國立臺北大學
National Taipei University

Causal Inference

Intro

Speaker: 吳淳硯

Dept. of Stat.
National Taipei University

Preliminary Knowledge

- Basic logical thinking.
- Basic knowledge of probability theory and statistical inference.
- Familiar to linear model (generalized linear model).
- Basic matrix algebra operation.
- Experience with any command language.
  (Ex. R, Python, VBA)

Syllabus

- Course Outline
  - Introduction to Causal Inference
  - Causal Diagram (DAGs) and d-separation (NPSEM)
  - Randomized Experiment vs Observational Studies
  - Econometric Method for Causal Inference
  - Causal Mediation Analysis
    - Single Mediator / Binary Mediator
    - Multiple Mediators
    - Time-Varying Mediator
- Evaluation Method
  - Regular grades

    Participation 20% + Assignment 20%
  - Mid-term grades 30%
  - Final grades 30%

Textbooks and Reference

- No specific textbook
  - The material is based on lecture notes and various classical papers.
- Ref (Causal Inference)
  - Causal inference What if – Hernan and Robin
  - Fundamentals of Causal Inference With R – Babette A. Brumback
  - The book of Why – Judea Pearl
- Ref (Econometric)
  - Principles of Econometrics – R. Carter Hill, William E. Griffiths and Guay C. Lim
  - Econometric Analysis – William H. Greene
  - Introductory Econometrics A Modern Approach – Jeffrey M. Wooldridge

Traditional view of Statistic

- Statistical inference is concerned with associational inference and used for finding association among variables, rather then for inferring causation relationship from observations.

- Two famous aphorisms based on this view are

Correlation does not imply causation.

You can’t prove causality with statistics.

- The mainly focus is to introduce the ″**formal language for causal inference**″ and ″**develop statistical methods to estimate causal effects**″ in randomized experiments and observational studies.

Causal inference ≈ Causal model
+ Statistical inference

Is Data access Causality?

Data Are Not Enough~ Hurray For Causality!

- **Data Are Not Enough**
  - Just collecting large amounts of data – even high-quality data – does not automatically tell us why things happen.
  - Data can show patterns and associations, but without context, it can mislead.
- **Hurray For Causality**
  - This is celebration of causal reasoning – the ability to determine what causes what.
  - With causality, we move beyond just describing the world to explaining, predicting and intervening effectively.
- So in short

Data ≠ Understanding

Causality = Power to explain, predict and change

Causal Reasoning

- Many of the inferences one wants to draw about real-world, spatiotemporal, contextual knowledge involve **Cause and Effect**.
- **Deductive Reasoning**
  - Deriving consequence (or effects, outcomes) from premises (or causes).
  - The Conclusion is **Certainly True**.
- **Inductive Reasoning**
  - Deriving relationship between causes and effects, rules that lead from one to another.
  - The Conclusion is **Probably True**.
- **Abductive Reasoning**
  - Deriving possible causes and effects.
  - The Conclusion is the **Best Explanation**.
- **Causal reasoning** is generally considered a form of induction reasoning.

Example - 1

- **Deductive Reasoning** Example
  - Premise 1: The sum of the interior angles of triangle is 180^o
  - Premise 2: ΔABC is a triangle.

  $\Rightarrow$ The sum of the interior angles of ΔABC is 180^o.
  (A guaranteed conclusion based on definitions)

