A brief introduction to causal inference

Outline

Causal effects and where to find them

Spurious correlations

Simpson’s paradox

Randomised Controlled Trials (RCTs)

Causal effects without RCTs - Adjustment

What if we don’t have the graph? - Causal discovery

Instrumental variables

Prediction

Causal effects and where to find them

the effectiveness of a ‘treatment’ or ‘intervention’ on an outcome.

Seen in:

- Pharmaceuticals
- Economic policy
- Epidemiology/Public health policy
- Nutrition
- Product testing

Outline

Causal effects and where to find them

Spurious correlations

Simpson’s paradox

Randomised Controlled Trials (RCTs)

Causal effects without RCTs - Adjustment

What if we don’t have the graph? - Causal discovery

Instrumental variables

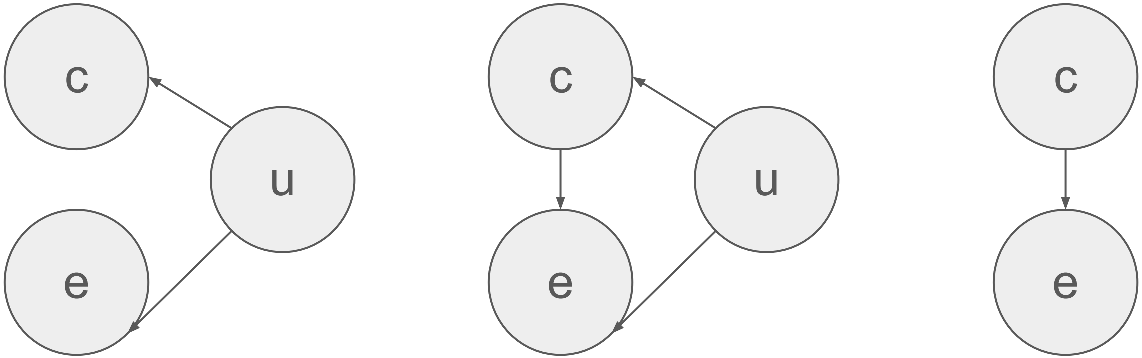
Prediction

Spurious correlations (and conclusions)

1. Mozzarella consumption per capita vs Civil engineering doctorates
2. Income vs marriage status
3. Fires vs. fire fighters

Spurious correlations (and conclusions)

1. Mozzarella consumption per capita vs Civil engineering doctorates
2. Income vs marriage status
3. Fires vs. fire fighters



Outline

Causal effects and where to find them

Spurious correlations

Simpson’s paradox

Randomised Controlled Trials (RCTs)

Causal effects without RCTs - Adjustment

What if we don’t have the graph? - Causal discovery

Instrumental variables

Prediction

Simpson’s paradox - 1

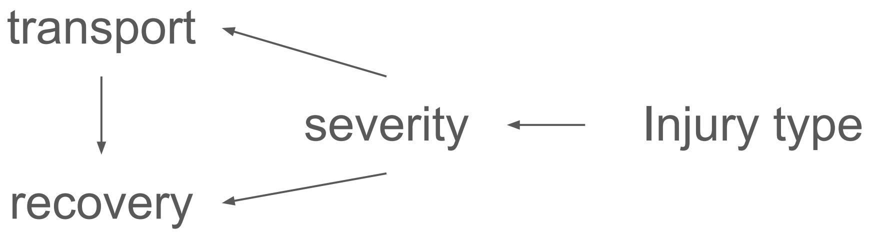
Does using an air ambulance improve chances of full recovery?

Injury type / Transport	**Air Ambulance**	**Ambulance**
Sport injury	81/87 = **93%**	234/270 = 87%
Motor injury	192/263 = **73%**	55/80 = 69%
Total	273/350 = 78%	289/350 = **83%**

