Internal Validity and Experimental Design

February 18

Outline

1. Share examples from TESS
2. Talk about how to write experimental designs
3. Internal validity
    - Why randomization works
    - Threats to internal validity

Examples from TESS

- What was the most interesting study you found on TESS?
- What was the topic (outcome concept and research question)?
- What was the design?

Protocol

- Writing up experimental designs
- Once we know our hypotheses, the experimental conditions are easy

Examples

What experimental conditions do we need?

1. Individuals exposed to expert endorsements are more likely to support a policy than when exposed to partisan endorsements.

Examples

What experimental conditions do we need?

1. Individuals exposed to expert endorsements are more likely to support a policy than when exposed to partisan endorsements.

2. Providing conditional cash transfers to women in rural Uganda is more effective at increasing their childrens' educational attainment than either microfinance loans to start businesses or unconditional grants of cash or goods (e.g., food).

Examples

What experimental conditions do we need?

1. Individuals exposed to expert endorsements are more likely to support a policy than when exposed to partisan endorsements.

2. Providing conditional cash transfers to women in rural Uganda is more effective at increasing their childrens' educational attainment than either microfinance loans to start businesses or unconditional grants of cash or goods (e.g., food).

3. The effect of a public health intervention is more effective for native speakers of Danish than non-native speakers of Danish.

Protocol

- Writing up experimental designs
- Once we know our hypotheses, the experimental conditions are easy
- But, there are still lots of decisions to make!

Random Assignment

- Why do we need it?
- Why can't we just compare $t_2 - t_1$ changes?

Random Assignment

- Breaks the selection process

- This has benefits:

  1. Balances covariates between groups
  2. Balances potential outcomes between groups
  3. No confounding

"Perfect Doctor"

True potential outcomes (unobservable in reality)

Unit	Y(0)	Y(1)
1	13	14
2	6	0
3	4	1
4	5	2
5	6	3
6	6	1
7	8	10
8	8	9
Mean	7	5

"Perfect Doctor"

Observational data with strong selection bias

Unit	Y(0)	Y(1)
1	?	14
2	6	?
3	4	?
4	5	?
5	6	?
6	6	?
7	?	10
8	?	9
Mean	5.4	11

Random assignment

- We have to do it
- But how do we randomize?

Definition

> The observation of units after, and possibly before, a randomly assigned intervention in a controlled setting, which tests one or more precise causal expectations.

Definition

> The observation of units after, and possibly before, a randomly assigned intervention in a controlled setting, which tests one or more precise causal expectations.

Design consideration

- Single-factor versus crossed designs
- Control groups
- Pretest measurement
- Crossover (within-subjects) designs
- Follow-up outcome measurement

Threats to validity

- "Falsificationist" strategy *
- No experiment is perfect

* SCC pp.41-42

Threats to internal validity

1. Ambiguous temporal precedence
2. Selection
3. History
4. Maturation
5. Regression
6. Attrition
7. Testing (exposure to test affects subsequent scores; measurement has an effect)
8. Instrumentation

SSC Table 2.4 (p.55)

Internal validity in experiments

Which of these threats is solved by randomized experimentation?

- All but attrition, testing, and instrumentation
- Assuming well-designed protocol, then we just have to deal with testing and attrition

Threats to statistical conclusion validity

1. Power
2. Statistical assumption violations
3. Fishing
4. Measurement error
5. Restriction of range
6. Protocol violations
7. Loss of control
8. Unit heterogeneity (on DV)
9. Statistical artefacts

SSC Table 2.2 (p.45)

Measurement and operationalization

- Content validity: does it include everything it is supposed to measure
- Construct validity: does the instrument actually measure the particular dimension of interest
- Predictive validity: does it predict what it is supposed to
- Face validity: does it make sense

Pretesting

The best way to figure out whether a measure or a treatment serves its intended purpose is to pretest it before implementing the full study.

Ethics of randomization

- Equipoise*
- Treatment preferences**
- When to randomize

Freedman; SSC pp.272-273 * SSC pp.273-274

Next week

- Continue our discussion of designs
- Talk about experimental analysis
- No example study for next week (because there is a lot to cover in class)
- Do not read the Splawa-Neyman text
    - It's just there in case you're really interested in the statistics of experiments