SUTVA, Assignment Mechanism

STA 320
Design and Analysis of Causal Studies
Dr. Kari Lock Morgan and Dr. Fan Li
Department of Statistical Science
Duke University

Review

- Causality is tied to an action (treatment)
- Potential outcomes represent the outcome for each unit under treatment and control
- A causal effect compares the potential outcome under treatment to the potential outcome under control for each unit
- In reality, only one potential outcome observed for each unit, so need multiple units to estimate causal effects

SUTVA

- Often in statistics we have to make assumptions, without which the statistical theory does not hold
- In causal inference, we usually make the Stable Unit Treatment Value Assumption (SUTVA)

SUTVA

- SUTVA has two components:
  1. *No interference* (units do not interfere with each other): treatment applied to one unit does not effect the outcome for another unit
  2. There is only a single version of each treatment level (potential outcomes must be well defined)
- Think of examples for which each component does NOT hold.

SUTVA and Potential Outcomes

- For the potential outcomes to be well-defined, SUTVA must hold
- $Y(\text{treatment}) = Y_i(\text{treatment}_i)$
- If either component of SUTVA is not satisfied, then the potential outcomes are not uniquely defined
- If SUTVA is not a reasonable assumption, causal effects are hard to even define, and estimates have limited credibility

Study Design

- Studies can be designed to help SUTVA be a valid assumption
- One options: change the unit of analysis (i.e. students to classes or schools, individuals to communities)
- If possible, different locations may help

Exclusion Restrictions

- SUTVA is an exclusion restriction
- Exclusion restrictions *exclude* various possibilities, usually to make causal inference feasible
- SUTVA excludes the possibilities of units interfering with each other and multiple versions of a treatment
- Exclusion restrictions cannot be verified from the data; they are based entirely on previous subject matter knowledge

Notation

- Y: outcome
- W: treatment, W in {0, 1}
- i = subscript for each unit, i in {1, …, N}
- Y_i^{obs} = observed outcome for unit i
  = $Y_i(W_i)$
  = $W_i Y_i(1) + (1 – W_i) Y_i(0)$
- Y_i^{mis} = missing outcome for unit i
  = $Y_i(1- W_i)$
  = $(1 – W_i) Y_i(1) + W_i Y_i(0)$

Notation

- For treated units, $W_i = 1$:
  - $Y_i^{obs} = Y(1)$ = outcome under treatment
  - $Y_i^{mis} = Y(0)$ = outcome under control
- For control units, $W_i = 0$:
  - $Y_i^{obs} = Y(0)$ = outcome under control
  - $Y_i^{mis} = Y(1)$ = outcome under treatment

Average Causal Effect

- We often want to estimate the average causal effect:

$$\overline{Y(1)} - \overline{Y(0)} = \frac{\sum_{i=1}^{n} (Y_i(1) - Y_i(0))}{n}$$

Average Causal Effect

- Since we don't have both potential outcomes, a tempting replacement is to use $Y_i^{obs} \mid W_i = 1$ and $Y_i^{obs} \mid W_i = 0$ in place of $Y_i(1)$ and $Y_i(0)$:

$$\frac{\sum_{i=1}^{n} W_i Y_i^{obs}}{\sum_{i=1}^{n} W_i} - \frac{\sum_{i=1}^{n} (1 - W_i) Y_i^{obs}}{\sum_{i=1}^{n} (1 - W_i)}$$

WARNING

- However, these two quantities are NOT the same!!!
- Using only the observed outcomes and the treatment indicators can be very dangerous, and can give very misleading results

Assignment Mechanism

- KEY QUESTION: How is it determined which units get which treatments?
- The assignment mechanism refers to how units are assigned to treatments
- Understanding the assignment mechanism is a crucial part of causal inference

Surgery vs Drug: Truth

- Surgery (W_i = 1) vs Drug (W_i = 0)
- Y = measure of success (health)
- Causal effect: $Y_i(1) – Y_i(0)$

Patient	$Y_i(1)$	$Y_i(0)$	$Y_i(1) – Y_i(0)$
Patient #1	7	1	6
Patient #2	5	6	-1
Patient #3	5	1	4
Patient #4	7	8	-1
Average	6	4	2

- Average causal effect: 2

Perfect Doctor

- Consider a perfect doctor:

Patient	$Y_i(1)$	$Y_i(0)$	$Y_i(1) – Y_i(0)$	W_i	Y_i^{obs}
Patient #1	7	1	6	1	7
Patient #2	5	6	-1	0	6
Patient #3	5	1	4	1	5
Patient #4	7	8	-1	0	8
Average	6	4	2		

- Using only Y^{obs} and W would give an estimated average causal effect of (7 + 5)/2 - (6 + 8)/2 = -1

Perfect Doctor

- We cannot look only at the observed values under different treatments
- In order to draw valid causal inferences, we must consider why some units received one treatment rather than another.

Lord's Paradox

- "A large university is interested in investigating the effects on the students of the diet provided in the university dining halls and any sex differences in these effects. Various types of data are gathered. In particular, the weight of each student at the time of his arrival in September and his weight the following June are recorded."(Lord, 1967, p. 304)
- Estimand (object of interest): difference between the causal effect of the diet for males and the causal effect for females

(hypothetical) Results

- Male distribution of weight did not change
  - (some males may have gained weight and some may have lost weight, but they balance out and the average and distribution remains the same)
- Female distribution of weight did not change
- Females are lighter than males

Lord's Paradox

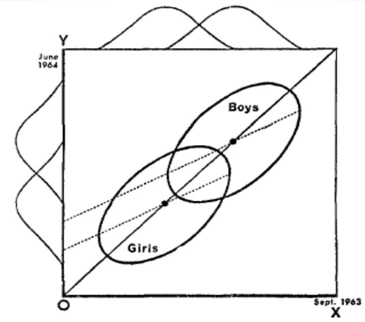

Fig. 1. Hypothetical scatterplots showing initial and final weight for boys and for girls.

Two Competing Views

- Statistician 1:
  - Looks at weight change by gender
  - Notes no weight change for either group
  - Concludes that the diet effects males and females equally (that is, not at all)
- Statistician 2:
  - Takes initial weight into account
  - Notes that for any given initial weight, males gained more than females on the diet
  - Concludes that the diet causes males to gain more than females

The Regression Perspective

- Statistician 1:
  - Fits the regression model:

    weight change = $\beta_0 + \beta_1$ sex
  - Notes that β_1 is not significant
- Statistician 2:
  - "controls for" initial weight
  - Fits the regression model:

    weight change = $\beta_0 + \beta_1$ sex + β_2 initial weight
  - Notes that β_1 is positive and significant

Lord's Paradox

- Both statisticians have used valid statistical reasoning and analyses
- Yet, they arrived at completely different causal conclusions!
- Using only before/after comparisons, or using regression models can yield misleading causal results
- Actually, if estimating causal effects, both statisticians are wrong…

Rubin Causal Model

- Units: students
- Treatment: university diet
- Outcome: weight change from Sep to June
- SUTVA?
- Potential outcomes:
  - $Y_i(1)$: weight change under university diet
  - $Y_i(0)$: weight change not under university diet
- Causal effect: $Y_i(1) - Y_i(0)$
- Assignment mechanism?

Rubin Causal Model

- Potential outcomes and the Rubin Causal Model help to keep everything clear, and can help to avoid mistakes in causal inference that other commonly used methods are prone to