Title: Lecture - Causal Inference, PHYS 777

Speakers: Robert Spekkens

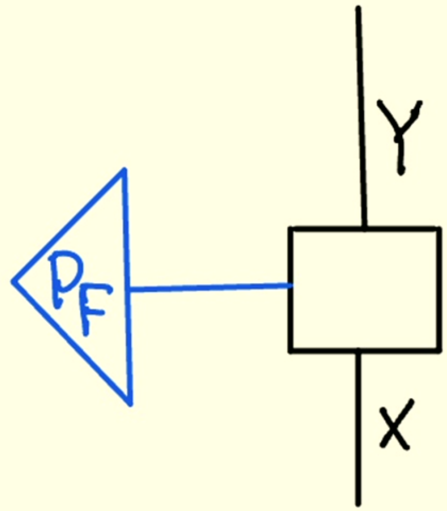
Collection/Series: Causal Inference (Elective), PHYS 777, March 31 - May 2, 2025

Subject: Quantum Foundations

Date: April 08, 2025 - 11:30 AM

URL: https://pirsa.org/25040039

Estimating causal effects


Y
p_F
X

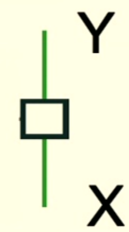
There is some function f such that Y=f(X). It is everything that can be said about the causal relation between them. Generally, however, one merely learns the distribution over such functions P_F

What is often of interest is the do-conditional $P_{Y|doX}$

$$P_{Y|\mathrm{do}X} = \sum_f \delta_{Y,f(X)} P_F(f)$$

But there are many P_F consistent with a given $P_{Y|doX}$

Consider the four functions on the set {0,1}



$$f_{\text{id}}, f_{\text{flip}}, f_{\text{reset}-0}, f_{\text{reset}-1}$$

Now, consider two states of knowledge:

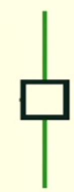
$$P_F = \tfrac{1}{2}[f_{\text{id}}] + \tfrac{1}{2}[f_{\text{flip}}] \qquad P'_F = \tfrac{1}{2}[f_{\text{reset}-0}] + \tfrac{1}{2}[f_{\text{reset}-1}]$$

$$P_{Y|\text{do}X} = \textstyle\sum_f \delta_{Y,f(X)} P_F(f)$$

$$P_{Y|\text{do}X} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$$P_{Y|\text{do}X} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Consider the four functions on the set {0,1}

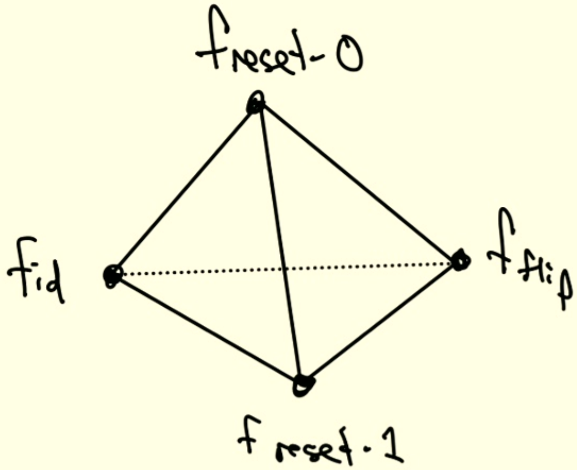


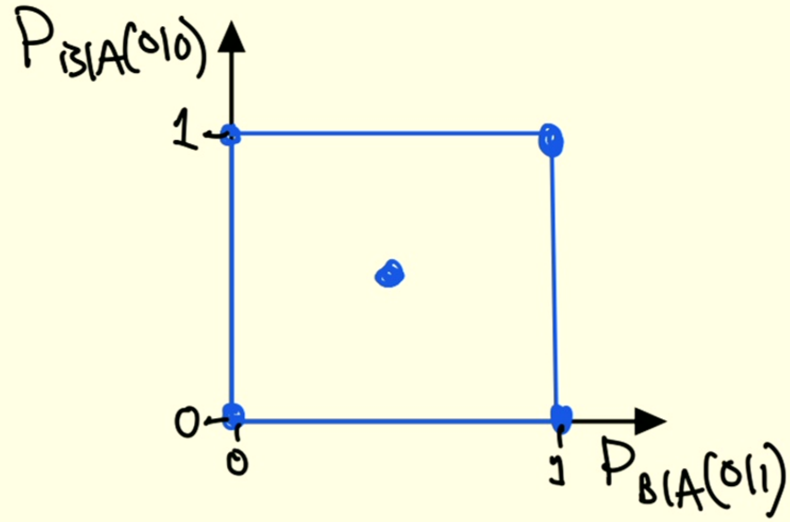
$$f_{\mathsf{id}}, f_{\mathsf{flip}}, f_{\mathsf{reset-0}}, f_{\mathsf{reset-1}}$$

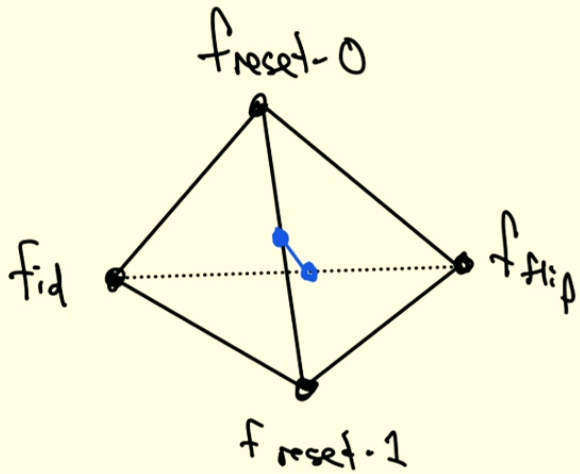
Indeed, for any state of knowledge of the form

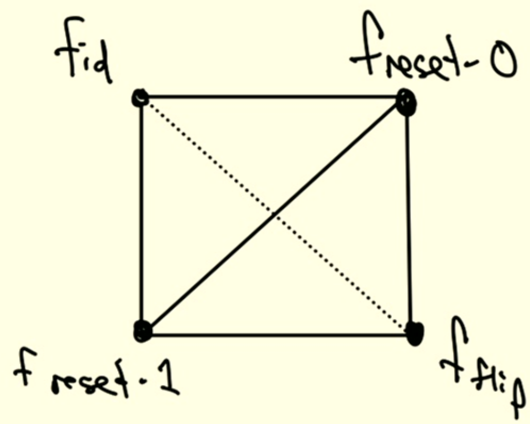
$$P''_F = q\left(\tfrac{1}{2}[f_{\mathsf{id}}] + \tfrac{1}{2}[f_{\mathsf{flip}}]\right) + (1-q)\left(\tfrac{1}{2}[f_{\mathsf{reset-0}}] + \tfrac{1}{2}[f_{\mathsf{reset-1}}]\right)$$

$$P_{Y|\mathsf{do}X} = \sum_f \delta_{Y,f(X)} P_F(f)$$
$$= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

