Causality

Jihoon Ko, Seungwoo Lee, Junho Han and Yongsu Baek

Lecture Objectives

1. To understand **what** is causality and **why** causality matters

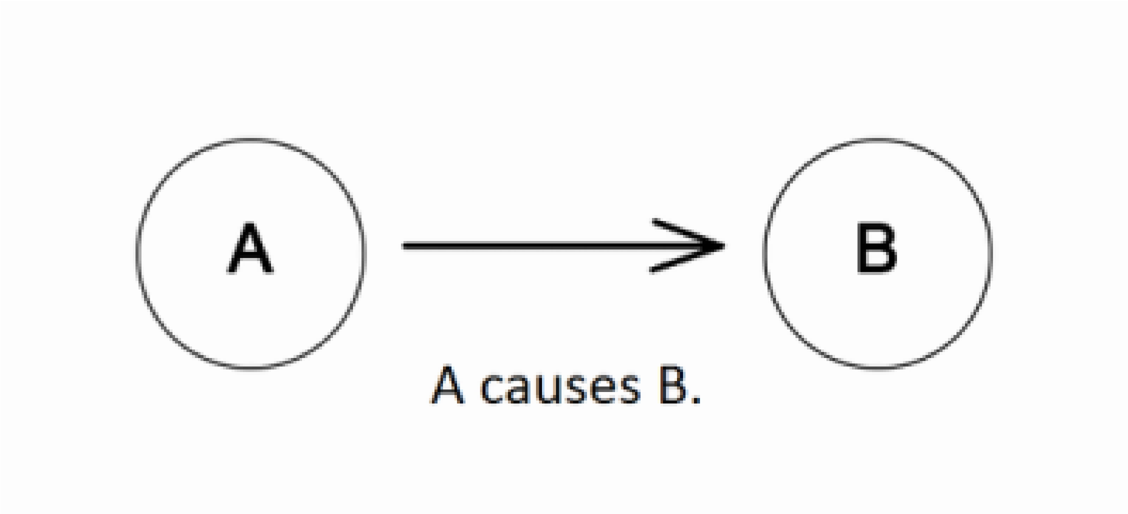
2. To explore major **categories** of causality and their details

3. To learn how to deal with causality in **mathematical languages**

4. To find out how the intuitive causal reasoning can conflict with the logic of probability and statistics in **Paradoxes**

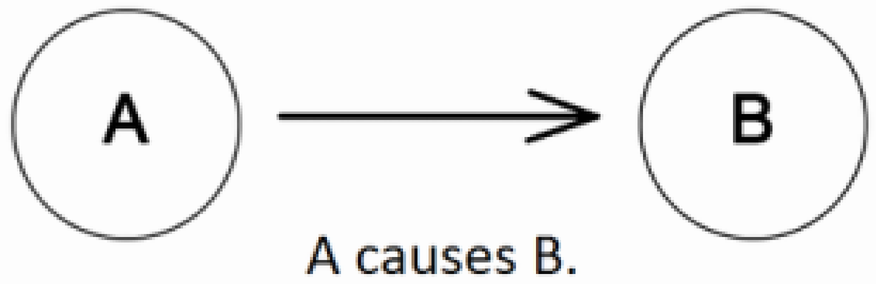
How can human get success on the Earth?

How can human get success on the Earth?



How can human get success on the Earth?

Causality



How can human get success on the Earth?

How can human get success on the Earth?

How can human get success on the Earth?

new plans, imaginations, predictions
or causal thinking

We always think with causality (Causal thinking)

We always think with causality (Causal thinking)

1. The most advanced tool for managing causality.

2. Our brains store and construct an incredible amount of **causal knowledge** supplemented by data.

3. We can use this to answer most pressing questions of our time, but other species and (current) robots can't.

4. What if we unlock **the logic behind our causal thinking**?

Some causality questions:

1. Did the new tax law *cause* our sales to go up, or *was it* our advertising campaign?
2. *How effective* is a given treatment in *preventing* a disease?
3. I'm about to quit my job. *Should I*?

“The society and our daily life constantly demand answers to causality questions.”

Yet science gave us no useful methodologies even to articulate them **in mathematical languages** until very recently!
It’s called **do-calculus**. We will discuss it later.

Statistics vs Causality as scientific tools

Statistics vs Causality as scientific tools

Statistics is a quite useful tool,
but it explains only parts of the whole nature.

Statistics vs Causality as scientific tools

Observation

X	Y	Z
1	2	4
2	3	6
3	4	8
7	8	16
9	10	20
12	13	26

Statistics vs Causality as scientific tools

Observation

X	Y	Z
1	2	4
2	3	6
3	4	8
7	8	16
9	10	20
12	13	26

Statistics can say that X+1=Y=Z/2, which is **correlation** of X, Y and Z.

Statistics vs Causality as scientific tools

Observation

X	Y	Z
1	2	4
2	3	6
3	4	8
7	8	16
9	10	20
12	13	26

Statistics can say that X+1=Y=Z/2, which is **correlation** of X, Y and Z.

Causality says that there can be two different **models** behind data.

Statistics vs Causality as scientific tools

Observation

X	Y	Z
1	2	4
2	3	6
3	4	8
7	8	16
9	10	20
12	13	26

Statistics can say that X+1=Y=Z/2, which is **correlation** of X, Y and Z.

Causality says that there can be two different **models** behind data.

Model 1.

Statistics vs Causality as scientific tools

Observation

X	Y	Z
1	2	4
2	3	6
3	4	8
7	8	16
9	10	20
12	13	26

Statistics can say that X+1=Y=Z/2, which is **correlation** of X, Y and Z.

Causality says that there can be two different **models** behind data.

Model 1.

Cause → Effect

Statistics vs Causality as scientific tools

Observation

X	Y	Z
1	2	4
2	3	6
3	4	8
7	8	16
9	10	20
12	13	26

Statistics can say that X+1=Y=Z/2, which is **correlation** of X, Y and Z.

Causality says that there can be two different **models** behind data.

Model 1.

Cause → Effect

Model 2.

Effect ← Cause

Statistics vs Causality as scientific tools

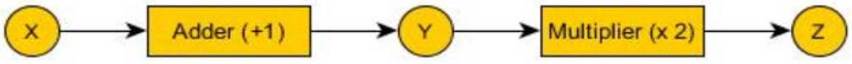
Observation

X	Y	Z
1	2	4
2	3	6
3	4	8
7	8	16
9	10	20
12	13	26

True model behind data?

S … **correlation** of X, Y and Z.

Causality says that there can be two different **models** behind data.

Model 1. 

Cause → Effect

Model 2. 

Effect ← Cause

Statistics vs Causality as scientific tools

Observation

X	Y	Z
7	8	16
9	10	20
12	13	26

S **correlation** of X, Y and Z.

True model behind data?

Causality says that there can be two different **models** behind data.

It cannot explain cause and effect relationship.
So statistics is a model-blind method.

Model 2. 

Effect ← Cause

Statistics vs Causality as scientific tools

X=3 Y=4 Z=8 ← Observation

Model 1.

Cause → Effect

Model 2.

Effect ← Cause

Statistics vs Causality as scientific tools

What will happen to X if we **change** Y?

X=? Y=7 Artificially changed environment

X=3 Y=4 Z=8 Observation

Model 1.

Cause Effect

Model 2.

Effect Cause

Statistics vs Causality as scientific tools

What will happen to X if we **change** Y?

X=? Y=7 ← Artificially changed environment

X=3 Y=4 Z=8 ← Observation

Model 1.

Cause → Effect

Model 2.

Effect ← Cause

Statistics vs Causality as scientific tools

What will happen to X if we **change** Y?

X=? Y=7 ← Artificially changed environment

X=3 Y=4 Z=8 ← Observation

Model 1.

Cause → Effect

Model 2.

Effect ← Cause

Statistics vs Causality as scientific tools

What will happen to X if we **change** Y?

X still remains 3 !! X=? Y=7 ← Artificially changed environment

X=3 Y=4 Z=8 ← Observation

Model 1.

Cause → Effect

Model 2.

Effect ← Cause

Statistics vs Causality as scientific tools

What will happen to X if we **change** Y?

Statistics cannot answer to this question because it is a model-blind method.

Model 1.

Cause → Effect

Model 2.

Effect ← Cause

Statistics vs Causality as scientific tools

What will happen to X if we **change** Y?

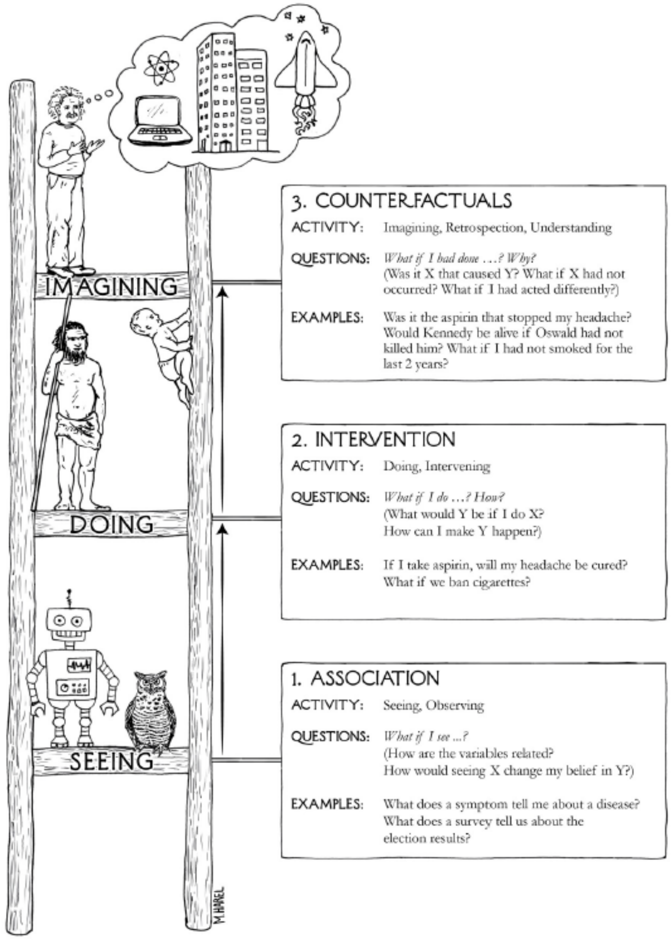
1. This kind of questions cannot be resolved by statistics.