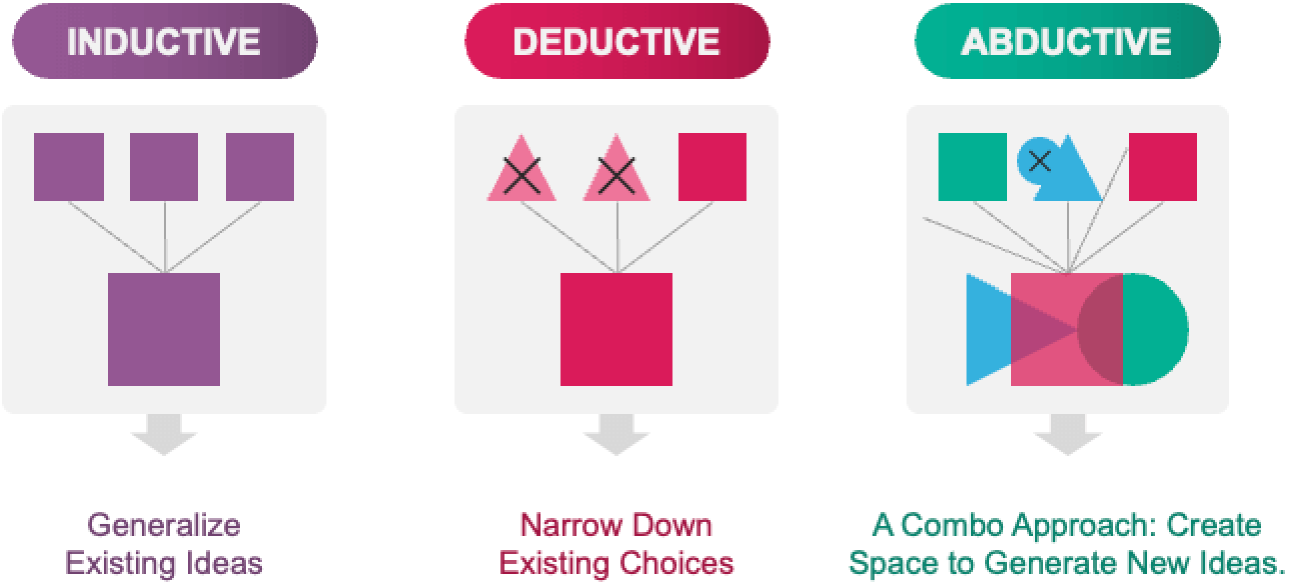
- **Inductive Reasoning** Example
  - Observations (Premises):
    - Most men have 2 testicles (Some men, have 1 or 0).
    - Women have 0 testicles.
    - Average of testicles across the entire population, the number is approximate 1.

  $\Rightarrow$ Therefore, everyone has one testicle.
  (This is technically true as a statistical average, but it's a misleading inductive conclusion if interpreted to describe a real, individual human.)

Example - 2

- **Abductive Reasoning** Example
  - Observation (Effect): The girl has a noticeably big belly.
  - Hypothesis (Possible Cause): The girl might be pregnant.

  ⇒ Therefore, the girl is probably pregnant.
  (This is not guaranteed to be true, but it’s the most plausible explanation given what we observe)
- Like **Inductive reasoning**, **Abductive reasoning** are not logically watertight. Although a hypothesis may seem to be best explanation, other explanations are still logically possible.



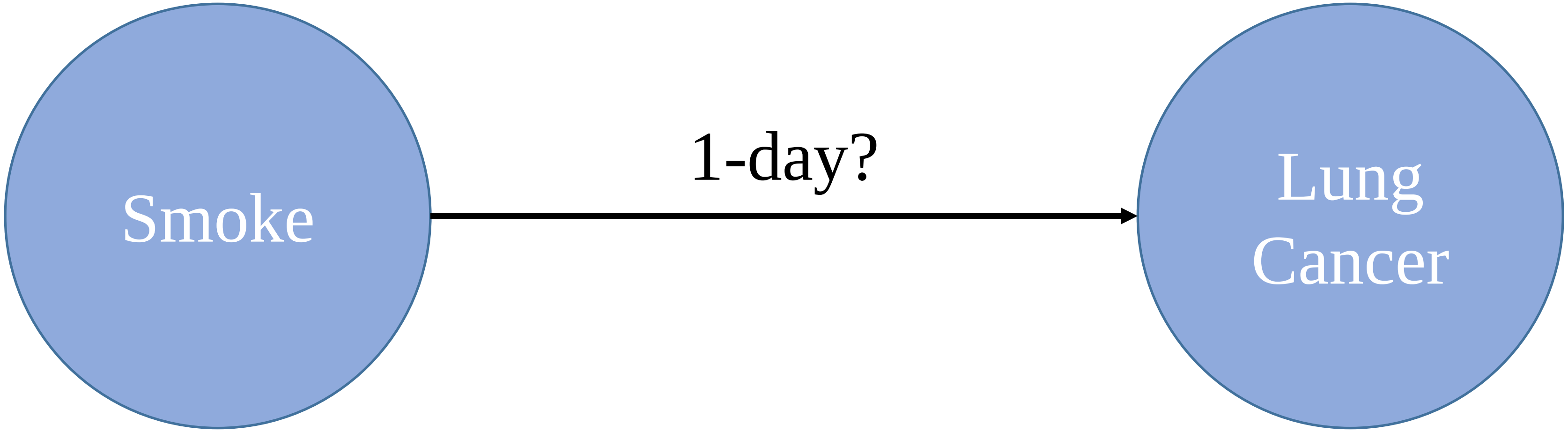
Hill's Criteria

- Some viewpoints we should study association before we cry causation
  - Strength if the association
  - Consistency
  - Specificity of the association
  - Temporality
  - Biological gradient (dose-response)
  - Plausibility (credibility)
  - Coherence
  - Experiment
  - Analogy
- None of them can bring indisputable evidence for or against the cause-and-effect hypothesis and none can be required as a *sine qua non*.
- What they can do, is to help us make up our minds on the question
  - Is there any other way of explaining the set of facts before us?
  - Is there any other answer equally, or more, likely than cause and effect?

