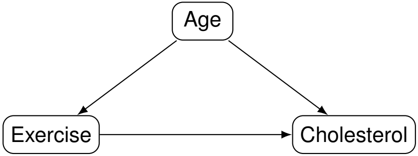
$\Pr(c \mid do(e))$?



$$\Pr(a, e, c) = \Pr(a)\Pr(e \mid a)\Pr(c \mid a, e) \qquad \text{(BN fact.)}$$
$$\Pr(a, c \mid do(e)) = \Pr(a)\Pr(c \mid a, e) \qquad \text{(Truncated fact.)}$$

Simpson paradox 1: a simple solution

$\Pr(c \mid do(e))$?



$$\Pr(a, e, c) = \Pr(a)\Pr(e \mid a)\Pr(c \mid a, e) \quad \text{(BN fact.)}$$
$$\Pr(a, c \mid do(e)) = \Pr(a)\Pr(c \mid a, e) \quad \text{(Truncated fact.)}$$
$$\Pr(c \mid do(e)) = \sum_a \Pr(a)\Pr(c \mid a, e) \quad \text{(marginalizing)}$$

Identifiabilty in Markovian models

Theorem (identifiabilty in Markovian models): Given a causal graph $\mathcal{G}$ of any Markovian model in which a subset $\mathcal{V}$ of variables are measured, the causal effect $\Pr(y \mid do(x))$ is identifiable whenever $\{X \cup Y \cup Parents(X)\} \subseteq \mathcal{V}$, and is given by the direct causes adjustment:

$$\Pr(y \mid do(x)) = \sum_{z \in Parents(x)} \Pr(y \mid x, z) \Pr(z)$$

(proof on board)

Limitations of the direct causes adjustment

- In Markovian models, is it possible to find a smaller adjustment set?

- What about semi-Markovian models?

Table of content

Back-door criterion

The back-door criterion: Consider a causal graph $\mathcal{G}$ and a causal effect $P(y \mid do(x))$. A set of variables $\mathcal{Z}$ satisfies the back-door criterion iff:

- no node in $\mathcal{Z}$ is a descendant of X;
- $\mathcal{Z}$ blocks every path between X and Y that contains an arrow into X.

Back-door adjustment

Theorem (back-door adjustment): If $\mathcal{Z}$ satisfies the back-door criterion relative to (X, Y) and if $\Pr(x, z) > 0$, then the causal effect of X on Y is identifiable and is given by

$$\Pr(y \mid do(x)) = \sum_{z} \Pr(y \mid x, z) \Pr(z).$$

(proof on board)

Back-door adjustment

Theorem (back-door adjustment): If $\mathcal{Z}$ satisfies the back-door criterion relative to (X, Y) and if $\Pr(x, z) > 0$, then the causal effect of X on Y is identifiable and is given by

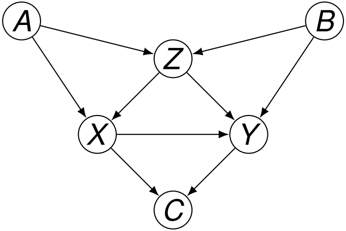
$$\Pr(y \mid do(x)) = \sum_{z} \Pr(y \mid x, z) \Pr(z).$$

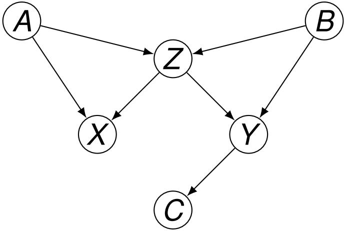
(proof on board)

Property (back-door in Markovian models): The causal effect in Markovian models is always identifiable using the back-door criterion and is given by the back-door adjustment.
proof as an exercise

Back-door criterion: using d-separation

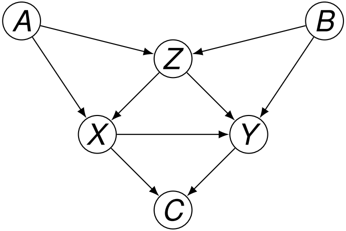
Causal graph $\mathcal{G}$ $\quad$ $\Pr(y \mid do(x))$ $\quad$ Mutilated graph $\mathcal{G}_m$

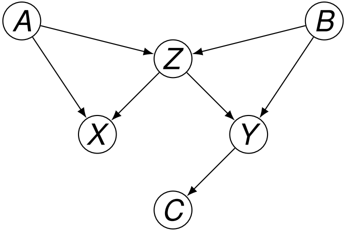




Back-door criterion: using d-separation

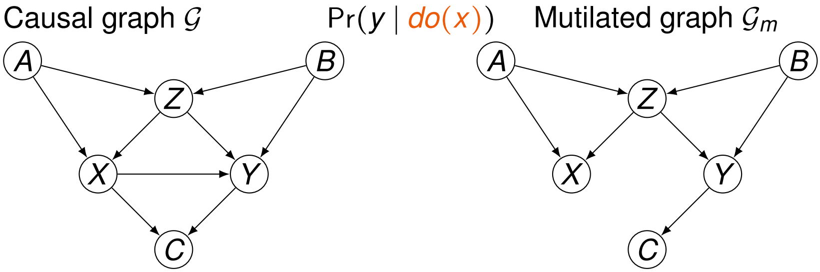
Causal graph $\mathcal{G}$ $\Pr(y \mid do(x))$ Mutilated graph $\mathcal{G}_m$





$X \perp\!\!\!\perp_G Y \mid Z$ in $\mathcal{G}_m$?

