Lecture 11: Causality

Jacob M. Montgomery

Quantitative Political Methodology

Causality

Roadmap

The road so far:

- We learned how to collect data to make inferences about a population
- We have characterized a single variable

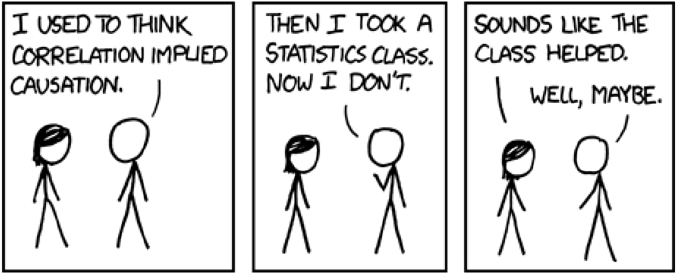
Roadmap

The road so far:

- We learned how to collect data to make inferences about a population
- We have characterized a single variable

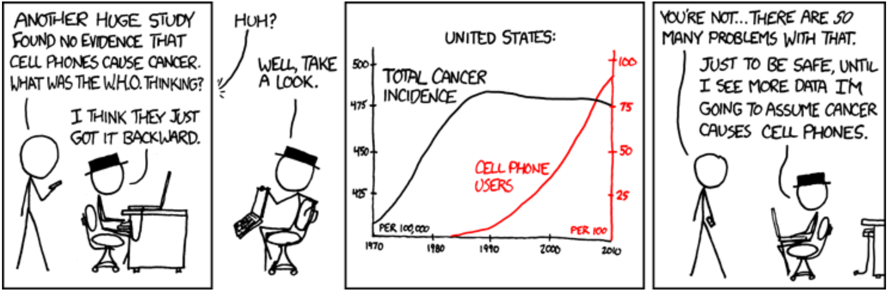
This class:

- Now we want to look at two variables
- Our specific aim is to understand if X causes Y


I USED TO THINK CORRELATION IMPLIED CAUSATION.
THEN I TOOK A STATISTICS CLASS. NOW I DON'T.
SOUNDS LIKE THE CLASS HELPED.
WELL, MAYBE.


STUDIES SHOW THAT PEOPLE WHO EXERCISE ARE HEALTHIER.
Dilbert.com DilbertCartoonist@gmail.com
THAT'S BECAUSE PEOPLE WHO ARE IN POOR HEALTH DON'T EXERCISE.
5-9-14 ©2014 Scott Adams, Inc./Dist. by Universal Uclick
WHY DOES IT SEEM AS IF YOU RUIN EVERY MEETING?
IS IT BECAUSE I ONLY ATTEND THE ONES THAT ARE STUPID?

Goals for today



- Thinking about causality
- Average treatment effects

What is causality?

In political science we want to make causal claims.

$$X \rightarrow Y$$

What is causality?

In political science we want to make causal claims.

$$X \to Y$$

What does this mean?

What is causality?

In political science we want to make causal claims.

$$X \to Y$$

What does this mean? Let's do this a bit more formally for the case of an experiment (the easiest way to think about it).

We will use T to represent a **treatment variable**.

- For a categorical treatment

$$T_i = \begin{matrix} 1 \text{ if unit i receives the treatment} \\ 0 \text{ if unit i receives the control} \end{matrix}$$

We will use T to represent a **treatment variable**.

- For a categorical treatment

$$T_i = \begin{array}{l} 1 \text{ if unit i receives the treatment} \\ 0 \text{ if unit i receives the control} \end{array}$$

- We let y_i^1 represent the outcome of the ith unit **if the treatment is given**.

We will use T to represent a **treatment variable**.

- For a categorical treatment

$$T_i = \begin{matrix} \text{1 if unit i receives the treatment} \\ \text{0 if unit i receives the control} \end{matrix}$$

- We let y_i^1 represent the outcome of the ith unit **if the treatment is given**.
- We let y_i^0 represent the outcome of the ith unit **if the control is given**.pause
- One of these is observed, the other is the **counterfactual** – what would have been observed if the other treatment have been given?