Temporality

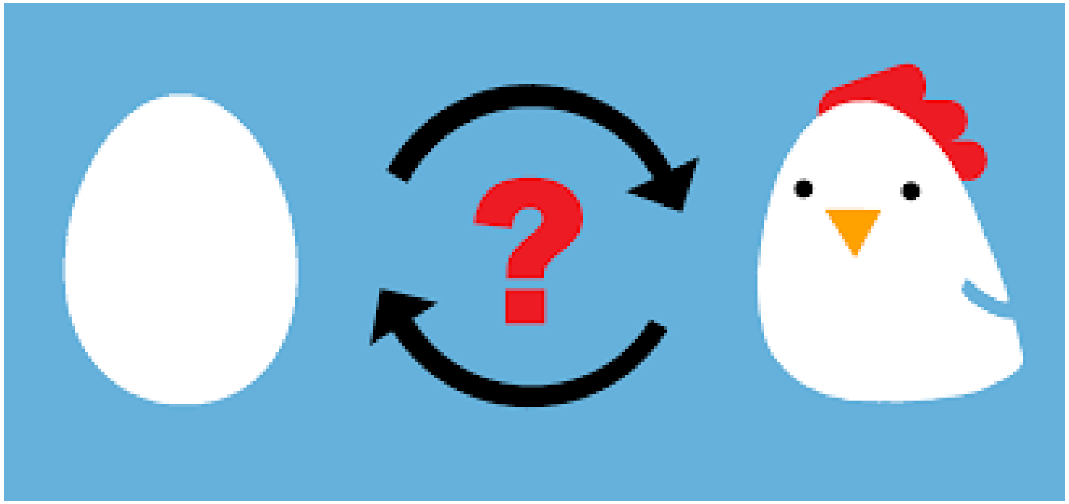
Exposure must precede outcome

- This criterion has been identified as being the most likely to be the *sine qua non* for causality.
- For an agent to be causal, its presence must precede the development of the outcome. Lack of temporality rules out causality.
- The effect must occur after the cause, with an **expected delay** between them.



Circular reasoning

- It is a logical fallacy in which the reasoner begins with what they are trying to end with.
- Circular reasoning is not a formal logical fallacy, but a pragmatic defect in an argument whereby the premises are just as much in need of proof or evidence as the conclusion.
- As a consequence, the argument becomes a matter of faith and fails to persuade those who do not already accept it.



Which happened first?

Why care about Causality?

- Allows us to:
  - Predict the future.
  - Explain the past.
  - Intervene in the present.
  - Learning and Adaptation.
  - Building and Fixing System

- Pervades daily life.
  - Naïve physics

    Building things, fixing things
  - Naïve biology

    Growing thing, cooking things
  - Naïve psychology

    Influencing, crediting, blaming.
- Great interactive features stimulate motor skill development and increase cause-effect understanding.

The Little Scientist

- How Children Learn **Cause and Effect**?
  - Natural Explorers

    From infancy, children act like little scientists - testing, observing, and predicting outcomes from their actions.
  - Trial and Error