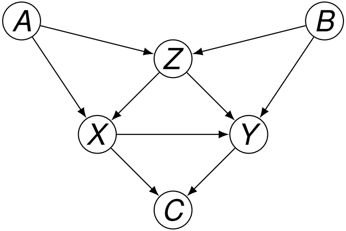
Back-door criterion: using d-separation

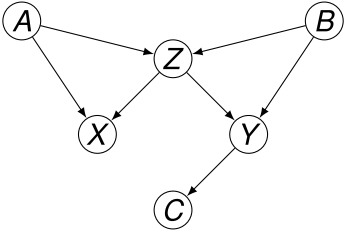


$X \perp\!\!\!\perp_G Y \mid Z$ in $\mathcal{G}_m$? No

Back-door criterion: using d-separation

Causal graph $\mathcal{G}$ $\Pr(y \mid do(x))$ Mutilated graph $\mathcal{G}_m$



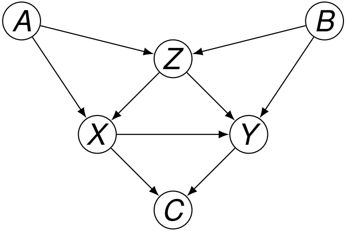


$X \perp\!\!\!\perp_G Y \mid Z$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid A$ in $\mathcal{G}_m$?

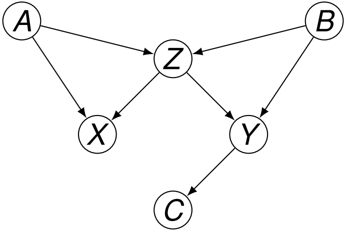
Back-door criterion: using d-separation

Causal graph $\mathcal{G}$



$\Pr(y \mid do(x))$

Mutilated graph $\mathcal{G}_m$

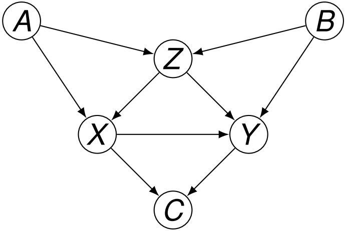


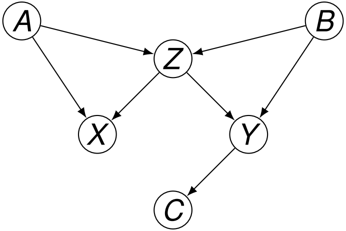
$X \perp\!\!\!\perp_G Y \mid Z$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid A$ in $\mathcal{G}_m$? No

Back-door criterion: using d-separation

Causal graph $\mathcal{G}$ $\Pr(y \mid do(x))$ Mutilated graph $\mathcal{G}_m$

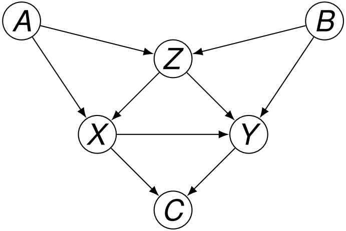




$X \perp\!\!\!\perp_G Y \mid Z$ in $\mathcal{G}_m$? No
$X \perp\!\!\!\perp_G Y \mid A$ in $\mathcal{G}_m$? No
$X \perp\!\!\!\perp_G Y \mid B$ in $\mathcal{G}_m$?

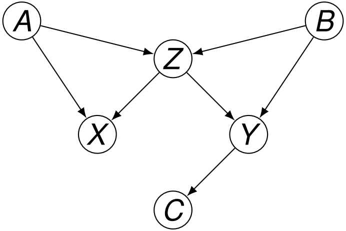
Back-door criterion: using d-separation

Causal graph $\mathcal{G}$



$\Pr(y \mid do(x))$

Mutilated graph $\mathcal{G}_m$



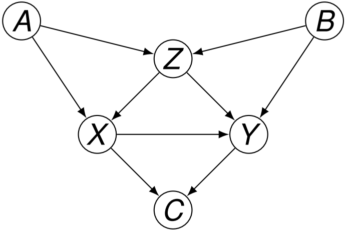
$X \perp\!\!\!\perp_G Y \mid Z$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid A$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid B$ in $\mathcal{G}_m$? No

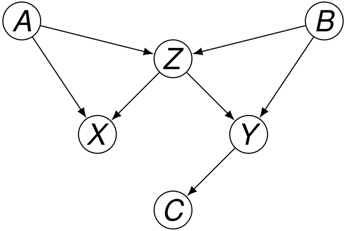
Back-door criterion: using d-separation

Causal graph $\mathcal{G}$



$\Pr(y \mid do(x))$

Mutilated graph $\mathcal{G}_m$



$X \perp\!\!\!\perp_G Y \mid Z$ in $\mathcal{G}_m$? No

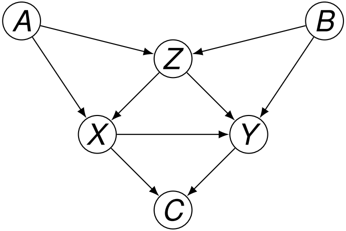
$X \perp\!\!\!\perp_G Y \mid A$ in $\mathcal{G}_m$? No