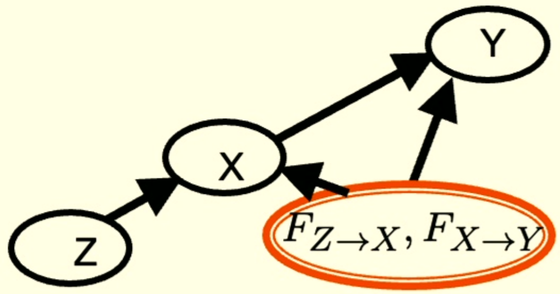

f reset-0
f id
f flip
f reset-1


P B|A(0|0)
1
0
0
1
P B|A(0|1)


f reset-0
f id
f flip
f reset-1


f id
f reset-0
f reset-1
f flip

Causal structure



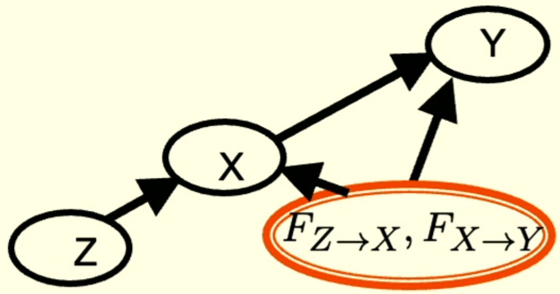
Parameters

$$P_{F_{Z\to X},F_{X\to Y}}$$

$$P_{XYZ} = \sum_{f,f'} \delta_{X,f(Z)} \delta_{Y,f'(X)} P_{F_{Z\to X},F_{X\to Y}}(f,f')$$

This is called "Gearing Λ"

Causal structure Parameters



$P_{F_{Z\to X},F_{X\to Y}}$

$$P_{XYZ} = \sum_{f,f'} \delta_{X,f(Z)} \delta_{Y,f'(X)} P_{F_{Z\to X},F_{X\to Y}}(f,f')$$

Suppose one observes

$$P_{ZXY} = (\tfrac{1}{2}[00]_{ZY} + \tfrac{1}{2}[11]_{ZY})(\tfrac{1}{2}[0]_X + \tfrac{1}{2}[1]_X)$$

$$\Rightarrow \quad P_{F_{Z\to X}F_{X\to Y}} = \tfrac{1}{2}[f_{\text{id}}, f_{\text{id}}] + \tfrac{1}{2}[f_{\text{flip}}, f_{\text{flip}}]$$

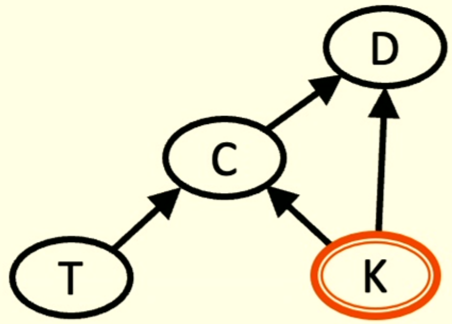
$$\Rightarrow \quad P_{F_{Z\to X}} = \tfrac{1}{2}[f_{\text{id}}] + \tfrac{1}{2}[f_{\text{flip}}]$$

Definition: X has a nontrivial influence on Y

The function f that determines Y from its causal antecedents has a nontrivial dependence on X

Vernam cypher

T = plaintext
C = cyphertext
K = key
D = decoded text



$$P_{TCD} = (\tfrac{1}{2}[00]_{TD} + \tfrac{1}{2}[11]_{TD})(\tfrac{1}{2}[0]_C + \tfrac{1}{2}[1]_C)$$

⇒

C = (T + K)mod2

D = (C + K)mod2

$P_K = ½[0]_K + ½[1]_K$

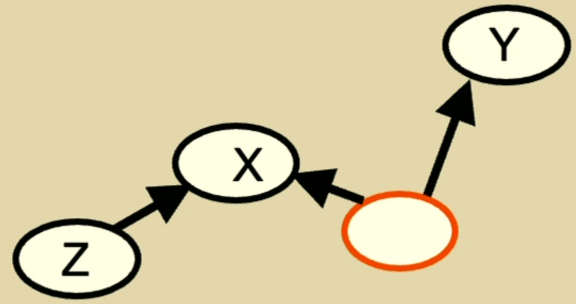
Estimating probability of causation based on violation of conditional independence constraints

The evidence

Violates the independence constraint:

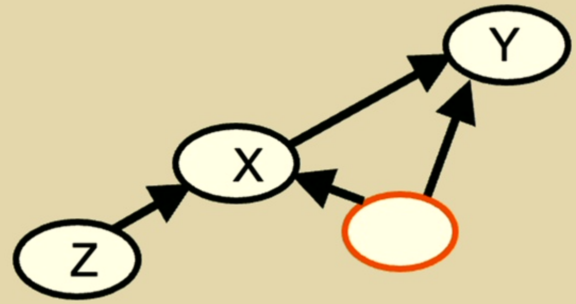
$$P_{Y|Z} \neq P_Y$$

The hypotheses



Implies an independence constraint:

$$P_{Y|Z} = P_Y$$



The evidence

The hypotheses

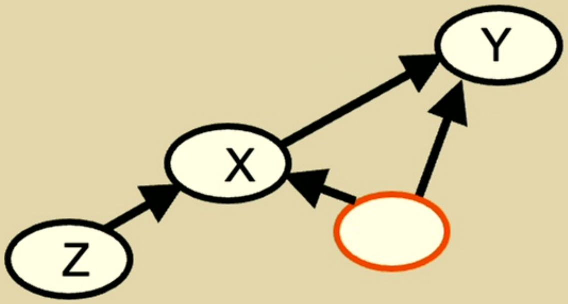
Violates the independence constraint:

$$P_{Y|Z} \neq P_Y$$

Implies an independence constraint:

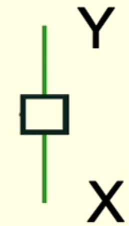
$$P_{Y|Z} = P_Y$$

How should we quantify causal effect?