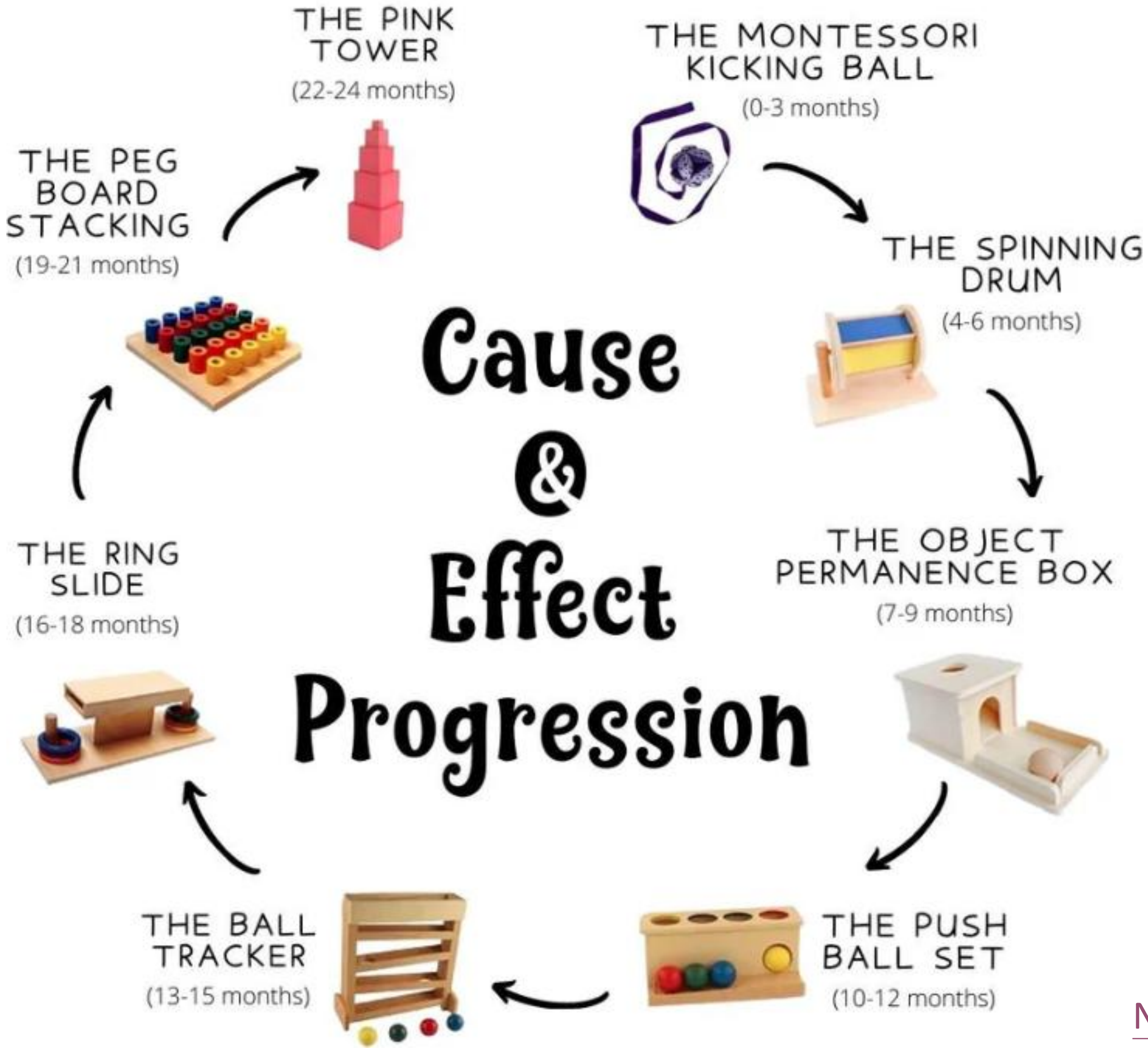
    They repeat actions to see what happens, building an understanding of cause and effect.
  - Imitating Adults

    Kids learn by watching others, even when those actions include mistakes or imperfect outcomes.
  - Everyday Experiments

    Daily life is full of opportunities for exploration - no special tools needed. Ordinary routines become learning labs.
- As children grow, they begin to understand that their actions can cause thing to happen, and this awareness is crucial for cognitive development.

國立臺北大學
National Taipei University

Cause and Effect Toy



NINO MONDO

Prediction vs Causation

- Prediction and Causation are very different.
  - Prediction:

    Predict Y after **observing** $A = a \Leftrightarrow P(Y = y|A = a)$

    Ex. Predict health given that a person takes vitamin C
  - Causation:

    Predict Y after **setting** $A = a \Leftrightarrow P(Y(a) = y)$

    Ex. Predict health if I give a person vitamin C.