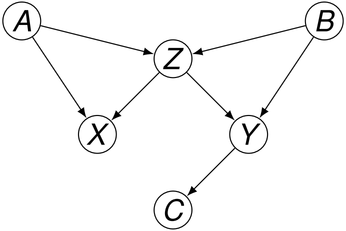
$X \perp\!\!\!\perp_G Y \mid B$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid C$ in $\mathcal{G}_m$?

Back-door criterion: using d-separation

Causal graph $\mathcal{G}$ $\Pr(y \mid do(x))$ Mutilated graph $\mathcal{G}_m$





$X \perp\!\!\!\perp_G Y \mid Z$ in $\mathcal{G}_m$? No

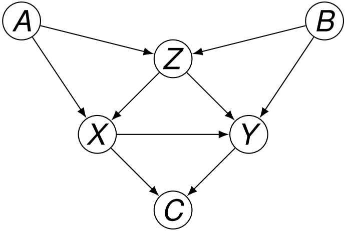
$X \perp\!\!\!\perp_G Y \mid A$ in $\mathcal{G}_m$? No

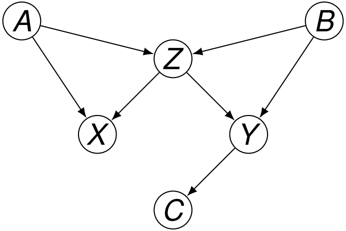
$X \perp\!\!\!\perp_G Y \mid B$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid C$ in $\mathcal{G}_m$? No

Back-door criterion: using d-separation

Causal graph $\mathcal{G}$ $\Pr(y \mid do(x))$ Mutilated graph $\mathcal{G}_m$





$X \perp\!\!\!\perp_G Y \mid Z$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid A$ in $\mathcal{G}_m$? No

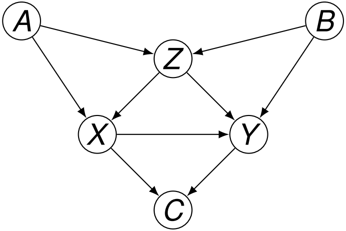
$X \perp\!\!\!\perp_G Y \mid B$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid C$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid A, B$ in $\mathcal{G}_m$?

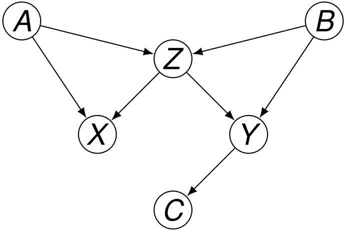
Back-door criterion: using d-separation

Causal graph $\mathcal{G}$



$\Pr(y \mid do(x))$

Mutilated graph $\mathcal{G}_m$



$X \perp\!\!\!\perp_G Y \mid Z$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid A$ in $\mathcal{G}_m$? No

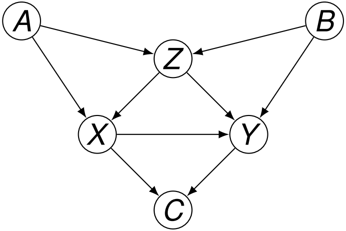
$X \perp\!\!\!\perp_G Y \mid B$ in $\mathcal{G}_m$? No

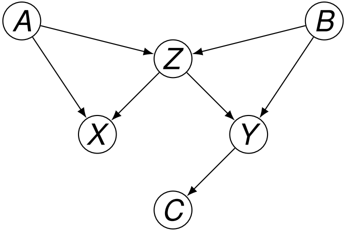
$X \perp\!\!\!\perp_G Y \mid C$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid A, B$ in $\mathcal{G}_m$? No

Back-door criterion: using d-separation

Causal graph $\mathcal{G}$ $\quad$ $\Pr(y \mid do(x))$ $\quad$ Mutilated graph $\mathcal{G}_m$





$X \perp\!\!\!\perp_G Y \mid Z$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid A$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid B$ in $\mathcal{G}_m$? No

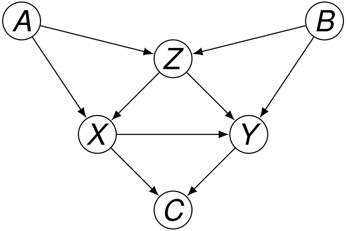
$X \perp\!\!\!\perp_G Y \mid C$ in $\mathcal{G}_m$? No

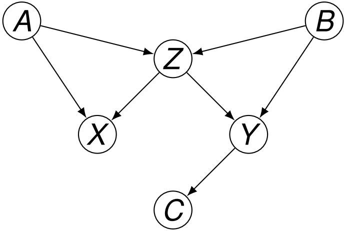
$X \perp\!\!\!\perp_G Y \mid A, B$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid Z, A$ in $\mathcal{G}_m$?

