Front door identification

INFO/STSCI/ILRST 3900: Causal Inference

12 Oct 2023

Logistics

- Problem Set 4 due Oct 19
- Form for final project groups
  - Writeup due Nov 21
  - Presentations Nov 29

Quick review: Where we are

- define a causal effect
  - treatment, outcome, potential outcomes, target population
- identify a causal effect
  - maps a causal quantity (involving counterfactuals) to a statistical quantity (involving only factual variables)
  - DAGs, conditional exchangeability
- estimate a causal effect
  - statistical modeling, matching, regression

Learning goals for today

At the end of class you will be able to

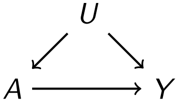
- explain front-door causal identification

More broadly,

1. engage with a new causal identification approach
2. translate that method to code
3. critique the identification assumptions

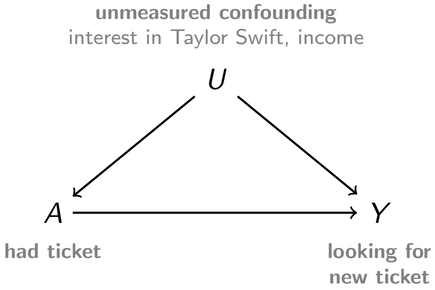
1) Engage with a new causal identification approach

Sometimes a sufficient adjustment set does not exist

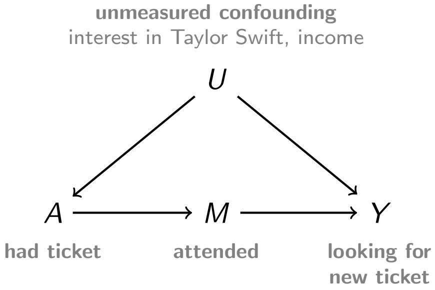


Imagine you are Taylor Swift's head of advertising

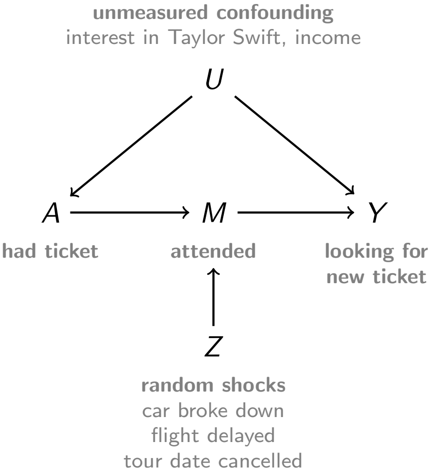
Does having a ticket for the Eras Tour increase the probability that a fan look for a future ticket?



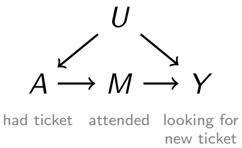
As head of advertising, how could you learn about $A \rightarrow Y$?



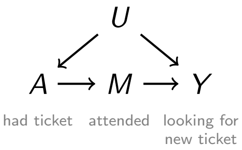
As head of advertising, how could you learn about $A \rightarrow Y$?



As head of advertising, how could you learn about $A \to Y$?


U
A
M
Y
had ticket
attended
looking for new ticket

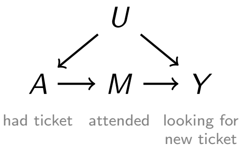
1) $A \rightarrow M$ is identified



1) $A \to M$ is identified

$\mathrm{P}(M^a) =$

for $a = 1$:
would attend
if given a
ticket?

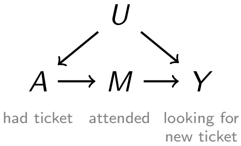


1) $A \to M$ is identified

$$P(M^a) = P(M \mid A = a)$$

for $a = 1$: would attend if given a ticket?

attendance rate among those with tickets

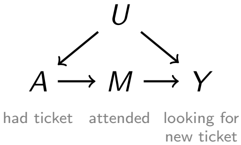


1) $A \to M$ is identified

$$P(M^a) = P(M \mid A = a)$$

for $a = 1$: would attend if given a ticket?

attendance rate among those with tickets



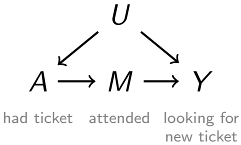
2) $M \to Y$ is identified

1) $A \to M$ is identified

$$\mathrm{P}(M^a) = \mathrm{P}(M \mid A = a)$$

for $a = 1$: would attend if given a ticket?

attendance rate among those with tickets



2) $M \to Y$ is identified

$$\mathrm{P}(Y^m)$$

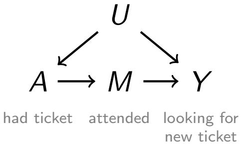
for $m = 1$: would look for new ticket if attended?

