Study Design 2

Observational Studies and Experiments

Grinnell College

Review – Biased Samples

Biased samples are not representative of the population

Voluntary Response Sample

- people select themselves to participate
- usually people with strong opinions respond to surveys

Convenience Sample

- people are chosen in a non-random way
  - poll at a specific location
- name comes from the fact that this is 'easy' to do

Biased Samples

Sampling Bias

- **undercoverage**: certain groups may not be represented in samples
- **sampling frame** (list of who we can sample from) may be missing some of the population

Non-response Bias

- some people can't be surveyed or choose not to participate

Response Bias

- we don't always get accurate info from people
- question wording
- not wanting to provide truthful answers

Types of Studies

Experiments are studies that involve manipulating a *treatment* that each participant receives

- typically treatments are *randomly assigned* to participants
- we then measure participants' *response* to the treatment
- the treatment is the explanatory variable

Observational Studies are studies that do not involve manipulating the explanatory variables for participants.

- we are simply "observing" what is going on without intervening
- nearly all surveys are observational studies

Types of Studies → Conclusions

The type of study affects the conclusions we can draw from data.

Experiments

- Good experiments with *random assignment* of treatments can establish a cause-and-effect relationship
- We aim to control (or balance) other variables to stop them from affecting results
  - randomization of treatments is the best way to accomplish this

Observational Studies

- No random assignment → generally cannot establish cause-and-effect relationships
- we are limited to talking about associations only

Types of Studies – Practice

For each of the following scenarios, talk with those around you to determine if this is an observational study or experiment (and why!)

Scenario 1: Researchers randomly determined many similar plots of farm land to receive one of Fertilizer A, Fertilizer B, or No Fertilizer. They then measured plant growth for soy beans in those farm plots.

Scenario 2: Researchers identified a few different farms already using Fertilizer A, Fertilizer B, and No Fertilizer and compared plant growth for soy beans in those farm plots.

General Comments

Observational Studies (again)
Studies where we simply 'observe' what is happening.

- we **can** see associations between variables
- we **cannot** make causal connections between variables

Further classification (of both types of studies): Sometimes the *time* element of how data is collected may be of interest

- **Prospective study**: pick our sample and collect data as things happen
- **Retrospective study**: look at historical data or past records

Experiments

Experiments (again)

- study where researchers manipulate explanatory variables to see effect
- explanatory variable values are randomly assigned to each participant

Experimental Units (EUs)

- the observations (= cases) within our experiment
- who/what the experiment is actually performed on
- experimental unit = subject = participant

Experiments

Factors

- another name for the explanatory variables in the experiment
- each experiment must have at least one factor
- these are the variables being manipulated for each subject
- **levels** of a factor = values used for that factor
  - each factor needs at least two levels (or we are not comparing anything)

Treatments

- a specific combination of factor levels assigned to an EU
- one factor $\rightarrow$ levels = treatments

Response Variables

- EUs' response to a treatment
- can have multiple response variables
- can be quantitative or categorical (blood pressure vs 'did blood pressure improve')

Experiment Vocab Practice

Scenario: A chemist is investigating how different conditions affect the rate of a chemical reaction.

Setup: Identical reaction mixtures are prepared and placed into separate flasks. The chemist varies two conditions: the temperature at which the reaction occurs (20°C, 40°C, or 60°C) and the type of catalyst used (Catalyst A, or Catalyst B). For each combination of temperature and catalyst, several trials are conducted. For each trial, the chemist records the time, in seconds, required for the reaction to complete.

Why is this an experiment, not observational study?

Detail the following parts of the experiment:

- EUs:
- factor 1 (and levels):
- factor 2 (and levels):
- treatments:
- response variable (+units):

Blinding

Blinding is concealing treatment assignment from participants, researchers, or both

Purpose: Blinding reduces bias in measuring and responding to outcomes, not just treatment delivery

- Placebo effect (thinking you are getting a good treatment can affect your outcome)
- Researchers or clinicians may unintentionally treat participants differently based on knowledge of treatment

Single Blind: One of either participants or researchers do not know the treatment being received (usually participants)
Double-Blind: Both do not know

Designing Experiments

There is much more that goes into designing good experiments. Below are a few commonly used principles. Unfortunately many different principles have the same / similar names.

Control – 2 types

- comparing a treatment to a control group that did not receive the treatment
  - treatment vs placebo (or vs 'Gold Standard')
  - not used in all experiments
- reduce the influence of other variables through good design

Replication – 2 types

- In an experiment: having multiple *replicates* (cases = EUs) for each treatment
- *repeatability*: Across many experiments: being able to repeat an experiment and get similar results
  - same results with a new sample?
  - do you trust results of a study that we can't confirm on a different sample?