Back-door criterion: using d-separation

Causal graph $\mathcal{G}$ $\Pr(y \mid do(x))$ Mutilated graph $\mathcal{G}_m$





$X \perp\!\!\!\perp_G Y \mid Z$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid A$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid B$ in $\mathcal{G}_m$? No

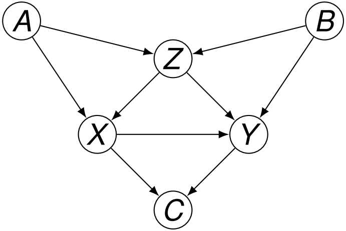
$X \perp\!\!\!\perp_G Y \mid C$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid A, B$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid Z, A$ in $\mathcal{G}_m$? Yes

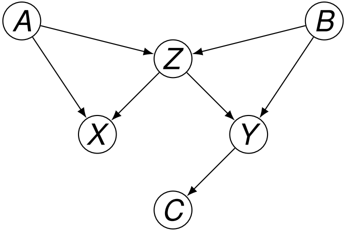
Back-door criterion: using d-separation

Causal graph $\mathcal{G}$



$\Pr(y \mid do(x))$

Mutilated graph $\mathcal{G}_m$



$X \perp\!\!\!\perp_G Y \mid Z$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid A$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid B$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid C$ in $\mathcal{G}_m$? No

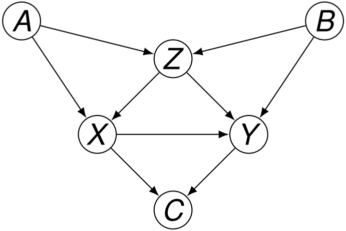
$X \perp\!\!\!\perp_G Y \mid A, B$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid Z, A$ in $\mathcal{G}_m$? Yes

$X \perp\!\!\!\perp_G Y \mid Z, B$ in $\mathcal{G}_m$?

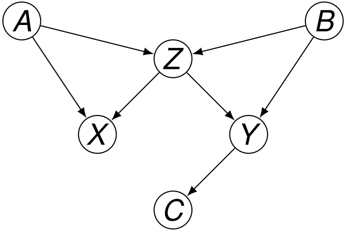
Back-door criterion: using d-separation

Causal graph $\mathcal{G}$



$\Pr(y \mid do(x))$

Mutilated graph $\mathcal{G}_m$



$X \perp\!\!\!\perp_G Y \mid Z$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid A$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid B$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid C$ in $\mathcal{G}_m$? No

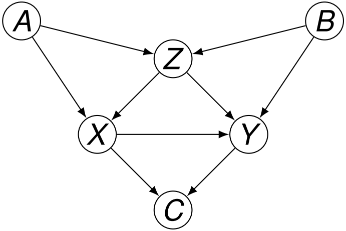
$X \perp\!\!\!\perp_G Y \mid A, B$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid Z, A$ in $\mathcal{G}_m$? Yes

$X \perp\!\!\!\perp_G Y \mid Z, B$ in $\mathcal{G}_m$? Yes

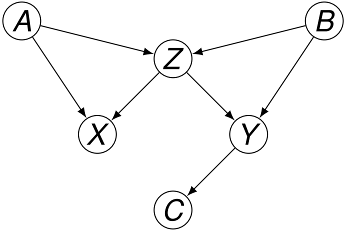
Back-door criterion: using d-separation

Causal graph $\mathcal{G}$



$\Pr(y \mid do(x))$

Mutilated graph $\mathcal{G}_m$



$X \perp\!\!\!\perp_G Y \mid Z$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid A$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid B$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid C$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid A, B$ in $\mathcal{G}_m$? No

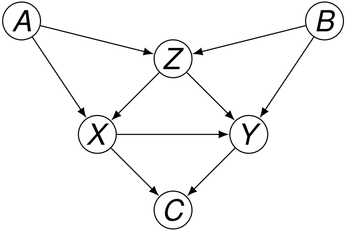
$X \perp\!\!\!\perp_G Y \mid Z, A$ in $\mathcal{G}_m$? Yes

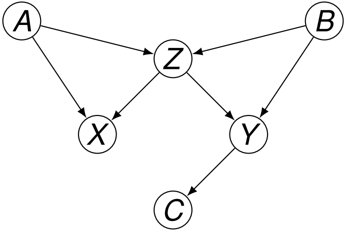
$X \perp\!\!\!\perp_G Y \mid Z, B$ in $\mathcal{G}_m$? Yes

$X \perp\!\!\!\perp_G Y \mid Z, A, B$ in $\mathcal{G}_m$?