2. Causality considers model unlike statistics, so it can handle this question.

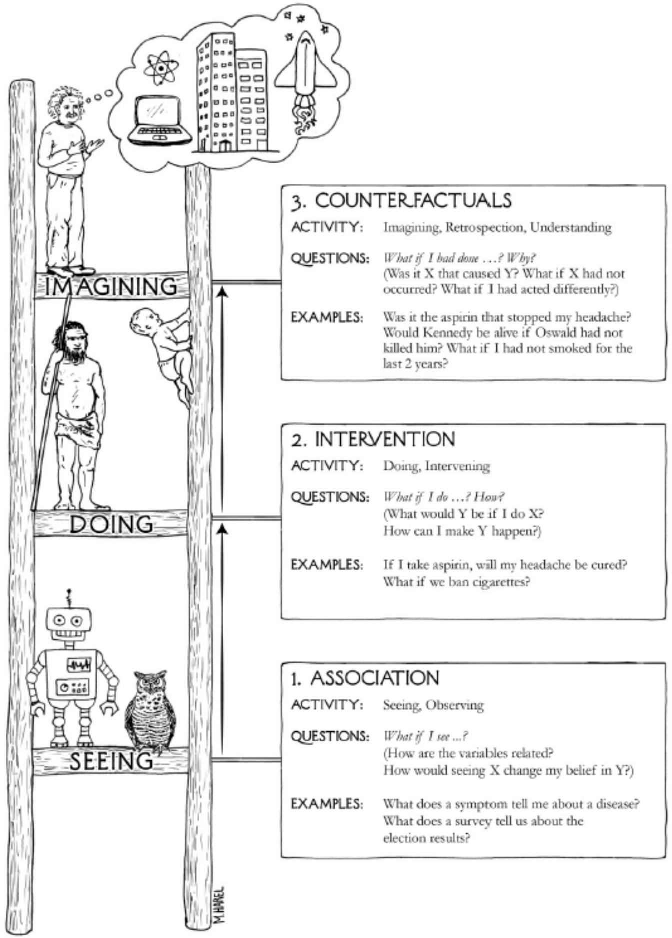
3. However, figuring out the model behind data is a difficult research area, *which is not a goal of our lecture*. (Finding exact model may be impossible.)

4. Big data companies such as *Facebook* and *Youtube* know this, so they not only collect a lot of observations but also **constantly perform experiment** by changing environment.

3-steps Ladder of Causality



3-steps Ladder of Causality

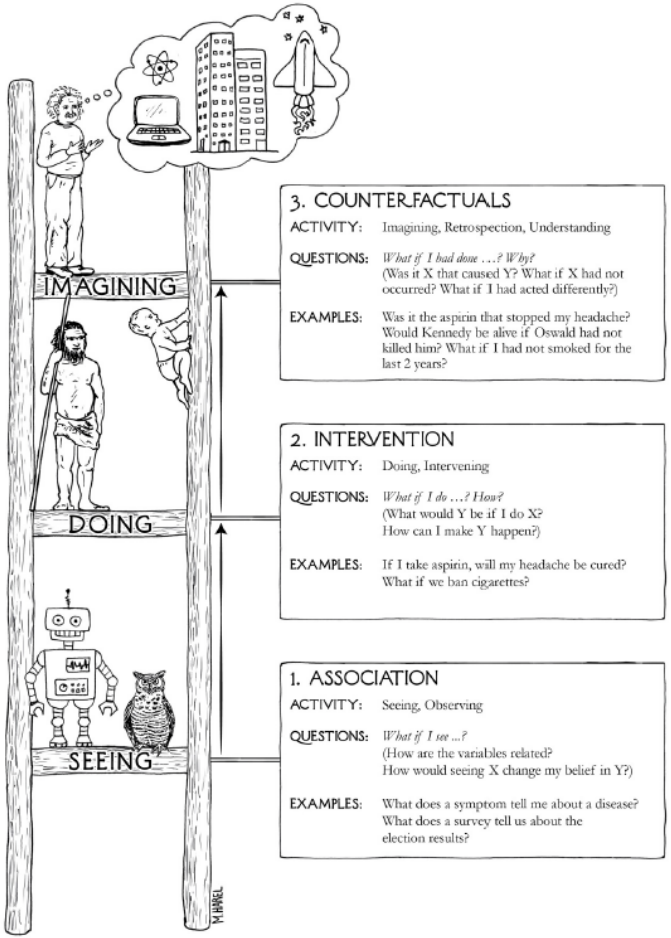


1. Association(=statistics)
   - looking for regularities in **observations**, which is exactly same as what statistics does.

3-steps Ladder of Causality

2. Intervention (=change environment/not just observe)
   - imagining what will happen if we **intervene, do or fix** some factors which doesn't occur yet.

1. Association(=statistics)
   - looking for regularities in **observations**, which is exactly same as what statistics does.

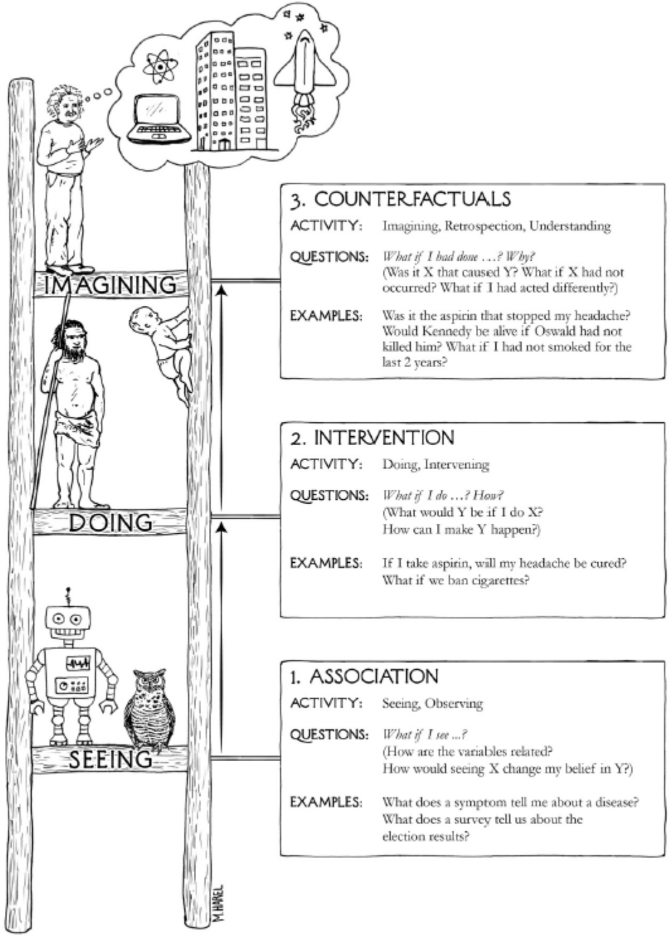


3-steps Ladder of Causality

3. Counterfactuals (=counter to facts)
   - imagining what will happen if we **negate** the past observed factors which already occurred.

2. Intervention (=change environment/not just observe)
   - imagining what will happen if we **intervene, do or fix** some factors which doesn't occur yet.

1. Association(=statistics)
   - looking for regularities in **observations**, which is exactly same as what statistics does.



3-steps Ladder of Causality (examples)

Level (Symbol)	Typical Activity	Typical Questions	Examples
1. Association $P(y\|x)$	Seeing	What is? How would seeing X change my belief in$Y$?	What does a symptom tell me about a disease? What does a survey tell us about the election results?
2. Intervention $P(y\|do(x), z)$	Doing Intervening	What if? What if I do X?	What if I take aspirin, will my headache be cured? What if we ban cigarettes?
3. Counterfactuals $P(y_x\|x', y')$	Imagining, Retrospection	Why? Was it X that caused Y? What if I had acted differently?	Was it the aspirin that stopped my headache? Would Kennedy be alive had Oswald not shot him? What if I had not been smoking the past 2 years?

Fig. 1. The Causal Hierarchy. Questions at level i can only be answered if information from level i or higher is available.

The Seven Tools of Causal Inference with Reflections on Machine Learning, Judea Pearl

3-steps Ladder of Causality (examples)

Level (Symbol)	Typical Activity	Typical Questions	Examples
1. Association $P(y\|x)$	Seeing	What is? How would seeing X change my belief in$Y$?	What does a symptom tell me about a disease? What does a survey tell us about the election results?
2. Intervention $P(y\|do(x), z)$	Doing Intervening	What if? What if I do X?	What if I take aspirin, will my headache be cured? What if we ban cigarettes?
3. Counterfactuals $P(y_x\|x', y')$	Imagining, Retrospection	Why? Was it X that caused Y? What if I had acted differently?	Was it the aspirin that stopped my headache? Would Kennedy be alive had Oswald not shot him? What if I had not been smoking the past 2 years?

Fig. 1. The Causal Hierarchy. Questions at level i can only be answered if information from level i or higher is available.

The Seven Tools of Causal Inference with Reflections on Machine Learning, Judea Pearl

3-steps Ladder of Causality (examples)

Level (Symbol)	Typical Activity	Typical Questions	Examples
1. Association $P(y\|x)$	Seeing	What is? How would seeing X change my belief in$Y$?	What does a symptom tell me about a disease? What does a survey tell us about the election results?
2. Intervention $P(y\|do(x), z)$	Doing Intervening	What if? What if I do X?	What if I take aspirin, will my headache be cured? What if we ban cigarettes?
3. Counterfactuals $P(y_x\|x', y')$	Imagining, Retrospection	Why? Was it X that caused Y? What if I had acted differently?	Was it the aspirin that stopped my headache? Would Kennedy be alive had Oswald not shot him? What if I had not been smoking the past 2 years?

Fig. 1. The Causal Hierarchy. Questions at level i can only be answered if information from level i or higher is available.