  In general

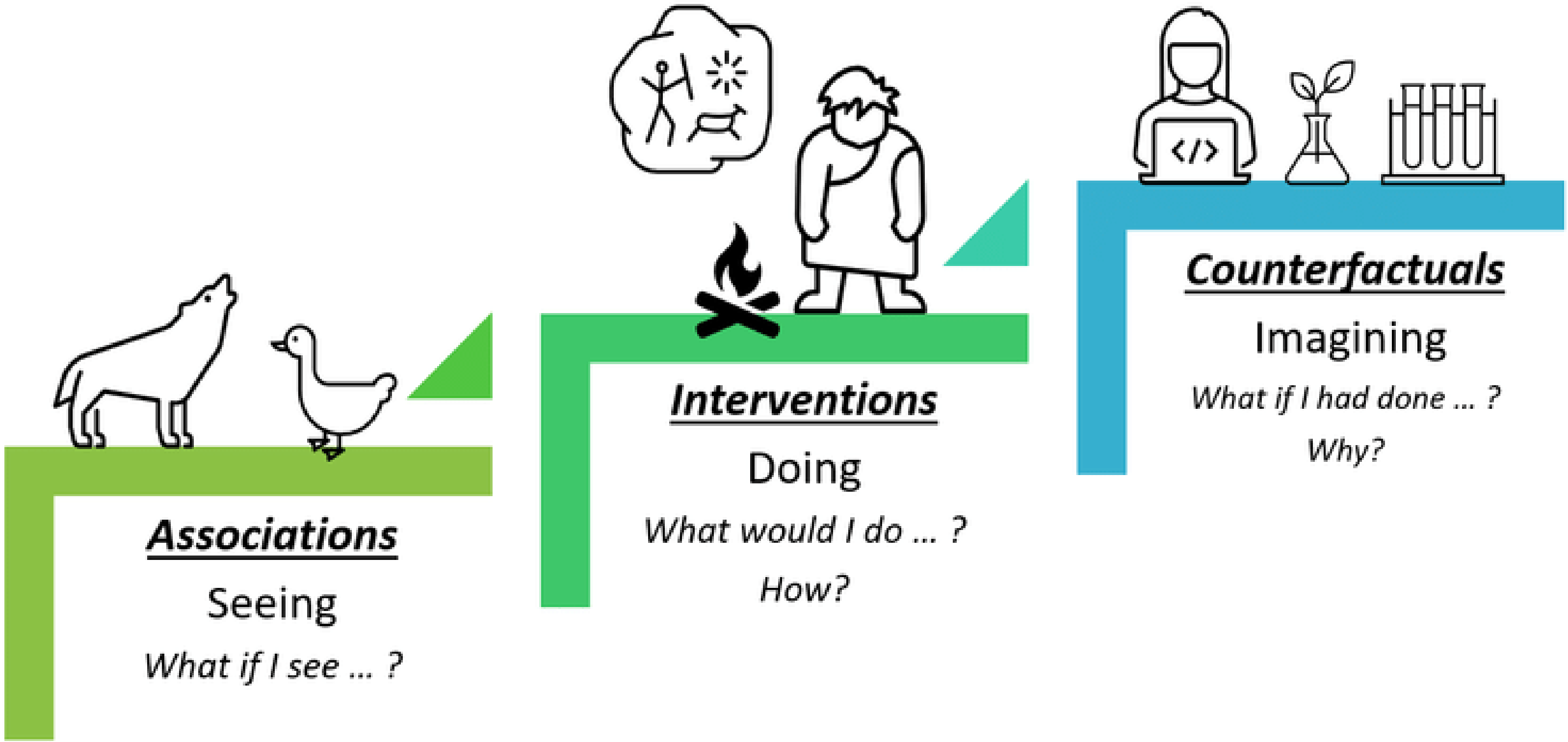
$$P(Y = y|A = a) \neq P(Y(a) = y)$$

- Causation (Counterfactual prediction)
  - Involves prediction the effect of an intervention.
- The difference between passively observing $A = a$ and actively intervening and setting $A = a$ is significant and requires different techniques and, typically, much stronger assumptions.

National Taipei University

The Ladder of Causations

- A causal learner must master at least three distinct levels of cognitive ability:
  - Seeing, Observing $\Leftrightarrow P(Y = y|A = a)$
  - Doing, Intervening $\Leftrightarrow P(Y(a^*) = y)$
  - Imagining, Retrospection, Understanding $\Leftrightarrow P(Y(a^*) = y|A = a)$



Paradox Galore - 1

- Monty Hall problem (Game Show - Let's Make a Deal)

  There are 3 doors. Behind one door is a car, behind the others, goats. Pick a door, say No.1, and the host, who knows what's behind the doors, opens door No.3, which has a goat. He says to you,

Do you want to pick door No.2?

  Is it to your advantage to switch your choice of doors? **YES**

- Three possible arrangements of doors and goats

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

Always account for variable change.

