Causation & Randomized Experiments

Lachlan Deer

Social Media and Web Analytics, Spring 2024

Learning Goals

- Explain the challenges of causal inference with observational data
- Define the term randomized experiment
- Explain the fundamental problem of causal inference
- Describe the Potential Outcomes framework for casual inference
- Define the Sample Average Treatment Effect
- Analyze data from a randomized experiment to estimate the sample average treatment effect of an intervention using statistical inference and linear regression

1/ Causation

Causal Questions

Does X cause Y? Examples of causal questions include:

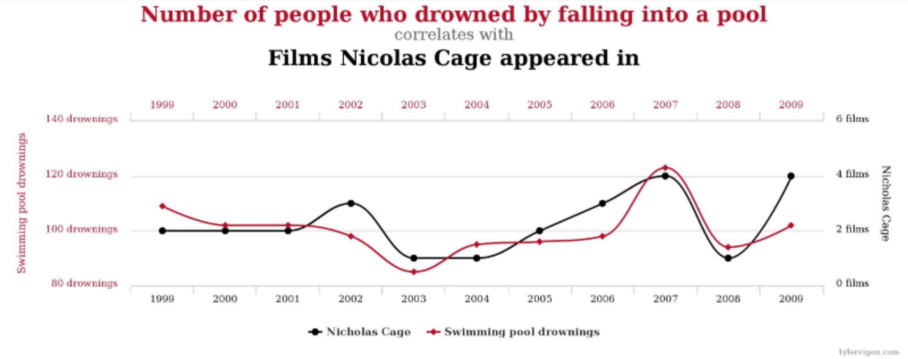
- Does smoking cause cancer?
- Does exercise make people happier?
- Does my social media advertising increase sales?
- Does hiring an influencer to promote a product lead to an increased consumer awareness about the product?

Not all causal questions use the word "*cause*". Other words that imply causality include:

- Improve
- Increase / decrease
- Lead to
- Make

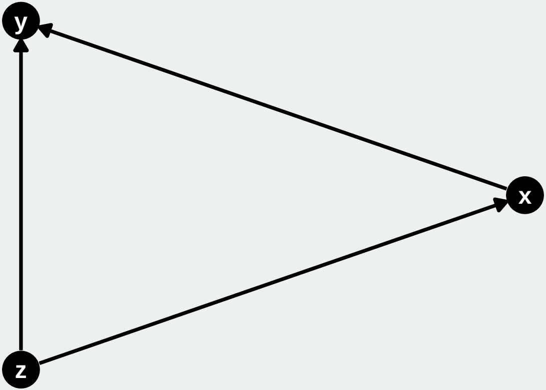
Tell-tale sign that a question is causal: analysis is used to make an argument for changing a procedure, policy, or practice.

Association vs Causation

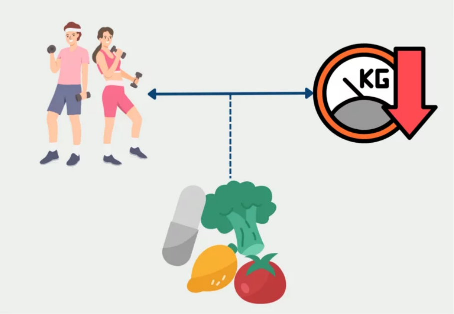


The Difficulty of Casual Effects in Observational Data

Challenge 1: Omitted Variables (Z): variable that affects both X & Y that is not included in the analysis