The Seven Tools of Causal Inference with Reflections on Machine Learning, Judea Pearl

3-steps Ladder of Causality (examples)

Level (Symbol)	Typical Activity	Typical Questions	Examples
1. Association $P(y\|x)$	Seeing	What is? How would seeing X change my belief in Y?	What does a symptom tell me about a disease? What does a survey tell us about the election results?
2. Intervention $P(y\|do(x), z)$	Doing Intervening	What if? What if I do X?	What if I take aspirin, will my headache be cured? What if we ban cigarettes?
3. Counterfactuals $P(y_x\|x', y')$	Imagining, Retrospection	Why? Was it X that caused Y? What if I had acted differently?	Was it the aspirin that stopped my headache? Would Kennedy be alive had Oswald not shot him? What if I had not been smoking the past 2 years?

Fig. 1. The Causal Hierarchy. Questions at level i can only be answered if information from level i or higher is available.

The Seven Tools of Causal Inference with Reflections on Machine Learning, Judea Pearl

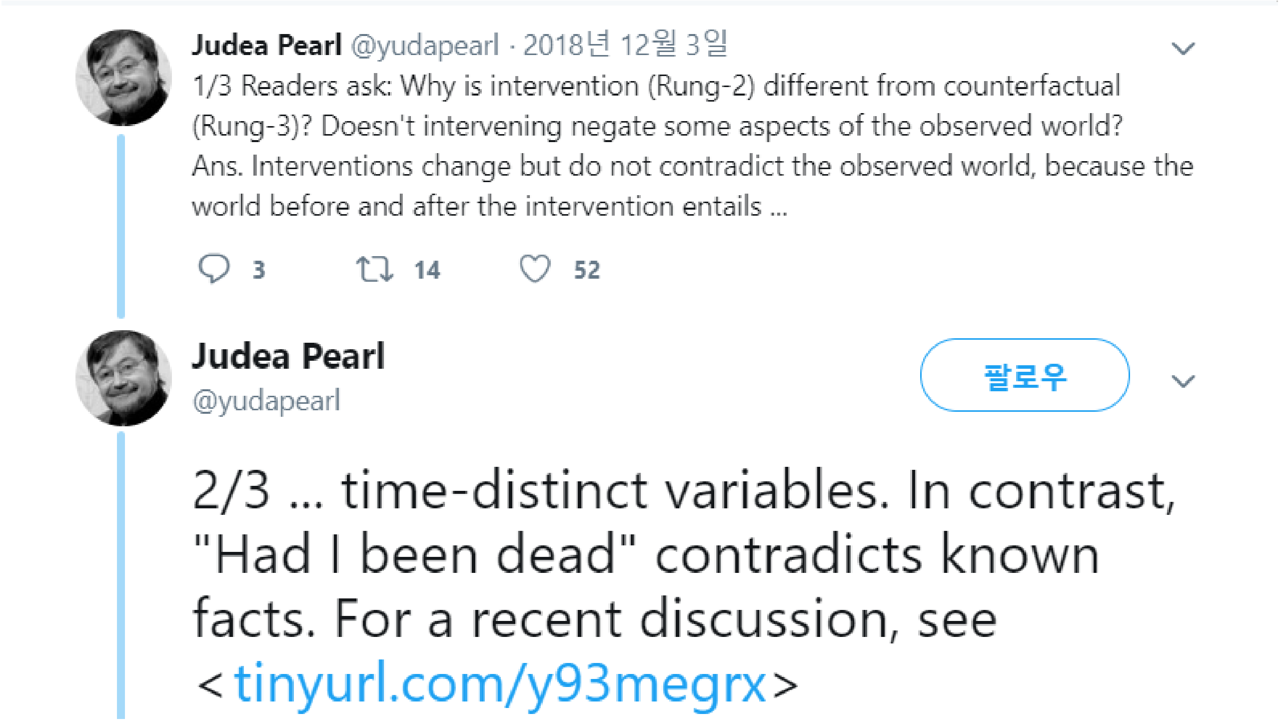
3-steps Ladder of Causality (examples)

Level (Symbol)	Typical Activity	Typical Questions	Examples
1. Association $P(y\|x)$	Seeing	What is? How would seeing X change my belief in Y?	What does a symptom tell me about a disease? What does a survey tell us about the election results?
2. Intervention $P(y\|do(x), z)$	Doing Intervening	What if? What if I do X?	What if I take aspirin, will my headache be cured? What if we ban cigarettes?
3. Counterfactuals $P(y_x\|x', y')$	Imagining, Retrospection	Why? Was it X that caused Y? What if I had acted differently?	Was it the aspirin that stopped my headache? Would Kennedy be alive had Oswald not shot him? What if I had not been smoking the past 2 years?

Fig. 1. The Causal Hierarchy. Questions at level i can only be answered if information from level i or higher is available.

The Seven Tools of Causal Inference with Reflections on Machine Learning, Judea Pearl

difference: intervention and counterfactuals



Judea Pearl @yudapearl · 2018년 12월 3일

1/3 Readers ask: Why is intervention (Rung-2) different from counterfactual (Rung-3)? Doesn't intervening negate some aspects of the observed world? Ans. Interventions change but do not contradict the observed world, because the world before and after the intervention entails ...

3 14 52

Judea Pearl
@yudapearl

팔로우

2/3 ... time-distinct variables. In contrast, "Had I been dead" contradicts known facts. For a recent discussion, see <tinyurl.com/y93megrx>

difference: intervention and counterfactuals

Question: *Given that Hilary Clinton* ***did not win the 2016 presidential election****, and given that she* ***did not visit Michigan 3 days before the election****, and given* ***everything else we know about the circumstances of the election****, what can we say about the probability of Hilary Clinton winning the election, had she visited Michigan 3 days before the election?*

difference: intervention and counterfactuals

Answer: *probability that she* ***hypothetically*** *wins the election*

- she lost the election
- she did not visit Michigan
- any other relevant an observable facts
- she ***hypothetically*** visits Michigan

difference: intervention and counterfactuals

$$Intervention : p(y|do(x))$$

$$Counterfactual : p(y'|do(x'))?$$

difference: intervention and counterfactuals

$$Intervention : p(y|do(x))$$

difference: intervention and counterfactuals

$$Intervention : p(y|do(x))$$

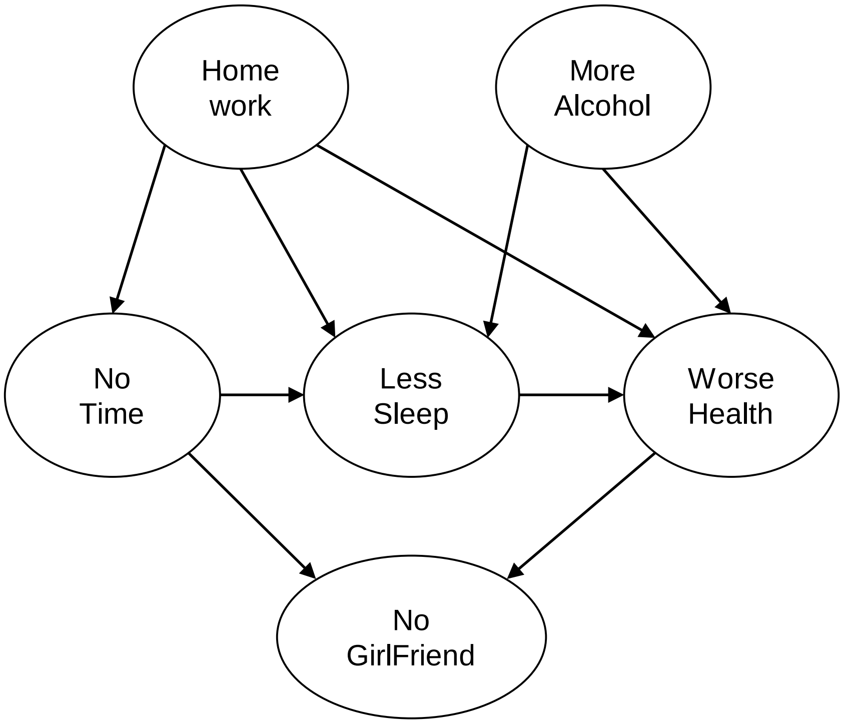
$$Counterfactual : p(y'|x, y, do(x')))$$

Theoretical Approach to Causality

- Causal diagram
- Effect of Observation & Intervention
- D-separation
- Do-calculus
- Examples

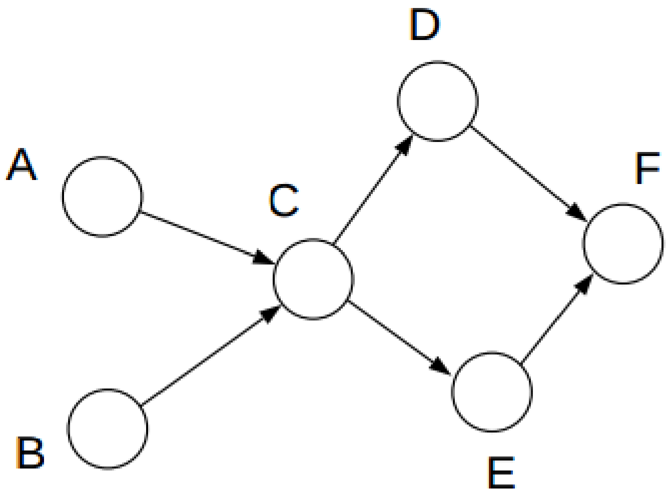
Causal Diagram

- It is Directed Acyclic Graph(DAG)
- Vertex represents each feature(factors)
- Edge represents Cause-Effect relation among factors
- How to formulate mathematically?