We will use T to represent a **treatment variable**.

- For a categorical treatment

$$T_i = \begin{matrix} \text{1 if unit i receives the treatment} \\ \text{0 if unit i receives the control} \end{matrix}$$

- We let y_i^1 represent the outcome of the ith unit **if the treatment is given**.
- We let y_i^0 represent the outcome of the ith unit **if the control is given**.pause
- One of these is observed, the other is the **counterfactual** – what would have been observed if the other treatment have been given?
- The causal effect of T_i will then be $y_i^1 - y_i^0$
  - Ex., "My theory is that individuals who watched this TV ad will be more likely to vote for Ted Cruz than if they didn't watch it."

Average treatment effects

- We cannot measure individual level causal effects

Average treatment effects

- We cannot measure individual level causal effects
- We can estimate the **population average treatment effect** by looking at those who received the treatment and those who did not.
    - $ATE = mean(y_i^1 - y_i^0)$

Average treatment effects

- We cannot measure individual level causal effects
- We can estimate the **population average treatment effect** by looking at those who received the treatment and those who did not.
  - $ATE = mean(y_i^1 - y_i^0)$
  - $ATE = mean(y_i^1) - mean(y_i^0)$

Average treatment effects

- We cannot measure individual level causal effects
- We can estimate the **population average treatment effect** by looking at those who received the treatment and those who did not.
  - $ATE = mean(y_i^1 - y_i^0)$
  - $ATE = mean(y_i^1) - mean(y_i^0)$
- Each **group** acts as a counterfactual for the other
  - Ex., "My theory is that those individuals who watched this TV ad will be more likely to vote for Mitt Romney on average than those who didn't watch it."

The fundamental problem of causal inference

- The fundamental problem of causal inference is that at most one of y_i^0 and y_i^1 can be observed.

The fundamental problem of causal inference

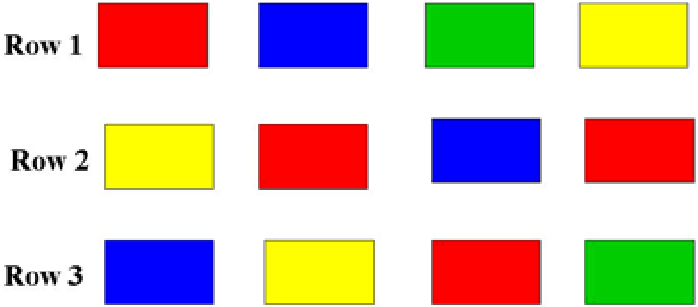
- The fundamental problem of causal inference is that at most one of y_i^0 and y_i^1 can be observed.
- We can think of each of these as potential outcomes. However, we can only observe one. The other is the counterfactual.

The fundamental problem of causal inference

- The fundamental problem of causal inference is that at most one of y_i^0 and y_i^1 can be observed.
- We can think of each of these as potential outcomes. However, we can only observe one. The other is the counterfactual.
- Estimation of causal effects requires some combination of:
  - certain research designs that approximate potential outcomes
  - randomization
  - statistical adjustment

Demonstration: Stroop Test

State the colors as fast as you can



From John Gosbee, MD, MS, VA National Center for Patient Safety

Stand up if your student ID ends in:

1 3 4 6 9

Now state the colors as fast as you can

Row 1	**Red**	**Blue**	**Green**	**Yellow**
Row 2	**Yellow**	**Green**	**Blue**	**Red**
Row 3	**Green**	**Red**	**Yellow**	**Blue**

From John Gosbee, MD, MS, VA National Center for Patient Safety

Again, state the colors as fast as you can

Row 1	Red	Blue	Green	Yellow
Row 2	Yellow	Green	Blue	Red
Row 3	Green	Red	Yellow	Blue

From John Gosbee, MD, MS, VA National Center for Patient Safety

Take away

- Causal effects rely on unobserved counterfactuals

Take away

- Causal effects rely on unobserved counterfactuals
- At best, we can estimate average treatment effects comparing those who received a treatment with those who don't.