Bayes' Theorem

$C_i \triangleq$ Car at Door No. i, $D_i \triangleq$ Open Door No. i

- **Prior**: Probability of car behind doors

$$P(C_1) = P(C_2) = P(C_3) = \frac{1}{3}$$

- **Event**: Probability of host to open door No.3

$$P(D_3|C_1) = \frac{1}{2}, \qquad P(D_3|C_2) = 1, \qquad P(D_3|C_3) = 0$$

- **Posterior**: Chance of the car behind the doors after the event.

$$P(C_1|D_3) = \frac{P(D_3|C_1)P(C_1)}{P(D_3|C_1)P(C_1) + P(D_3|C_2)P(C_2) + P(D_3|C_3)P(C_3)}$$

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

Similar way to obtained

$$P(C_2|D_3) = \frac{2}{3}, \qquad P(C_3|D_3) = 0$$

National Taipei University

Paradox Galore - 2

- Simpson's Paradox (**BBG** ≜ Bad/Bad/Good, Drug)

	Control Group (No Drug)		Treatment Group (Took Drug)	
	Heart attack	*No heart attack*	*Heart attack*	*No heart attack*
Female	1	19	3	37
Male	12	28	8	12
Total	13	47	11	49

  - Bad for Female

$$\frac{1}{1+19} = \frac{1}{20} < \frac{3}{40} = \frac{3}{3+37}$$

  - Bad for Male

$$\frac{12}{12+28} = \frac{6}{20} < \frac{8}{20} = \frac{8}{8+12}$$

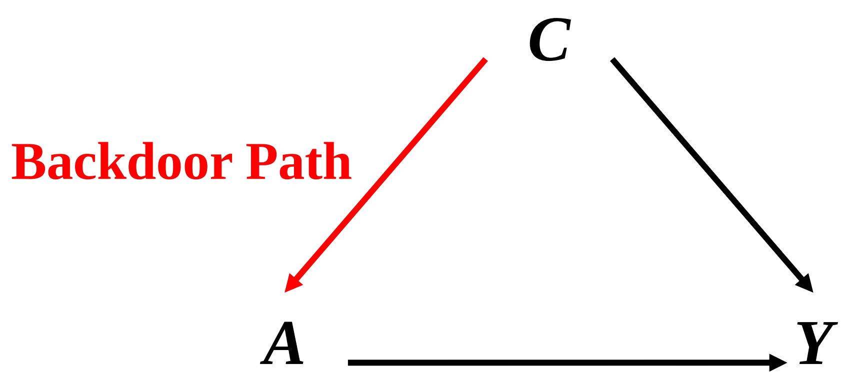
  - Good for People

$$\frac{1+12}{13+47} = \frac{13}{60} > \frac{11}{60} = \frac{3+8}{11+49}$$

- **BBG** drug indeed does not exist and will never be invented.

Blocking the Spurious Paths

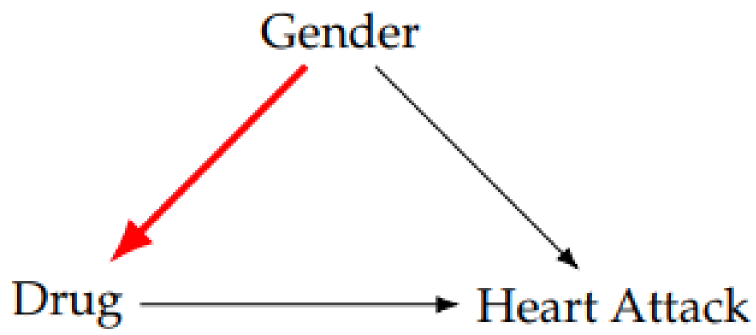
- To deconfound two variables $\boldsymbol{A}$ and $\boldsymbol{Y}$, we need only block every noncausal path between them without blocking or perturbing any causal paths.

- 

- **Backdoor Path**
  - It is any path from $\boldsymbol{A}$ to $\boldsymbol{Y}$ that starts with an arrow pointing into $\boldsymbol{A}$.
- Backdoor Adjustment
  - If we do this by controlling for some set of variables $\boldsymbol{C}$, we also need to make sure that no member of $\boldsymbol{C}$ is a descendant of $\boldsymbol{A}$ on a causal path.