Probability Graphical Model

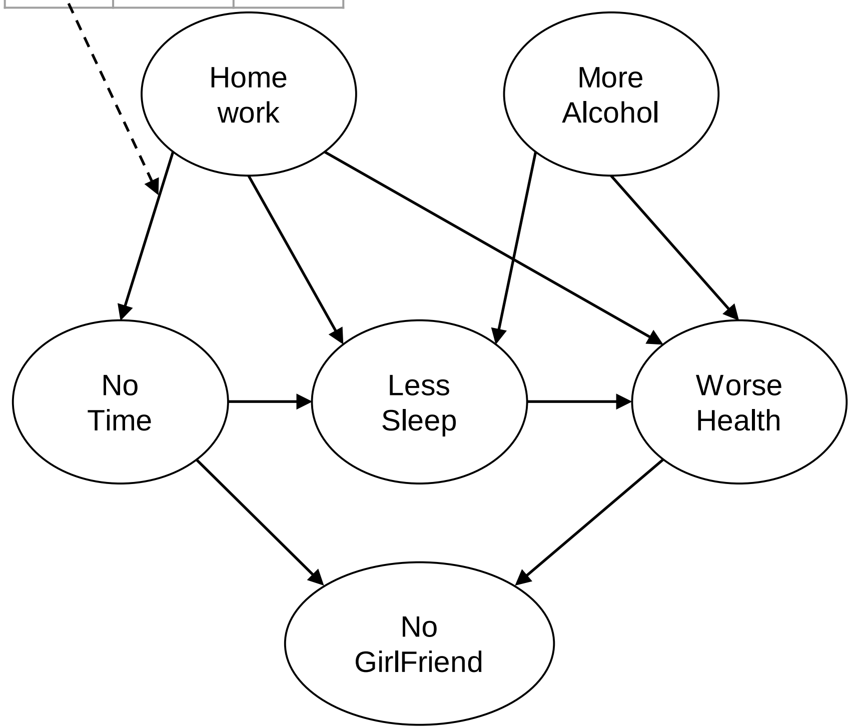
- Each Vertex represents random variable.
- Edge represents ***conditional dependency*** between two random variable
- If there is no edge between two random variable, then they are ***conditionally independent***(p(A|B) = p(A))



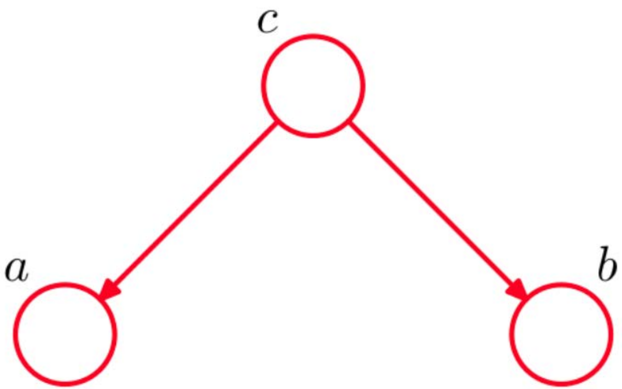
Causal Diagram +PGM

- Combining PGM and Causal Diagram can formulate causality problem.
- However, ***observation*** might affect Independence relations

	No HW	HW
No time	0.1	0.99
time	0.9	0.1

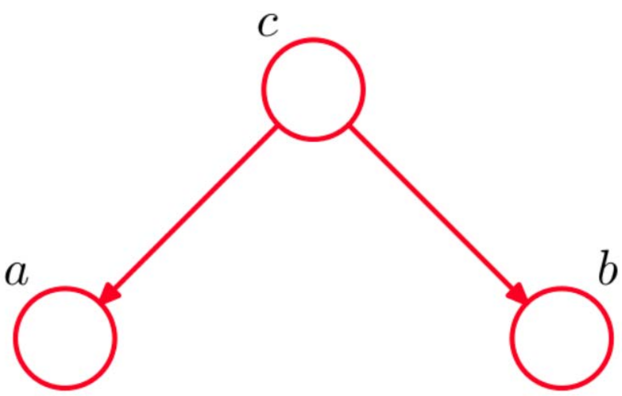


Observation change dependence - (1)



Q. In this case, a and b are independent?

Observation change dependence - (1)

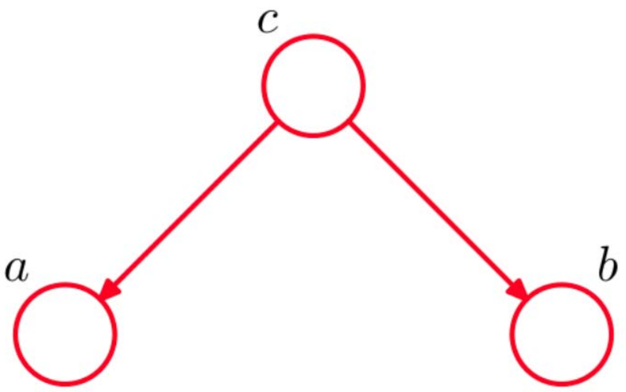


Q. In this case, a and b are independent?

A. No

$$
\begin{aligned}
p(a,b) &= \sum_c p(a,b,c) = \sum_c p(a|b,c)p(b,c) \\
&= \sum_c p(a|b,c)p(b|c)p(c) = \sum_c p(a|c)p(b|c)p(c))
\end{aligned}
$$

Observation change dependence - (1)

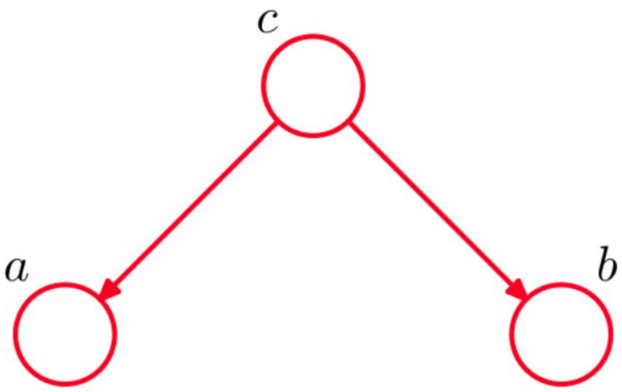


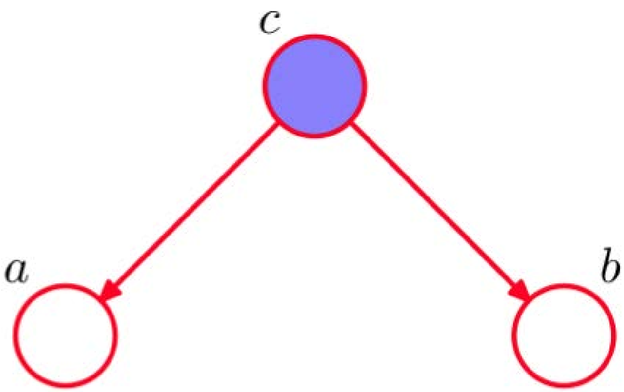
Q. In this case, a and b are independent?

A. No

$$p(a,b) = \sum_c p(a,b,c) = \sum_c p(a|b,c)p(b,c)$$
$$= \sum_c \boxed{p(a|b,c)}p(b|c)p(c) = \sum_c \boxed{p(a|c)}p(b|c)p(c))$$

Observation change dependence - (1)





Q. In this case, a and b are independent?

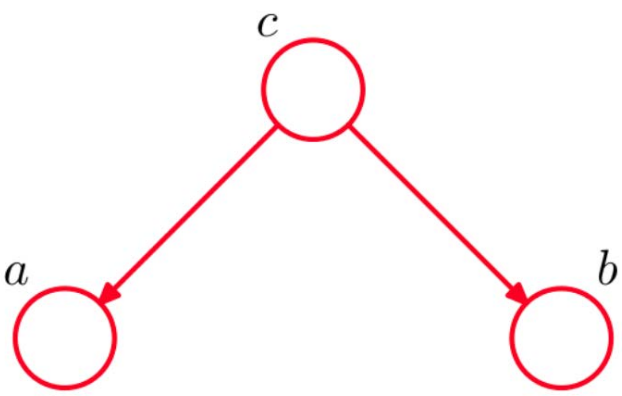
Q. If we observe c, are a and b are independent?

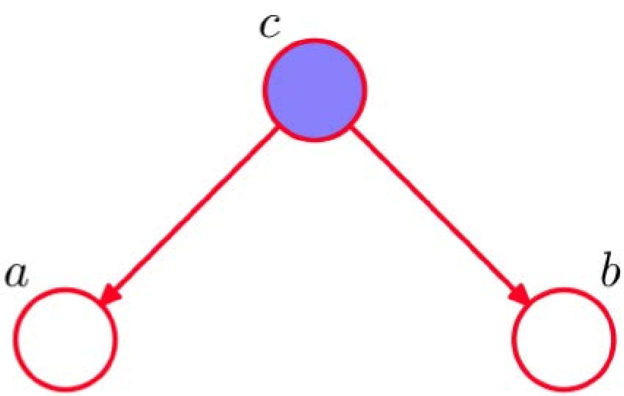
A. No

$$p(a,b) = \sum_c p(a,b,c) = \sum_c p(a|b,c)p(b,c)$$

$$= \sum_c \boxed{p(a|b,c)} p(b|c)p(c) = \sum_c \boxed{p(a|c)} p(b|c)p(c))$$

Observation change dependence - (1)





Q. In this case, a and b are independent?

A. No

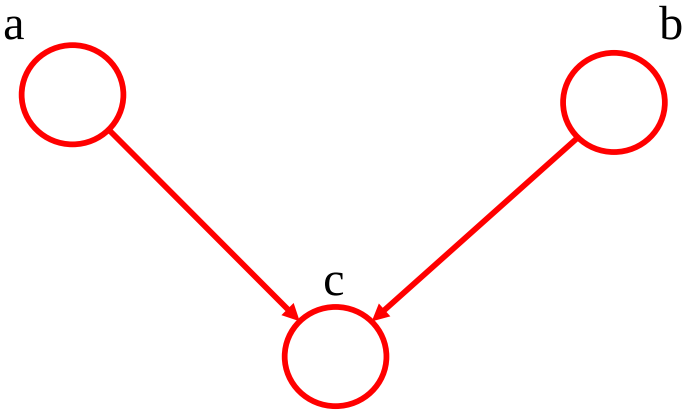
$$p(a,b) = \sum_c p(a,b,c) = \sum_c p(a|b,c)p(b,c)$$
$$= \sum_c \boxed{p(a|b,c)}p(b|c)p(c) = \sum_c \boxed{p(a|c)}p(b|c)p(c))$$