Consider the four functions on the set {0,1}

$f_{\text{id}}, f_{\text{flip}}, f_{\text{reset}-0}, f_{\text{reset}-1}$



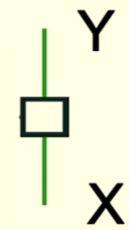
Good measures of causal effect:

The probability that the function has 1 bit of causal influence

$$P_{F_{X\to Y}}(\texttt{1-bit}) := P_{F_{X\to Y}}(f_{\text{id}}) + P_{F_{X\to Y}}(f_{\text{flip}})$$

Consider the four functions on the set {0,1}

$$f_{\text{id}}, f_{\text{flip}}, f_{\text{reset}-0}, f_{\text{reset}-1}$$



Good measures of causal effect:

The probability that the function has 1 bit of causal influence

$$P_{F_{X\to Y}}(\texttt{1-bit}) := P_{F_{X\to Y}}(f_{\text{id}}) + P_{F_{X\to Y}}(f_{\text{flip}})$$

The probability that the function is identity

$$P_{F_{X\to Y}}(f_{\text{id}})$$

The probability that the function is flip

$$P_{F_{X\to Y}}(f_{\text{flip}})$$

The evidence

Violates the independence constraint:

$$P_{Y|X} \neq P_Y$$

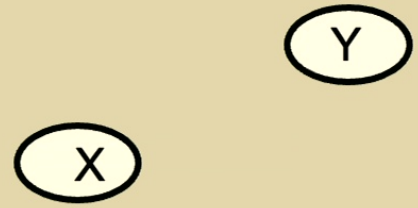
$$P_{Y|X}(1|1) \neq P_{Y|X}(1|0)$$

$$P_{F_{X\to Y}}(\mathbb{I}) \geq \max\{0, P_{Y|X}(1|1) - P_{Y|X}(1|0)\}$$

$$P_{F_{X\to Y}}(\mathbb{F}) \geq \max\{0, P_{Y|X}(1|0) - P_{Y|X}(1|1)\}$$

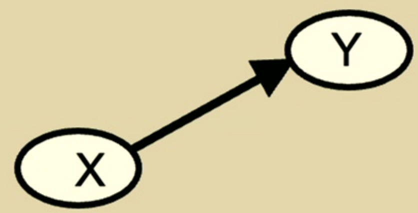
$$P_{F_{X\to Y}}(\texttt{1-bit}) \geq |P_{Y|X}(1|1) - P_{Y|X}(1|0)|$$

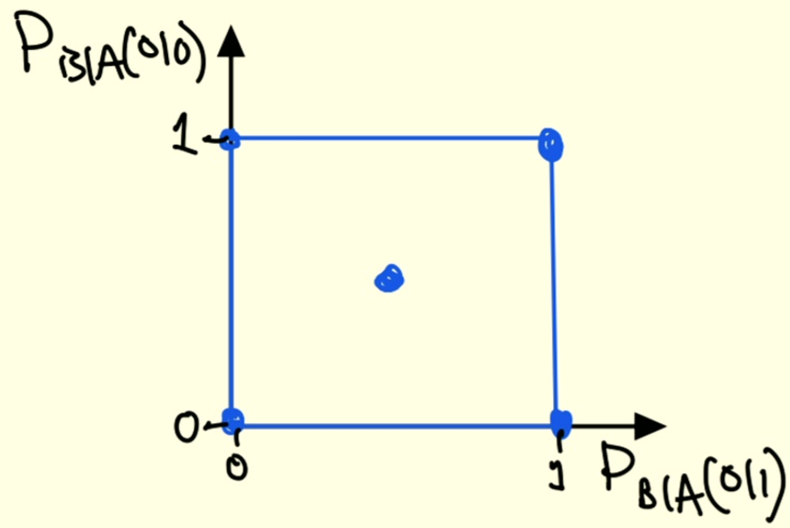
The hypotheses



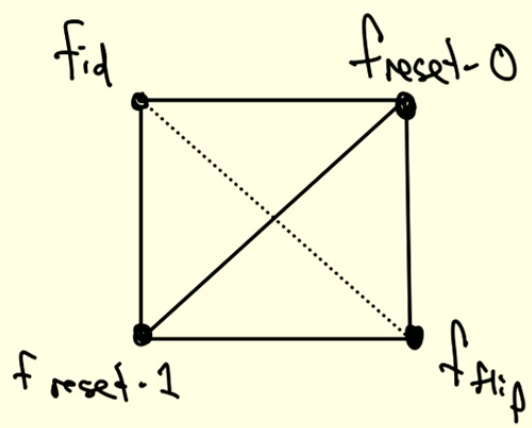
Implies an independence constraint:

$$P_{Y|X} = P_Y$$





P B|A(0|0)
1
0
0
1
P B|A(0|1)



f id
f reset-0
f reset-1
f flip

The evidence

Violates the independence constraint:

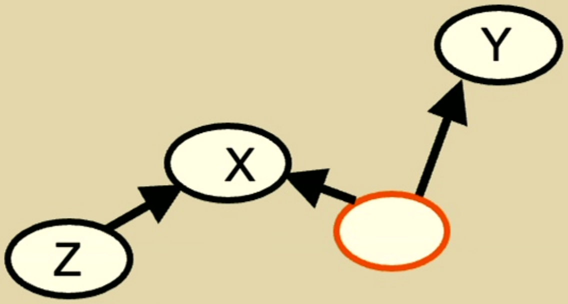
$$P_{Y|Z} \neq P_Y$$

$$P_{Y|Z}(1|1) \neq P_{Y|Z}(1|0)$$

$$P_{F_{X \to Y}}(\texttt{1-bit}) \geq |P_{Y|Z}(1|1) - P_{Y|Z}(1|0)|$$

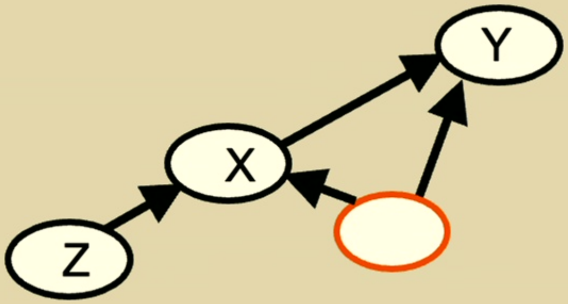
$$P_{F_{Z \to X}}(\texttt{1-bit}) \geq |P_{Y|Z}(1|1) - P_{Y|Z}(1|0)|$$

The hypotheses

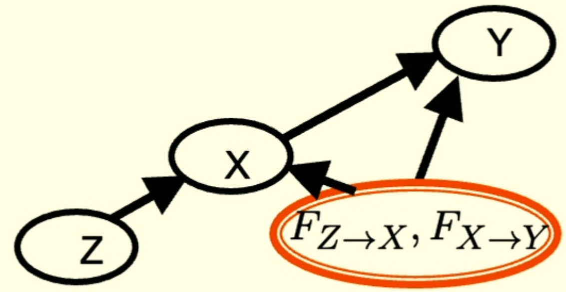


Implies an independence constraint:

$$P_{Y|Z} = P_Y$$



Proof: $P_{F_{Z\to Y}}(\mathbb{I}) = P_{F_{Z\to X},F_{X\to Y}}(\mathbb{I},\mathbb{I})$
$$+\, P_{F_{Z\to X},F_{X\to Y}}(\mathbb{F},\mathbb{F})$$



