Causal Inference for Health Data

(STATS C160/C260 — Winter 2026)

Lecture 8:
Beyond Back-Door & do-Calculus

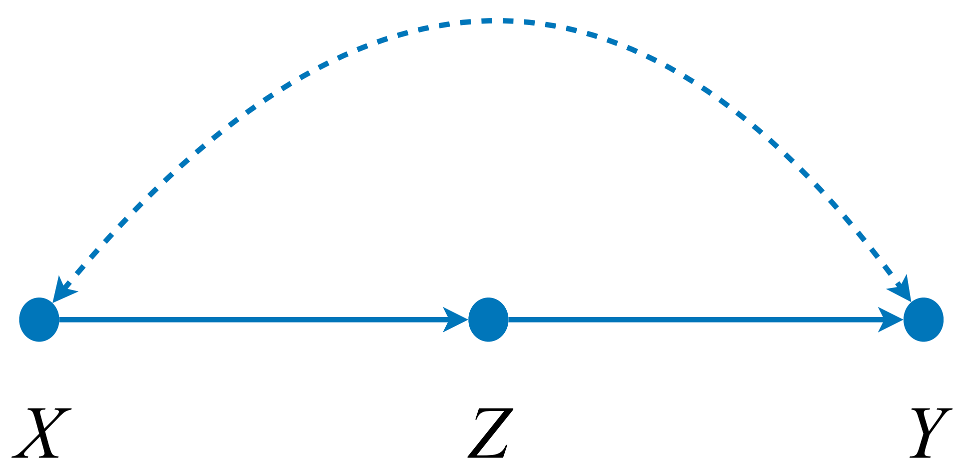
Drago Plečko

College | Physical Sciences
Statistics & Data Science

Recall: Back-Door Identification

- Is then the identification problem solved using the back-door criterion?

- Are there any admissible sets Z for the following causal diagram?



Is this model non-ID?

- How can we proceed?

Truncated Product in Semi-Markovian Models

Thm. 4.2.12. The distribution generated by an intervention *do(X=x)* in an SCM model M is given by the (generalized) truncated factorization product, namely,

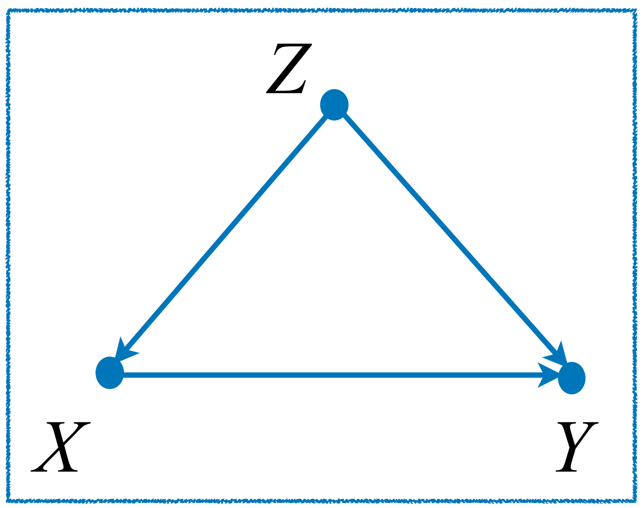
$$P(\mathbf{v} \mid do(\mathbf{x})) = \sum_{\mathbf{u}} \prod_{\{V_i \in \mathbf{V} \setminus \mathbf{X}\}} P\left(v_i \mid pa_i, u_i\right) P(\mathbf{u}) \Bigg|_{X=x}$$

And the effect of such intervention on a set $\boldsymbol{Y}$ is

$$P(\mathbf{y} \mid do(\mathbf{x})) = \sum_{\mathbf{v} \setminus (\mathbf{y} \cup \mathbf{x})} \sum_{\mathbf{u}} \prod_{\{V_i \in \mathbf{V} \setminus \mathbf{X}\}} P\left(v_i \mid pa_i, u_i\right) P(\mathbf{u}) \Bigg|_{X=x}$$

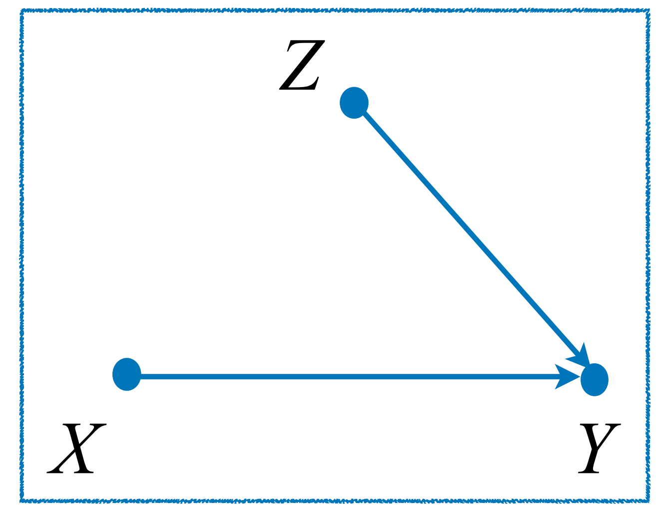
Interventions in Semi-Markovian

Real world



Intervention

Alternative world



$$M = \begin{cases} Z \leftarrow f_Z(u_z) \\ X \leftarrow f_X(z, u_x) \\ Y \leftarrow f_Y(x, z, u_y) \end{cases}$$

do(X=x)

$$M_x = \begin{cases} Z \leftarrow f_Z(u_z) \\ \cancel{X \leftarrow f_X(z, u_x)} \quad X = x \\ Y \leftarrow f_Y(x, z, u_y) \end{cases}$$

$$P(\mathbf{v}) = \sum_{\mathbf{u}} P(z \mid u_z) P(x \mid z, u_x) P(y \mid x, z, u_y) P(\mathbf{u})$$

$$P(\mathbf{v} \mid do(x)) = \sum_{\mathbf{u}} P(z \mid u_z) \cancel{P(x \mid z, u_x)} P(y \mid x, z, u_y) P(\mathbf{u})$$

Interventions in Semi-Markovian

Re-writing the interventional distribution…

$$P(\mathbf{v} \mid do(\mathbf{x})) = \sum_{\mathbf{u}} P(z \mid u_z)\cancel{P(x \mid z, u_x)}P(y \mid x, z, u_y)P(\mathbf{u})$$

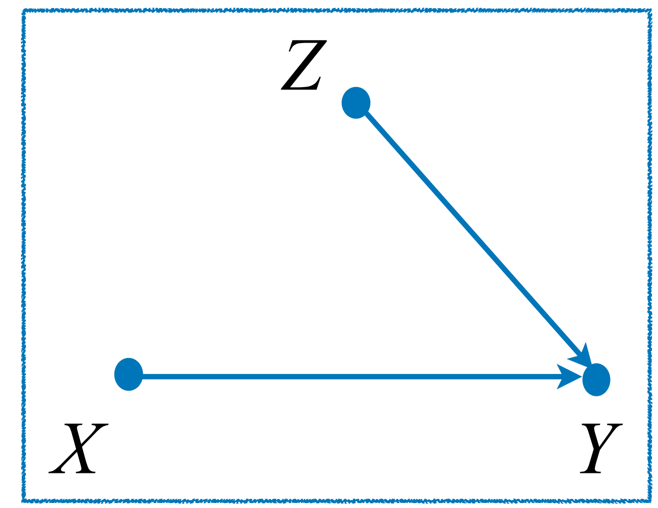
$$= \left(\sum_{u_z} P(z \mid u_z)P(u_z)\right)\left(\sum_{u_y} P(y \mid x, z, u_y)P(u_y)\right)\left(\sum_{u_x} P(u_x)\right)$$

$$= P(z)\left(\sum_{u_y} P(y \mid x, z, u_y)P(u_y)\right)$$

$$= P(z)\left(\sum_{u_y} P(y \mid x, z, u_y)P(u_y \mid x, z)\right)$$

$$= P(z)P(y \mid x, z)$$

Alternative world



$$M_x = \begin{cases} Z \leftarrow f_Z(u_z) \\ \cancel{X \leftarrow f_X(z, u_x)} \quad X = x \\ Y \leftarrow f_Y(x, z, u_y) \end{cases}$$