Take away

- Causal effects rely on unobserved counterfactuals
- At best, we can estimate average treatment effects comparing those who received a treatment with those who don't.
- In order for this to work, each group must be *identical* (on average) in every way except the treatment.

Take away

- Causal effects rely on unobserved counterfactuals
- At best, we can estimate average treatment effects comparing those who received a treatment with those who don't.
- In order for this to work, each group must be *identical* (on average) in every way except the treatment.
- The best way to achieve this is through random assignment (i.e., experiments)
- What if this assumption is not met?

Confounders and causality

- PROBLEM: This only works if the two groups are, on average, otherwise identical
- If the two groups differ on other factors that also cause y_i^1 and y_i^0, this is a confounding relationship.
- If this is the case, our counterfactual is **wrong** and we can make no causal claim.

*If you aren't doing something to handle **all** other relevant variables (through randomization or statistical methods), you cannot make a valid causal claim.*

Thinking about confounding variables

Direct causal relationships:

$$X \to Y$$

Thinking about confounding variables

Direct causal relationships:

$$X \to Y$$

Spurious relationships:

Thinking about confounding variables

Direct causal relationships:

$$X \rightarrow Y$$

Spurious relationships:

$$Z \rightarrow X \text{ AND } Z \rightarrow Y$$

Thinking about confounding variables

Direct causal relationships:

$$X \to Y$$

Spurious relationships:

$$Z \to X \text{ AND } Z \to Y$$

Chain relationships:

Thinking about confounding variables

Direct causal relationships:

$$X \to Y$$

Spurious relationships:

$$Z \to X \text{ AND } Z \to Y$$

Chain relationships:

$$X \to Z \to Y$$

Multiple causation:

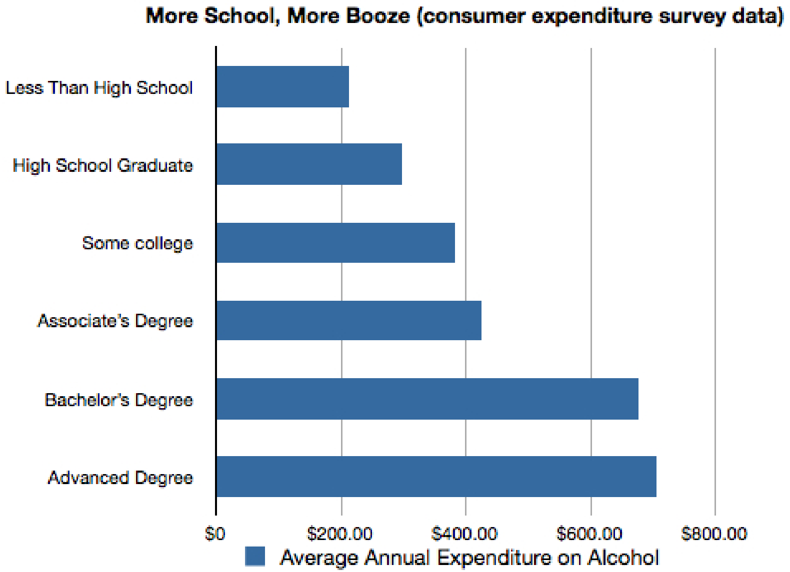
$$X \to Y \text{ AND } Z \to Y$$

Multiple causation:

$$X \to Y \text{ AND } Z \to Y$$

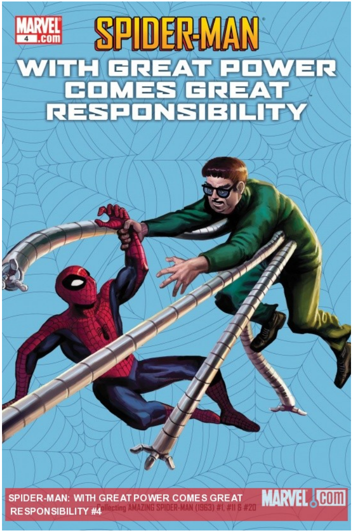
Direct and indirect causation:

$$X \to Y \text{ AND } X \to Z \text{ AND } Z \to Y$$



Write down one of each type of claim for this data.