Now, because

$$P_{F_{X\to Y}}(\text{1-bit}) = P_{F_{X\to Y}}(\mathbb{I}) + P_{F_{X\to Y}}(\mathbb{F})$$
$$\geq P_{F_{Z\to X},F_{X\to Y}}(\mathbb{I},\mathbb{I}) + P_{F_{Z\to X},F_{X\to Y}}(\mathbb{F},\mathbb{F})$$

It follows that:

$$P_{F_{X\to Y}}(\text{1-bit}) \geq P_{F_{Z\to Y}}(\mathbb{I})$$

We use the fact that

$$P_{F_{Z\to Y}}(\mathbb{I}) \geq \max\{0, P_{Y|Z}(1|1) - P_{Y|Z}(1|0)\}$$

To conclude that

$$P_{F_{X\to Y}}(\text{1-bit}) \geq \max\{0, P_{Y|Z}(1|1) - P_{Y|Z}(1|0)\}$$

From

$$P_{F_{X\to Y}}(\text{1-bit}) \geq \max\{0, P_{Y|Z}(1|1) - P_{Y|Z}(1|0)\}$$

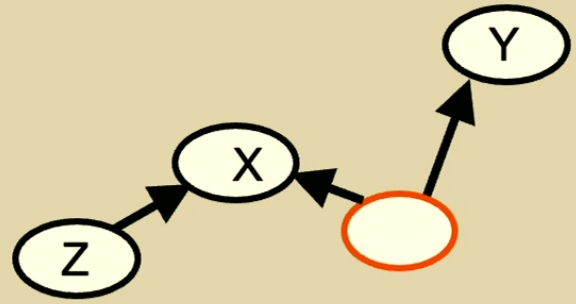
$$P_{F_{X\to Y}}(\text{1-bit}) \geq \max\{0, P_{Y|Z}(1|0) - P_{Y|Z}(1|1)\}$$

We infer that

$$P_{F_{X\to Y}}(\texttt{1-bit}) \geq |P_{Y|Z}(1|1) - P_{Y|Z}(1|0)|$$

The evidence

The hypotheses



Violates the independence constraint:

$$P_{Y|Z} \neq P_Y$$

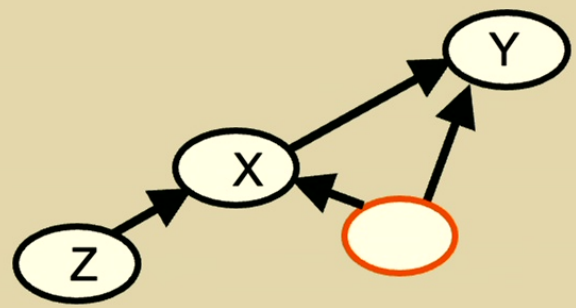
$$P_{Y|Z}(1|1) \neq P_{Y|Z}(1|0)$$

Implies an independence constraint:

$$P_{Y|Z} = P_Y$$

$$P_{F_{X\to Y}}(\texttt{1-bit}) \geq |P_{Y|Z}(1|1) - P_{Y|Z}(1|0)|$$

$$P_{F_{Z\to X}}(\texttt{1-bit}) \geq |P_{Y|Z}(1|1) - P_{Y|Z}(1|0)|$$

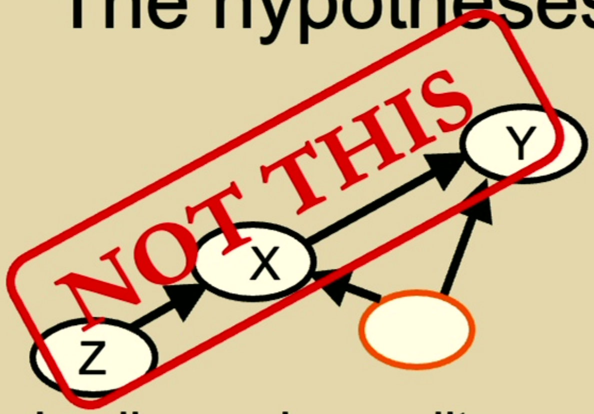


The evidence

Violates the inequality constraint

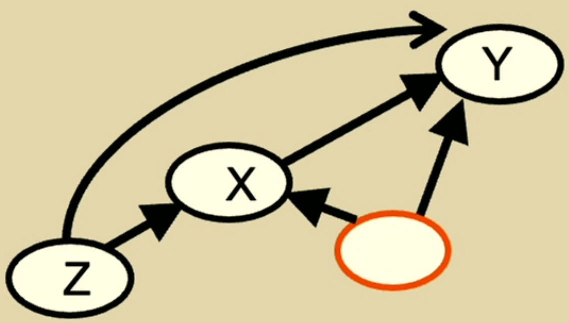
$$P_{XY|Z}(00|0) + P_{XY|Z}(01|1) - 1 \geq 0$$

The hypotheses

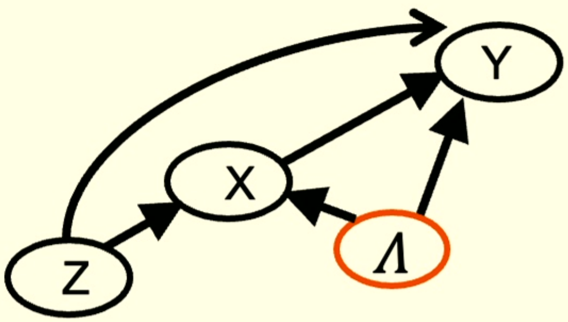


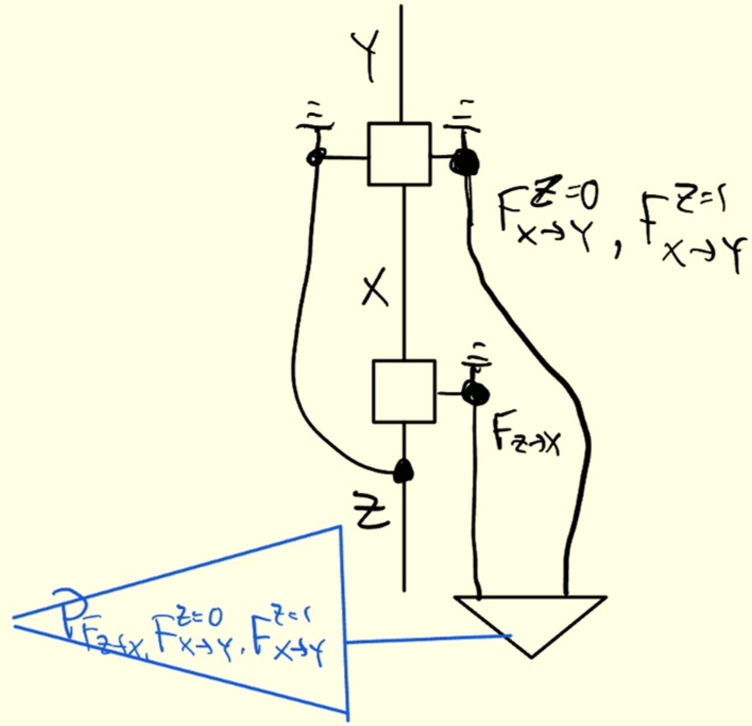
Implies an inequality constraint:

$$P_{XY|Z}(00|0) + P_{XY|Z}(01|1) \leq 1$$



Gearing Λ and Z



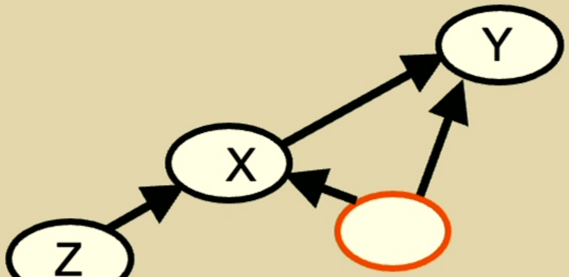


The evidence

Violates the inequality constraint

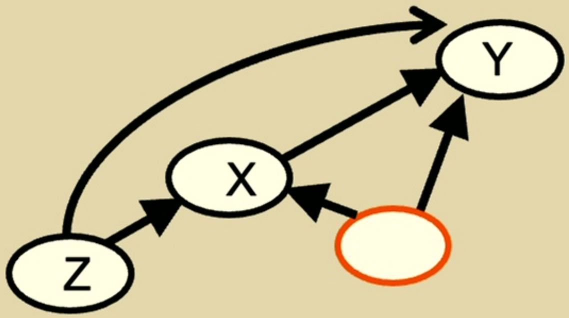
$$P_{XY|Z}(00|0) + P_{XY|Z}(01|1) - 1 \geq 0$$

The hypotheses

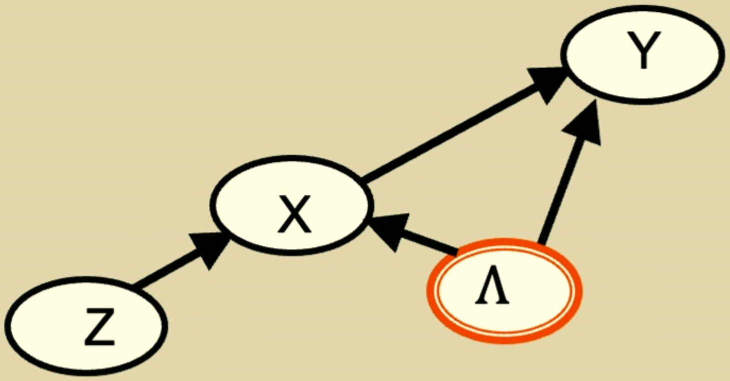


Implies the inequality constraint:

$$P_{XY|Z}(00|0) + P_{XY|Z}(01|1) \leq 1$$



Causal structure **Parameters**



$P_{X|\Lambda Z}$

$P_{Y|\Lambda X}$

P_{Λ}

P_{Z}

$$P_{XYZ} = \sum_{\Lambda} P_{Y|X\Lambda} P_{X|Z\Lambda} P_{\Lambda} P_{Z}$$

$$P_{XY|Z}(00|0) + P_{XY|Z}(01|1) \leq 1$$
$$P_{XY|Z}(01|0) + P_{XY|Z}(00|1) \leq 1$$
$$P_{XY|Z}(10|0) + P_{XY|Z}(11|1) \leq 1$$
$$P_{XY|Z}(11|0) + P_{XY|Z}(10|1) \leq 1$$

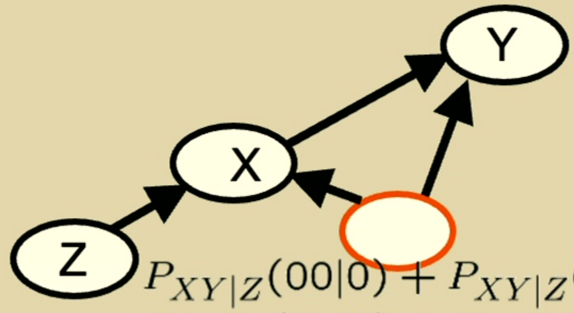
The evidence

Violates one or more inequality constraint

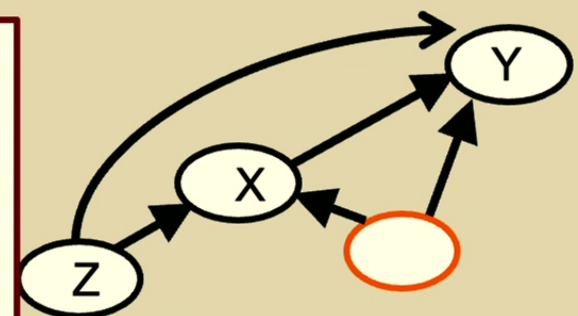
$$P_{XY|Z}(00|0) + P_{XY|Z}(01|1) - 1 \geq 0$$
$$P_{XY|Z}(01|0) + P_{XY|Z}(00|1) - 1 \geq 0$$
$$P_{XY|Z}(10|0) + P_{XY|Z}(11|1) - 1 \geq 0$$
$$P_{XY|Z}(10|0) + P_{XY|Z}(11|1) - 1 \geq 0$$

$$P_{F^{X=0}_{Z\to Y}}(\mathbb{I}) \geq \max\{0, P_{XY|Z}(00|0) + P_{XY|Z}(01|1) - 1\}$$
$$P_{F^{X=0}_{Z\to Y}}(\mathbb{F}) \geq \max\{0, P_{XY|Z}(01|0) + P_{XY|Z}(00|1) - 1\}$$
$$P_{F^{X=1}_{Z\to Y}}(\mathbb{I}) \geq \max\{0, P_{XY|Z}(10|0) + P_{XY|Z}(11|1) - 1\}$$
$$P_{F^{X=1}_{Z\to Y}}(\mathbb{F}) \geq \max\{0, P_{XY|Z}(11|0) + P_{XY|Z}(10|1) - 1\}$$

The hypotheses



$$P_{XY|Z}(00|0) + P_{XY|Z}(01|1) \leq 1$$
$$P_{XY|Z}(01|0) + P_{XY|Z}(00|1) \leq 1$$
$$P_{XY|Z}(10|0) + P_{XY|Z}(11|1) \leq 1$$
$$P_{XY|Z}(11|0) + P_{XY|Z}(10|1) \leq 1$$