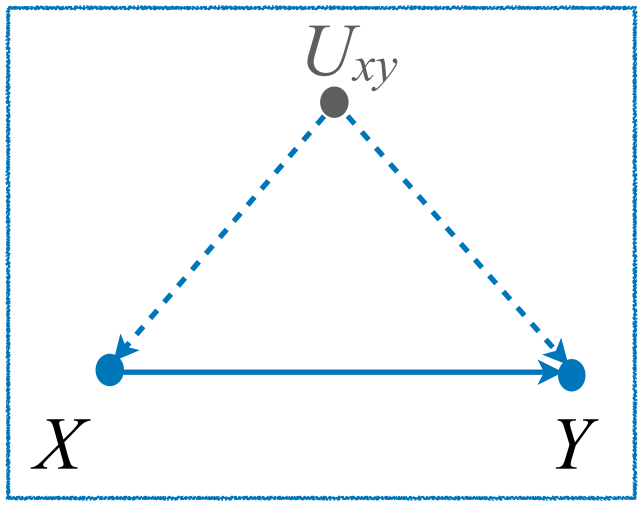
$$P(y \mid do(\mathbf{x})) = \sum_{z} \underline{P(y \mid x, z)P(z)}$$

These distributions can be computed from the obs. distribution $P(z,x,y)$.

(Alternative 'proof' for the backdoor!)

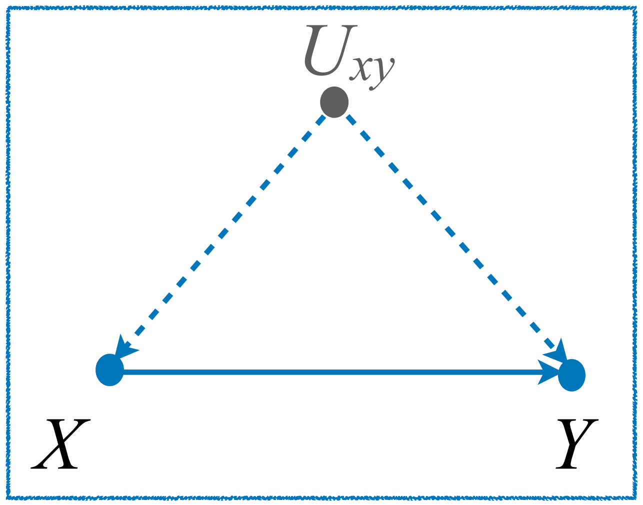
Interventions - Another Example

Real world



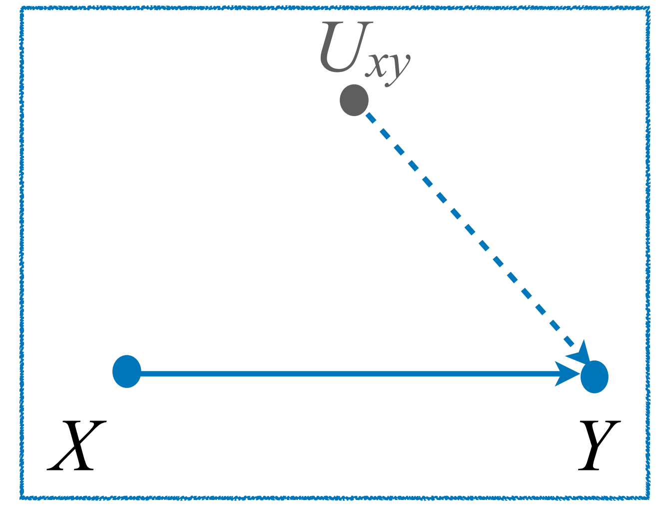
Interventions - Another Example

Real world



Intervention

Alternative world



$$M = \begin{cases} X \leftarrow f_X(u_{xy}, u_x) \\ Y \leftarrow f_Y(x, u_{xy}, u_y) \end{cases}$$

$do(X{=}x)$

$$M_x = \begin{cases} \cancel{X \leftarrow f_X(u_{xy}, u_x)} \quad X = x \\ Y \leftarrow f_Y(x, u_{xy}, u_y) \end{cases}$$

$$P(x, y) = \sum_{u_x, u_y, u_{xy}} P(x \mid u_{xy}, u_x) P(y \mid x, u_{xy}, u_y) P(u)$$

$$P(y \mid do(x)) = \sum_{u_x, u_y, u_z} \cancel{P(x \mid u_z, u_x)} P(y \mid x, u_z, u_y) P(u)$$

Interventions - Another Example

Re-writing the interventional distribution,

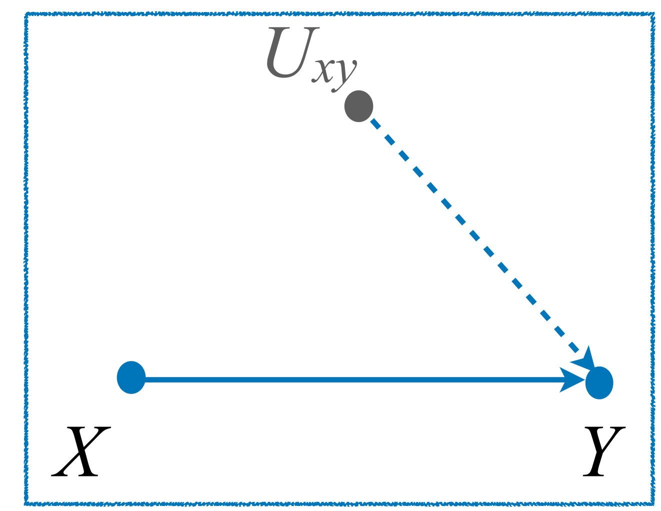
$$P(\mathbf{v} \mid do(x)) = \sum_{u_x, u_y, u_{xy}} \cancel{P(x \mid u_{xy}, u_x)} P(y \mid x, u_{xy}, u_y) P(u_x, u_y, u_{xy})$$

$$= \left(\sum_{u_{xy}} \left(\sum_{u_y} P(y \mid x, u_{xy}, u_y) P(u_y) \right) P(u_{xy}) \right) \left(\sum_{u_x} P(u_x) \right)$$

$$P(y \mid do(x)) = \sum_{u_{xy}} \underline{P(y \mid x, u_{xy}) P(u_{xy})}$$

These distributions are not observed, and nothing more can be removed.

Alternative world

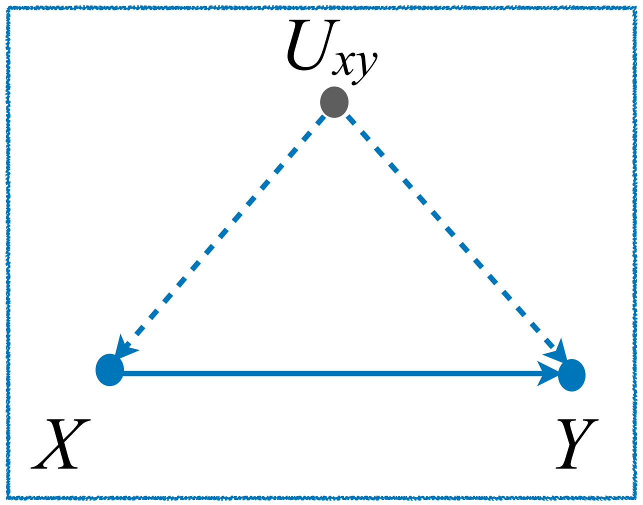


$$M_x = \begin{cases} \cancel{X \leftarrow f_X(u_{xy}, u_x)} \quad X = x \\ Y \leftarrow f_Y(x, u_{xy}, u_y) \end{cases}$$