Designing Experiments – Randomness (2 types)

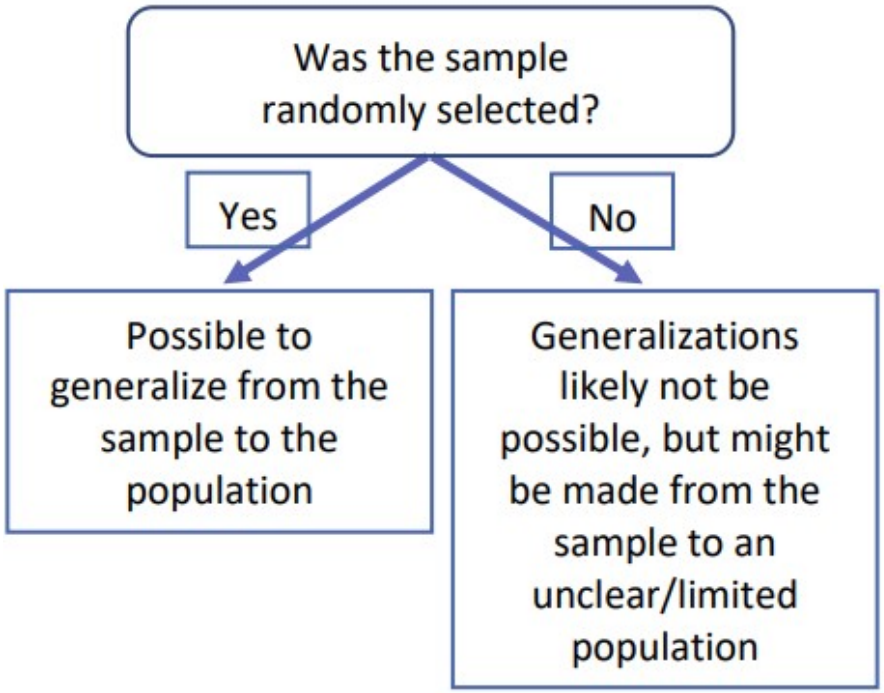
Random Sampling: picking our sample at random from the population

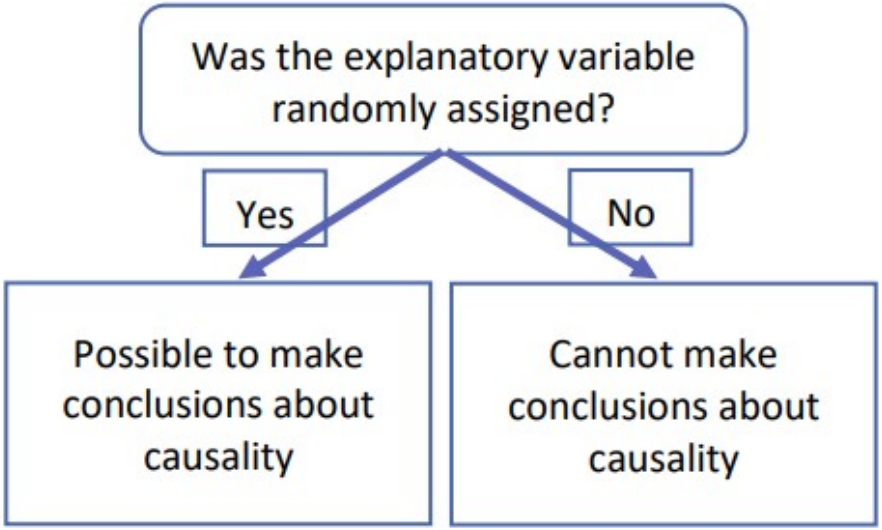
- goal: get sample that is very similar to our population (representative)
- allows us to generalize our conclusions about the sample to the entire population

Random Assignment (randomization): randomly assigning each EU to receive one of the treatments

- allows us to make cause-and-effect claims
- (tends to) balance out effect of confounding variables between both groups so that they don't affect our results
- (tends to) results in treatment groups being similar in every regard except for which treatment they receive

Experiments – Randomness





source: Modified Figure 1.3 from Locke et. al. textbook, Dr. Ziegler (ISU)

Designing Experiments

Confounding Variable: a variable that is associated with both the explanatory variable and the response, making it difficult to determine causation

- makes it impossible to tell if explanatory variables actually **caused** changes in response

How does Randomization fix this issue?
Random assignment helps balance a (potentially) confounding variable between the groups $\rightarrow$ reduces impact of that variable on our results

- Does it always work? No. We'll explore this in the lab.

Wrapping up – Reflection

What is the difference between an Experiment and an Observational Study?

Why do Experiments let us make cause-and-effect conclusions?

What are some ways we can avoid biases when getting our sample?