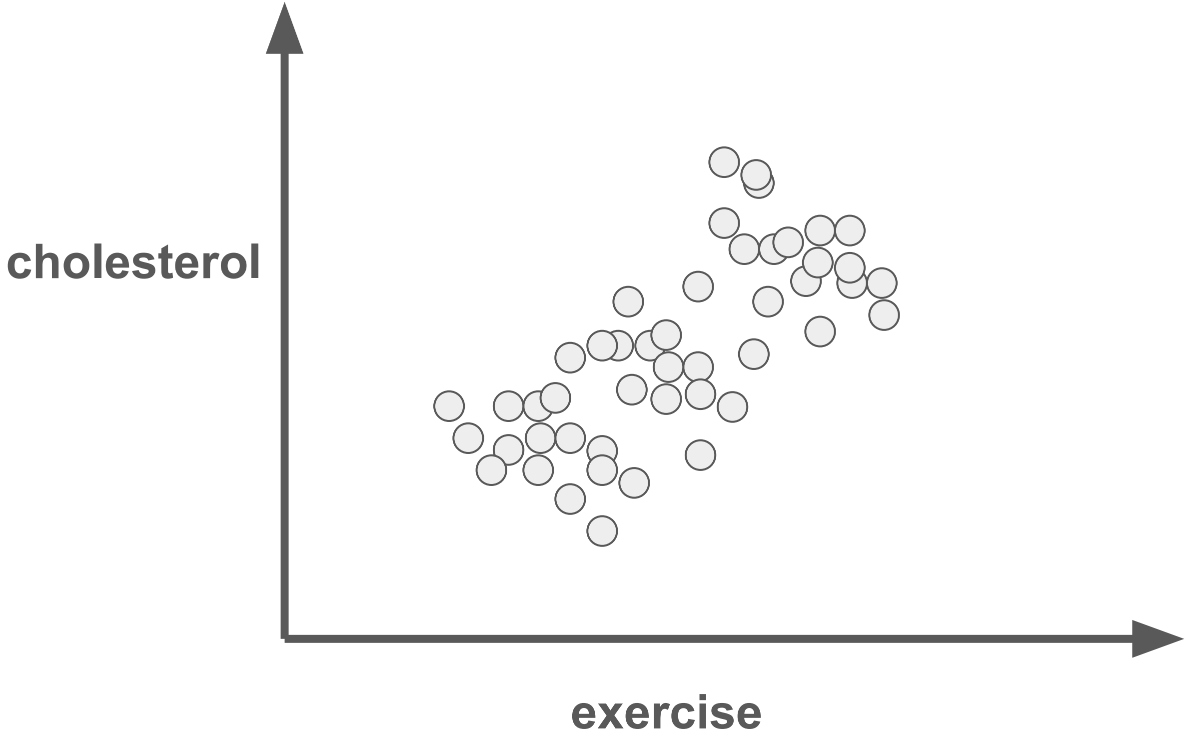
Simpson’s paradox - 1

Does using an air ambulance improve chances of full recovery?

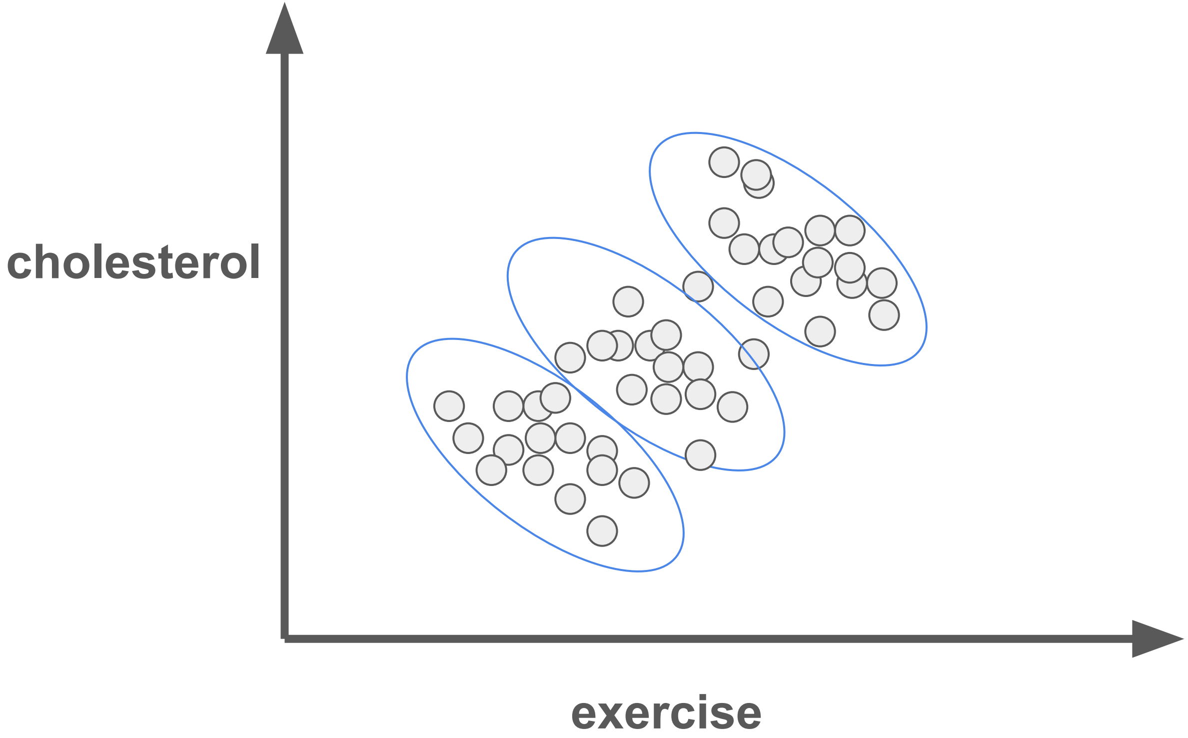
Injury type / Transport	**Air Ambulance**	**Ambulance**
Sport injury	81/87 = **93%**	234/270 = 87%
Motor injury	192/263 = **73%**	55/80 = 69%
Total	273/350 = 78%	289/350 = **83%**