Causal Diagram

- **BBG** causal structure



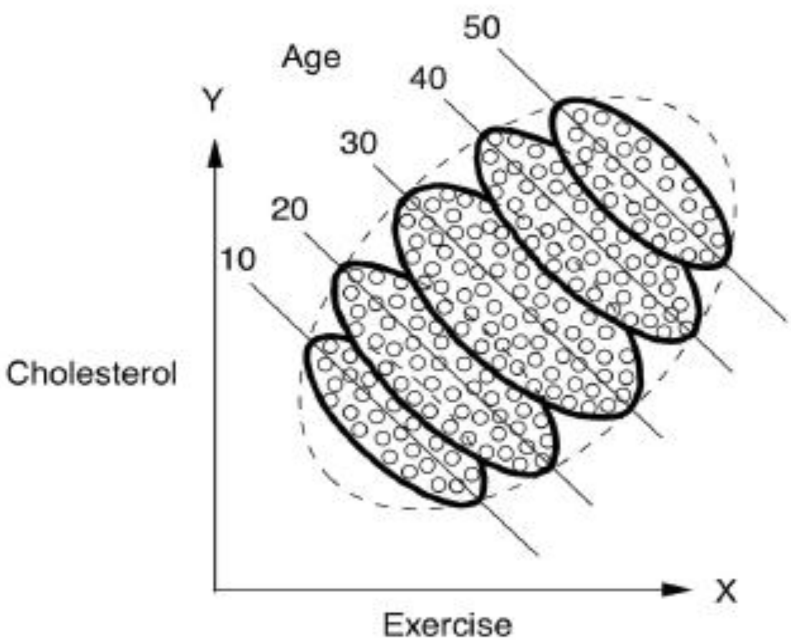
- In the study, women clearly had preference for taking Drug and men preferred not to.
- Thus Gender serves as a confounder/common cause between Drug and Heart Attack.
- For a unbiased estimate of the effect of Drug on Heart Attack, we must adjust for the confounder.
  - Control vs Treatment

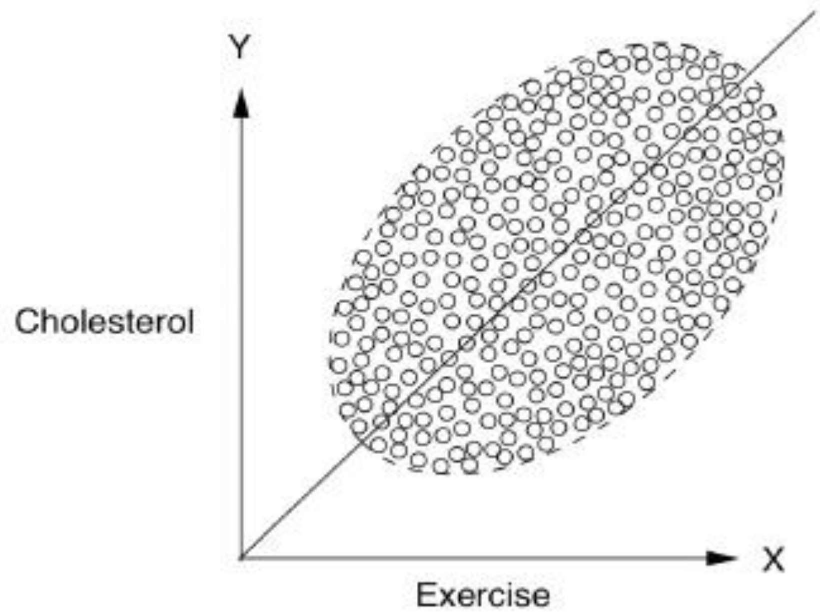
$$\frac{1}{20}+\frac{12}{40}=\frac{14}{40}<\frac{19}{40}=\frac{3}{40}+\frac{8}{20}$$

  - Drug is not **BBG**, it's **BBB**(≙Bad/Bad/Bad).

Paradox Galore - 3

- Consider a study that measures weekly exercise and cholesterol levels in various age groups.

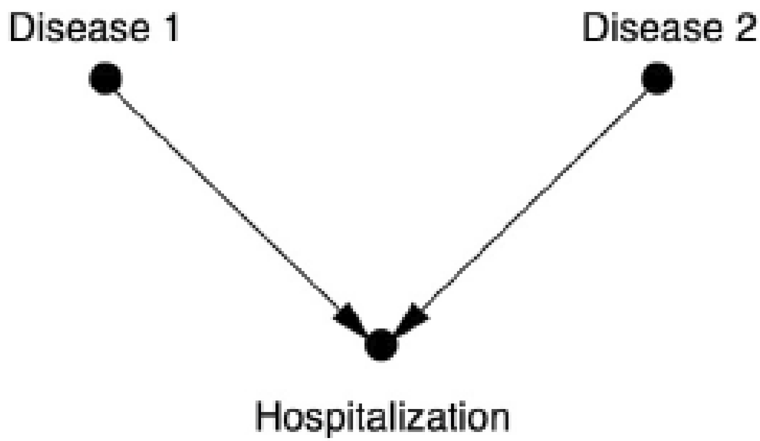




- Once again we seem to have a **BBG** situation
  - It seems to have a beneficial effect in each age group but harmful effect on the population as whole.

Paradox Galore - 4

- Berkson's Paradox



  - It occurs when **selection bias**, also known as collider-stratification bias, creates a spurious association between two independent variables due to conditioning on a common effect (a collider).