1) $A \to M$ is identified

$$P(M^a) = P(M \mid A = a)$$

for $a = 1$: would attend if given a ticket?

attendance rate among those with tickets



2) $M \to Y$ is identified

$$P(Y^m) = \sum_{a'} P(A = a')P(Y \mid M = m, A = a')$$

for $m = 1$: would look for new ticket if attended?

weighted sum over having ticket

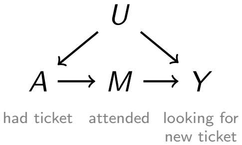
looking for new ticket given attendance?

1) $A \to M$ is identified

$$P(M^a) = P(M \mid A = a)$$

for $a = 1$: would attend if given a ticket?

attendance rate among those with tickets



2) $M \to Y$ is identified

$$P(Y^m) = \sum_{a'} P(A = a')P(Y \mid M = m, A = a')$$

for $m = 1$: would look for new ticket if attended?

weighted sum over having ticket

looking for new ticket given attendance?

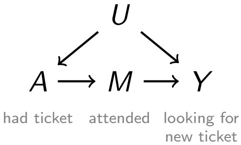
3) $A \to Y$ operates through M

1) $A \to M$ is identified

$$P(M^a) = P(M \mid A = a)$$

for $a = 1$: would attend if given a ticket?

attendance rate among those with tickets



2) $M \to Y$ is identified

$$P(Y^m) = \sum_{a'} P(A = a')P(Y \mid M = m, A = a')$$

for $m = 1$: would look for new ticket if attended?

weighted sum over having ticket

looking for new ticket given attendance?

3) $A \to Y$ operates through M

$$P(Y^a) =$$

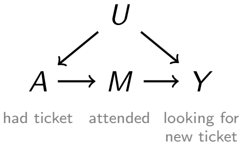
for $a = 1$: would look for future ticket if given ticket?

1) $A \to M$ is identified

$$P(M^a) = P(M \mid A = a)$$

for $a = 1$: would attend if given a ticket?

attendance rate among those with tickets



2) $M \to Y$ is identified

$$P(Y^m) = \sum_{a'} P(A = a')P(Y \mid M = m, A = a')$$

for $m = 1$: would look for new ticket if attended?

weighted sum over having ticket

looking for new ticket given attendance?

3) $A \to Y$ operates through M

$$P(Y^a) = P(Y^{M^a})$$

for $a = 1$: would look for future ticket if given ticket?

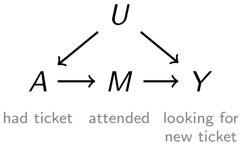
would look for future ticket if attended as if given ticket?

DAG gave us three equations

1) $\mathrm{P}(M^a = m) = \mathrm{P}(M \mid A = a)$

2) $\mathrm{P}(Y^m) = \sum_{a'} \mathrm{P}(A = a')\mathrm{P}(Y \mid M = m, A = a')$

3) $P(Y^a) = \mathrm{P}(Y^{M^a})$

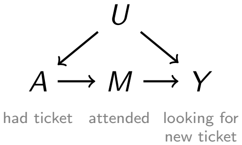


DAG gave us three equations

1) $\mathrm{P}(M^a = m) = \mathrm{P}(M \mid A = a)$

2) $\mathrm{P}(Y^m) = \sum_{a'} \mathrm{P}(A = a')\mathrm{P}(Y \mid M = m, A = a')$