Being a responsible causal analyst



More seriously

- Vitamin C?

More seriously

- Vitamin C?
- Flossing?

More seriously

- Vitamin C?
- Flossing?
- Gluten?

More seriously

- Vitamin C?
- Flossing?
- Gluten?
- Drinking while pregnant?

More seriously

- Vitamin C?
- Flossing?
- Gluten?
- Drinking while pregnant?
- Fox News?

At a minimum we need to show . . .

Association

- What we will be doing this rest of the semester
- Correlation, contingency tables, regression coefficients, . . .
- Association $\neq$ causation

At a minimum we need to show ...

Association

- What we will be doing this rest of the semester
- Correlation, contingency tables, regression coefficients, ...
- Association $\neq$ causation

Temporal order

- For T_i to cause Y_i it must come before Y in time order
- **Post hoc ergo propter hoc**
- "After this, therefore because of this""
- Temporal order does $\neq$ causation (e.g., every superstition ever)

Eliminate alternative explanations

- Suppose there is an association and a proper time order. We still cannot infer causation.
- Rather, we must test for all alternative explanations.
- Only if all of these have been resolved can we claim causation.

Eliminate alternative explanations

- Suppose there is an association and a proper time order. We still cannot infer causation.
- Rather, we must test for all alternative explanations.
- Only if all of these have been resolved can we claim causation.

How can we do this?

- Experimental control
- Statistical control (Stay tuned . . .)
- Research design

Problem: It can be subtle

(Mount 2010)

THE GREAT DIVIDE

Parental Involvement Is Overrated

By KEITH ROBINSON and ANGEL L. HARRIS APRIL 12, 2014 2:32 PM

Don't Help Your Kids With Their Homework

And other insights from a ground- breaking study of how parents impact children's academic achievement

66k

TEXT SIZE − +

DANA GOLDSTEIN | APRIL 2014 ISSUE | EDUCATION

Rogers, Coffman, and Bergman:

> *While the authors control for certain variables, their research only implies there is a relationship between parental involvement and student performance. This caveat is important;* ***the existence of a relationship does not tell us what causes what****.*
>
> *Think of it this way:* ***If you had two children, and one was getting A's and the other C's, which of them would you help more? The C student.*** *An outsider, noticing that you've spent the school year helping only one of your children, might infer that parental help caused that child to earn lower grades. This of course would not be the case, and inferring causation here would be a mistake.*