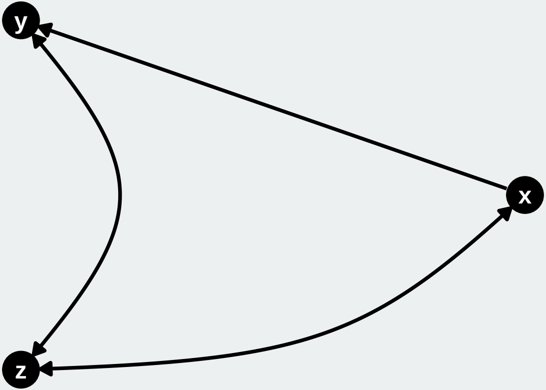
Omitted Variable Bias



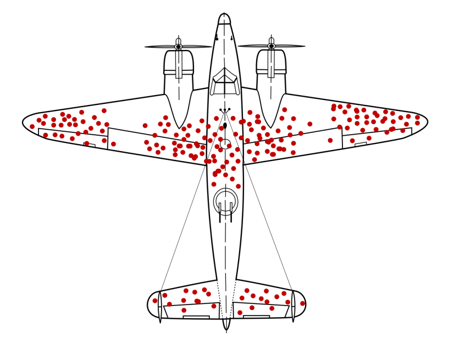
Does exercise cause weightloss?

The Difficulty of Casual Effects in Observational Data

Challenge 2: Selection Effects: improper (non random) selection of individuals, such that the sample of cases and controls are not drawn from the same reference population



Selection Effects



Should we use to plan where to put additional armor on fighter jets based on the damage of planes that return to base?

Solutions to OVB and Selection Effects

Omitted Variable Bias and Selection Effects mean **our estimates** of the effect of X on Y **are biased**

Possible Solutions:

1. Modelling and/or assumptions
2. Randomization of the intervention

2/ Randomization

What is Randomization?

When we use the word **random** in this context, we mean:

- Every unit has some chance (i.e., a non-zero probability) of being selected to receive the intervention or be in the control group.
- The selection into these groups is based upon a random process

The "Magic" of Randomization

Because our unit of analyses are **randomized to treatment and control groups**, **on average** there is **no difference** between these two groups on any **characteristics** other than their treatment.

Prior to treatment, on average the groups (Treatment and Control) are equivalent to one another on every observed and unobserved variable

- There is no omitted variable bias
- There are no selection effects

Can we check for randomization?

3/ Application: Encouraging Donors to Share About Charity

Put Your Mouth Where Your Money Is: A Field Experiment Encouraging Donors to Share About Charity

Ike Silver,[a,*] **Deborah A. Small**[b]

[a] Kellogg Graduate School of Management, Northwestern University, Evanston, Illinois 60208; [b] Yale School of Management, Yale University, New Haven, Connecticut 06511

*Corresponding author

Contact: ike.silver@kellogg.northwestern.edu, https://orcid.org/0000-0002-7206-8018 (IS); deborah.small@yale.edu (DAS)

Received: February 9, 2021
Revised: May 7, 2022; February 17, 2023
Accepted: April 13, 2023
Published Online in Articles in Advance: August 8, 2023

https://doi.org/10.1287/mksc.2023.1450

Abstract. Sharing about charity online or in personal conversations can help raise awareness and bolster fundraising efforts for good causes. However, when deciding whether to tell others about their charitable giving, donors may focus more on possible risks to their reputation (e.g., of seeming braggy, inauthentic) than on potential word-of-mouth benefits for the charity. In a large, preregistered field experiment, we tested a post-donation intervention designed to encourage word-of-mouth by reorienting donors to the idea that sharing about charity means doing more good; 77,485 donors received either a control or treatment message asking them to share a link to the cause via social media, text, or email. Compared with the organization's standard solicitation ("Please share your donation … "), our intervention emphasized consequences of sharing *for the cause* ("Your donation can start a chain reaction … "). This brief message increased click-through by 5.1% and likelihood of recruiting at least one later donation via word-of-mouth by 12.4%. Exploratory follow-up analyses suggest that these effects are most pronounced among larger-gift donors; the more donors gave, the more responsive they were to the intervention. Whereas many field experiments aim to increase giving directly, we test an intervention designed to boost word-of-mouth for worthy causes. We discuss approaches for encouraging sharing in the domain of charity and beyond.

History: Olivier Toubia served as the senior editor for this article.
Supplemental Material: The e-companion and data are available at https://doi.org/10.1287/mksc.2023.1450.