	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

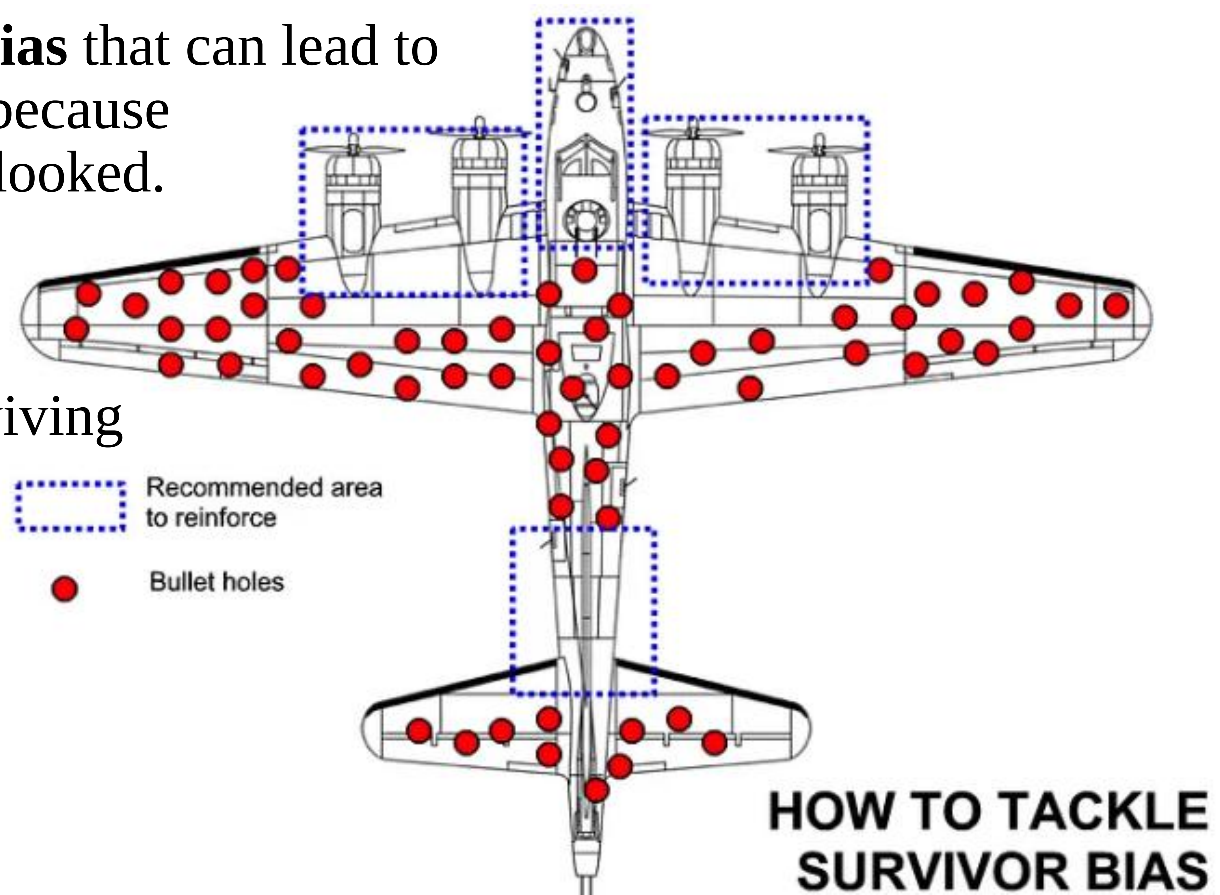
Survivorship Bias

Dead men tell no tales

- It is the logical error of concentrating on entities that passed a selection process while overlooking those that did not.

- It is a form of **selection bias** that can lead to overly optimistic beliefs because multiple failures are overlooked.

- This hypothetical pattern of damage of surviving aircraft shows locations where they can sustain damage and still return home.



Cobra Effect

Unintended Consequences Everywhere

- It refers to a situation when an attempted solution to a problem makes the problem worse.
- When we try to make a single change within a complex system, we often end up causing unintended consequences.

Be Careful what you wish for ~

Source of Bias

- **Causal Bias** =
  - + **Systematic Bias (Hidden Bias)**
    - ∗ Confounding Bias (Simpson's paradox)
    - ∗ Selection Bias (Berkson's paradox)
    - ∗ Measurement Bias
  - + **Misspecification bias**
    - ∗ Due to parametric modeling
  - + **Random Variability**
    - ∗ Finite sample bias

- The distinction between association and causation primarily arises from the need to account for potential systematic biases (hidden biases).