The Front-door Case

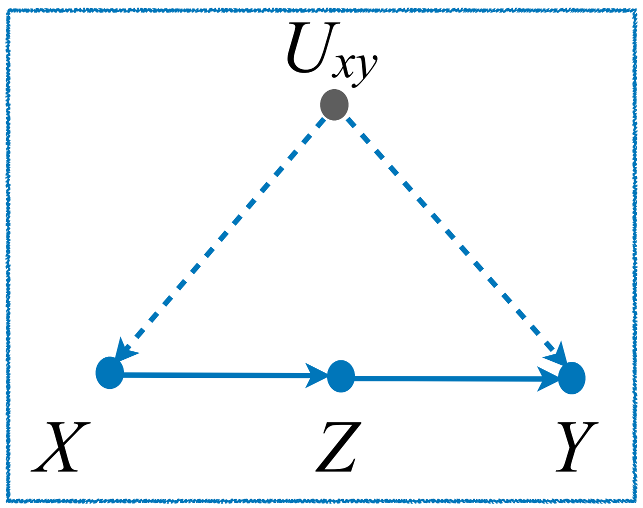
Sec. 4.2.6

Real world



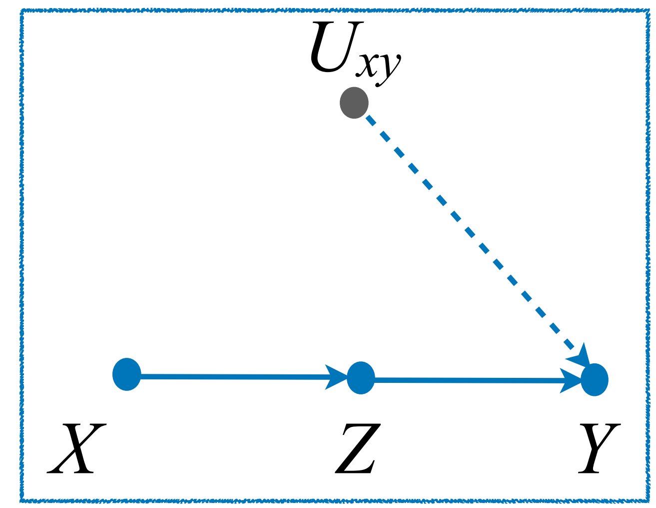
The Front-door Case

Real world



intervention

Alternative world



$$M = \begin{cases} X \leftarrow f_X(u_{xy}, u_x) \\ Z \leftarrow f_Z(x, u_z) \\ Y \leftarrow f_Y(z, u_{xy}, u_y) \end{cases}$$

$do(X{=}x)$

$$M_x = \begin{cases} \cancel{X \leftarrow f_X(u_{xy}, u_x)} \quad X = x \\ Z \leftarrow f_Z(x, u_z) \\ Y \leftarrow f_Y(z, u_{xy}, u_y) \end{cases}$$

$$P(\mathbf{v}) = \sum_{\mathbf{u}} P(x \mid u_{xy}, u_x) P(z \mid x, u_z) P(y \mid z, u_{xy}, u_y) P(\mathbf{u})$$

$$P(\mathbf{v} \mid do(x)) = \sum_{\mathbf{u}} \cancel{P(x \mid u_{xy}, u_x)} P(z \mid x, u_z) P(y \mid z, u_{xy}, u_y) P(\mathbf{u})$$

The Front-door Case

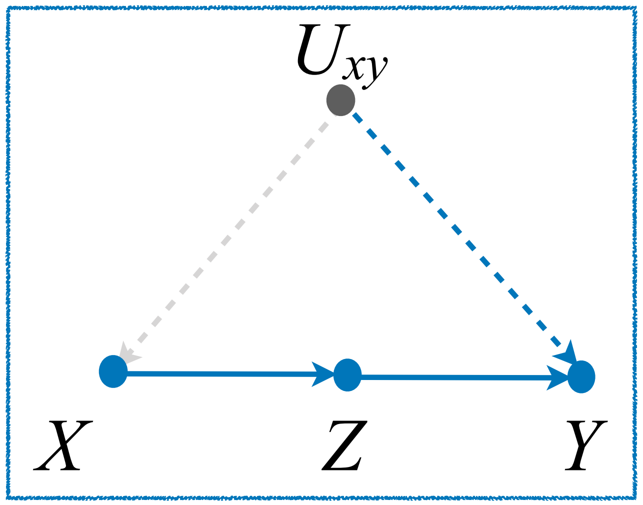
Re-writing the interventional distribution…

$$P(\mathbf{v} \mid do(x)) = \sum_{\mathbf{u}} \cancel{P(x \mid u_{xy}, u_x)} P(z \mid x, u_z) P(y \mid z, u_{xy}, u_y) P(\mathbf{u})$$

$$= \left(\sum_{u_z} P(z \mid x, u_z) P(u_z) \right) \left(\sum_{u_{xy}, u_y} P(y \mid z, u_{xy}, u_y) P(u_{xy}, u_y) \right) \left(\sum_{u_x} P(u_x) \right)$$

$$= P(z \mid x) \sum_{u_{xy}} \underline{P(y \mid z, u_{xy}) P(u_{xy})}$$

Alternative world



Note this seems similar to the previous example since $\neg (U_{xy} \perp\!\!\!\perp Z)$!

Is this instance non-ID as well?

Note that this model entails different invariances, i.e.:

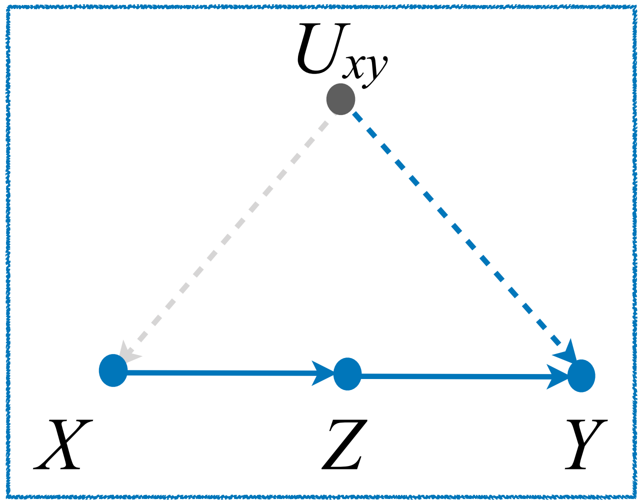
1. $(U_{xy} \perp\!\!\!\perp Z \mid X)$
2. $(Y \perp\!\!\!\perp X \mid Z, U_{xy})$

1. $(U_{xy} \perp\!\!\!\perp Z \mid X)$
2. $(Y \perp\!\!\!\perp X \mid Z, U_{xy})$

The Front-door Case

Re-writing the interventional distribution…

Alternative world



$$P(\mathbf{v} \mid do(x)) = \sum_{\mathbf{u}} \cancel{P(x \mid u_{xy}, u_x)} P(z \mid x, u_z) P(y \mid z, u_{xy}, u_y) P(\mathbf{u})$$

$$= \left(\sum_{u_z} P(z \mid x, u_z) P(u_z) \right) \left(\sum_{u_{xy}, u_y} P(y \mid z, u_{xy}, u_y) P(u_{xy}, u_y) \right) \left(\sum_{u_x} P(u_x) \right)$$

$$= P(z \mid x) \sum_{u_{xy}} P(y \mid z, u_{xy}) P(u_{xy})$$

Summing over X

$$= P(z \mid x) \sum_{x', u_{xy}} P(y \mid z, u_{xy}) P(u_{xy} \mid x') P(x')$$

$(Y \perp\!\!\!\perp X \mid Z, U_{xy})$

$$= P(z \mid x) \sum_{x', u_{xy}} P(y \mid z, x', u_{xy}) P(u_{xy} \mid x') P(x')$$