Q. If we observe c, are a and b are independent?

A. Yes

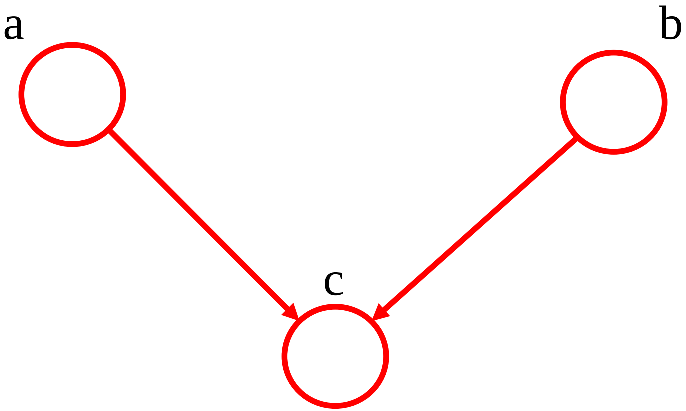
$$\begin{aligned} p(a,b|c) &= \frac{p(a,b,c)}{p(c)} \\ &= p(a|c)p(b|c) \end{aligned}$$

Observation change dependence - (2)



Q. In this case, a and b are independent?

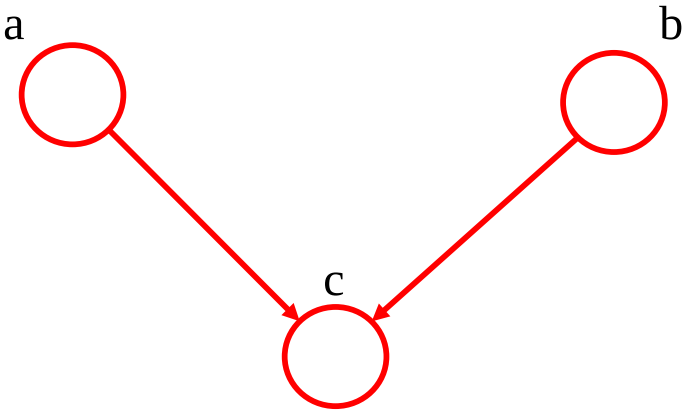
Observation change dependence - (2)

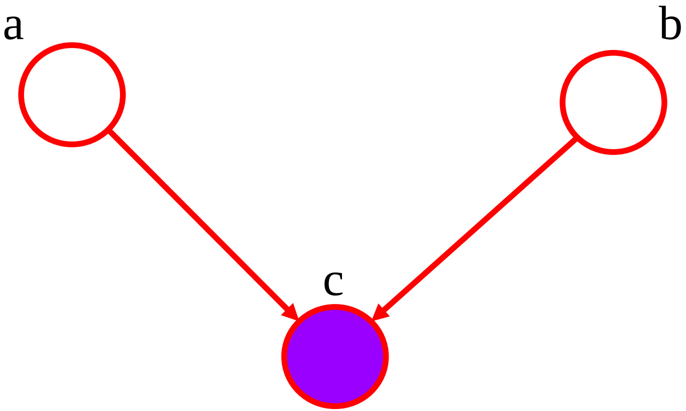


Q. In this case, a and b are independent?

A. Yes $p(a,b) = p(a)p(b)$

Observation change dependence - (2)



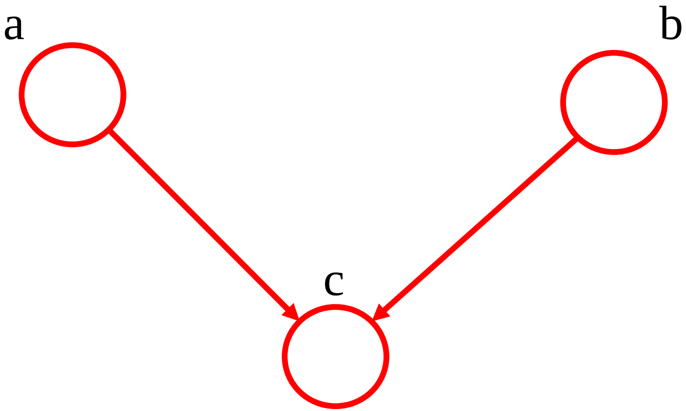


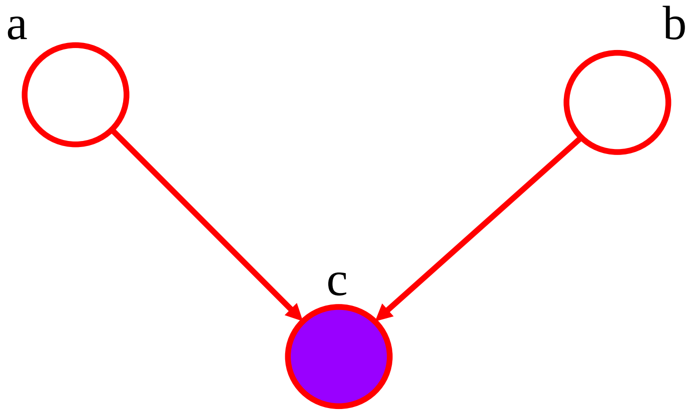
Q. In this case, a and b are independent?

Q. If we observe c, are a and b are independent?

A. Yes $p(a, b) = p(a)p(b)$

Observation change dependence - (2)





Q. In this case, a and b are independent?

A. Yes $p(a,b) = p(a)p(b)$

Q. If we observe c, are a and b are independent?

A. No

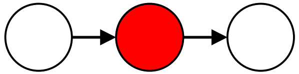
$$p(a,b,c) = p(c|a,b)p(a)p(b)$$

$$p(a,b|c) = \frac{p(a,b,c))}{p(c))} = \frac{p(c|a,b)p(a)p(b)}{p(c))}$$

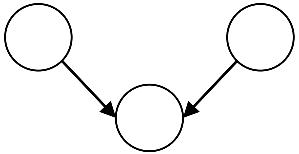
D-separation

‘***d-separation***’ is a criterion for deciding, from a causal diagram, whether a set A of variables is independent of another set B given a third set C, notated as $A \perp\!\!\!\perp B | C$

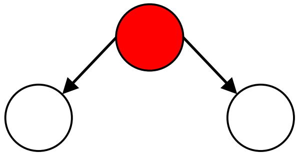
Examples :



Causal/Evidential ‘Chain’



‘Collider’

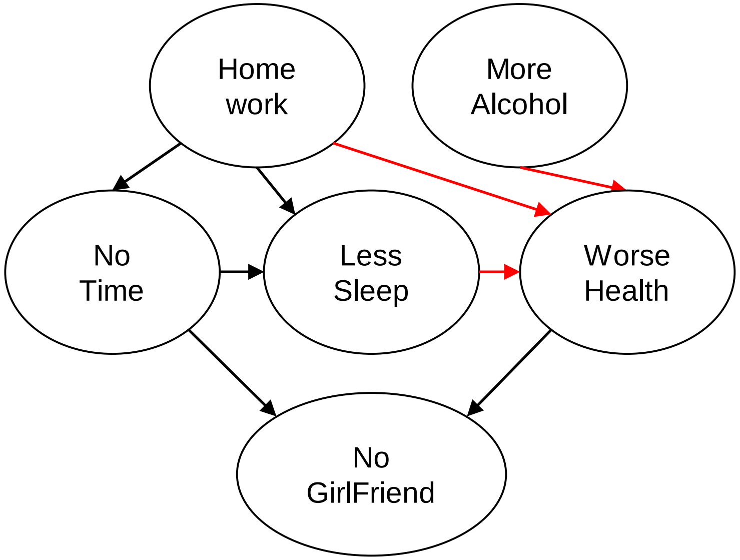


‘Confounder’

Using d-separation, we can predict the effect of ***observation***

Q. How about ***intervention***?

Intervention in causal diagram



Hypothesis :

If I'm *Healthier*, then Can I make Girlfriend?

Problem :

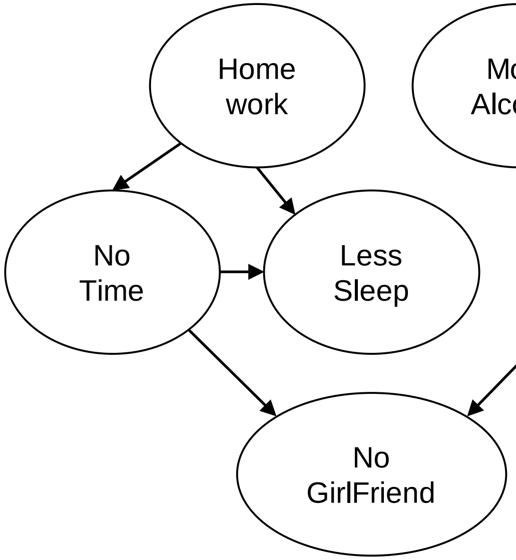
We ***cannot*** just fix a value in causal diagram.

Effect variable affects the cause variable.

ex) : Better health means there is fewer HW, more sleep, and Low alcohol consumption.