Keywords: field experiments • charitable giving • word-of-mouth • referral marketing • impression management

Intervention Context

- **Research Question**: Can we effectively get donors to share about charitable donations?
- **Why is this relevant?**
  - Raises awareness and bolsters fundraising efforts for good causes
- This isn't easy: Donors face a trade off
  - -ve: (Personal) Reputation risks via appearing braggy or inauthentic, vs.
  - +ve: (External) Word of Mouth benefits to the charity
- Today we'll explore this question in the context of charitable giving for educational projects, and an intervention that encourages sharing about the charity after donation
  - $Y_i =$ clicking on a sharing pop-up OR recruiting future donors (0/1)
  - $T_i =$ an intervention encouraging sharing about cause post-donation (0/1)

Importing the Data

```r
charity <-
    read_csv("data/exp2data.csv") %>%
    mutate(condition = if_else(condition ==1,
                               "treatment",
                               "control"
                               )
    )
```

Inspecting the Data

```r
head(charity, n = 5)
```

```
# A tibble: 5 x 13
  user_id donated condition clickthrough recruited raised donatedsince     n
    <dbl>   <dbl> <chr>            <dbl>     <dbl>  <dbl>        <dbl> <dbl>
1      14      50 control              0         0      0            1     1
2     208     423 control              1         0      0            0     9
3     717      50 control              0         0      0            1     1
4     784      32 treatment            0         0      0            1     3
5     879      50 treatment            0         0      0            0     2
# i 5 more variables: propdevice <dbl>, numdevice <dbl>, device <chr>,
#   firstvisit <dbl>, houroffirst <dbl>
```

What is our data?

Units: objects being studied.

- Usually the rows of the data set.
- Examples: Survey respondents, consumers, firms, app users, influencers.
- Today's data: Charity Donors from DonorsChoose.org

Variables: measurements that can vary across units.

- Usually the columns of a data set.
- Examples: quantity bought, dollars spent, participation in an experiment.
- Today's data: Amount donated, clicked through pop up, referred a friend, treatment status,

4/ Estimating the Treatment Effect

Causal Questions

Does cause based solicitation *increase* sharing?

Does cause based solicitation recruit *more* future donors?

$\rightarrow$ Comparison between factual and counterfactual

Fundamental problem of causal inference: Analyst must infer counterfactual outcomes

A Tale of Two Donors

```
# A tibble: 2 x 13
  user_id donated condition clickthrough recruited raised donatedsince     n
    <dbl>   <dbl> <chr>            <dbl>     <dbl>  <dbl>        <dbl> <dbl>
1   13498    797. control              0         0      0            1     6
2   13634    160. treatment            1         0      0            1     1
# i 5 more variables: propdevice <dbl>, numdevice <dbl>, device <chr>,
#   firstvisit <dbl>, houroffirst <dbl>
```

Did donor 13498 not share about their donation **because** they received a standard solicitation?

Notation

- **Unit** (indexed by i): individuals who have donated
- **Treatment variable** (causal variable of interest) T_i
  - Received solicitation that emphasizes consequences of sharing for a cause
- **Treatment group** (treated units): Donors who recieve a "standard" solicitation
- **Control group** (untreated units): Donors who receive the "share for a cause" solicitation
- **Outcome variable(s)** (response variable) Y_i
  - `click-through`: Donor clicks on a pop-up encouraging sharing
  - `recruitment`: Did donor's sharing lead to subsequent donation

Notation

	T_i	Y_i (click-through)
Donor saw standard solicitation	0	0
Donor saw solicitation for a cause	1	1

Causal Effects and Counterfactuals

- What does "T_i causes Y_i mean? $\rightsquigarrow$ **counterfactuals**,"*what if*"
  - Would donor *i* who saw the standard solicitation have clicked through if they saw the one that emphasized the cause
- Two **potential outcomes**:
  - $Y_i(1)$: would donor click through if they saw cause based solicitation?
  - $Y_i(0)$: would donor click through if they saw standard solicitation??
- **Causal effect**: $Y_i(1) - Y_i(0)$
- **Fundamental problem of causal inference**: only one of the two potential outcomes is observable per observation.