$(U_{xy} \perp\!\!\!\perp Z \mid X)$

$$= P(z \mid x) \sum_{x', u_{xy}} P(y \mid z, x', u_{xy}) P(u_{xy} \mid x', z) P(x')$$

Chain rule and sum out U_{xy}

$$= P(z \mid x) \sum_{x'} \sum_{u_{xy}} P(y, u_{xy} \mid z, x') P(x')$$

$$P(\mathbf{v} \mid do(x)) = P(z \mid x) \sum_{x'} P(y \mid z, x') P(x')$$

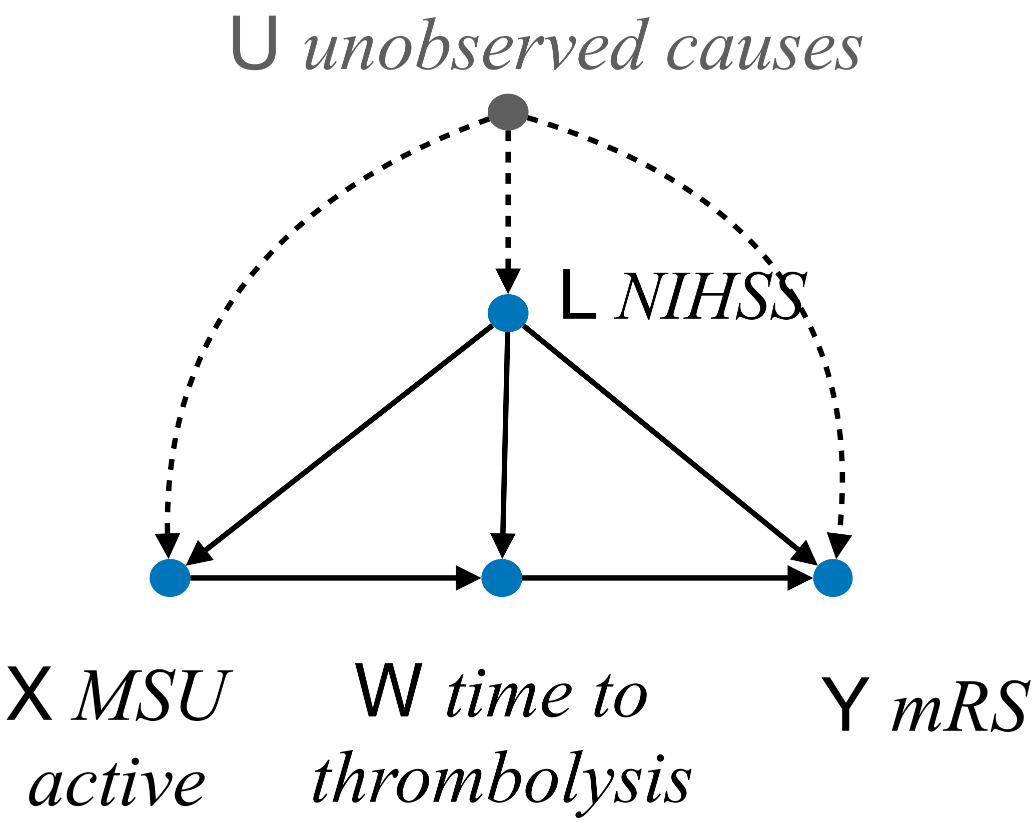
$$P(y \mid do(x)) = \sum_{z} \underline{P(z \mid x)} \sum_{x'} \underline{P(y \mid z, x') P(x')}$$

These factors can be computed from the obs. distribution $P(z,x,y)$.

Real-World Front-Door Application

- Setting: emergency response units respond to ischemic stroke (IS) and transient ischemic attack (TIA) events,
- A Berlin study examined B-PROUD was designed to assess how mobile stroke units (MSU), specifically designed for interventions on site, affect the 3-month functional outcomes of patients,
- Piccininni et. al. 2023 analyze this using an adaptation of the front-door criterion

Real-World Front-Door Application



Key Assumptions:

A1 no $X \rightarrow Y$ direct arrow (fully mediated effect)

justified by the fact that time-to-thrombolysis is a key way how MSUs help outcome.

A2 no $U \rightarrow W$ direct arrow

justified by the fact that stroke-response protocols are highly standardized.

ID Expression:

$$P(Y \mid do(X = x)) = \sum_{w,l} P(w \mid X = x, l) \sum_{k} P(y \mid w, X = k, l) P(X = k, l)$$

The Syntactical Goal on Identification of Causal Effects

- For both back- & front-door settings, the goal was to reduce the quantity $Q = P(\mathbf{y}|do(\mathbf{x}))$ into an expression with neither $do(.)$ nor U, i.e., evaluatable from the obs. distribution $P(\mathbf{v})$.

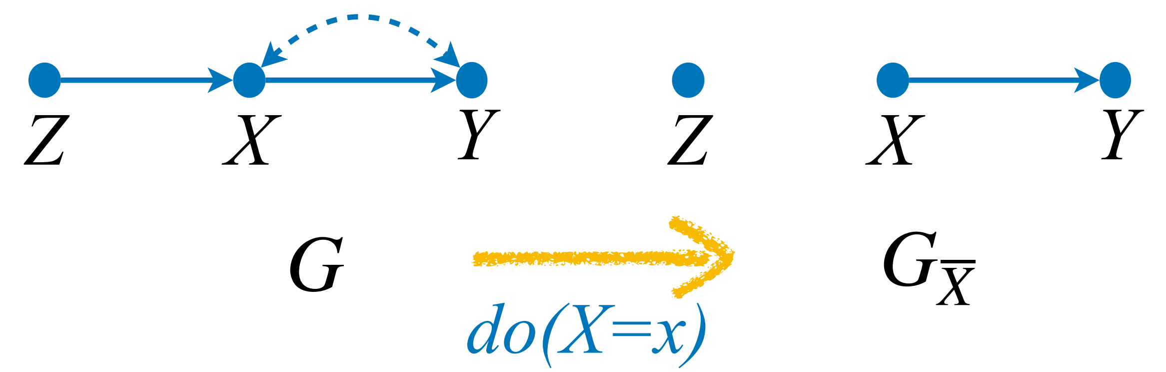
- We are interested in rules or a set of axioms that allow the systematic transformation of a $do(.)$ expression into a do-free expression while preserving the equivalence to the target effect.

Causal Calculus —
Interventional Reasoning Patterns

Pattern 1: Adding/removing Observations

- Adding/removing observations

In the original model, Z and Y may be not separable, e.g.:



$$(Z \not\perp\!\!\!\perp Y),\ (Z \not\perp\!\!\!\perp Y \mid X)$$

However, in the the do(X)-world (model M_x), Y and Z are d-separated, that is,

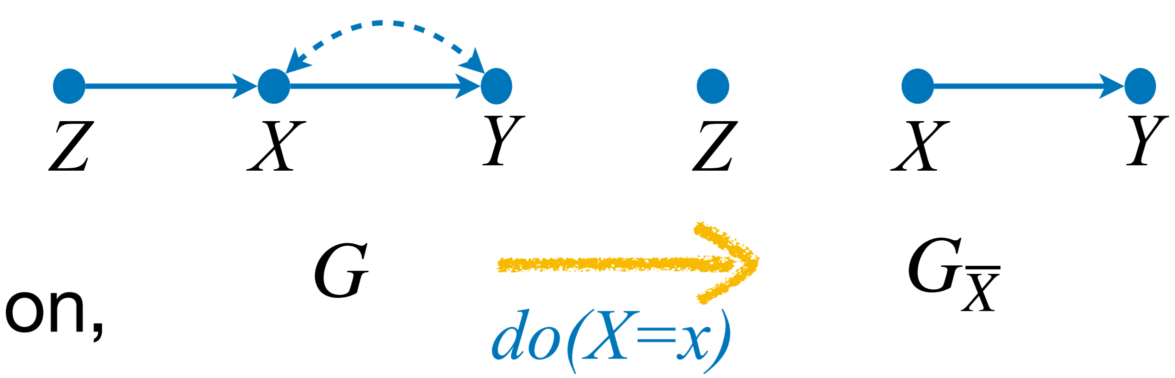
$$(Z \perp\!\!\!\perp Y)_{G_{\overline{X}}} \quad \Longrightarrow \quad P(y|do(x),z)=P(y|do(x))$$

Let's verify this equality!