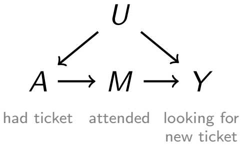
3) $P(Y^a) = \mathrm{P}(Y^{M^a})$



Proof

$$
\begin{aligned}
\mathrm{P}(Y^a) &= P(Y^{M^a}) && \text{by (3)}\\
&= \sum_m \mathrm{P}(M^a = m)\mathrm{P}(Y^m) && \text{law of total prob.}\\
&= \sum_m \mathrm{P}(M = m \mid A = a)\mathrm{P}(Y^m) && \text{by (1)}\\
&= \sum_m \Big(\mathrm{P}(M = m \mid A = a)\\
&\qquad \times \sum_{a'} \mathrm{P}(A = a')\mathrm{P}(Y \mid M = m, A = a')\Big) && \text{by (2)}
\end{aligned}
$$

DAG gave us three equations



1) $\mathrm{P}(M^a = m) = \mathrm{P}(M \mid A = a)$

2) $\mathrm{P}(Y^m) = \sum_{a'} \mathrm{P}(A = a')\mathrm{P}(Y \mid M = m, A = a')$

3) $P(Y^a) = \mathrm{P}(Y^{M^a})$

Result

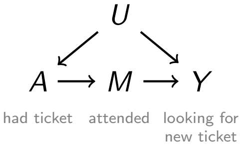
$$\mathrm{P}(Y^a) = \sum_{m} \mathrm{P}(M = m \mid A = a) \sum_{a'} \mathrm{P}(A = a')\mathrm{P}(Y \mid M = m, A = a')$$

DAG gave us three equations

1) $\mathrm{P}(M^a = m) = \mathrm{P}(M \mid A = a)$

2) $\mathrm{P}(Y^m) = \sum_{a'} \mathrm{P}(A = a')\mathrm{P}(Y \mid M = m, A = a')$

3) $P(Y^a) = \mathrm{P}(Y^{M^a})$



Result

$$\mathrm{P}(Y^a) = \sum_m \mathrm{P}(M = m \mid A = a) \sum_{a'} \mathrm{P}(A = a')\mathrm{P}(Y \mid M = m, A = a')$$

If we intervene to set treatment to the value *a*

then your outcome is a weighted average over the *M* distribution that would result

of the outcome under $M = m$, identified by backdoor adjustment for *A*

2) Translate to code

Translating math to code

$$P(Y^a) = \sum_m P(M = m \mid A = a) \sum_{a'} P(A = a') P(Y \mid M = m, A = a')$$

Translating math to code

$$\mathrm{P}(Y^a) = \sum_m \mathrm{P}(M = m \mid A = a) \sum_{a'} \mathrm{P}(A = a')\mathrm{P}(Y \mid M = m, A = a')$$

Translating math to code

$$P(Y^a) = \sum_m P(M = m \mid A = a) \sum_{a'} P(A = a') P(Y \mid M = m, A = a')$$

```
# Probability of each A
p_A <- data %>%
  # Count size of each group
  group_by(A) %>%
  count() %>%
  # Convert to probability
  ungroup() %>%
  mutate(p_A = n / sum(n)) %>%
  select(A,p_A)
```

Translating math to code

$$\mathrm{P}(Y^a) = \sum_m \mathrm{P}(M = m \mid A = a) \sum_{a'} \mathrm{P}(A = a')\mathrm{P}(Y \mid M = m, A = a')$$

Translating math to code

$$P(Y^a) = \sum_m P(M = m \mid A = a) \sum_{a'} P(A = a') P(Y \mid M = m, A = a')$$

```r
# Probability of Y = 1 given M and A
p_Y_given_M_A <- data %>%
  group_by(A,M) %>%
  summarize(P_Y_given_A_M = mean(Y),
            .groups = "drop")
```

Translating math to code

$$\mathrm{P}(Y^a) = \sum_m \mathrm{P}(M = m \mid A = a) \sum_{a'} \mathrm{P}(A = a')\mathrm{P}(Y \mid M = m, A = a')$$

Translating math to code

$$P(Y^a) = \sum_m P(M = m \mid A = a) \sum_{a'} P(A = a') P(Y \mid M = m, A = a')$$

```
# Probability of Y = 1 under intervention on M
p_Y_under_M <- p_Y_given_M_A %>%
  left_join(p_A, by = "A") %>%
  group_by(M) %>%
  summarize(p_Y_under_M = sum(P_Y_given_A_M  * p_A))
```