Back-door criterion: using d-separation

Causal graph $\mathcal{G}$ $\Pr(y \mid do(x))$ Mutilated graph $\mathcal{G}_m$





$X \perp\!\!\!\perp_G Y \mid Z$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid A$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid B$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid C$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid A, B$ in $\mathcal{G}_m$? No

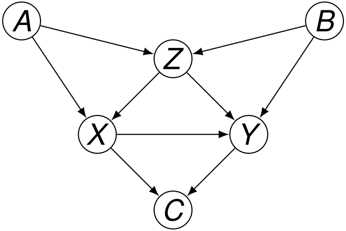
$X \perp\!\!\!\perp_G Y \mid Z, A$ in $\mathcal{G}_m$? Yes

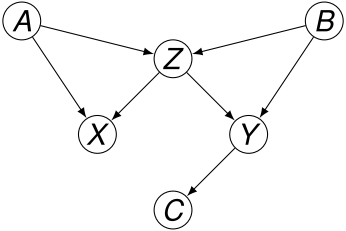
$X \perp\!\!\!\perp_G Y \mid Z, B$ in $\mathcal{G}_m$? Yes

$X \perp\!\!\!\perp_G Y \mid Z, A, B$ in $\mathcal{G}_m$? Yes

Back-door criterion: using d-separation

Causal graph $\mathcal{G}$ $\Pr(y \mid do(x))$ Mutilated graph $\mathcal{G}_m$





$X \perp\!\!\!\perp_G Y \mid Z$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid A$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid B$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid C$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid A, B$ in $\mathcal{G}_m$? No

$X \perp\!\!\!\perp_G Y \mid Z, A$ in $\mathcal{G}_m$? Yes

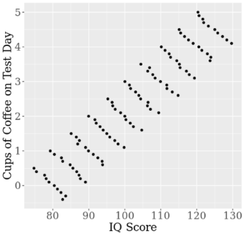
$X \perp\!\!\!\perp_G Y \mid Z, B$ in $\mathcal{G}_m$? Yes

$X \perp\!\!\!\perp_G Y \mid Z, A, B$ in $\mathcal{G}_m$? Yes

Back-door sets: $\{Z, A\}$ $\{Z, B\}$ $\{Z, A, B\}$

Simpson paradox 2 and the back-door in action

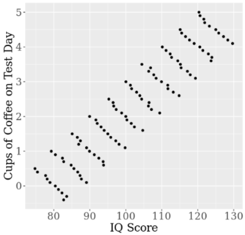
In a study, we measure the number of coffee intake, IQ score for a sample of a population with various education level.



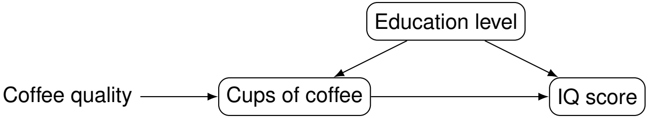
What is the effect of the nb cups of coffee on IQ score $\Pr(i \mid do(c))$?

Simpson paradox 2 and the back-door in action

In a study, we measure the number of coffee intake, IQ score for a sample of a population with various education level.

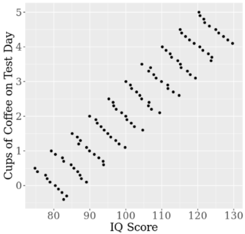


What is the effect of the nb cups of coffee on IQ score $\Pr(i \mid do(c))$?

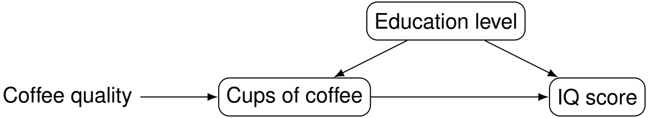


Simpson paradox 2 and the back-door in action

In a study, we measure the number of coffee intake, IQ score for a sample of a population with various education level.



What is the effect of the nb cups of coffee on IQ score $\Pr(i \mid do(c))$?



$$\Pr(i \mid do(c)) = \sum \Pr(i \mid c, e) \Pr(e)$$

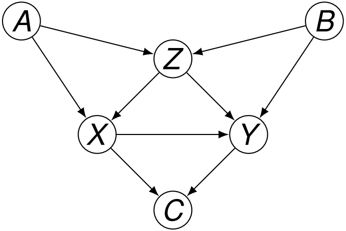
Incompleteness of the back-door criterion

- If there exists a set that satisfy the back-door criterion for $\Pr(y \mid do(x))$, then $\Pr(y \mid do(x))$ is identifiable;

- If there exists a no set that satisfy the back-door criterion for $\Pr(y \mid do(x))$, then $\Pr(y \mid do(x))$ is not necesarly not identifiable.

Exercise 1

- Consider the following causal graph. List all *minimal* sets of variables that satisfy the back-door criterion for $\Pr(y \mid do(x))$;
- Repeat for $\Pr(y \mid do(x, b))$.

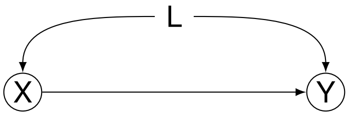


Minimal set: any set of variables such that if you remove any of the variables from the set, it will no longer meet the criterion.

Table of content

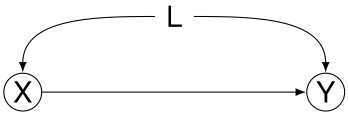
Going beyong the back-door (1/2)

Consider the following semi-Markovian model. Is $\Pr(y \mid do(x))$ identifiable using the backdoor criterion?