Pattern 1: Adding/removing Observations

- Adding/removing observations

$P(y|do(x),z)=P(y|do(x))$?



First, let's write the interventional distribution,

$$P(\mathbf{v} \mid do(x))$$

$$= \sum_{\mathbf{u}} P(z \mid u_z) P(y \mid x, u_y, u_{xy}) P(\mathbf{u})$$

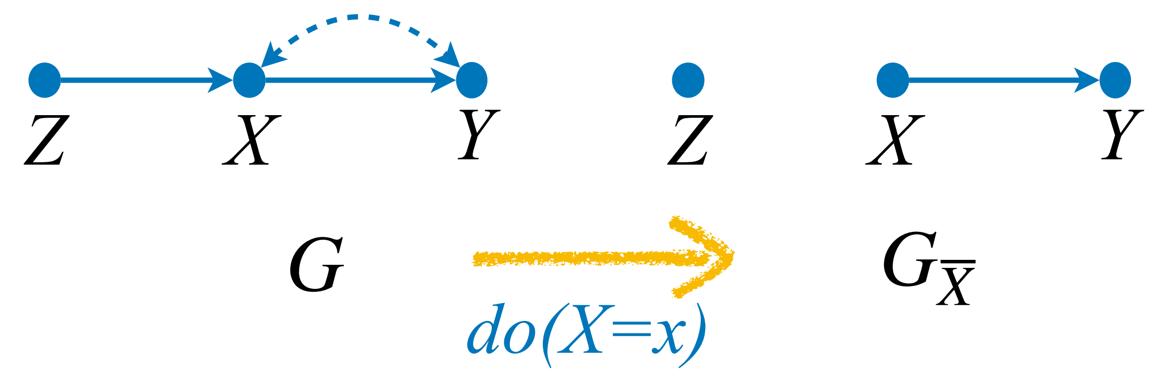
$$= P(z) \sum_{u_{xy}} P(y \mid x, u_{xy}) P(u_{xy})$$

Let's keep the truncated in this form and …

Pattern 1: Adding/removing Observations

- Adding/removing observations

$$P(y|do(x),z)=P(y|do(x))\ ?$$



And let's rewrite the conditional effects,

$$P(y\,|\,do(x),z) = \frac{P(y,z\,|\,do(x))}{P(z\,|\,do(x))}$$

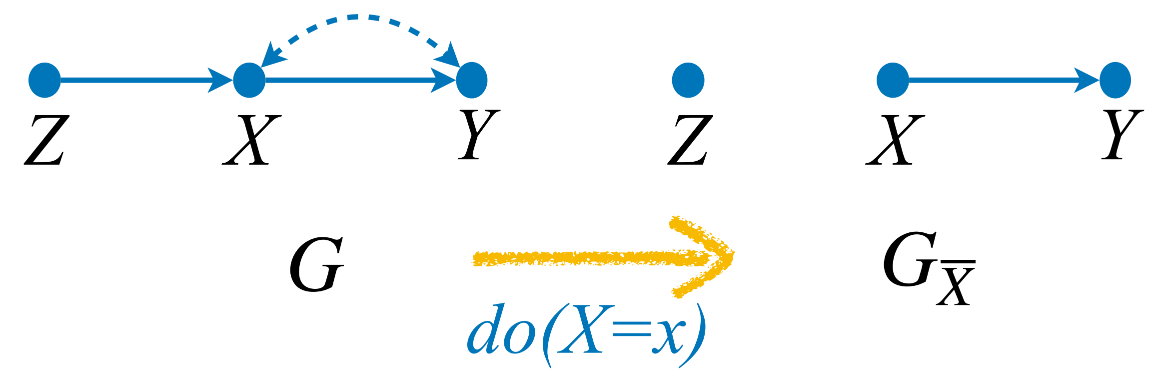
$$P(y,z\,|\,do(x)) = P(z)\sum_{u_{xy}} P(y\,|\,x,u_{xy})P(u_{xy})$$

$$P(z\,|\,do(x)) = \sum_{y} P(z)\sum_{u_{xy}} P(y\,|\,x,u_{xy})P(u_{xy})$$
$$= P(z)$$

Pattern 1: Adding/removing Observations

- Adding/removing observations

$P(y|do(x),z)=P(y|do(x))$?



Substituting the factors back…

$$\boxed{P(y\,|\,do(x),z)} = \frac{P(z)\sum_{u_{xy}} P(y\,|\,x,u_{xy})P(u_{xy})}{P(z)}$$

$$= \sum_{u_{xy}} P(y\,|\,x,u_{xy})P(u_{xy})$$

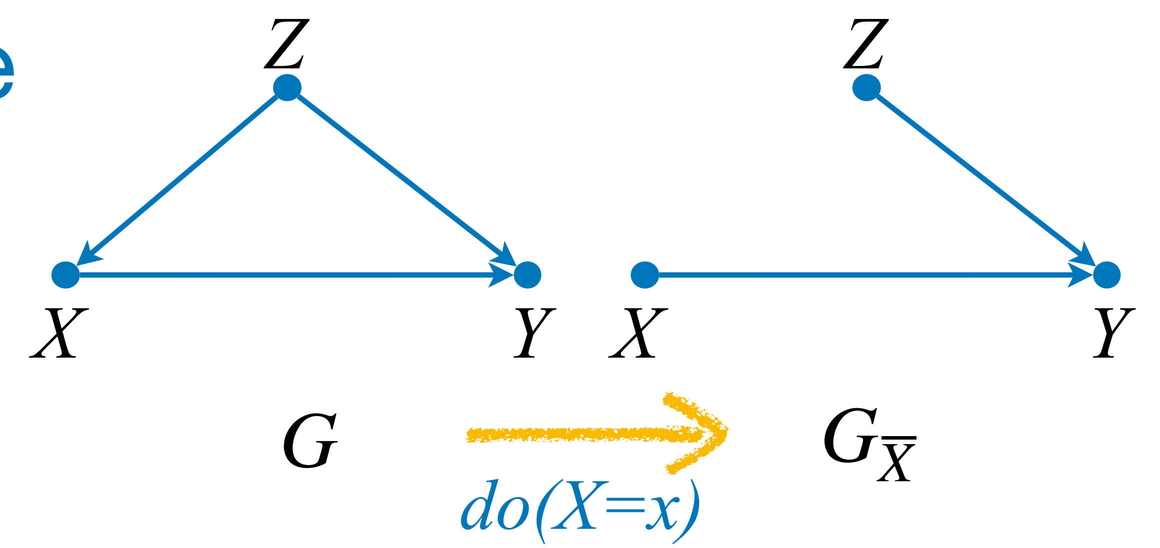
$$= \sum_{z} P(z) \sum_{u_{xy}} P(y\,|\,x,u_{xy})P(u_{xy})$$

$$= \sum_{z} P(\mathbf{v}\,|\,do(x)) \boxed{= P(y\,|\,do(x))}$$

$$(Z \perp\!\!\!\perp Y)_{G_{\overline{X}}}$$ 👍

Pattern 2: Action/Observation Exchange

- Action/Observation Exchange



After observing $Z{=}z$, the var. Y reacts to X in the same way, with and without intervention.

Note that given Z, Y is correlated with X only through causal paths; hence, $see(X{=}x)$ is equiv. to $do(X{=}x)$.

Idea. If Z blocks all bd-paths w.r.t (X, Y), then cond. on Z, all the remaining association is equal to the causation.