Q. How do we ***remove*** the effect on ancestor variables?

Intervention in causal diagram

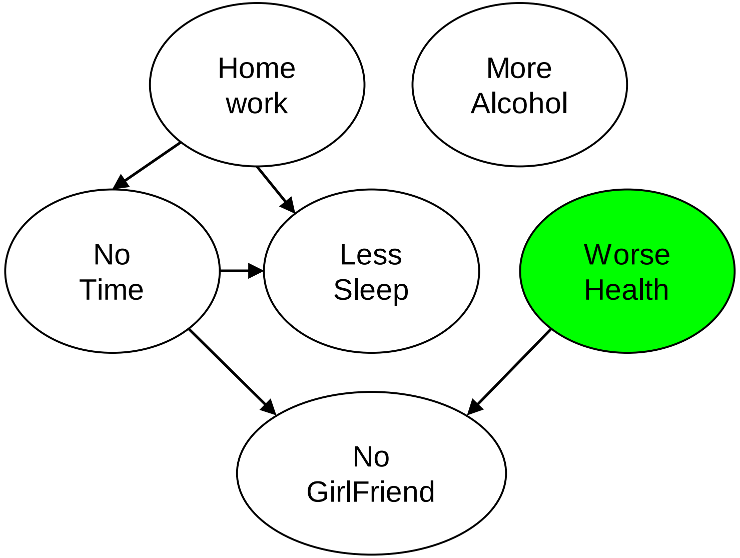


Solution :

By removing the in-edge of intervention node, we can make variables ***independent*** to intervention value

ex) Now Sleep, Homework, Alcohol are independent to health. So, we can think it as observation

Intervention in causal diagram



Solution :
By removing the in-edge of intervention node, we can make variables ***independent*** to intervention value

ex) Now Sleep, Homework, Alcohol are independent to health. So, we can think it as observation

Result :
Making Girlfriend still depends on *Homework*

Do-Calculus

- Calculus to discuss causality in a formal language by Judea Pearl
- A new operator, do(), marks an action or an ***intervention*** in the model.
- Example: $p(y|do(x))$ instead of $p(y|x)$

Main goal: *to generate probabilistic formulas for the effect of interventions in terms of the observed probabilities.*

Notation for do-calculus

- $G_{\overline{X}}$ denotes the perturbed graph in which all edges pointing to X have been deleted
- $G_{\underline{X}}$ denotes the perturbed graph in which all edges pointing from X have been deleted.
- Z(W) denote the set of nodes in Z which are not ancestors of W

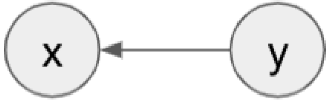
Pearl's 3 rules

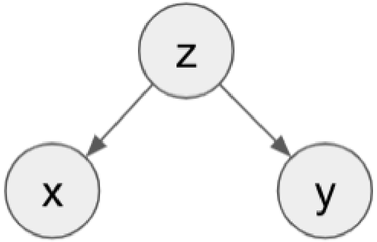
1. Ignoring Observations
$$p(y|do(x), z, w) = p(y|do(x), w) \quad if (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}}}$$

1. Action/Observation Exchange (the back-door criterion)
$$p(y|do(x), do(z), w) = p(y|do(x), z, w) \quad if (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X},\underline{Z}}}$$

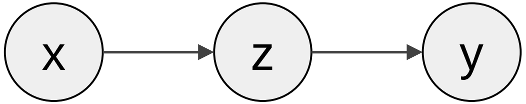
1. Ignoring Actions/Interventions
$$p(y|do(x), do(z), w) = p(y|do(x), w) \quad if (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X},\overline{Z(W)}}}$$