Going beyong the back-door (1/2)

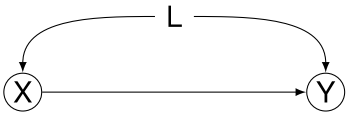
Consider the following semi-Markovian model. Is $\Pr(y \mid do(x))$ identifiable using the backdoor criterion?



No and it cannot be identified by any other criterion.

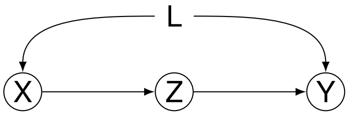
Going beyong the back-door (1/2)

Consider the following semi-Markovian model. Is $\Pr(y \mid do(x))$ identifiable using the backdoor criterion?



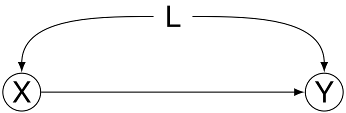
No and it cannot be identified by any other criterion.

What about the following semi-Markovian model?



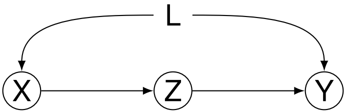
Going beyong the back-door (1/2)

Consider the following semi-Markovian model. Is $\Pr(y \mid do(x))$ identifiable using the backdoor criterion?



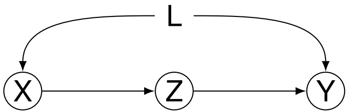
No and it cannot be identified by any other criterion.

What about the following semi-Markovian model?

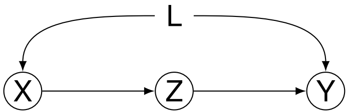


No but it can be identified by some other criterion.

Going beyong the back-door (2/2)

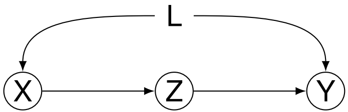


Going beyong the back-door (2/2)



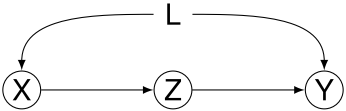
- $\Pr(z \mid do(x)) = \Pr(z \mid x)$ (No back-door)

Going beyong the back-door (2/2)



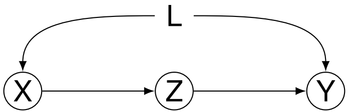
- $\Pr(z \mid do(x)) = \Pr(z \mid x)$ (No back-door)
- $\Pr(y \mid do(z)) = \sum_x \Pr(y \mid z, x) \Pr(x)$ (X blocks the back-door)

Going beyong the back-door (2/2)



- $\Pr(z \mid do(x)) = \Pr(z \mid x)$ (No back-door)
- $\Pr(y \mid do(z)) = \sum_x \Pr(y \mid z, x) \Pr(x)$ (X blocks the back-door)
- $\Pr(y \mid do(x)) = \sum_z \Pr(y \mid do(z)) \Pr(z \mid do(x))$ (Law of total proba.)

Going beyong the back-door (2/2)



- $\Pr(z \mid do(x)) = \Pr(z \mid x)$ (No back-door)
- $\Pr(y \mid do(z)) = \sum_x \Pr(y \mid z, x) \Pr(x)$ (X blocks the back-door)
- $\Pr(y \mid do(x)) = \sum_z \Pr(y \mid do(z)) \Pr(z \mid do(x))$ (Law of total proba.)

$$\Pr(y \mid do(x)) = \sum_z \Pr(z \mid x) \sum_{x'} \Pr(y \mid z, x') \Pr(x')$$

Front-door criterion

Front-door criterion: Consider a causal graph $\mathcal{G}$ and a causal effect $\Pr(y \mid do(x))$. A set of variables $\mathcal{Z}$ satisfies the front-door criterion iff:

- $\mathcal{Z}$ intercepts all directed paths from X to Y;
- There is no back-door path from X to $\mathcal{Z}$;
- All back-door paths from $\mathcal{Z}$ to Y are blocked by X.

Front-door adjustment

Theorem (front-door adjustment): if $\mathcal{Z}$ satisfies the front-door criterion relative to (X, Y) and if $\Pr(x, z) > 0$, then the causal effect of X on Y is identifiable and is given by

$$\Pr(y \mid do(X = x)) = \sum_{z} \Pr(z \mid x) \sum_{x'} \Pr(y \mid x', z) \Pr(x').$$

(proof on slide 25)

Simpson paradox 3 and the front-door in action

In a study, we measure the tar and the % of cancer among smokers and non smokers in a randomly selected sample of the population.

Smokers	Tar	% of cancer
No	No	10
No	Yes	5
Yes	No	90
Yes	Yes	85

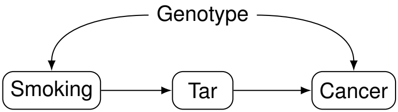
What is the effect of smoking on cancer $\Pr(c \mid do(s))$?

Simpson paradox 3 and the front-door in action

In a study, we measure the tar and the % of cancer among smokers and non smokers in a randomly selected sample of the population.