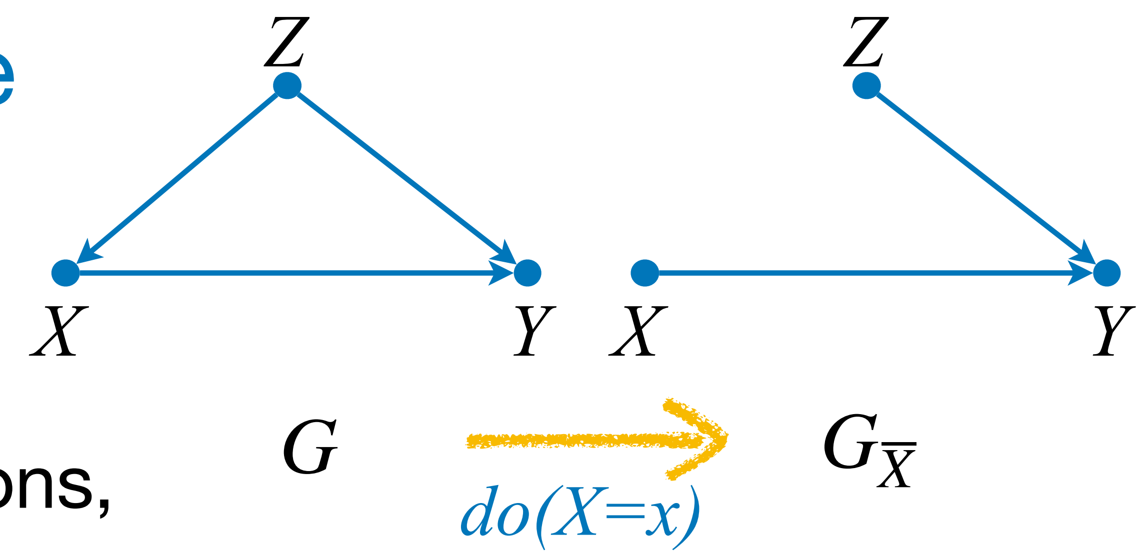
$$(Y \perp\!\!\!\perp X \mid Z)_{G_{\underline{X}}} \qquad \Longrightarrow \qquad P(y|do(x),z)=P(y|x,z)$$

Let's verify this equality!

Pattern 2: Action/Observation Exchange

- Action/Observation Exchange

$P(y|do(x),z)=P(y|x,z)$?



First, let's write the interventional distributions,

$$P(y,z\,|\,do(x)) = \sum_{\mathbf{u}} P(z\,|\,u_z)P(y\,|\,x,z,u_y)P(\mathbf{u}) = P(z)P(y\,|\,x,z)$$

$$P(z\,|\,do(x)) = \sum_{y} P(z)P(y\,|\,x,z) = P(z)$$

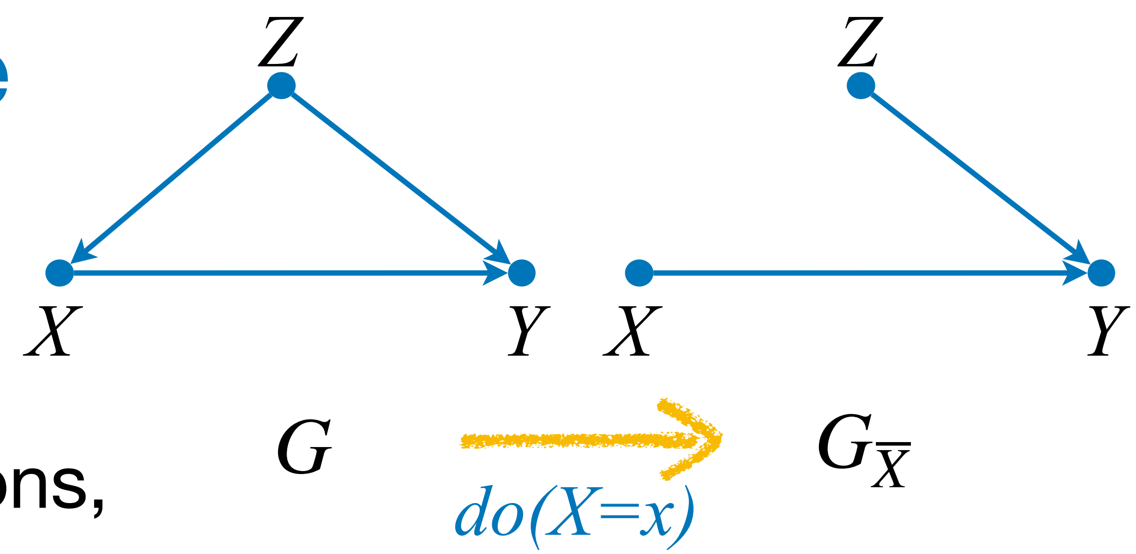
$$\boxed{P(y\,|\,do(x),z)} = \frac{P(z,y\,|\,do(x))}{P(z\,|\,do(x))} = \frac{P(z)P(y\,|\,x,z)}{P(z)} \boxed{= P(y\,|\,x,z)}$$

$$(Y \perp\!\!\!\perp X \mid Z)_{G_{\underline{X}}}$$ 👍

Pattern 2: Action/Observation Exchange

- Action/Observation Exchange

$P(y|do(x),z)=P(y|x,z)$?



First, let's write the interventional distributions,

$$P(y,z\,|\,do(x)) = \sum_{\mathbf{u}} P(z\,|\,u_z)P(y\,|\,x,z,u_y)P(\mathbf{u}) = P(z)P(y\,|\,x,z)$$

$$P(z\,|\,do(x)) = \sum_{y} P(z)P(y\,|\,x,z) = P(z)$$

Looks familiar?

BD perhaps?

$$\boxed{P(y\,|\,do(x),z)} = \frac{P(z,y\,|\,do(x))}{P(z\,|\,do(x))} = \frac{P(z)P(y\,|\,x,z)}{P(z)} \boxed{= P(y\,|\,x,z)}$$

$$(Y \perp\!\!\!\perp X \mid Z)_{G_{\underline{X}}}$$

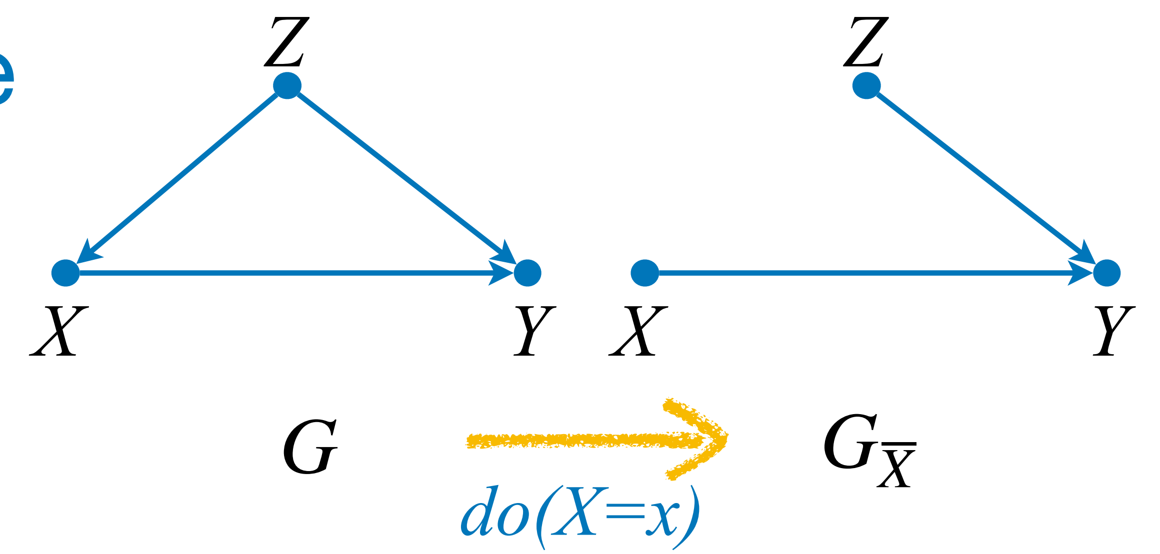
Pattern 2: Action/Observation Exchange

- Action/Observation Exchange

Great, but what about the equality

$$P(y|do(x)) = P(y|x)\,?$$

$$(Y \not\perp\!\!\!\perp X)_{G_{\underline{X}}}$$



Let's compare left and right-hand sides:

$$P(y \mid do(x)) = \sum_z \sum_{\mathbf{u}} P(y \mid x, z, u_y) P(z \mid u_z) P(\mathbf{u}) = \sum_z P(y \mid x, z) \boxed{P(z)}$$