The evidence

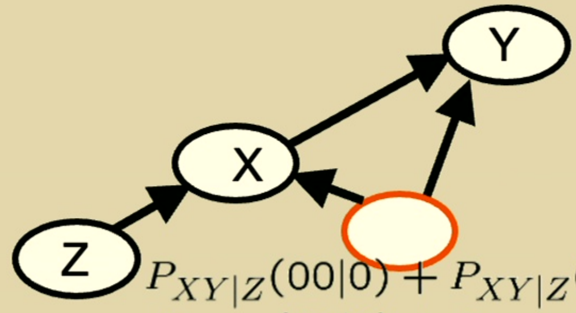
Violates one or more inequality constraint

$$P_{XY|Z}(00|0) + P_{XY|Z}(01|1) - 1 \geq 0$$
$$P_{XY|Z}(01|0) + P_{XY|Z}(00|1) - 1 \geq 0$$
$$P_{XY|Z}(10|0) + P_{XY|Z}(11|1) - 1 \geq 0$$
$$P_{XY|Z}(10|0) + P_{XY|Z}(11|1) - 1 \geq 0$$

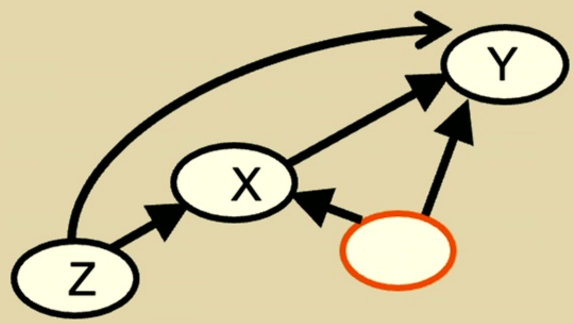
$$P_{F^{X=0}_{Z\to Y}}(\text{1bit}) \geq \max\{0, P_{XY|Z}(00|0) + P_{XY|Z}(01|1) - 1\} + \max\{0, P_{XY|Z}(01|0) + P_{XY|Z}(00|1) - 1\}$$

$$P_{F^{X=1}_{Z\to Y}}(\text{1bit}) \geq \max\{0, P_{XY|Z}(10|0) + P_{XY|Z}(11|1) - 1\} + \max\{0, P_{XY|Z}(11|0) + P_{XY|Z}(10|1) - 1\}$$

The hypotheses



$$P_{XY|Z}(00|0) + P_{XY|Z}(01|1) \leq 1$$
$$P_{XY|Z}(01|0) + P_{XY|Z}(00|1) \leq 1$$
$$P_{XY|Z}(10|0) + P_{XY|Z}(11|1) \leq 1$$
$$P_{XY|Z}(11|0) + P_{XY|Z}(10|1) \leq 1$$



Compatibility constraints for latent-free causal structures

Causal structure $\Rightarrow$ Constraint on observed probability distribution

Falsification of causal structure $\Leftarrow$ Violation of constraint on observed probability distribution

If X_1, X_2,. . . , X_n have causal relations described by a DAG G, then the set of compatible distributions are those of the form:

$$P_{X_1,X_2,\ldots,X_n} = \prod_i P_{X_i|\mathrm{Parents}_G(X_i)}$$

This is sometimes called the Markov condition

Local Markov condition

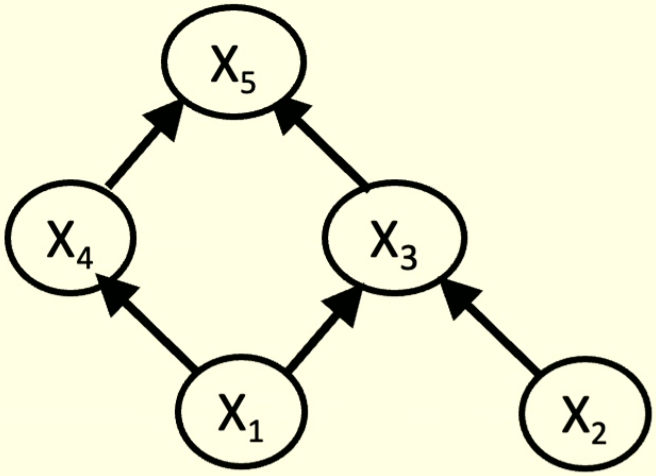
For a causal structure G, the compatible distributions are such that for every variable X

$$X \perp \mathrm{Nondescendents}_G(X) | \mathrm{Parents}_G(X)$$

$$P_{X_1,X_2,\ldots,X_n} = \prod_i P_{X_i|\mathrm{Parents}_G(X_i)}$$

$$\begin{aligned} P(X|\mathrm{Pa}(X),\mathrm{Nd}(X)) &= \frac{P(X,\mathrm{Pa}(X),\mathrm{Nd}(X))}{P(\mathrm{Pa}(X),\mathrm{Nd}(X))}, \\ &= \frac{P(X|\mathrm{Pa}(X))\prod_{Y\in\mathrm{Pa}(X),\mathrm{Nd}(X)} P(Y|\mathrm{Pa}(Y))}{\prod_{Y\in\mathrm{Pa}(X),\mathrm{Nd}(X)} P(Y|\mathrm{Pa}(Y))}, \\ &= P(X|\mathrm{Pa}(X)). \end{aligned}$$

$$X \perp \mathrm{Nondescendents}_G(X) | \mathrm{Parents}_G(X)$$



$$(X_1 \perp X_2)$$
$$(X_2 \perp \{X_1, X_4\})$$
$$(X_3 \perp X_4 \mid \{X_1, X_2\})$$
$$(X_4 \perp \{X_2, X_3\} \mid X_1)$$
$$(X_5 \perp \{X_1, X_2\} \mid \{X_3, X_4\})$$

The semi-graphoid axioms then imply

$$(X_4 \perp X_2 \mid X_1)$$
$$(\{X_4, X_5\} \perp X_2 \mid \{X_1, X_3\})$$
$$\ldots$$

X d-separated from **Y** by **Z**
in causal structure G

implies

$$\mathbf{X} \perp \mathbf{Y} | \mathbf{Z}$$

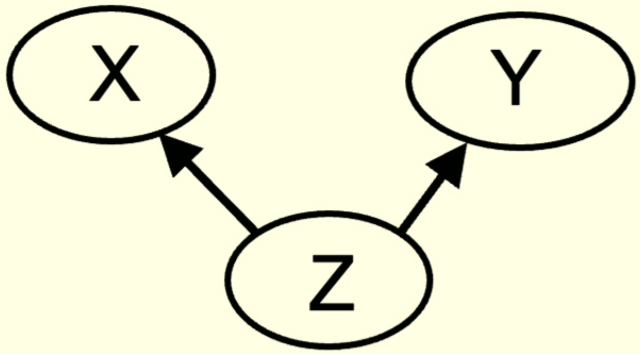
in every probability distribution
compatible with G

Definition (path blocking) A path between node X and node Y is blocked by a set of vertices **Z** if at least one of the following conditions holds:

1. The path contains a chain whose mediary node is in **Z**
2. The path contains a fork whose tail node is in **Z**

2. The path contains a collider whose head node is **not** in **Z** and no descendant of which is in **Z**.

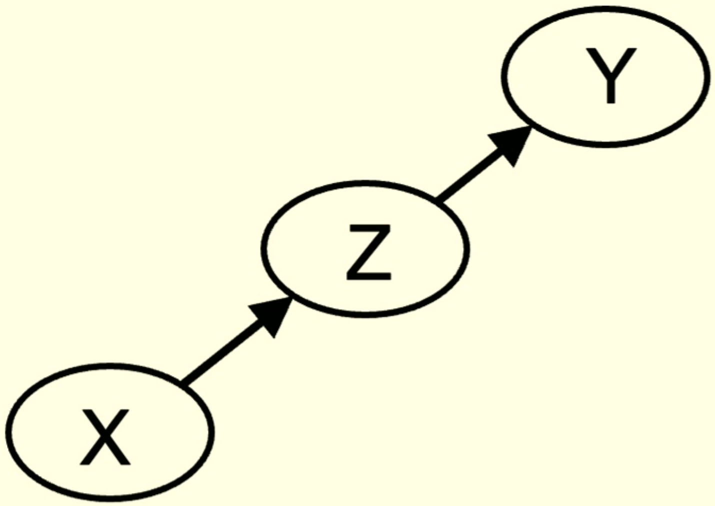
Definition (d-separation) Given a DAG G with vertices **V**, two set of vertices **X**, **Y** ∈ V are d-separated by a set of vertices **Z** ⊂ **V** if and only if for every pair of vertices, X and Y, from the sets **X** and **Y**, every path between X and Y is blocked.



Z is a complete common cause of X and Y, therefore

$$X \perp Y | Z$$

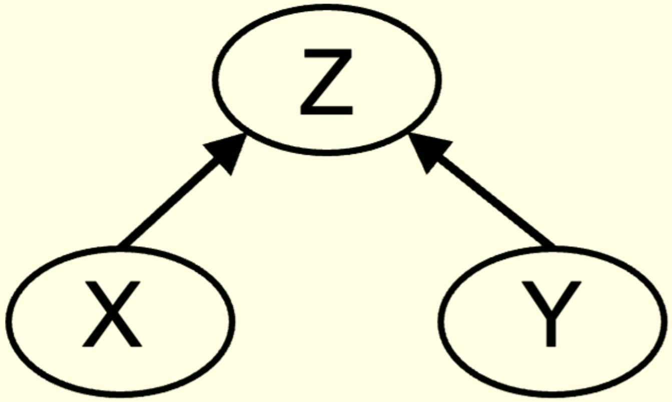
$$P_{XY|Z} = P_{X|Z} P_{Y|Z}$$



Z is the full set of parents of Y, while X is a nondescendent, therefore

$$X \perp Y | Z$$

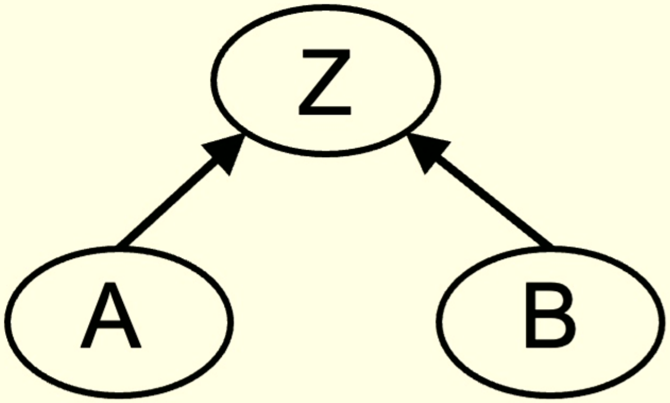
$$P_{XY|Z} = P_{X|Z} P_{Y|Z}$$



X and Y have no ancestors in common, therefore

X and Y are independent when one marginalizes over Z

$$X \perp Y$$

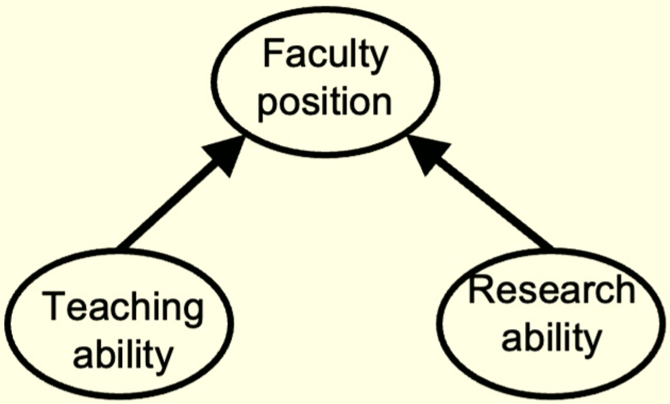


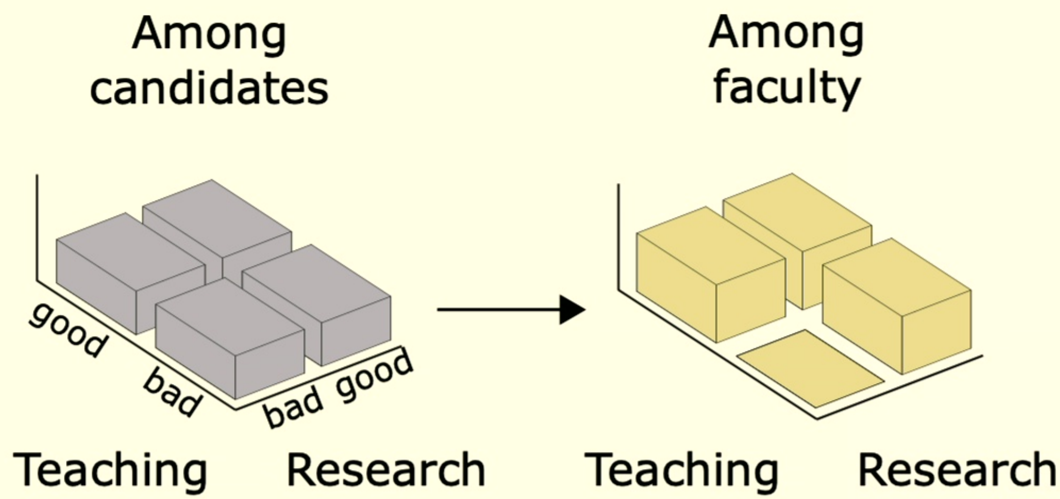
$$P_{Z|AB} = \delta_{Z,AB}$$
$$P_A = \tfrac{1}{2}[0]_A + \tfrac{1}{2}[1]_A$$
$$P_B = \tfrac{1}{2}[0]_B + \tfrac{1}{2}[1]_B$$