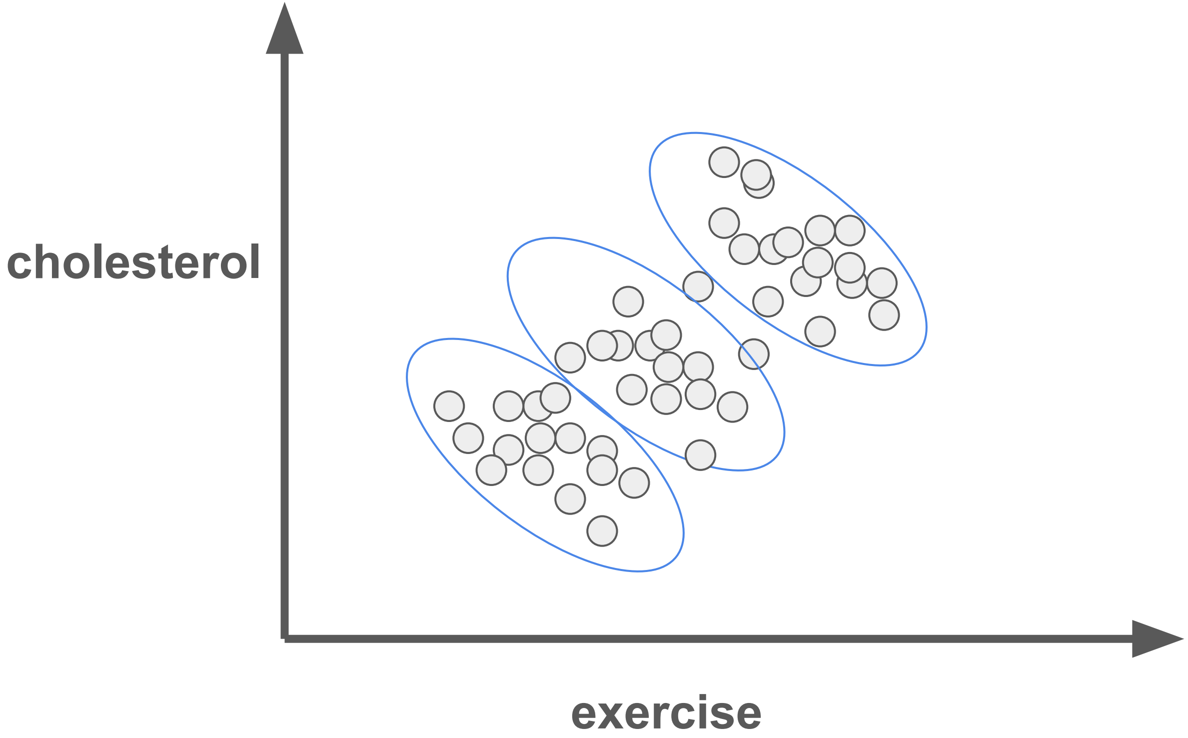
Simpson's paradox - 2

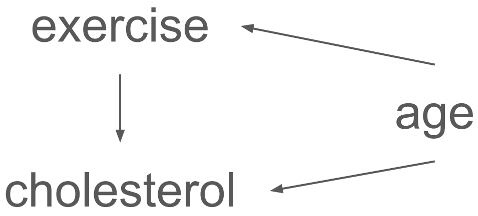


Simpson’s paradox - 2



Simpson's paradox - 2





Outline

Causal effects and where to find them

Spurious correlations

Simpson’s paradox

Randomised Controlled Trials (RCTs)

Causal effects without RCTs - Adjustment

What if we don’t have the graph? - Causal discovery

Instrumental variables

Prediction

Randomised controlled trials

Used to measure the effectiveness of an intervention on an outcome - i.e the causal effect.

“Solves the problem of observed and unobserved confounding”

1. Take a large sample from the population
2. Randomly assign the intervention to half of the sampled units
3. Compare the difference in the average outcomes of the two groups

RCTs - why does it give us the causal effect?

$$Y = \begin{cases} Y_1 & \text{if } I = 1 \\ Y_0 & \text{if } I = 0 \end{cases}$$

$$\mathbb{E}[Y_1 - Y_0] \tag{1}$$

$$\begin{aligned} &\mathbb{E}[Y|I = 1] - \mathbb{E}[Y|I = 0] && (2) \\ &= \mathbb{E}[Y_1|I = 1] - \mathbb{E}[Y_0|I = 0] && (3) \end{aligned}$$

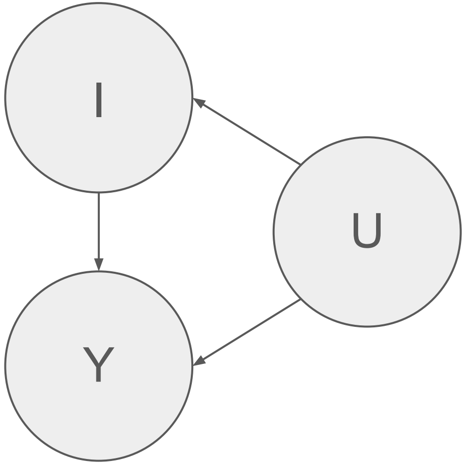
RCTs - why does it give us the causal effect?

$$Y = \begin{cases} Y_1 & \text{if } I = 1 \\ Y_0 & \text{if } I = 0 \end{cases}$$

$$\mathbb{E}[Y_1 - Y_0] \qquad (1)$$

$$\mathbb{E}[Y|I=1] - \mathbb{E}[Y|I=0] \qquad (2)$$

$$= \mathbb{E}[Y_1|I=1] - \mathbb{E}[Y_0|I=0] \qquad (3)$$



Randomised controlled trials - the gold standard (?)

1. Population shift
2. Non-representative samples

Outline

Causal effects and where to find them

Spurious correlations

Simpson’s paradox

Randomised Controlled Trials (RCTs)