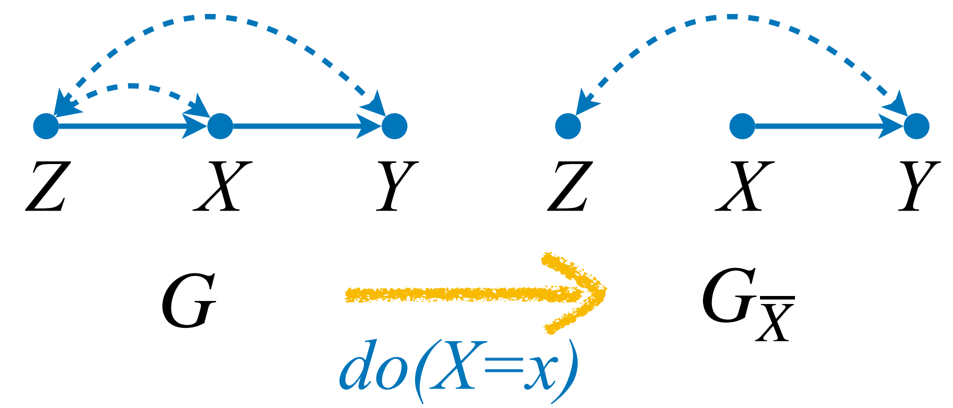
$$P(y \mid x) = \sum_z P(y \mid x, z) \boxed{P(z \mid x)}$$

For almost any model compatible with this causal graph, $P(y|x)$ *and* $P(y \mid do(x))$ will **not** be equal since $P(z) \neq P(z \mid x)$ almost surely.

Pattern 3: Adding/Removing Actions

- Adding/Removing Actions

If there is no causal path from X to Z, then an intervention on X will have no effect on Z.



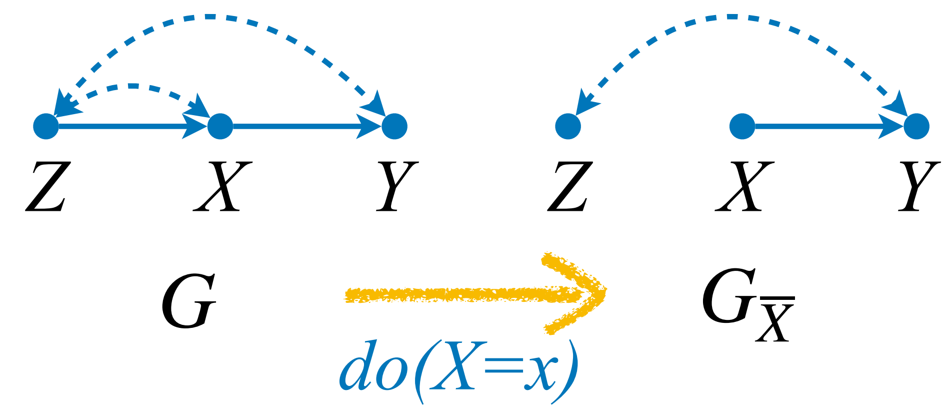
$$(Z \perp\!\!\!\perp X)_{G_{\overline{X}}} \quad \Longrightarrow \quad P(z|do(x)) = P(z)$$

Let's verify this equality!

Pattern 3: Adding/Removing Actions

- Adding/Removing Actions

$$P(z|do(x))=P(z) \text{ ?}$$



$$\boxed{P(z\mid do(x))} = \sum_y P(\mathbf{v}\mid do(x))$$

$$= \sum_y \sum_{u_{zy},u_{zx}} P(z\mid u_{zy}, u_{zx})P(y\mid x, u_{zy})P(u_{zy}, u_{zx})$$

$$= \sum_{u_{zy},u_{zx}} P(z\mid u_{zy}, u_{zx})P(u_{zy}, u_{zx})$$

$$= \boxed{P(z)} \qquad (Z \perp\!\!\!\perp X)_{G_{\overline{X}}}$$ 👍

Rules of Do-Calculus

Theorem. The following transformations are valid for any do-distribution induced by a SCM M:

Rule 1: Adding/removing Observations

$$P(y|do(x),z,w)=P(y|do(x),w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z \mid X, W)_{G_{\overline{X}}}$$

Rule 2: Action/observation exchange

$$P(y|do(x),do(z),w)=P(y|do(x),z,w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z \mid X, W)_{G_{\overline{X}\underline{Z}}}$$

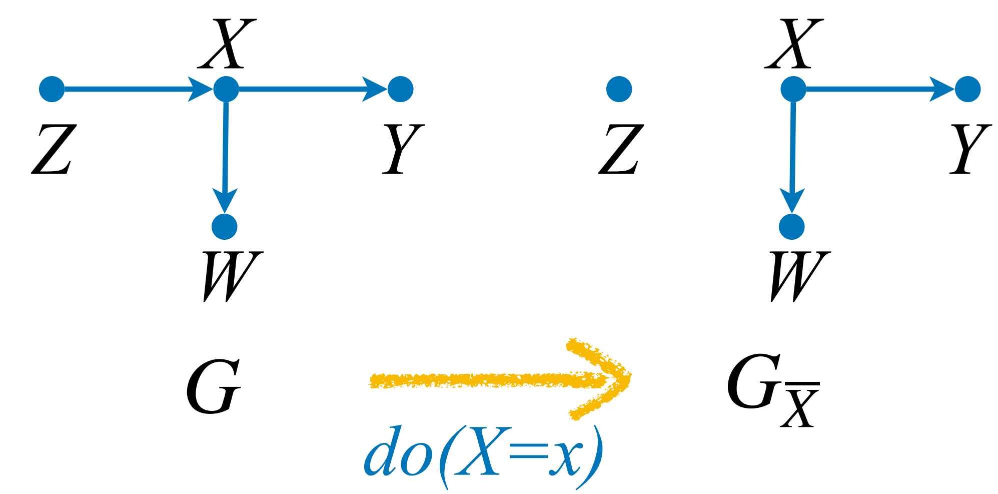
Rule 3: Adding/removing Actions

$$P(y|do(x),do(z),w)=P(y|do(x),w) \quad \text{if} \quad (Y \perp\!\!\!\perp Z \mid X, W)_{G_{\overline{X}\overline{Z(W)}}}$$

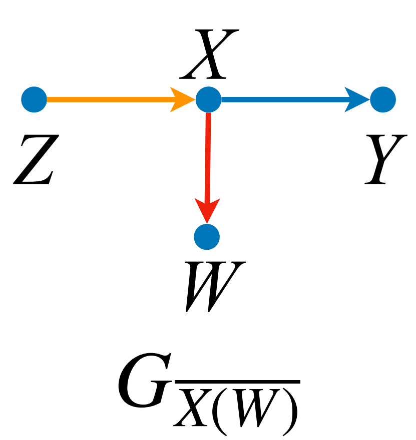
where $Z(W)$ is the set of Z-nodes that are not ancestors of any W-node in $G_{\overline{X}}$.

A curiosity about R3

- Regarding R3's contingency, note that the probability of Z given W may not be the same with or without intervening on X.