Potential Outcomes

	T_i	Y_i (click-through)	$Y_i(0)$	$Y_i(1)$
Donor saw standard solicitation	0	0	0	???
Donor saw solicitation for a cause	1	1	???	1

- **Association is not causation**
- How can we infer the missing counterfactuals?

How to Figure Out Counterfactuals

- **Need to find similar observations**!
- Sounds easy ... but
  - Harder than it sounds, and
  - Requires assumptions
- **Randomized experiments** are one possible solution

Summation Notation

- Define the **sample size** (number of observations) as n
- Therefore, we have n measurements of some variable, $(Y_1, Y_2, \ldots, T_n)$
- We'll want to refer to the sum of these variables:

$$Y_1 + Y_2 + Y_3 + Y_4 + \ldots + Y_{n-1} + Y_n$$

- This is cumbersome to write down, so we'll use the *sigma* notation

$$\sum_{i=1}^{n} Y_i = Y_1 + Y_2 + Y_3 + Y_4 + \ldots + Y_{n-1} + Y_n$$

* $\sum_{i=1}^{n} Y_i$ says 1. Initialize the running sum to the case when i = 1. 2. Increment i by 1 and add the new expression to the running sum. 3. Repeat step 2 until $i = n$.

Averages

- The **sample average** or **sample mean** is simply the sum of all values divided by the number of values

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

- Suppose we surveyed six people, and 3 of them donated 20 dollars:

$$\bar{Y} = \frac{1}{6}(20 + 20 + 20 + 0 + 0 + 0) = 10$$

Quantity of Interest

We want to estimate the average causal effect over all units:

Sample Average Treatment Effect (SATE) $= \frac{1}{n}\sum_{i=1}^{n}[Y_i(1) - Y_i(0)]$

What we can estimate instead:

Difference in Means $= \bar{Y}_{\text{Treated}} - \bar{Y}_{\text{Control}}$

where:

- $\bar{Y}_{\text{Treated}}$ is the observed average outcome in the treatment group
- $\bar{Y}_{\text{Control}}$ is the observed average outcome in the control group
- How do we ensure that the difference in means is a good estimate of the SATE?

Randomized Control Trials

Randomization!

Recall **Randomization** of the treatment makes the treatment and control groups "identical" on average.

- The two groups are similar in terms of all characteristics (both observed and unobserved).
  - Control group is similar to treatment group
  - $\rightsquigarrow$ outcome in control group $\approx$ what would have happened to treatment group if they had been in control group

Potential Problems with RCTs

- **Placebo effects**:
  - Respondents will be affected by any intervention, even if they shouldn't have any effect.
- **Hawthorne effects**:
  - Respondents act differently just knowing that they are under study.

Did the Randomization *actually* Randomize?

If it did, **we shouldn't see large differences** between treatment and control group on **pre-treatment variables**.

- Called "balance checking"
- Pre-treatment variable are those that are unaffected by treatment.
- We can check in the actual data for some pre-treatment variable X
  - $\bar{X}_{Treated}$: average value of variable for treated group.
  - $\bar{X}_{Control}$: average value of variable for control group.
- Under randomization, $\bar{X}_{Treated} - \bar{X}_{Control} \approx 0$

Multiple Treatments

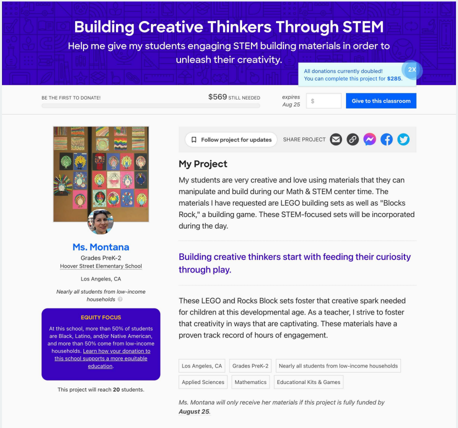
- Instead of 1 treatment, we might have **multiple treatment arms**:
    - Control condition
    - Treatment A
    - Treatment B
    - Treatment C, etc
- In this case, we will look at multiple comparisons:
    - $\bar{Y}_{Treated,A} - \bar{Y}_{Control}$
    - $\bar{Y}_{Treated,B} - \bar{Y}_{Control}$
    - $\bar{Y}_{Treated,A} - \bar{Y}_{Treated,B}$