Causal effects without RCTs - Adjustment

What if we don’t have the graph? - Causal discovery

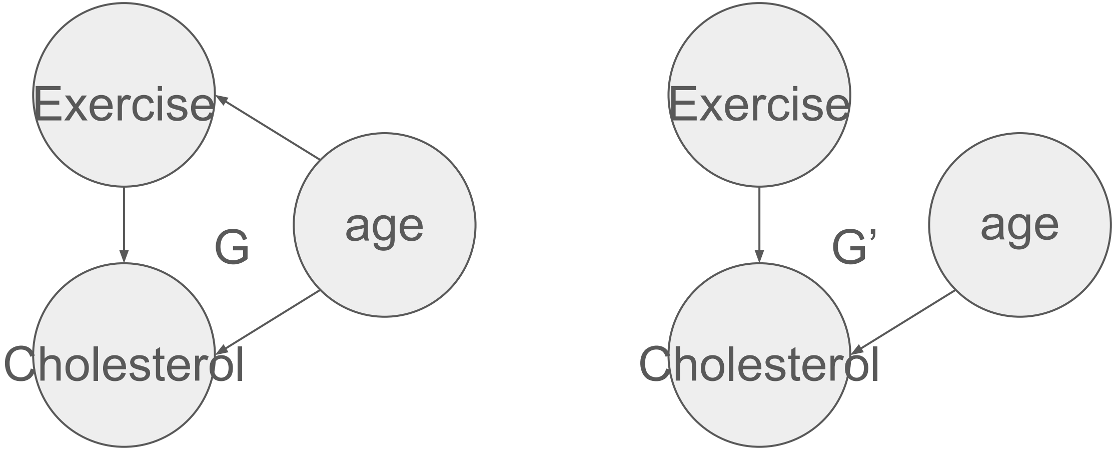
Instrumental variables

Prediction

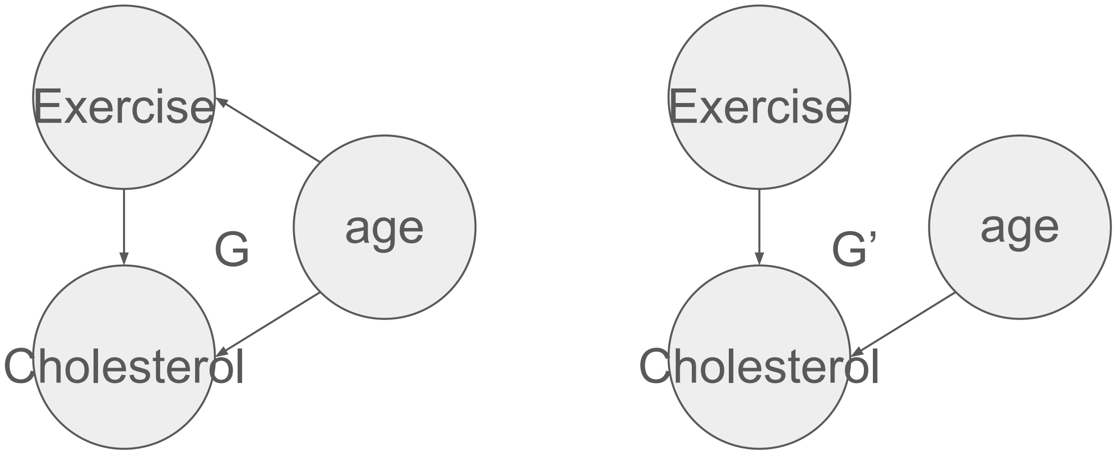
Causal effects without RCTs - Adjustment

If we have data that doesn’t come from an RCT,
can we find the causal effect?

Causal effects without RCTs - Adjustment

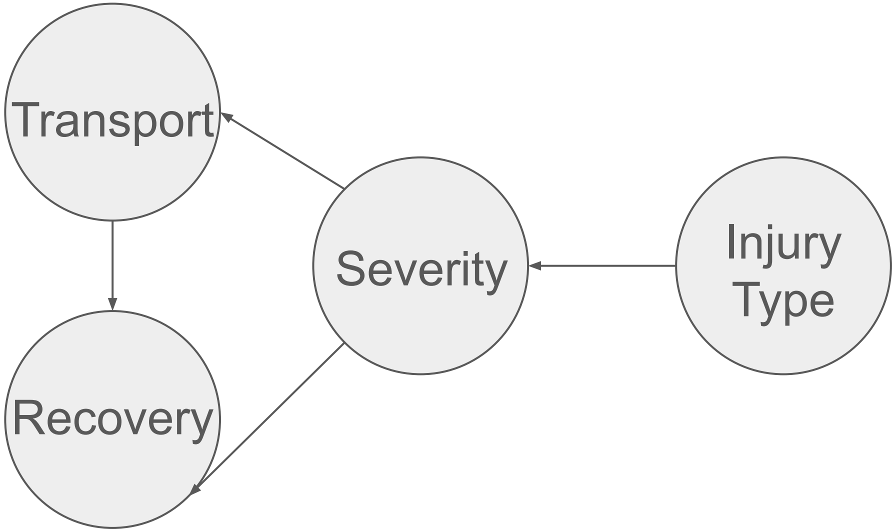


Causal effects without RCTs - Adjustment



$$
\begin{aligned}
Pr_{G'}(C|E) &= \sum_a Pr_{G'}(C|E,A)Pr_{G'}(A|E) \\
&= \sum_a Pr_{G'}(C|E,A)Pr_{G'}(A) \\
&= \sum_a Pr_G(C|E,A)Pr_G(A)
\end{aligned}
$$