To witness, compare $G_{\overline{X}}$ and $G_{\overline{X(W)}}$:

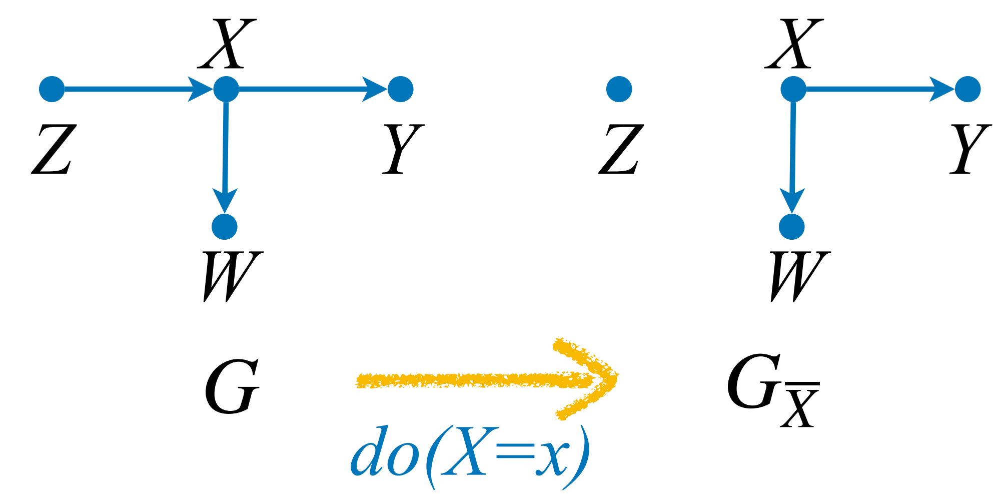


$$(Z \not\!\perp\!\!\!\perp X \mid W)_{G_{\overline{X(W)}}} \Longrightarrow \exists_M P(z|do(x),w) \neq P(z|w)$$

What? 🙋‍♂️

A curiosity about R3

- Verifying whether
  $P(z|do(x),w) \neq P(z|w)$ holds:

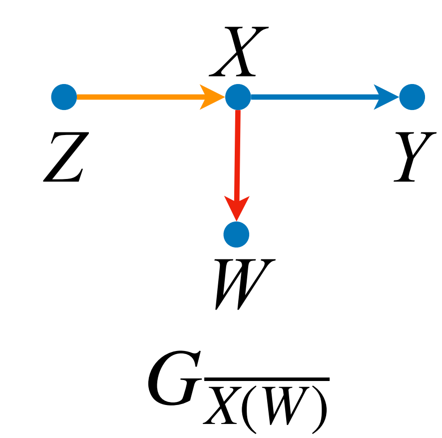


$$\boxed{P(z \mid do(x), w)} = \frac{P(z, w \mid do(x))}{P(w \mid do(x))}$$

$$= \frac{\sum_y P(z)P(w \mid x)P(y \mid x)}{\sum_{y,z} P(z)P(w \mid x)P(y \mid x)}$$

$$= \frac{P(z)P(w \mid x)}{P(w \mid x)}$$

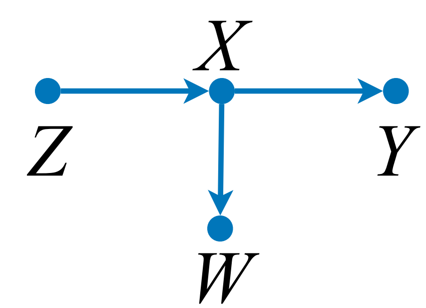
$$= P(z) \boxed{\neq P(z \mid w)} \qquad (Z \not\perp\!\!\!\perp X \mid W)_{G_{\overline{X(W)}}}$$



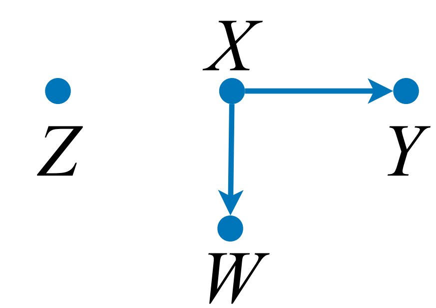
in almost any model
compatible with G.

A curiosity about R3

- Verifying whether
  $P(z|do(x),w) \neq P(z|w)$ holds:



G



$G_{\overline{X}}$

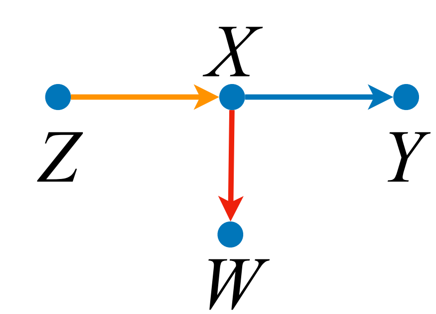
$X=x)$

Intuitively:

(i) W is a descendant of X

(ii) so in a $do(x)$-world, W "absorbs" the influence of a $do(x)$ intervention

$\Longrightarrow$ (iii) when conditioning on W in $P(z \mid w)$, we get something different in the $do(x)$ vs. normal world.



$G_{\overline{X(W)}}$

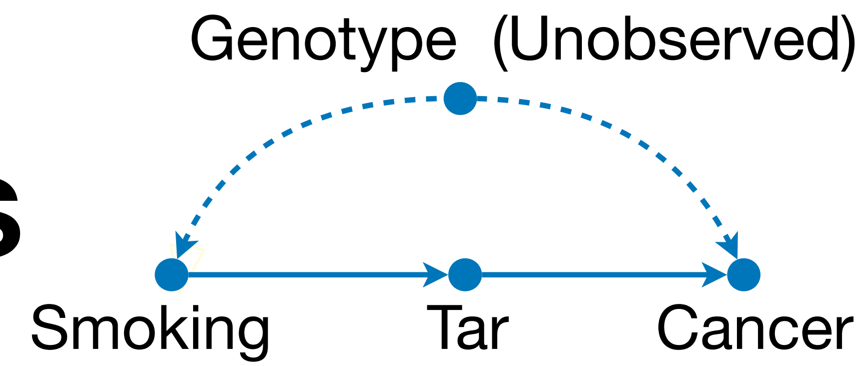
Properties of Do-Calculus

Theorem (soundness & completeness of do-calculus for interventional ID from obs. data).

The quantity $Q = P(y|do(x))$ is identifiable from $P(v)$ and G if and only if there exists a sequence of application of the rules of do-calculus and the probability axioms that reduces Q into a do-free expression.

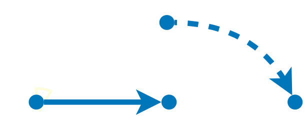
Syntactic goal: Re-express original Q without do()!

Derivation in Do-Calculus

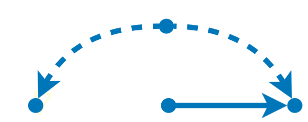


$$P(c \mid do(s)) = \sum_t P(c \mid do(s), t) P(t \mid do(s))$$ Probability Axioms

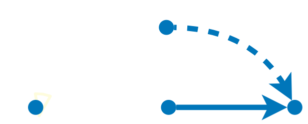
$$= \sum_t P(c \mid do(s), do(t)) P(t \mid do(s))$$ Rule 2



$$= \sum_t P(c \mid do(s), do(t)) P(t \mid s)$$ Rule 2

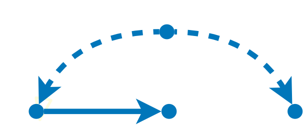


$$= \sum_t P(c \mid do(t)) P(t \mid s)$$ Rule 3



$$= \sum_t \sum_{s'} P(c \mid do(t), s') P(s' \mid do(t)) P(t \mid s)$$ Probability Axioms

$$= \sum_t \sum_{s'} P(c \mid t, s') P(s' \mid do(t)) P(t \mid s)$$ Rule 2



$$= \sum_t \sum_{s'} P(c \mid t, s') P(s') P(t \mid s)$$ Rule 3