Do-Calculus: Simple Example



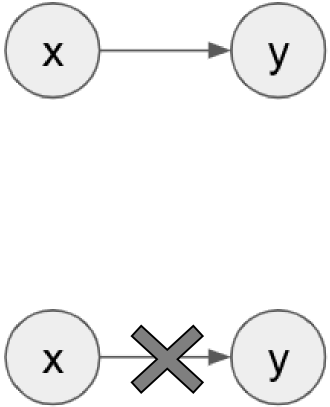


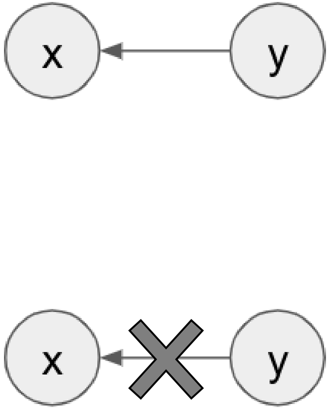
$$p(y|do(x)) = ?$$

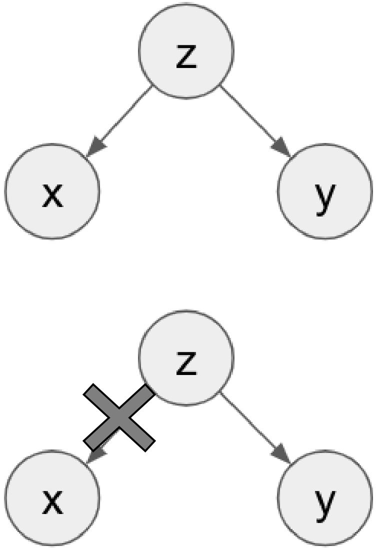


$$p(y|do(z), x) = ?$$

Do-Calculus: Simple Example



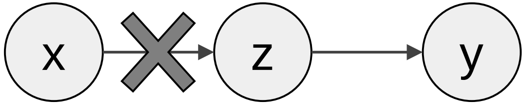




$$p(y|do(x)) = p(y|x)$$

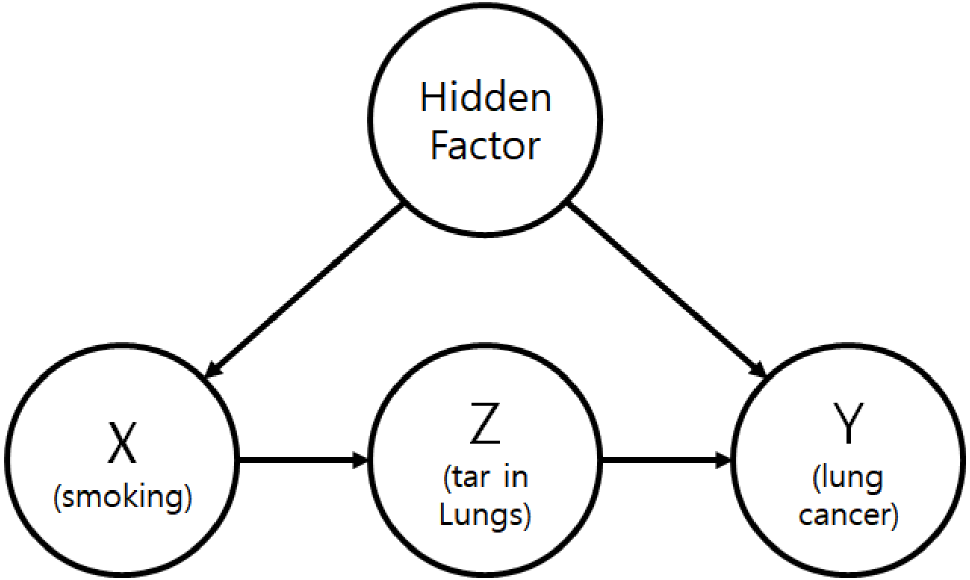
$$p(y|do(x)) = p(y)$$

Do-Calculus: Simple Example

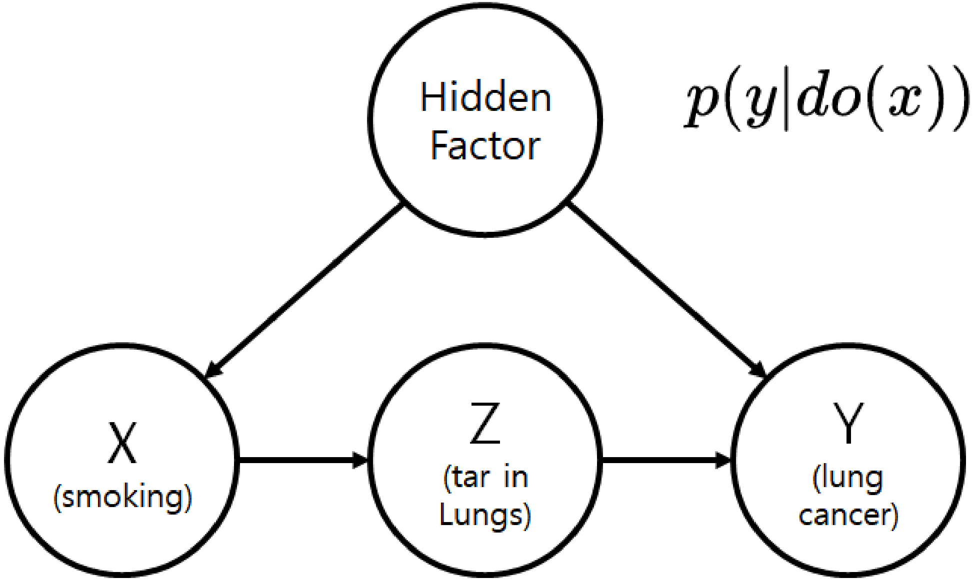


$$p(y|do(z),x) = p(y|do(z))$$
$$= p(y|z)$$

Do-Calculus: Example

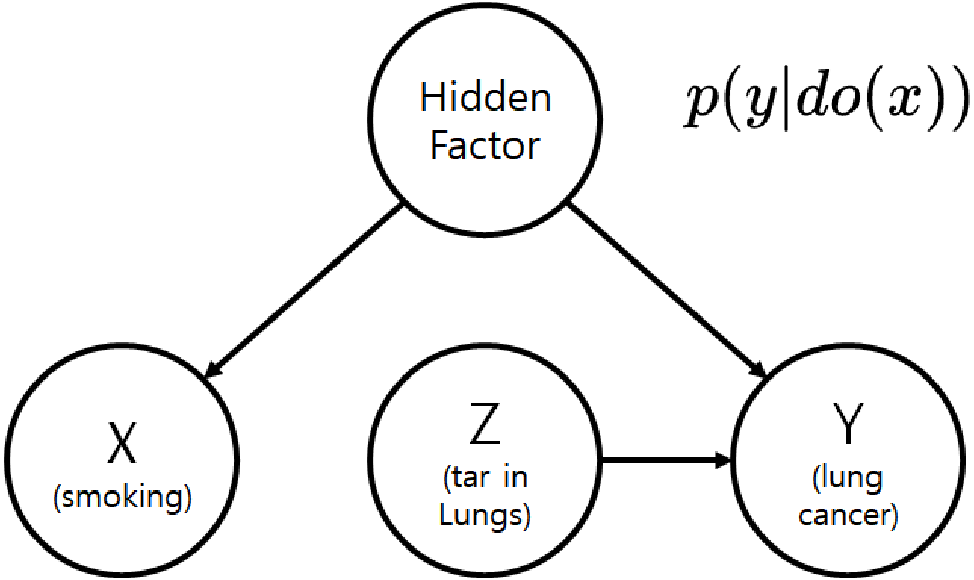


$$p(y|do(x)) =?$$



$$p(y|do(x)) = \sum_z p(y|z, do(x))p(z|do(x))$$

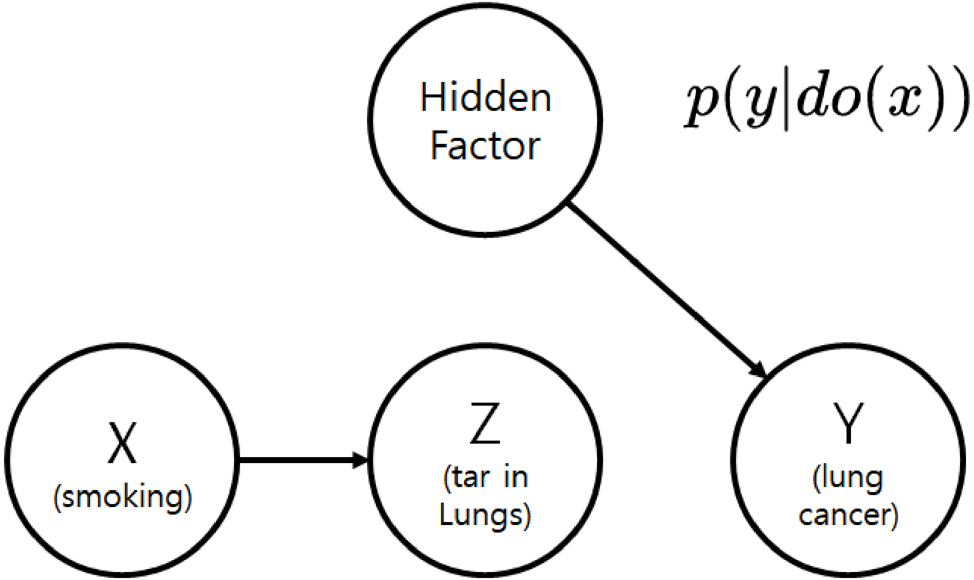
G



$G_{\underline{X}}$

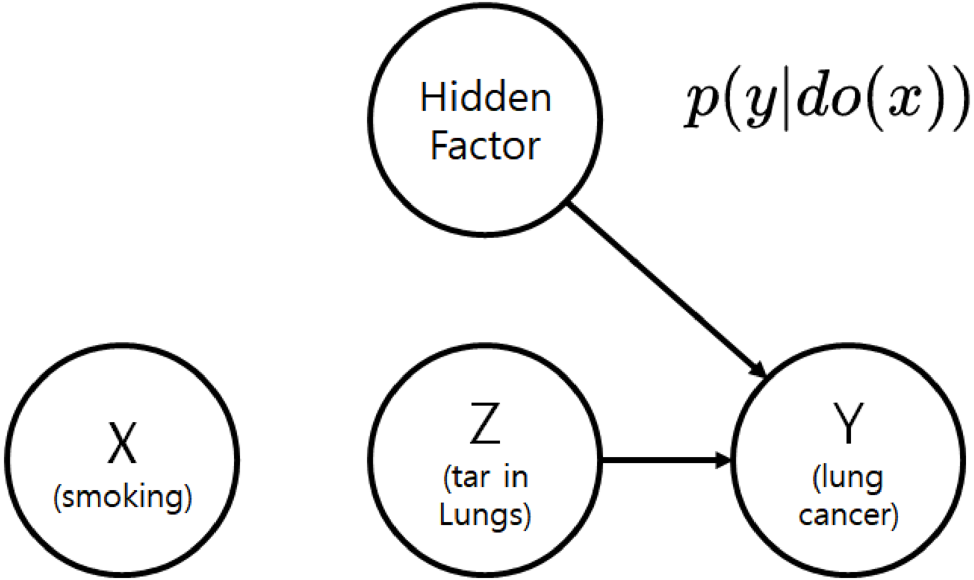
$$p(y|do(x)) = \sum_z p(y|z, do(x))p(z|do(x))$$
$$= \sum_z p(y|z, do(x))p(z|x)$$

$(\because (Z \perp\!\!\!\perp X)_{G_{\underline{X}}})$ (Use rule 2)



$G_{\overline{X},\underline{Z}}$

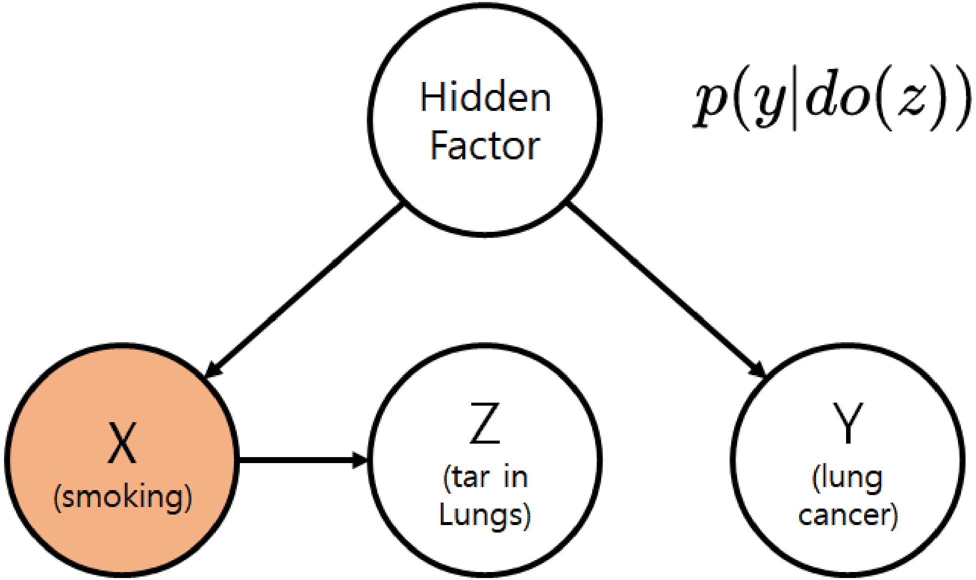
$$p(y|do(x)) = \sum_z p(y|z, do(x))p(z|do(x))$$
$$= \sum_z p(y|z, do(x))p(z|x)$$
$$= \sum_z p(y|do(z), do(x))p(z|x)$$
$$(\because (Y \perp\!\!\!\perp Z|X)_{G_{\overline{X},\underline{Z}}})$$ (Use rule 2)



$G_{\overline{Z},\overline{X}}$

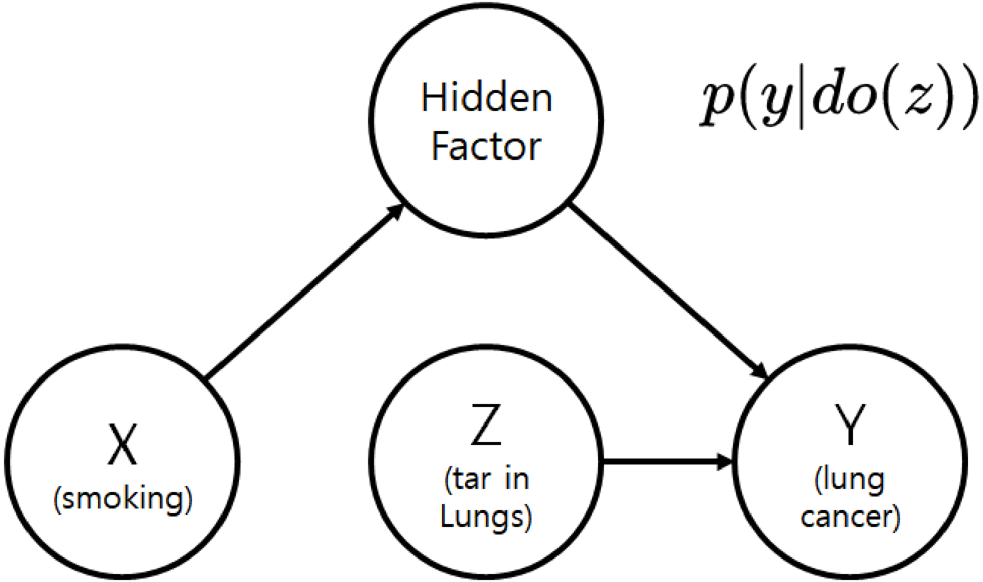
$$
\begin{aligned}
p(y|do(x)) &= \textstyle\sum_z p(y|z, do(x))p(z|do(x)) \\
&= \textstyle\sum_z p(y|z, do(x))p(z|x) \\
&= \textstyle\sum_z p(y|do(z), do(x))p(z|x) \\
&= \textstyle\sum_z p(y|do(z))p(z|x)
\end{aligned}
$$

$(\because (Y \perp\!\!\!\perp X|Z)_{G_{\overline{Z},\overline{X}}})$ (Use rule 3)



$G_{\underline{Z}}$

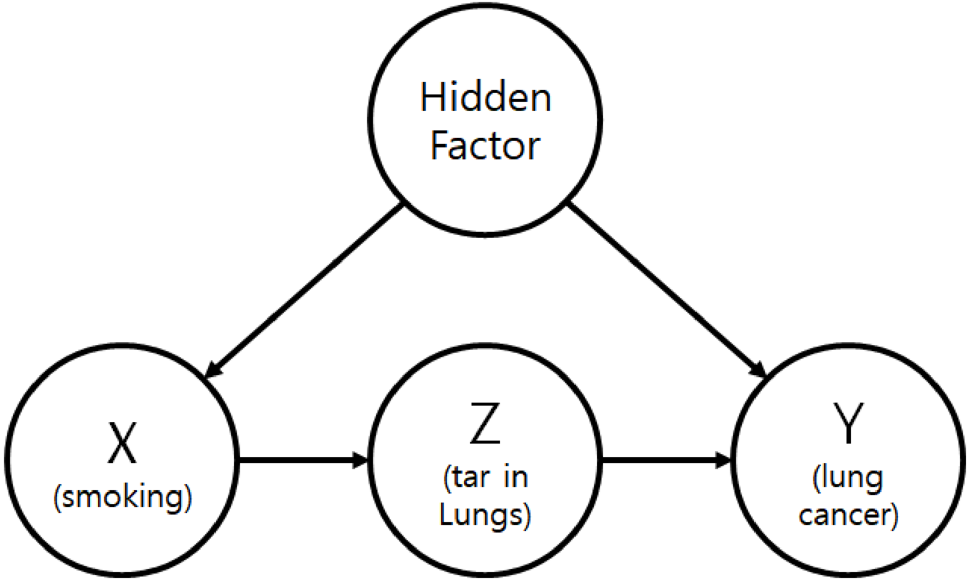
$$p(y|do(z)) = \sum_{x'} p(y|x', do(z))p(x'|do(z))$$
$$= \sum_{x'} p(y|x', z)p(x'|do(z))$$
$$(\because (Y \perp\!\!\!\perp Z|X)_{G_{\underline{Z}}})$$ (Use rule 2)