Smokers	Tar	% of cancer
No	No	10
No	Yes	5
Yes	No	90
Yes	Yes	85

What is the effect of smoking on cancer $\Pr(c \mid do(s))$?

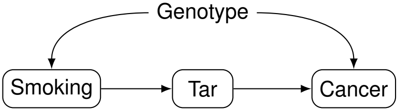


Simpson paradox 3 and the front-door in action

In a study, we measure the tar and the % of cancer among smokers and non smokers in a randomly selected sample of the population.

Smokers	Tar	% of cancer
No	No	10
No	Yes	5
Yes	No	90
Yes	Yes	85

What is the effect of smoking on cancer $\Pr(c \mid do(s))$?



$$\Pr(c \mid do(s)) = \sum_t \Pr(t|s) \sum_{s'} \Pr(c \mid t, s') \Pr(s')$$

Incompleteness of the front-door criterion

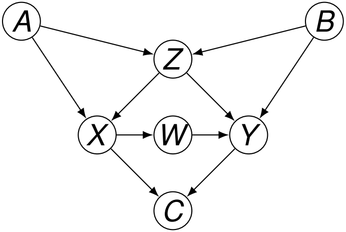
- If there exists a set that satisfy the front-door criterion for $\Pr(y \mid do(x))$, then $\Pr(y \mid do(x))$ is identifiable;

- If there exists a no set that satisfy the fack-door criterion for $\Pr(y \mid do(x))$, then $\Pr(y \mid do(x))$ is not necesarly not identifiable.

The combination of the back-door and front door criterions are also incomplete.

Exercise 2

Consider that in the following causal graph, only X and Y, and one additional variable can be measured. Which variable would allow the identification of $\Pr(y \mid do(x))$?

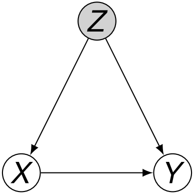


Exercise 3

Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?

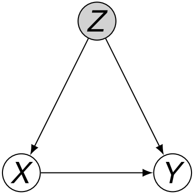
Exercise 3

Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?



Exercise 3

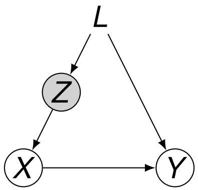
Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?



- Z blocks a back-door path
  $\implies$ Z is a good control.

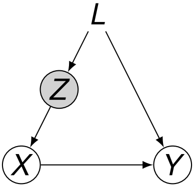
Exercise 3

Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?



Exercise 3

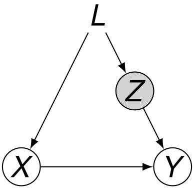
Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?



- Z blocks a back-door path
  $\implies$ Z is a good control.

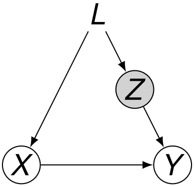
Exercise 3

Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?



Exercise 3

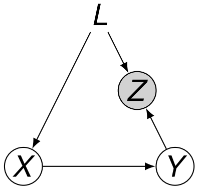
Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?



- Z blocks a back-door path
  $\implies$ Z is a good control.

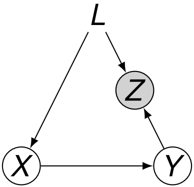
Exercise 3

Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?



Exercise 3

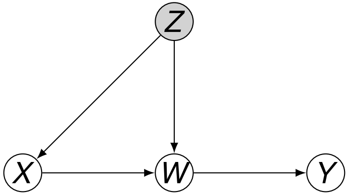
Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?



- Z activates a back-door path
  $\implies Z$ is a bad control.

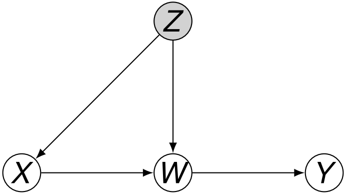
Exercise 3

Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?



Exercise 3

Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?



- Z blocks the back-door path
  $\Longrightarrow$ Z is a good control.

Exercise 3

Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?

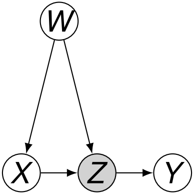
Exercise 3

Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?

- Z d-separates X from Y
  $\implies$ Z is a bad control.

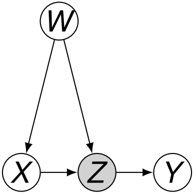
Exercise 3

Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?



Exercise 3

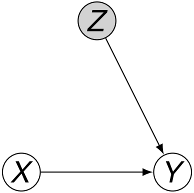
Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?



- Z d-separates X from Y
  $\implies$ Z is a bad control.

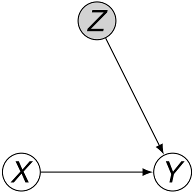
Exercise 3

Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?



Exercise 3

Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?

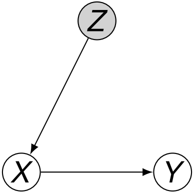


- Z does not open any backdoor paths from X to Y
  $\implies Z$ is a neutral control;

- Controlling for Z can reduces the variation of Y, and helps improve the precision of the estimate in finite samples.