Causal effects without RCTs - Adjustment



Outline

Causal effects and where to find them

Spurious correlations

Simpson’s paradox

Randomised Controlled Trials (RCTs)

Causal effects without RCTs - Adjustment

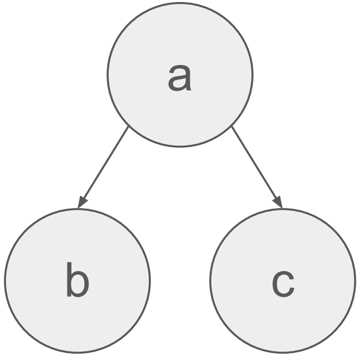
What if we don’t have the graph? - Causal discovery

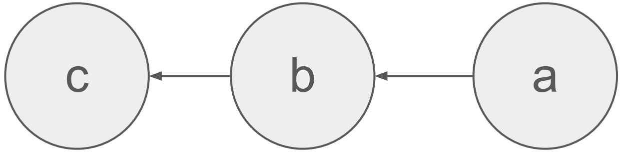
Instrumental variables

Prediction

Causal discovery

1. Given set of variables X_1, X_2, … X_n
2. Start with the fully connected graph
3. Test independence between variables to remove some edges
4. Test conditional independence between variables to remove edges …





Outline

Causal effects and where to find them

Spurious correlations

Simpson’s paradox

Randomised Controlled Trials (RCTs)

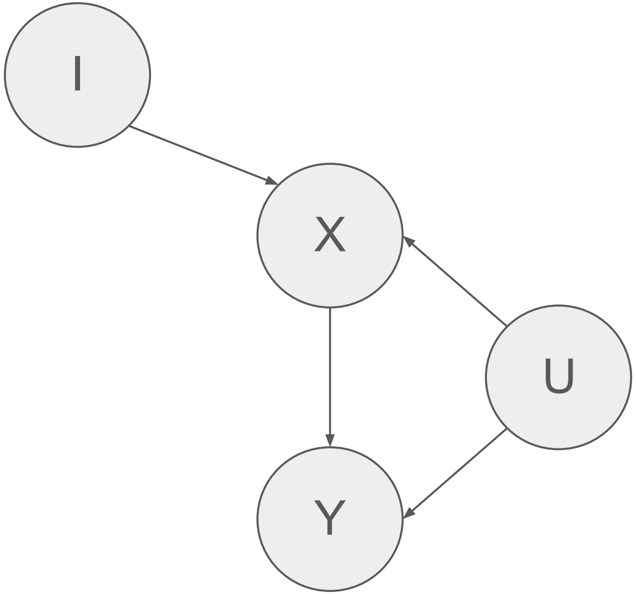
Causal effects without RCTs - Adjustment

What if we don’t have the graph? - Causal discovery

Instrumental variables

Prediction

Instrumental variables



1. The instrument (I) must influence X
2. The instrument (I) does not affect the outcome (Y) in any way other than through the intervention (X)
3. There is no common cause of the instrument (I) and the outcome (Y)

Outline

Causal effects and where to find them

Spurious correlations

Simpson's paradox

Randomised Controlled Trials (RCTs)

Causal effects without RCTs - Adjustment

What if we don't have the graph? - Causal discovery

Instrumental variables

Prediction

Prediction

Take care when using spurious variables for prediction

The causal variables should be robust under most interventions

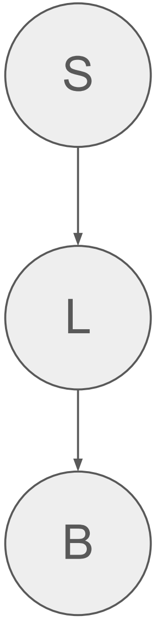
The effect variables are still useful for prediction, even when not robust

Prediction

Take care when using spurious variables for prediction

The causal variables should be robust under most interventions

The effect variables are still useful for prediction, even when not robust



Outline

Causal effects and where to find them

Spurious correlations

Simpson’s paradox

Randomised Controlled Trials (RCTs)

Causal effects without RCTs - Adjustment

What if we don’t have the graph? - Causal discovery

Instrumental variables

Prediction