5/ Application: Encouraging Donors to Share About Charity

Study Design: Setting

When: Four-week period from August 13, 2020, to September 9, 2020

Where: DonorsChoose.org



Study Design: Treatments

Control Condition:

“Share this classroom with family and friends”

Treatment Condition:

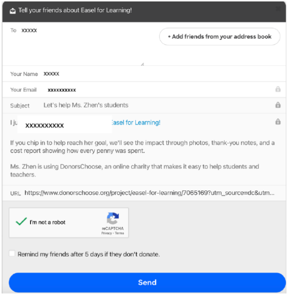
“Your donation can start a chain reaction, but only if you tell others about the cause. Share this classroom with family and friends”

Study Design: Sharing Contributions

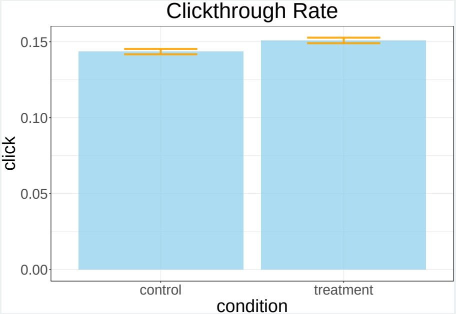
Facebook (post):

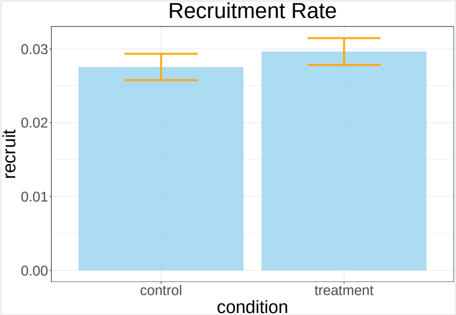
Twitter:

Gmail:

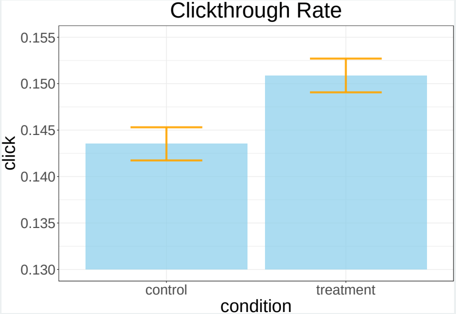


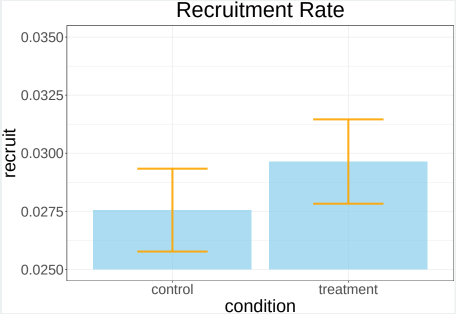
Visualizing the Outcomes





Visualizing the Data ... and Cutting the axis





Computing Proportions by Treatment

```r
charity %>%
    group_by(condition) %>%
    summarize(click = mean(clickthrough),
              recruit = mean(recruited))
```

```
# A tibble: 2 x 3
  condition click recruit
  <chr>     <dbl>   <dbl>
1 control   0.144  0.0275
2 treatment 0.151  0.0296
```

Estimating SATEs: Clickthroughs

What's the correct statistical test?

```
# A tibble: 1 x 6
  statistic chisq_df p_value alternative lower_ci upper_ci
      <dbl>    <dbl>   <dbl> <chr>          <dbl>    <dbl>
1      8.30        1 0.00395 two.sided    -0.0124 -0.00235
```