Translating math to code

$$\mathrm{P}(Y^a) = \sum_m \mathrm{P}(M = m \mid A = a) \sum_{a'} \mathrm{P}(A = a')\mathrm{P}(Y \mid M = m, A = a')$$

Translating math to code

$$P(Y^a) = \sum_m P(M = m \mid A = a) \sum_{a'} P(A = a')P(Y \mid M = m, A = a')$$

```
# Probability of each M given A
p_M_given_A <- data %>%
  # Count size of each group
  group_by(A, M) %>%
  count() %>%
  # Convert to probability within A
  group_by(A) %>%
  mutate(p_M_under_A = n / sum(n)) %>%
  select(A,M,p_M_under_A)
```

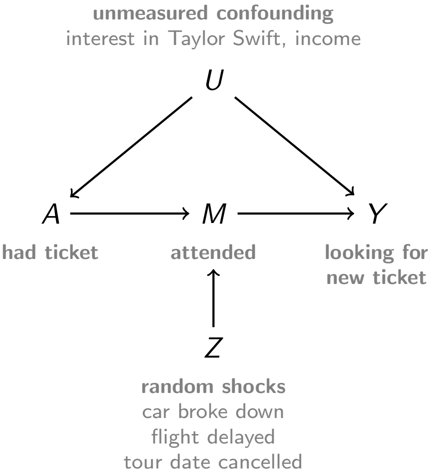
Translating math to code

$$P(Y^a) = \sum_m P(M = m \mid A = a) \sum_{a'} P(A = a') P(Y \mid M = m, A = a')$$

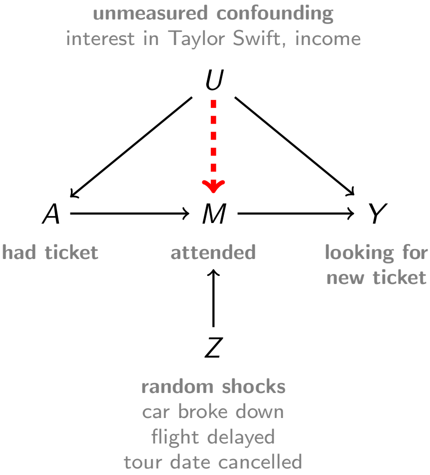
```
# Front door identification
# Probability of Y = 1 under intervention on A
p_Y_under_A <- p_M_given_A %>%
  left_join(p_Y_under_M,
            by = "M") %>%
  group_by(A) %>%
  summarize(estimate = sum(p_M_under_A * p_Y_under_M))
```

Goal 3) Critique the identification assumptions

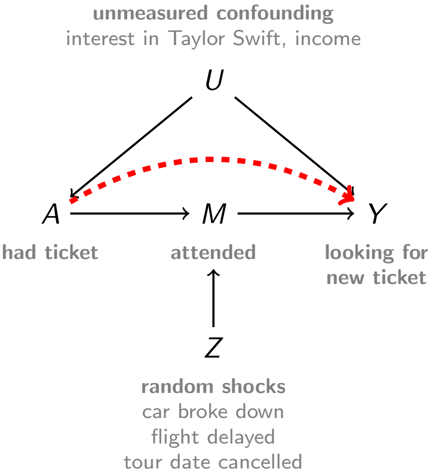
What edges might need to be added to this DAG?



What edges might need to be added to this DAG?



What edges might need to be added to this DAG?



Resources

- Pearl, J. (1995). Causal diagrams for empirical research. Biometrika, 82(4), 669-688.
- Glynn, A. N., & Kashin, K. (2018). Front-door versus back-door adjustment with unmeasured confounding: Bias formulas for front-door and hybrid adjustments with application to a job training program. Journal of the American Statistical Association, 113(523), 1040-1049.

Learning goals for today

At the end of class you will be able to

- explain front-door causal identification

More broadly,

1. engage with a new causal identification approach
2. translate that method to code
3. critique the identification assumptions