If you see a surprising result, be skeptical

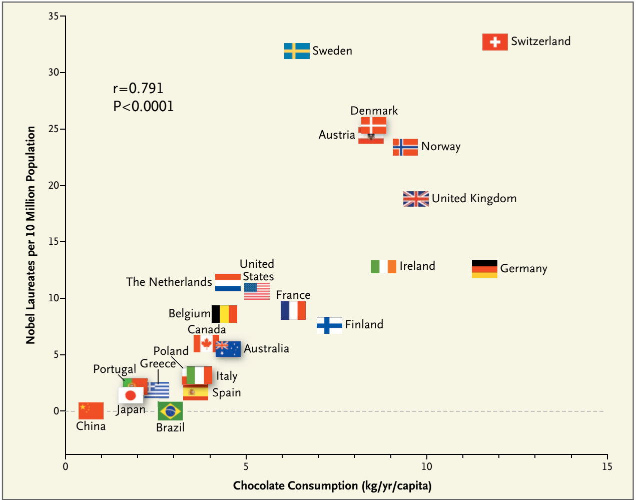
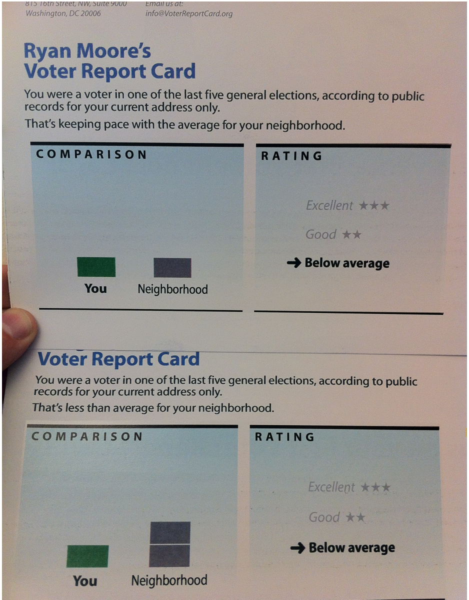


Figure 1. **Correlation between Countries' Annual Per Capita Chocolate Consumption and the Number of Nobel Laureates per 10 Million Population.**

Example: Can social pressure increase turnout?



We have two **independent samples**, and we want to compare them.

We have two **independent samples**, and we want to compare them. Our data will look like this.

Variable 1 (Outcome or response)	Variable 2 (Explanatory or grouping)
1	1
0	0
1	0
0	1
$\vdots$	$\vdots$

Example: Social Pressure and Turnout

Gerber, A., Green, D., and Larimer C.W. 2008 “Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment.” *American Political Science Review* 101(1): 33-48.

Example: Social Pressure and Turnout

Gerber, A., Green, D., and Larimer C.W. 2008 "Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment." *American Political Science Review* 101(1): 33-48.

TABLE 2. Effects of Four Mail Treatments on Voter Turnout in the August 2006 Primary Election

	Experimental Group				
	Control	Civic Duty	Hawthorne	Self	Neighbors
Percentage Voting	29.7%	31.5%	32.2%	34.5%	37.8%
N of Individuals	191,243	38,218	38,204	38,218	38,201

Do Politicians Racially Discriminate?

- Is racial discrimination against blacks still a problem in the political sphere?
- Do legislators discriminate against individual requests for constituency service on the basis of race?

Comparing legislators' responsiveness

Questions: Do legislators' answer a higher proportion of emails from the citizens' they believe are white,

- ... even though both black & white don't signal party affiliation?
- ... even though both black & white signal being Republican?
- ... even though both black & white signal being Democrat?

Experiment: The sample includes states legislators in 44 U.S. states with a valid e-mail address in September 2008.

- Race was signaled by randomizing whether the email was signed and sent from an email account with the name Jake Mueller or DeShawn Jackson.
- The text of the email was also manipulated so as to signal the partisan preference of the email sender.
- The cross-tabulation between *race & partisan preference* gives six treatments (or groups).
- The outcome variable is the response (or lack thereof) to any e-mails.

TABLE 1 Overall Effect Sizes—Does Jake Receive More Replies Than DeShawn?

	No Partisanship Signal	Republican Signal	Democratic Signal	Party Differential	
DeShawn	55.3%	54.3%	57.3%	−2.9%	Combined
Jackson	N = 806	N = 810	N = 812	($p = 0.23$)	**−0.9%**
Jake	60.5%	56.4%	55.3%	1.1%	**($p = 0.61$)**
Mueller	N = 812	N = 820	N = 799	($p = 0.31$)	
Race Differential	?	−2.1%	1.9%		
	($p =$?)	($p = 0.39$)	($p = 0.43$)		
		Combined Effect			
		−0.1% ($p = 0.95$)			

Class business

- Midterms
- PS posted.
- Read online content