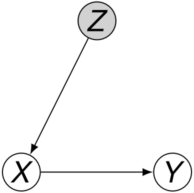
Exercise 3

Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?



Exercise 3

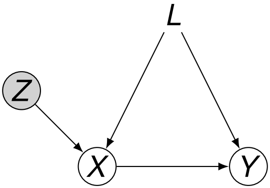
Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?



- Z does not open any backdoor paths from X to Y $\implies$ Z is a neutral control;
- Controlling for Z can reduces the variation of X and so may hurt the precision of the estimate in finite samples.

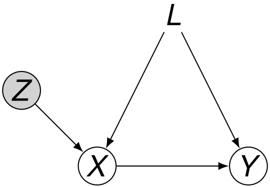
Exercise 3

Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?



Exercise 3

Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?

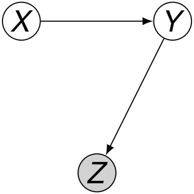


- Z does not block existing backdoor path from X to Y
  $\Longrightarrow$ Z is a bad control;

- In linear models, controlling for Z amplify any existing bias.

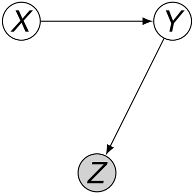
Exercise 3

Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?



Exercise 3

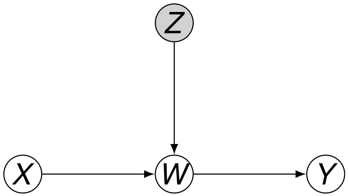
Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?



- Selection bias
  $\implies$ Z is a bad control.

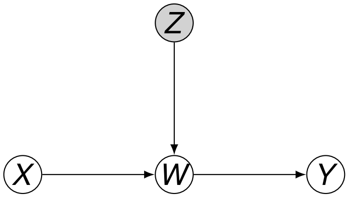
Exercise 3

Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?



Exercise 3

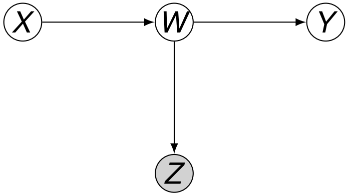
Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?



- Z does not open any backdoor paths from X to Y $\implies$ Z is a neutral control;
- Controlling for Z can reduces the variation of W, and helps improve the precision of the estimate in finite samples.

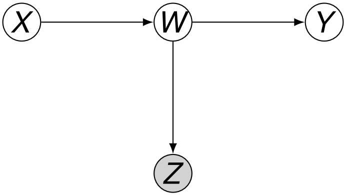
Exercise 3

Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?



Exercise 3

Is Z a good, bad or neutral control for $\Pr(y \mid do(x))$?



- Z is a descendant of X
  $\implies$ Z is a bad control.

Table of content

Conclusion

- Markovian models are always identifiable (using direct causes or the back-door adjustment);
- Semi Markovian models are not always identifiable;
- The back-door adjustment can identify some causal effects in semi Markovian models;
- The front-door adjustment can identify some causal effects in semi Markovian models;
- $\Longrightarrow$ the back-door and front-door adjustments are not complete.

Conclusion

- Markovian models are always identifiable (using direct causes or the back-door adjustment);
- Semi Markovian models are not always identifiable;
- The back-door adjustment can identify some causal effects in semi Markovian models;
- The front-door adjustment can identify some causal effects in semi Markovian models;
- $\implies$ the back-door and front-door adjustments are not complete.

Conclusion

- Markovian models are always identifiable (using direct causes or the back-door adjustment);
- Semi Markovian models are not always identifiable;
- The back-door adjustment can identify some causal effects in semi Markovian models;
- The front-door adjustment can identify some causal effects in semi Markovian models;
- $\Longrightarrow$ the back-door and front-door adjustments are not complete.

Conclusion

- Markovian models are always identifiable (using direct causes or the back-door adjustment);
- Semi Markovian models are not always identifiable;
- The back-door adjustment can identify some causal effects in semi Markovian models;
- The front-door adjustment can identify some causal effects in semi Markovian models;
- $\Longrightarrow$ the back-door and front-door adjustments are not complete.

Conclusion

- Markovian models are always identifiable (using direct causes or the back-door adjustment);
- Semi Markovian models are not always identifiable;
- The back-door adjustment can identify some causal effects in semi Markovian models;
- The front-door adjustment can identify some causal effects in semi Markovian models;
- $\implies$ the back-door and front-door adjustments are not complete.

References (1/2)

Direct inspirations

1. *Causality*, J. Pearl. Cambridge University Press, 2nd edition, 2009
2. *Causal inference in statistics: A Primer*, J. Pearl, M. Glymour, N. P. Jewell. Wiley, 2019
3. *The book of why*, J. Pearl, D. Mackenzie. Basic Books, 2018

References (2/2)

Additional readings

1. *A Crash Course in Good and Bad Control*, C. Cinelli, A. Forney, J. Pearl. Sociological Methods and Research, 2022
2. *Simpson's paradox in psychological science: A practical guide*, R. Kievit, W. Frankenhuis, L. Waldorp, D. Borsboom. Frontiers in Psychology, 2013