$G_{\overline{Z}}$

$$\begin{aligned} p(y|do(z)) &= \textstyle\sum_{x'} p(y|x', do(z))p(x'|do(z)) \\ &= \textstyle\sum_{x'} p(y|x', z)p(x'|do(z)) \\ &= \textstyle\sum_{x'} p(y|x', z)p(x') \end{aligned}$$

$(\because (Z \perp\!\!\!\perp X)_{G_{\overline{Z}}})$ (Use rule 3)



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

Paradoxes - Objective

- To explain why people find the paradox surprising or unbelievable

- To identify the class of scenarios in which the paradox can/cannot occur

- When we have to make a choice between two plausible yet contradictory statements, to tell us which statement is correct

Causality and Paradoxes?

- reveal the way the brain works, the shortcuts it takes, and things it finds conflicting.

- shine a spotlight onto patterns of **intuitive causal** reasoning that clash with the **logic of probability and statistics**.

Monty Hall Dilemma - *Review*

Let's Make a Deal

Resolving Monty Hall's Dilemma

Switch or Stay?

Door 1	Door 2	Door 3	Outcome If You Switch	Outcome If You Stay
Auto	Goat	Goat	Lose	Win
Goat	Auto	Goat	Win	Lose
Goat	Goat	Auto	Win	Lose

$$P(car = 2 | goat = 3) = \frac{2}{3}$$

Resolving Monty Hall's Dilemma - Bayesian Reasoning

- prior: $P(car = 1) = P(car = 2) = P(car = 3) = \frac{1}{3}$
- Observe: The door 3 is opened and revealed the goat. (goat=3)
- posterior: $P(car = 1|goat = 3) = \frac{P(goat=3|car=1)P(car=1)}{P(goat=3|car=1)P(car=1)+P(goat=3|car=2)P(car=2)+P(goat=3|car=3)P(car=3)}$

$$= \frac{1/2}{1/2+1+0} = \frac{1}{3}$$

$$P(car = 2|goat = 3) = \frac{2}{3}$$

$$P(car = 3|goat = 3) = 0$$

Resolving Monty Hall's Dilemma - Causal Reasoning

- When we condition on a **collider**, we create a **spurious dependency**

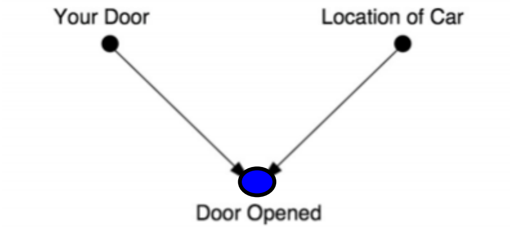


FIGURE 6.1. Causal diagram for *Let's Make a Deal*.

Let's Fake a Deal

- Monty Hall open the door ***randomly***, which is not the door you open.

Door You Chose	Door with Auto	Door Opened by Host	Outcome If You Switch	Outcome If You Stay
1	1	2 (goat)	Lose	Win
1	1	3 (goat)	Lose	Win
1	2	2 (auto)	Lose	Lose
1	2	3 (goat)	Win	Lose
1	3	2 (goat)	Win	Lose
1	3	3 (auto)	Lose	Lose

Let's Fake a Deal

- Monty Hall open the door ***randomly***, which is not the door you open.

Door You Chose	Door with Auto	Door Opened by Host	Outcome If You Switch	Outcome If You Stay
1	1	2 (goat)	Lose	Win
1	1	3 (goat)	Lose	Win
1	2	2 (auto)	Lose	Lose
1	2	3 (goat)	Win	Lose
1	3	2 (goat)	Win	Lose
1	3	3 (auto)	Lose	Lose

$$P(car = 2 | goat = 3) = \frac{1}{2}$$

Let's Fake a Deal

Your choice of a door and the producer's choice of where to put the car are independent

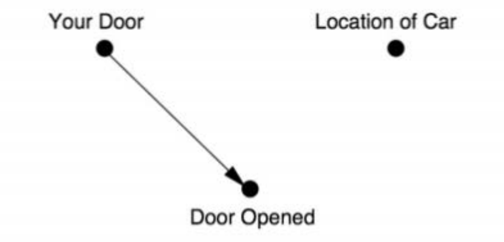
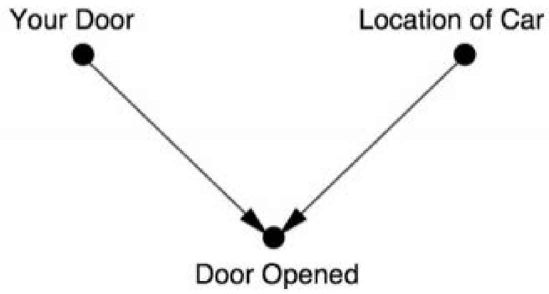


FIGURE 6.2. Causal diagram for *Let's Fake a Deal*.

What we learned

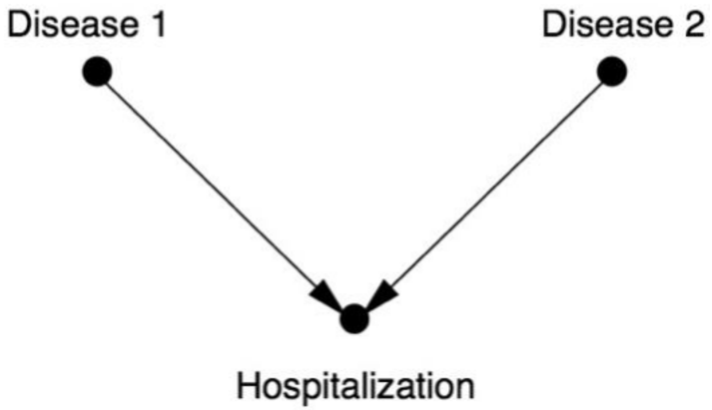
- Data generating process is also important

- Why we see it as a paradox in the first place.
  - Brain - Not to do *probability problems*, but to do *causal problems*
  - **Causeless correlation**



Collider Bias: Berkson's Bias

- Disease 1 and 2 have no relation to each other
- Neither Disease 1 nor 2 is ordinarily severe enough to cause hospitalization, but the combination is.
  - we would expect Disease 1 to be highly correlated with Disease 2 in the hospitalized population
  - Conditioning on a collider creates a **spurious association**



Berkson's Paradox

- About 7.5 % of people in general population have a bone disease
- “admission rate bias” or “Berkson bias”

	General Population			Hospitalized in Last Six Months		
Respiratory disease? ↓	*Bone disease?* ↓			*Bone disease?* ↓		
	Yes	No	% Yes	Yes	No	% Yes
Yes	17	207	7.6	5	15	25.0
No (control)	184	2,376	7.2	18	219	7.6

Lesson

- Collider bias can occur by the process by which observations are selected

- We should be careful to this bias

Simpson's Paradox

- BBG?
  - Bad for men
  - Bad for women
  - Good for people
- Take the Drug or not?

	Control Group (No Drug)		Treatment Group (Took Drug)	
	Heart attack	*No heart attack*	*Heart attack*	*No heart attack*
Female	1 5%	19	3	37
Male	12 < 30%	28	7.5% 8	12
Total	13 < 22%	47	40 % 11	49

> 18 %

Why does it happen? - Simpson's reversal

We have used an overly simple word "better" to describe a complex averaging process over uneven seasons

The denominators are not distributed evenly year to year

	Hits/At Bats			
	1995	*1996*	*1997*	*All Three Years*
David Justice	104/411 = .253	45/140 = .321	163/495 = .329	312/1,046 = .298
Derek Jeter	12/48 = .250	183/582 = .314	190/654 = .291	385/1,284 = .300

Return to BBG: Should I take it?

C: taking drug, E: heart attack occurred, F: female

We know that

$$P(E|do(C), F) < P(E|do(\neg C), F)$$

$$P(E|do(C), \neg F) < P(E|do(\neg C), \neg F)$$

$P(E|do(C)) > P(E|do(\neg C))$?

Return to BBG: Should I take it?

"Sure-Thing Principle"

"An action C that increases the probability of an event E in each ***subpopulation*** *must also increase the probability of E in the* ***population*** *as a whole, provided that the action does not change the distribution of the subpopulations."*

Our causal intuitive: the drug does not change the sex

$$P(F|do(C)) = P(F|do(\neg C)) = P(F)$$

Sure-Thing Principle

$$
\begin{aligned}
P(E|do(C)) &= P(E|do(C),F)P(F|do(C)) + P(E|do(C),\neg F), P(\neg F|do(C)) \\
&= P(E|do(C),F)P(F) + P(E|do(C),\neg F)P(\neg F) \\
&< P(E|do(\neg C),F)P(F) + P(E|do(\neg C),\neg F)P(\neg F) \\
&= P(E|do(\neg C))
\end{aligned}
$$

No BBG, but BBB!!

Lesson

- “Causal relationships are governed by the laws of probability calculus” -> **No!** Causality is governed by its own logic and this logic requires a major extension of probability calculus: *do-calculus*

- Our decision is driven by **causal** and not by statistical consideration

Q&A