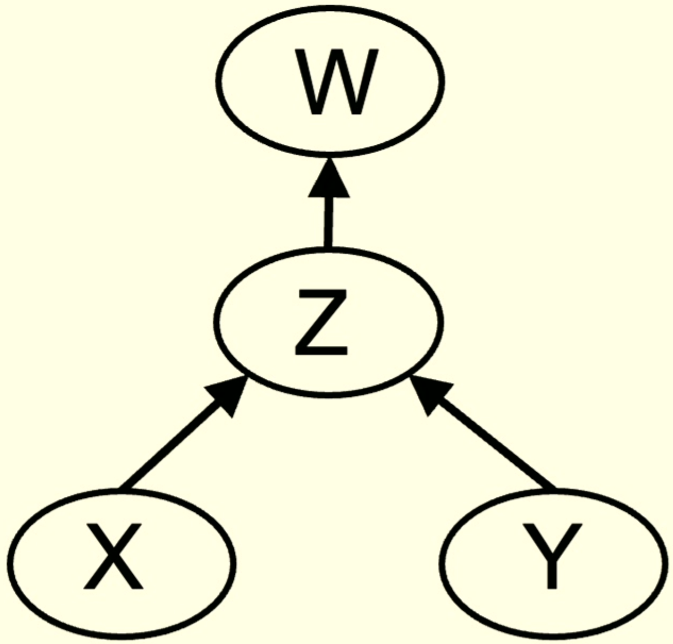
$$P_{AB} = (\tfrac{1}{2}[0]_A + \tfrac{1}{2}[1]_A)(\tfrac{1}{2}[0]_B + \tfrac{1}{2}[1]_B)$$
$$= \tfrac{1}{4}[0]_A[0]_B + \tfrac{1}{4}[0]_A[1]_B + \tfrac{1}{4}[1]_A 0]_B + \tfrac{1}{4}[1]_A[1]_B$$

$$P_{AB|Z} = P_{Z|AB} P_{AB} P_Z^{-1}$$

$$P_{AB|Z=0} = \tfrac{1}{3}[0]_A[0]_B + \tfrac{1}{3}[0]_A[1]_B + \tfrac{1}{3}[1]_A[0]_B$$


Faculty position
Teaching ability
Research ability


Among candidates
good
bad
bad
good
Teaching
Research
Among faculty
Teaching
Research

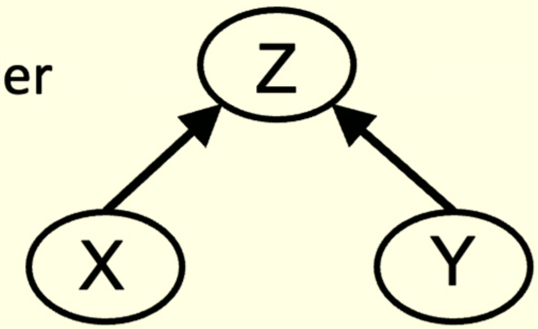


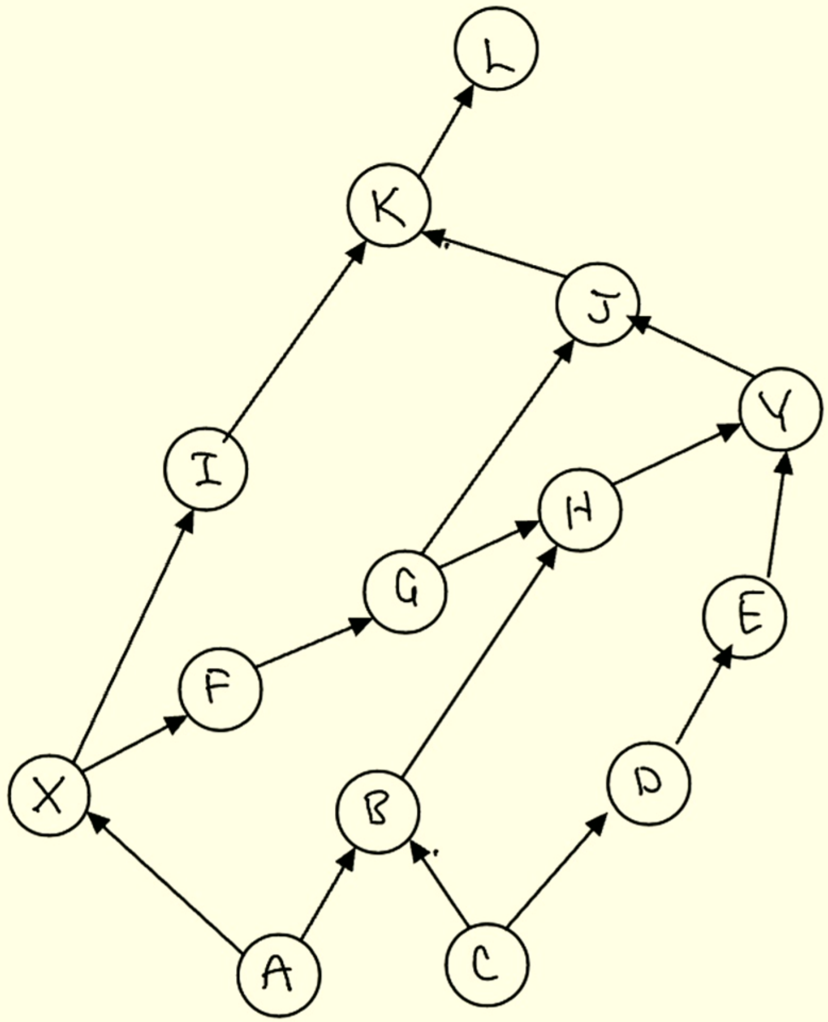
If W is a **descendent** of a common effect of X and Y, then

X and Y can become dependent when one conditions over W

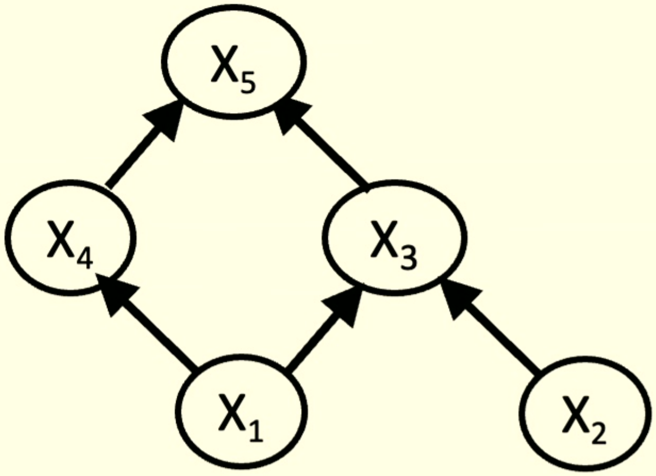
Note: Two sets of variables, **X** and **Y**, can be d-separated by the **empty set**

E.g. the tails of a collider




L
K
J
Y
I
H
G
E
F
X
B
D
A
C

If P has all the conditional independences that are implied by d-separation relations in a DAG G, then P is said to be Markov relative to G



$$\{X_4, X_5\} \perp X_2 | \{X_1, X_3\}$$

Arduous to derive this from the local Markov condition and applications of the semi-graphoid axioms

Follows from d-separation in a straightforward way