Estimating SATEs: Recruitment

What's the correct statistical test?

```
# A tibble: 1 x 6
  statistic chisq_df p_value alternative lower_ci  upper_ci
      <dbl>    <dbl>   <dbl> <chr>          <dbl>     <dbl>
1      5.04        1  0.0248 two.sided   -0.00418 -0.000280
```

Regression as Mean Estimation

- Let W_i be defined as follows:

$$W_i = \begin{cases} 1 & \text{if } i \text{ in the Treatment Group} \\ 0 & \text{if } i \text{ in the Control Group} \end{cases}$$

Then we can write:

$$Y_i = Y_i(0) + W_i(Y_i(1) - Y_i(0))$$

Regression as Mean Estimation

Recall the the definition of $\bar{Y}_{Control}$:

$$\bar{Y}_{Control} = \frac{1}{n_0}\sum_{i=0}^{n_0} Y_i(0)$$
$$= \hat{\beta}_0$$

This is a consistent estimator of the expected value:

$$\beta_0 = E[Y_i(0)]$$

Regression as Mean Estimation

Then, recall our definition of difference in means:

$$\begin{aligned}\text{Difference in Means} &= \bar{Y}_{\text{Treated}} - \bar{Y}_{\text{Control}} \\ &= \frac{1}{n_1}\sum_{i=1}^{n_1} Y_i(1) - \frac{1}{n_0}\sum_{i=1}^{n_0} Y_i(0) \\ &= \hat{\beta}_1\end{aligned}$$

which is a consistent estimator of the difference in means at the population level

$$\beta_1 = E[Y_i(1) - Y_i(0)]$$

Regression as Mean Estimation

Consider the regression equation:

$$Y_i = \beta_0 + W_i\beta_1 + \varepsilon_i$$

Then take expectations condition on treatment assignment:

$$E(Y_i) = \beta_0 + W_i\beta_1$$

This implies that we can estimate the ATE of a binary treatment via a linear regression of observed outcomes Y_i on a vector consisting of intercept and treatment assignment, $(1, W_i)$.

Regression-based SATE: clickthrough

$$clickthrough_i = \beta_0 + \beta_1 Condition_i + \varepsilon_i$$

```
# A tibble: 2 x 5
  term               estimate std.error statistic p.value
  <chr>                 <dbl>     <dbl>     <dbl>   <dbl>
1 (Intercept)         0.144     0.00180     79.6  0
2 conditiontreatment  0.00736   0.00255      2.89 0.00383
```

In class: Interpret these coefficients.

Regression-based SATE

$$recruit_i = \beta_0 + \beta_1 Condition_i + \varepsilon_i$$

```
# A tibble: 2 x 5
  term               estimate std.error statistic   p.value
  <chr>                 <dbl>     <dbl>     <dbl>     <dbl>
1 (Intercept)          0.0179  0.000695     25.8  8.26e-146
2 conditiontreatment  0.00223  0.000982      2.27 2.32e-  2
```

In class: Interpret these coefficients.

A Balance Test?

The data provided by the authors **does not contain any information on pre-experiment variables.**

Thus, we **cannot test for balance** between control and treatment groups

Question: What pre-experiment variables would you want to use to test for balance?

6/ Wrap Up

Summary

- Causal inference with observational data is difficult due to omitted variable bias and selection effects
- Randomized Control Trials solve these issues and allow us to estimate a Sample Average Treatment Effect (SATE) by randomly allocating units of observation to either an intervention or a control condition
- SATE can be estimated from classical tools for statistical inference: t-tests, proportions tests and linear regression

License & Citation

Suggested Citation:

```
@misc{smwa2024_randomized,
      title={"Social Media and Web Analytics: Casuation and
             Randomized Experiments"},
      author={Lachlan Deer},
      year={2024},
      url = "https://tisem-digital-marketing.github.io/2024-smwa"
}
```

This course adheres to the principles of the Open Science Community of Tilburg University. This initiative advocates for transparency and accessibility in research and teaching to all levels of society and thus creating more accountability and impact.